مشاوره آماری

خاویر کابررا، اندرو مک‌دوگال

ترجمهٔ

محمدقاسم وحیدی اصل، الهه وحیدی اصل

مرکز نشر دانشگاهی

مرکز نشر دانشگاهی
۱۴۵۲

ریاضی، آمار، و رایانه
۱۷۰

مرکز نشر دانشگاهی

Statistical Consulting
Javier Cabrera, Andrew McDougall
Springer, 2010

مشاوره آماری

تألیف خاویر کابررا، اندرو مک‌دوگال

ترجمهٔ محمدقاسم وحیدی اصل، الهه وحیدی اصل

ویراستهٔ فرزانه صفوی‌منش
نسخه‌پرداز: احمد میهن‌جامه
طراح جلد: علیرضا دریائی
حروفچین و صفحه‌آرا: منیژه دیارمند
مرکز نشر دانشگاهی
چاپ اول ۱۳۹۲
تعداد ۵۰۰
لیتوگرافی: دریا
چاپ و صحافی: سامان
۱۸۰۰۰ تومان

نشانی فروشگاه مرکزی: خیابان انقلاب، روبه‌روی سینما سپیده، پاساژ خبیری، تلفن: ۶۶۴۰۸۸۹۱ ، ۶۶۴۱۰۶۸۶
نشانی فروشگاه و نمایشگاه دائمی: خیابان دکتر بهشتی، خیابان خالد اسلامبولی، نبش خیابان دهم، شماره ۵۰ تلفن: ۸۸۷۲۵۹۵۴

فروش اینترنتی: www.bookiup.ir

حق چاپ برای مرکز نشر دانشگاهی محفوظ است

فهرست‌نویسی پیش از انتشار کتابخانهٔ ملی جمهوری اسلامی ایران

سرشناسه: کابررا، خاویر، ۱۹۵۵ـ م. Cabrera, Javier
عنوان و نام پدیدآور: مشاوره آماری/ خاویر کابررا، اندرو مک‌دوگال؛ مترجم محمدقاسم وحیدی‌اصل، الهه وحیدی‌اصل.
مشخصات نشر: تهران: مرکز نشر دانشگاهی، ۱۳۹۲.
مشخصات ظاهری: ده، ۲۳۲ ص.: مصور، جدول، نمودار.
فروست: مرکز نشر دانشگاهی؛ ۱۴۵۲. ریاضی، آمار، و رایانه؛ ۱۷۰.
شابک: 978-964-01-1452-0
وضعیت فهرست‌نویسی: فیپا
یادداشت: عنوان اصلی: Statistical consulting, c2002.
یادداشت: کتابنامه.
موضوع: علوم اجتماعی ـ آمار
موضوع: مشاوران آماری
شناسه افزوده: مک‌دوگال، اندرو، ۱۹۶۱ـ م.
شناسه افزوده: McDougall, Andrew
شناسه افزوده: وحیدی‌اصل، محمدقاسم، ۱۳۲۶ـ ، مترجم
شناسه افزوده: وحیدی‌اصل، الهه، ۱۳۶۲ـ ، مترجم
رده‌بندی کنگره: ۱۳۹۲ ۵م۷۲/۲۹ HA
رده‌بندی دیویی: ۰۰۱/۴۲۲
شماره کتابشناسی ملی: ۳۲۵۳۱۲۸

بسم‌الله‌الرحمن‌الرحیم

فهرست

عنوان	صفحه
پیشگفتار	۱
I روش‌شناسی مشاورهٔ آماری	۳
۱ آشنایی با مشاورهٔ آماری	۵
۱.۱ تاریخ روش علمی	۷
۲.۱ پیدایش آمار	۱۱
۳.۱ مروری بر مشاورهٔ آماری	۱۴
۴.۱ محیط‌های مشاورهٔ آماری	۱۷
۱.۴.۱ داروسازی	۱۷
۲.۴.۱ مخابرات	۲۰
۳.۴.۱ بازرگانی	۲۲

ن		صفحه
۴.۴.۱	حکومت	۲۷
۵.۴.۱	دانشگاه	۲۹
	ارتباط	۳۲
۱.۲	تعامل شفاهی	۳۳
۲.۲	دیگر جنبه‌های تعامل شفاهی	۴۰
۳.۲	چگونه گزارش بنویسیم؟	۴۴
۴.۲	دستورالعمل‌های اساسی برای نوشتن	۴۹
۵.۲	چگونه ارائه‌کردهای مؤثر انجام دهیم؟	۵۳
۶.۲	اهمیت نمودارهای باکیفیت	۵۸
	جنبه‌های روش‌شناختی	۶۸
۱.۳	گردآوری داده‌ها	۶۸
۲.۳	پردازش داده‌ها	۷۷
۳.۳	موضوعات آماری	۸۲
۴.۳	روش‌های آماری مورد استفاده در مشاوره	۹۰
۵.۳	روش‌های استاندارد	۹۱
۶.۳	روش‌های عمومی	۱۴۰
۷.۳	طرح آزمایش‌ها	۱۵۰
۸.۳	نرم‌افزارهای آماری	۱۵۸
	ک پروژهٔ مشاوره از الف تا ی	۱۶۷
۱.۰	اطلاعات پیشین	۱۶۷
۲.۰	مسائل مالی	۱۶۹
۳.۰	جلسهٔ I: نخستین دیدار	۱۷۰
۴.۰	تهیهٔ مستندات	۱۸۵
۵.۰	تحلیل پروژه	۱۸۷
۶.۰	نشست II: ارائهٔ نتایج	

عنوان	صفحه
۷.۴ گزارش نهایی	۲۰۳
۸.۴ پس‌نوشت	۲۱۴
IIمطالعه‌های موردی	۲۱۹
۵ آشنایی با مطالعه‌های موردی	۲۲۱
۱.۵ فرمت ارائهٔ مطالعه‌های موردی	۲۲۲
۲.۵ جزئیات مطالعهٔ موردی	۲۲۳
۶ مطالعه‌های موردی گروه I	۲۲۸
۱.۶ تبعیض در ارتقای شغلی	۲۲۹
۱.۱.۶ ادعایی مبنی بر شواهد آماری	۲۲۹
۲.۱.۶ تحلیل جدول پیشایندی	۲۳۱
۳.۱.۶ تعبیر معنی‌داری	۲۳۲
۴.۱.۶ تحلیل مقدماتی	۲۳۴
۵.۱.۶ خلاصه	۲۳۵
۲.۶ مورد مرسولهٔ گمشده	۲۳۶
۱.۲.۶ تحلیل بررسی نمونه‌ای	۲۳۶
۲.۲.۶ طرحوارهٔ بررسی	۲۳۷
۳.۲.۶ تحلیل مقدماتی	۲۳۹
۴.۲.۶ خلاصه	۲۴۲
۳.۶ دستگاهی برای کاهش آلاینده‌های موتور	۲۴۳
۱.۳.۶ آزمون‌کردن ادعای یک تولیدگر	۲۴۳
۲.۳.۶ آزمون‌های t	۲۴۵
۳.۳.۶ تحلیل واریانس	۲۴۵
۴.۳.۶ تحلیل مقدماتی	۲۴۵
۵.۳.۶ خلاصه	۲۴۸
۴.۶ روان‌شناسی معکوس	۲۴۹

عنوان			صفحه
	۱.۴.۶	یک آزمایش مشاهده‌ای	۲۴۹
	۲.۴.۶	آمار: بازاری با هر نوع کالا	۲۵۲
	۳.۴.۶	داده‌های ترتیبی	۲۵۵
	۴.۴.۶	تحلیل مقدماتی	۲۵۶
	۵.۴.۶	خلاصه	۲۶۳
۷ مطالعه‌های موردی از گروه II			۲۶۴
۱.۷ بررسی دم‌تکاندن			۲۶۵
	۱.۱.۷	آمار پیش‌بالینی	۲۶۵
	۲.۱.۷	رگرسیون لوژستیک	۲۶۶
	۳.۱.۷	رگرسیون لوژستیک چندمتغیره	۲۶۸
	۴.۱.۷	تحلیل مقدماتی	۲۶۹
	۵.۱.۷	خلاصه	۲۷۰
۲.۷ آیا طعم خوبی دارد؟			۲۷۰
	۱.۲.۷	طرح‌های عاملی در علم تغذیه	۲۷۱
	۲.۲.۷	روش‌شناسی رویهٔ پاسخ	۲۷۳
	۳.۲.۷	طرح‌های مرتبهٔ اول و دوم	۲۷۵
	۴.۲.۷	ملاحظات عملی	۲۷۹
	۵.۲.۷	تحلیل مقدماتی	۲۸۰
	۶.۲.۷	خلاصه	۲۸۶
۳.۷ هزینه‌کرد در شهرداری‌های نیویورک			۲۸۷
	۱.۳.۷	مدلبندی رگرسیونی	۲۸۷
	۲.۳.۷	تحلیل رگرسیونی	۲۸۹
	۳.۳.۷	تحلیل مقدماتی	۲۹۴
	۴.۳.۷	خلاصه	۲۹۷
۴.۷ اندازه‌گیری زمان کیفیت			۲۹۸
	۱.۴.۷	تحلیل سری‌های زمانی	۲۹۸

عنوان			صفحه
	۲.۴.۷	مدل‌های ARIMA	۳۰۰
	۳.۴.۷	تحلیل مقدماتی	۳۰۳
	۴.۴.۷	خلاصه	۳۰۵
۸	**مطالعه‌های موردی از گروه III**		۳۰۶
۱.۸	داستان دو دزدک		۳۰۷
	۱.۱.۸	تحلیل واریانس با اثرهای آمیخته	۳۰۷
	۲.۱.۸	تحلیل مدل آمیخته	۳۱۲
	۳.۱.۸	تحلیل مقدماتی	۳۱۴
	۴.۱.۸	خلاصه	۳۱۸
۲.۸	ردیابی مواد منفجرهٔ پلاستیکی		۳۱۹
	۱.۲.۸	تشخیص الگو	۳۱۹
	۲.۲.۸	مروری کوتاه بر تحلیل مسیزی	۳۲۰
	۳.۲.۸	تحلیل مقدماتی	۳۲۳
	۴.۲.۸	خلاصه	۳۲۴
۳.۸	یک مطالعهٔ تحقیق بازار		۳۲۵
	۱.۳.۸	مروری کوتاه بر تحلیل مؤلفه‌های اصلی	۳۲۶
	۲.۳.۸	مروری کوتاه بر تحلیل عاملی	۳۲۸
	۳.۳.۸	تحلیل مقدماتی	۳۳۰
	۴.۳.۸	خلاصه	۳۳۳
۴.۸	فروش تجهیزات ارتوپدی		۳۳۳
	۱.۴.۸	کاربردهای داده‌کاوی در تحقیقات بازار	۳۳۴
	۲.۴.۸	داده‌کاوی	۳۳۵
	۳.۴.۸	مروری کوتاه بر تحلیل خوشه‌ای	۳۳۹
	۴.۴.۸	افرازبندی و رده‌بندی بازگشتی و درخت‌های رگرسیونی	۳۴۰
	۵.۴.۸	تحلیل مقدماتی	۳۴۰
	۶.۴.۸	خلاصه	۳۴۴

عنوان | صفحه

۹ چند مطالعهٔ موردی دیگر ۳۴۶
۱.۹ بهبود آموزش ۳۴۷
۲.۹ نمونه‌گیری تصادفی؟ ۳۴۸
۳.۹ چپ یا راست؟ ۳۵۰
۴.۹ برانگیختن حس در اسب‌ها ۳۵۱
۵.۹ سرخ‌موی بلندقد ۳۵۳
۶.۹ انتقام بنتلی ۳۵۵
۷.۹ هر چه می‌خواهید می‌پوشید؟ ۳۵۶
۸.۹ مطالعه‌ای در مورد ایدز ۳۵۷

پیوست الف ۳۵۹
الف.۱ مراجع ۳۵۹
الف.۲ مجموعه داده‌های مطالعه‌های موردی قسمت II ۳۶۴
الف.۳ درس مشاورهٔ آماری ۳۶۴
الف.۳.۱ توصیف درس ۳۶۴
الف.۳.۲ فهرست مباحث برحسب هفته ۳۶۷
الف.۳.۱ فهرست مراجع ۳۷۰

پیوست ب ۳۷۱
ب.۱ SAS ۳۷۱
ب.۱.۱ تجهیزات SAS ۳۷۱
ب.۱.۲ جزئیات در مرحلهٔ DATA ۳۷۴
ب.۱.۳ شیوه‌های SAS ۳۷۶
ب.۱.۴ جزئیات بیشتری از SAS ۳۸۲
ب.۲ S-PLUS ۳۸۴
ب.۲.۱ مقدمات S-PLUS ۳۸۵
ب.۲.۲ تنظیمات S-PLUS ۳۸۷
ب.۲.۳ فرمان‌های مقدماتی S-PLUS ۳۸۹

عنوان	صفحه
ب.۴.۲ استفادهٔ کارآمد از S-PLUS	۳۸۹
ب.۵.۲ شیوه‌های آماری S-PLUS	۳۹۷
ب.۶.۲ فرهنگ اصطلاحات S-PLUS	۳۹۹
پیوست ج	۴۰۵
ج.۱ توزیع‌های تک‌متغیره	۴۰۶
ج.۲ توزیع‌های چندمتغیره	۴۰۹
ج.۳ آزمونهای استاندارد	۴۱۳
ج.۴ اندازهٔ نمونه	۴۱۷
مراجع	۴۲۰
نمایه	۴۳۰

پیشگفتار

درسی در مشاورهٔ آماری که چند سالی به‌طور منظم تدریس کرده‌ایم، انگیزهٔ نوشتن این کتاب شد. در آن درس تمرکز ما، بر مطالب زیر است:

- درک «فرایند» مشاورهٔ آماری
- پروردن مهارت‌های ارتباطی
- کسب تجربه از راه مطالعه‌های موردی.

در واقع امر، هیچ جایگزینی برای مشاورهٔ مستقیم با مشتری وجود ندارد و برای آنکه این تعامل کارایی داشته باشد، مهارت‌های ارتباطی خوب از ضروریات اصلی‌اند. متأسفانه این جنبه از آموزش آماردانان اغلب مورد غفلت قرار می‌گیرد و دانش‌آموختگان آمار راهی جز این ندارند که این مهارت‌ها را ضمن کار بیاموزند. یکی از هدف‌های این کتاب پرداختن به این نیاز است.

مشاورهٔ آماری در حیطه‌های گوناگونی پیش می‌آید و برای دست یازیدن به مسئله‌های آماری دنیای واقع، آماردان نیاز به آن دارد که علاقهٔ زیادی به روش علمی داشته باشد. تاریخ خود بهترین مثال‌ها را برای پروردن این علاقه در اختیار می‌گذارد و بنابراین موضوع را با سفر تاریخی کوتاهی در فصل ۱ آغاز می‌کنیم. البته هیچ عصری مثل عصر حاضر نیست و ما در باقی این فصل برخی از محیط‌هایی را که مشاورهٔ آماری نقش عمده دارد، توصیف می‌کنیم.

بحثی مفصل از مهارت‌های ارتباطی گفتاری و نوشتاری که در یک محیط مشاوره مورد نیاز است، در فصل ۲ ارائه شده است. دیداری کردن یکی از نیرومندترین ابزارهای موجود در مشاورهٔ آماری است و به اهمیت نمودارهای با کیفیت و ارائه‌کردهای تأثیرگذار نیز در فصل ۲ پرداخته شده است.

روش‌شناسی آماری در اصل ۳ ارائه شده است. در توصیف روش شناسی‌یی که مشاور می‌تواند به‌کار گیرد، بر اهمیت دخالت دادن درک مشتری از هدف و تعبیر و تفسیر شیوهٔ آماری تأکید شده است. در حالی که فرض می‌شود که خواننده مهارت‌های فنی لازم را برای به‌کار بستن روش‌های آماری استفاده‌شده در این کتاب داراست، منظور ما ارائهٔ روش‌شناسی آماری از منظر مشتری است و نه آماردان.

قدم بعدی به هم پیوستن همهٔ مؤلفه‌های فرایند مشاوره است و در فصل ۴ مثالی از یک پروژهٔ مشاورهٔ آماری به‌طور کامل ارائه شده است. این فصل، قسمت I کتاب را به پایان می‌برد.

طیف گسترده‌ای از مطالعه‌های موردی با درجه‌های دشواری متفاوت در قسمت IIی کتاب ارائه شده است که به خواننده در درک و فهم تنوع پروژه‌هایی که ممکن است در مشاورهٔ آماری مطرح باشند، یاری می‌رساند. مطالعه‌های موردی مطابق با سطح دشواری فنی مرتبط با روش‌های آماری مورد نیاز برای تحلیل، گروهبندی شده‌اند. این کار تا حدی میزان انتظارات دانشجو و خواننده را از سطح روش‌شناس آماری و پیچیدگی تحلیلی مطرح در یک مطالعهٔ موردی خاص مشخص می‌کند. فصل ۹ چندین تمرین مطالعهٔ موردی را ارائه می‌کند که در آنجا پروژه‌ها و مسائل مورد نظر یکسان‌اند، اما تحلیل به خواننده واگذار شده است.

پیوست‌ها، اطلاعاتی دربارهٔ منابع (پیوست الف)، جزئیاتی دربارهٔ بسته‌های نرم‌افزاری SAS و S-PLUS (پیوست ب)، و گردایه‌ای از جدول‌های مرجع سودمند (پیوست ج) در اختیار می‌گذارند. این کتاب، کتاب درسی مناسبی برای درسی دربارهٔ مشاورهٔ آماری در سطح تحصیلات تکمیلی است و شرح برنامه‌ای درسی برای مدرسان در پیوست الف تهیه شده است. همهٔ مجموعه داده‌ها از نشانی وب‌سایت داده شده در پیوست الف قابل دسترسی‌اند.[۱]

هاویر گابررا — اندرو مک‌دوگال
پیسکاتاوی، نیوجرزی[۲] — مانت کلیر، نیوجرزی[۳]

---

۱. چندسطر مربوط به سپاسگزاری نویسندگان از کسانی که در پدیدآوردن کتاب سهم داشته‌اند ترجمه نشده‌اند.
2. Piscataway 3. Montclair

قسمت I

روش‌شناسی مشاورهٔ آماری

۱

آشنایی با مشاورهٔ آماری

«آمار چیست؟» پرسشی است که اغلب در آغاز کتاب‌های درسی آماری به آن پرداخته می‌شود و شاید تا حدی توضیحی برای این باشد که چرا به مشاوران آماری نیاز داریم: همه به فصل بعدی می‌روند تا در آنجا با نظریهٔ احتمال از آن‌ها استقبال شود. پس در این صورت مشاورهٔ آماری چیست؟ مشاورهٔ آماری در نهایت دربارهٔ ارتباط است، زیرا پاسخ نخستین پرسش این است که آمار دربارهٔ هر مسئله‌ای، در هر حوزه‌ای از علم، است که دربرگیرندهٔ تغییرات تصادفی باشد. البته، این بدان معنا نیست که هر زمینه‌ای از علم و هر مسئله‌ای لزوماً در این مقوله می‌گنجد، اما زندگی همین است[۱] ــ می‌توانیم آن را بپذیریم.

یک مشاور آماری، یک مسئله‌حل‌کن است و هدف این کتاب پرداختن به مهارت‌های مورد نیاز برای یک مشاور کارآمد است. مشاور آماری باید برای درک کامل این مهارت‌ها علاقه‌ای شدید به علم و به‌ویژه هنر اکتشاف علمی داشته باشد. در این فصل، اهمیت مشاورهٔ آماری را در زمینهٔ

۱. c'est la vie ـ اصطلاحی فرانسوی است ـ م.

فرایند علمی توضیح می‌دهیم. اگر این مطالب جاذبهٔ خاصی برای شما داشته باشد، در این صورت این کتاب برای شماست! با یک مثال شروع می‌کنیم.

مثال ۱.۱ آیا جامعه متقارن است؟

در یک جلسهٔ مشاوره، سؤالی دربارهٔ ماشین پرکنندهٔ شیشه‌های قرص مطرح شد. ماهیت بحث این بود که آیا می‌توان توزیعی متقارن را از لحاظ نظری برای مدلبندی تعداد قرص‌های هر شیشه موجه دانست یا خیر. پیشنهاد شد که تقارن بر این اساس قابل توجیه است که یک خطِ نظری در دهانهٔ بطری وجود داشته باشد و بتوان بطری را به یک اندازه تا بالا یا پایین این خط پر کرد.

موضوع در اینجا این نیست که استدلال نظری لزوماً درست یا نادرست است، بلکه صرفاً آن است که «راه حل» به‌عنوان یک «نظریه» باقی می‌ماند، در حالی که در واقع آنچه که به آن نیاز داریم مشاهدهٔ مستقیمی از واقعیت‌هاست. چرا امتحان نکنیم؟ مقداری داده گردآوری و بافت‌نگاری رسم کنید.

مثال ۲.۱ ابعاد ماه

آیا ماه وقتی در افق است، نسبت به زمانی که درست بالای سر ماست، بزرگتر به نظر می‌رسد؟

چرا دوباره امتحان نکنیم؟ البته، مقاومت در برابر وسوسهٔ شفافیت‌بخشی بر نظریه‌ای جدید که تفاوت ظاهری در اندازه را توضیح می‌دهد ممکن است دشوار باشد. برای افراد علم‌گرا شاید انکسار نور از طریق حجم بزرگتری از جوّ زمین (مانند نوعی تأثیر عدسی جوی) ممکن است به‌عنوان یک توضیح جذابیت داشته باشد، اما چرا باز هم آن را به کمک مشاهدهٔ مستقیم امتحان نکنیم؟ (از یک خط‌کش استفاده کنید) پاسخ صحیح این است که آنچه می‌بینیم یک خطای بصری است و در هر دو مورد اندازهٔ ماه یکسان است.

نکتهٔ قابل توجه در مثال‌های بالا این است که بهترین منبع دانشی که داریم داده‌هاست و این همزیستی بین داده‌ها و نظریه است که دانش علمی در اختیار ما قرار می‌دهد. این مثال‌ها نشان می‌دهند که نظریه به‌تنهایی همیشه ما را به نتایج معتبر سوق نمی‌دهد و بسیار مهم است که به داده‌ها توجه کنیم.

همین ترکیب مطالعهٔ مشاهداتی و دانش نظری **روش علمی** نامیده می‌شود و این شیوه‌ای است که در اکتشاف علمی دنبال می‌شود. قابل ذکر است که این سرمشق تا اواخر قرن هفدهم مطرح نشد. در بخش بعدی مثال‌هایی از تاریخ علم در نظر می‌گیریم که در آن‌ها آمار و روش علمی نقش بزرگی ایفا کرده‌اند.

۱.۱ تاریخ روش علمی

خود را در قرون وسطی، مثلاً اروپای قرن پانزدهم، تجسم کنید، حدوداً ۲۰۰ سال پس از اینکه متون یونان باستان در اسپانیای تحت تسلط اعراب به لاتین ترجمه شده بودند. این متون ترجمه‌شده به‌آرامی در دیگر نقاط اروپای غربی نفوذ کردند و دانش علمی و فلسفی اساسی آن زمان را به نمایش گذاشتند. در این جهان، دانش بنیادی بر پایهٔ سنن یونان باستان بنا شده بود. نقش دانشمندان یا فیلسوفان (در آن زمان اصطلاحاتی مترادف بودند) مطالعهٔ متون یونانی قدیم بود که بدنهٔ کامل دانش را تشکیل می‌داد ــ چیزی نبود که به آن اضافه شود.

از سوی دیگر همه چیز در سکون کامل نبود. پیشرفت‌های فنی از جانب صنعتگرانی که دانش علمی عملی را با قدم‌هایی آهسته پیش می‌بردند در دست بود. با این حال، دانش کاربردی که به‌وسیلهٔ صنعتگران به‌دست می‌آمد همواره یک فرایند به‌هم‌افزاینده نبود، زیرا آن‌ها یک سنت شفاهی را دنبال می‌کردند و اسناد و نشریات مکتوب را نگهداری نمی‌کردند. این بود که دانش و مهارت‌ها گاهی از دست می‌رفتند.

باید خاطرنشان کنیم که علم یونانی تنها پایهٔ دانش علمی نبود. در واقع، بسیاری از فرهنگ‌های باستانی دستگاه‌های نوشتار و تقویم‌ها را به وجود آوردند و بعضی از آن‌ها به پیشرفت‌هایی فنی مانند ریخته‌گری برنز نایل شدند. به‌ویژه، علم و فلسفهٔ چینی به‌طور مستمر از زمان کنفسیوس (۵۵۱ـ۴۷۹ ق.م) در نشو و نما بودند. با این حال، به چالش گرفته‌شدن دانش سنتی یونانی بود که به فرایند اکتشاف علمی مبتنی بر روش علمی منجر شد.

نخستین چالش مشهور بر علم یونانی از جانب گالیله (۱۵۶۴ـ۱۶۴۲) با آزمایش معروف برج پیزا بود که نشان داد که سرعت یک جسم در حال سقوط به وزن آن بستگی ندارد، اما جدی‌ترین و بنیادی‌ترین آن‌ها به چالش گرفته‌شدن کیهان‌شناسی ارسطو از سوی کپلر (۱۵۷۱ـ۱۶۳۰) و گالیله بود. به‌طور مشخص، این کیهان‌شناسی، جهان را به‌عنوان یک کرهٔ آسمانی توصیف می‌کرد که زمین در مرکز آن بود و ماه و خورشید و پنج سیارهٔ شناخته‌شده با فاصله‌هایی متناسب با تارهای چنگ در اطراف زمین در حال گردش‌اند. خارج کرهٔ سماوی آتش بود که می‌شد آن را از درون سوراخ‌های کوچکی در کره مشاهده کرد. امروزه عموماً به این‌ها ستاره اطلاق می‌شود.

اگرچه پیش از گالیله و کپلر چالش‌هایی بر اعتقادات ارسطویی مطرح شده بود، جستار علمی پیوند بسیار نزدیکی با ارزش‌های ماورای طبیعی و باورهای الوهی آن عصر داشت. این باور سنتی که افلاک و زمین تفاوت اساسی دارند، زمانی که گالیله با تلسکوب[1] اقمار مشتری را مشاهده کرد

۱. تلسکوب برای مقاصد نظامی و بازرگانی به‌دست تجار هلندی به وجود آمده بود، گرچه گالیله مدعی بود که نمونه‌ای اولیه از آن را پدید آورده است.

و دریافت که این‌ها حول زمین نمی‌چرخند، فروباشید. در اینجا شاید لازم باشد که توجه کنیم که علم ممکن است چیز خطرناکی نیز باشد ــ این کشف تقریباً به بهای جان گالیله به دست دستگاه تفتیش عقاید تمام شد! حدوداً در همین زمان، کپلر مشاهداتی از مدارهای سیاره‌ها انجام داده بود که عقیدهٔ عمومی مدل خورشید مرکزی را که کپرنیک (۱۴۷۳ـ۱۵۴۳) مطرح کرده بود تأیید می‌کرد: خورشید در مرکز منظومهٔ شمسی قرار دارد و زمین و دیگر سیاره‌ها حول خورشید در گردش‌اند.

این دو کشف مشاهده‌ای کوبنده آشکارا خطابودن مدل یونانی را تصدیق کرد و به‌سرعت به گسترش انفجاری مطالعه‌های علمی منجر شد. در زیر خلاصهٔ مختصری از تفکر علمی و برخی بدفهمی‌های رایج که به‌وسیلهٔ داده‌های آزمایشی در طی قرن‌های شانزدهم و هفدهم به چالش کشیده شدند آورده می‌شود.

تفکر علمی پیش از رنسانس

- سنت علمی یونانیان باستان،
- تمامی دانش از طریق کتاب‌ها فراگرفته می‌شد،
- نظریهٔ ریاضی برای درک طبیعت، با آزمایشگری بی‌ارتباط بود.
  مدارهای سیاره‌ها مدوّر بودند زیرا دایره‌ها از نظر ریاضی کامل‌اند. کره‌ها نیز کامل‌اند.
  اجسام با سرعتی متناسب با وزنشان سقوط می‌کنند.
  گرمای خورشید با گرمای آتش متفاوت است.
- دانش کاربردی که صنعتگران کسب می‌کردند از دست می‌رفت، زیرا آن‌ها اسناد و نوشته‌های منظمی نگهداری نمی‌کردند.

بسط روش علمی

کشفیات مهمی همچون صنعت چاپ، باروت، و قطب‌نمای مغناطیسی اهمیت دانش نوین را به نمایش گذاشتند. گیلبرت[1] (۱۵۴۰ـ۱۶۰۳) در زمینهٔ مغناطیس کارهای گسترده‌ای انجام داد و نخستین کسی بود که دریافت که دانش «کتاب‌ها» برای توضیح بسیاری از پدیده‌های فیزیکی کفایت نمی‌کند:

«سنت جدید آدمیانی که در جستجوی دانش نه در کتاب‌ها بلکه در خود چیزها هستند.»

کپلر، گالیله، و گیلبرت در زمرهٔ نخستین کسانی بودند که از روش علمی استفاده کردند و آن را بدون درکی «رسمی» از این سرمشق به‌کار بردند. نخستین افرادی که روش علمی را رسمیت بخشیدند و استلزامات آن را واقعاً درک کردند سر فرانسیس بیکن[1] (۱۵۶۱ـ۱۶۲۶) و رنه دکارت[2] (۱۵۹۶ـ۱۶۵۰) بودند. نقل قول زیر از بیکن تأکیدی بر عمق درک او از روش علمی است:

«اتحاد تفسیر نظری و مهار عملی طبیعت سلسله‌ای از اختراعات را به وجود خواهد آورد که ممکن است تا حدی بر نیازها و تیره‌روزی‌های بشری فائق آیند.»

دکارت رویکردی برای حل مسائل بر مبنای روش علمی به‌صورت زیر فرمولبندی کرد:

تحلیل شیوه‌ای که چیزهای عملی کشف می‌شوند.

ترکیب شیوه‌ای نظری که بتوان چیزهای مشابهی را از اصول اولیه استنتاج کرد.

وی همچنین نکتهٔ مهم زیر را یادآوری کرده است:

«هیچ چیز بی‌حاصل‌تر از آن نیست که خود را با اعداد مطلق و ارقام خیالی سرگرم کنیم.»

در ارتباط با اهداف این کتاب گفتهٔ بالا را چنین تعبیر می‌کنیم **داده‌های بی‌زمینه معنای چندانی ندارند.**

اکتشاف علمی

مشاهدات کپلر از موقعیت‌های سیاره‌ها او را بر آن داشت که نتیجه بگیرد خورشید در مرکز منظومهٔ شمسی قرار دارد. با این حال، برای پیش‌بردن دانش علمی نیازمند به توضیح فرایندی که مشاهدات از آن نشأت گرفته بودند نیز هستیم. این امر رهنمون کپلر در فرمولبندی سه قانون مشهور حرکت سیاره‌ای وی شد. ارزش ذکر دارد که کپلر، در آغاز، دومین قانون خود را این‌گونه فرمولبندی کرد که «دوره‌های تناوب با مربع فاصله‌ها از خورشید متناسب‌اند.» او پس از تحلیل بیشتر داده‌های موجود، دومین قانون خود را پس از ۱۰ سال مجدداً به این صورت فرمولبندی کرد که «مربع دورهٔ تناوب با مکعب فاصله از خورشید متناسب است». این تکرار بین مشاهده و فرمولبندی مدل، پس‌زمینهٔ روش علمی است:

1. Sir Francis Bacon 2. Rene Descartes

مطرح‌کردن نظریه مدلی نظری برای توضیح‌دادن پدیدهٔ مورد نظر مطرح می‌شود.

آزمایش مبتنی بر نظریه آزمایشی به‌منظور آزمون‌کردن موجه‌بودن مدل طراحی می‌شود.

نظریه به کمک داده‌های آزمایشی اصلاح می‌شود نظریه و داده‌ها مقایسه می‌شوند و اصلاحاتی انجام می‌گیرد تا دلیل تفاوت‌ها آشکار شود.

آزمایش مبتنی بر نظریهٔ اصلاح‌شده نظریهٔ بازنگری‌شده از طریق آزمایش مورد آزمون قرار می‌گیرد.

این فرایند، به‌طور بالقوه نامحدود یا تا زمانی که برای مقاصد مطالعه به تقریبی کافی دست یابیم، ادامه می‌یابد. این تکرار داده و نظریه همان است که منجر به پیشرفت‌هایی در علم و درک می‌شود. ما برخی از اکتشافات علمی مهم را در طی سه قرن اخیر فهرست می‌کنیم. برای اطلاعات بیشتر خواننده را به میسون (۱۹۶۲) ارجاع می‌دهیم.

قرن ۱۸

مکانیک نیوتونی. گرانش عمومی. حسابان.
مسئلهٔ دریانوردی. نجوم و زمان‌سنج.
زیست‌شناسی: تکوین موجودات زنده. گردش خون. نظریهٔ سلولی.
شیمی: نظریهٔ شیمی جدید لاووازیه. ترکیب مواد.

قرن ۱۹

داروین و نظریهٔ تکامل.
تکوین زمین‌شناسی.
نظریهٔ اتمی ماده.
نظریهٔ موجی نور.
برق و مغناطیس.
ریاضیات: آمار و هندسه.
مهندسی و علم.

قرن ۲۰ ــ فهرستی بسیار کوتاه

زیست‌شناسی نوین.
نظریهٔ نسبیت.
مکانیک کوانتومی.
تحلیل داده‌ها و آمار.
علوم اجتماعی.

۲.۱ پیدایش آمار

آمار نوین، آن‌گونه که به‌روشنی از نقش بنیادی تحلیل داده‌های تجربی در اکتشافات علمی مهم آشکار است، ریشه در روش علمی دارد. از سوی دیگر، سابقه‌ای طولانی از گردآوری داده‌های ناشی از مطالعهٔ آمارهای حیاتی جمعیت مانند سرشماری‌های روم باستان و آمارهای اقتصادی دولتی در دست است. دو مثال جالب از مورد اخیر، جدول‌های مرگ و میر[1] ساخته‌شده به دست جان گرانت[2] در ۱۶۶۲ و شاخص سالانهٔ قیمت گندم فروخته‌شده در بازارهای اروپایی، ۱۵۰۰ـ۱۸۶۹، ساخته‌شده به دست بِوریج[3] در ۱۹۲۱[4] است. با این حال، پایهٔ تفکر آماری تا اواخر قرن نوزدهم و اوایل قرن بیستم، یعنی مدت‌ها پس از اینکه نظریهٔ احتمال به‌خوبی تثبیت شده بود، به ظهور نپیوست.

ارزش ذکر دارد که ظهور علاقه به مسائل مرتبط با احتمال در ابتدا عمدتاً از پیدایش بیمه[5] نشأت گرفت. آنچه واقعاً اخگری برای علاقه به احتمال بود تقاضاهای اشراف‌زادگانی بود که به بازی‌های شانسی مبادرت می‌کردند! در واقع، شوالیه دومره[6] بود که به پاسکال (۱۶۲۳ـ۱۶۶۲) دربارهٔ مسئلهٔ موسوم به مسئلهٔ امتیازها مراجعه کرد که از موارد اولیهٔ مشاورهٔ آماری است. پاسکال پس از آن‌ها مکاتبات خود را با فرما (۱۶۰۱ـ۱۶۶۵) آغاز کرد و آن‌ها هر دو برخی از اصول احتمال را پایه‌ریزی کردند.

با وجود آنکه نظریهٔ ریاضی احتمال رونق یافت، حرکت به سوی تفکر آماری به واقع مقارن با اواخر قرن نوزدهم آغاز شد. صورت‌های اولیهٔ قضیهٔ حدی مرکزی و قانون اعداد بزرگ[7] پیش‌تر ظاهر شده بودند و کتله[8] (۱۷۹۶ـ۱۸۷۴) مفهوم جامعه‌های نرمال را در ۱۸۳۵ معرفی کرد. تفکر رسمی آماری از سوی گالتون[9] (۱۸۲۲ـ۱۹۱۱) و پیرسون[10] (۱۸۵۷ـ۱۹۳۶) مطرح شد که آن‌ها مفاهیم رگرسیون و همبستگی در ارتباط با تحلیل روابط بین اندازه‌گیری‌ها را، زمانی که یک مؤلفهٔ تصادفی ذاتی (خطا) حضور داشته باشد، به وجود آوردند. هر دوی آن‌ها نقشی اساسی در به رسمیت شناخته‌شدن اهمیت تحلیل رسمی آماری از سوی جامعهٔ علمی برای تثبیت اعتبار نتایج علمی به‌عنوان بخش مهمی از روش علمی داشتند. تا آن زمان دانشمندان تحلیل داده‌ها را به شیوه‌ای غیر رسمی به‌کار برده بودند.

---

1. tables of mortality 2. John Graunt 3. Beveridge

۴. شاخص بِوریج در اندرسن (۱۹۷۱) بازسازی شده است.

۵. جدول‌های سالانه‌ای بوسیلهٔ یوهان دویت(Johan De Witt) (۱۶۲۵ـ۱۶۷۲) در سال ۱۶۷۱ ساخته شد.

6. Chevalier de Mere

۷. گرچه قانون اعداد بزرگ عموماً به پواسون (۱۷۸۱ـ۱۸۴۰) نسبت داده می‌شود، میزان سهم پواسون در این قانون مورد مناقشه است. استیگلر (۱۹۸۶، صفحات ۱۸۲ـ۱۸۵) بحث جالبی را دربارهٔ این موضوع مطرح کرده است.

8. Quetelet 9. Galton 10. Pearson

مثالی برای تشریح مطلب، کشف قانون $PV = K$ توسط بویل[1] است که فشار هوا را با حجم مرتبط می‌سازد[2]. بویل (۱۶۲۷ـ۱۶۹۱) آزمایش‌های متعددی با تلمبهٔ هوا انجام داد بدون اینکه از شیوهٔ استنباط رسمی بهره گیرد و بر مبنای این شواهد تجربی رابطهٔ ریاضی این قانون را فرمولبندی کرد (ای کاش همهٔ مدل‌ها این‌قدر ساده بودند!) روابطی از این نوع نسبتاً به‌سرعت کشف شدند و یک رویکرد پیچیده‌تر تحلیل داده‌ای و استنباطی مورد نیاز بود تا روابط شامل خطای تصادفی مورد بررسی قرار گیرند.

گالتون و پی‌یرسون در معرفی روش‌هایی برای تحلیل این نوع از روابط توفیق داشتند. هرچند پی‌یرسون (۱۹۰۰) در اصل آمارهٔ خی‌دوی خود را برای جدول‌های پیشایندی $I \times J$ با $IJ - 1$ درجه‌آزادی معرفی کرد. تا سال ۱۹۲۲ کسی جز فیشر (۱۸۹۰ـ۱۹۶۲) این خطا را خاطرنشان نکرد که این امر موجب رنجیدگی پی‌یرسون شد و به این ادعای فیشر که او درجه‌های آزادی را نادرست بیان کرده است با عصبانیت واکنش نشان داد. شرح مختصری از این جدال در اگرستی (۱۹۹۰؛ صفحهٔ ۶۹) آمده است و جزئیات بیشتر را می‌توان در زندگینامهٔ فیشر نوشتهٔ باکس (۱۹۷۸) یافت.

کشف توزیع‌های t به‌وسیلهٔ گوست[3] (۱۸۷۶ـ۱۹۳۷) منظر مهم دیگری از استنباط آماری را نمایان ساخت: آزمایشگری به‌صرفه از لحاظ هزینه. در این زمان گوست در شرکت مخمرسازی گینس[4] مسئول تخمیر بود و از آزمون t برای شناسایی بهترین گونهٔ جو استفاده کرد. در واکنش به این کار، گینس بی‌درنگ موجودی غله را کلاً خریداری کرد. گوست اجازهٔ انتشار همهٔ اکتشافات خود را بدون ذکر نام به‌دست آورد. در نتیجه، او از نام مستعار «Student» استفاده کرد و توزیع نمونه‌ای آمارهٔ t در استیودنت (۱۹۰۸) منتشر شد.

فیشر ضمن کار طی سالیان دراز روی آزمایش‌های میدانی کشاورزی در راتمستد[5] انگلستان مطمئناً یکی از پیشگامان استفاده از روش‌های آماری در طراحی آزمایش‌ها بود. وی در آنجا شیوهٔ تحلیل واریانس را در تحلیل‌های آماری به وجود آورد و برای نخستین بار به‌کار برد. این کار به کتاب تأثیرگذار وی با عنوان روش‌های آماری برای پژوهشگران[6] (فیشر ۱۹۲۵) منجر شد. فیشر ناخودآگاه رسم استفاده از ۵٪ به‌عنوان سطح مخاطرهٔ قراردادی برای همهٔ آزمون‌های شناخته‌شده را باب کرد! با این حال، نوشته‌های او به سازمان‌دهی آمار به‌عنوان یک حوزهٔ مستقل علمی کمک کرد و پایهٔ استنباط آماری را در مطالعهٔ مسائل عملی مستحکم کرد. در ایالات متحده نیز آمار نقش

1. Boyle

۲. این قانون بعداً نسبت به دما اصلاح و قانون بویل‌ماریوت (Mariotte) از آن نتیجه شد. $PV/T=K$

3. Gosset 4. Guinness brewing company 5. Rothamsted 6. Statistical methods for research workers

مهمی در آزمایش‌های کشاورزی ایفا کرد و فعالیت آزمایشگاه آمار[1] در دانشگاه ایالتی آیووا[2,3] به‌وضوح در کتاب روش‌های آماری از اسنه‌دکور (۱۹۳۷) مشهود است.

در بخش صنعت، روش‌شناسی آماری نقشی بنیادی در پیدایش و فلسفهٔ کنترل کیفیت در ساخت و تولید انبوه ایفا کرد. به دنبال معرفی اصول مدیریت علمی که بنابه ضرورت طی انقلاب صنعتی به وقوع پیوست، دو رویکرد مشخص برای کنترل کیفیت توسط دانشمندان در آزمایشگاه‌های بل AT&T به وجود آمدند. شوهارت (۱۹۳۱) نمودار کنترل[4] را در ۱۹۲۴ برای بررسی تغییرپذیری فرایند برخط[5] به‌منظور بهبودبخشیدن به کیفیت محصول معرفی کرد، در حالی که در ۱۹۲۸ روش‌های نمونه‌گیری برای پذیرش[6] توسط داج و رومیگ (۱۹۵۹) به‌عنوان سیستمی برای ارزیابی کیفیت کل یک محصول بر مبنای بازرسی نمونه‌ای اتخاذ مطرح شدند.

طرح‌های نمونه‌گیری برای پذیرش طی چندین سال بر شیوه‌های کنترل کیفیت در ایالات متحده غلبه داشتند، اما به‌طور عمده صرفاً برای «نظارت» بر خط تولید به‌کار می‌رفتند. تا پیش از موفقیت خیره‌کنندهٔ صنعت خودروی ژاپن روش‌های آماری برای بهبود کیفیت در میان صنعتگران[7] سرتاسر آمریکای شمالی و اروپا پذیرش عام نیافتند. فلسفهٔ دمینگ[8] و روش‌های تاگوچی[9] که الهام‌بخش صنعت ژاپن شد تأثیر شگرفی بر نحوهٔ نگرش بر کیفیت از منظر تولید داشته است. مفاهیم جدیدتر مدیریت کیفیت کلی و تضمین کیفیت بر نیاز به تدارک یک «بسته»ی کیفی کامل تأکید دارند زیرا روش‌های آماری به‌تنهایی قادر به حل مسائل کیفی نخواهند بود.

از نظر عموم، شاید مهمترین سهم آمار در زمینهٔ سلامت بوده باشد. نخستین آزمایش‌های بیماری فلج اطفال، روش‌شناسی آماری را به‌عنوان مؤثرترین ابزار برای ارزیابی اثربخشی واکسیناسیون به ثبت رساند. رویکرد آماری در برقراری رابطهٔ بین سیگارکشیدن و سرطان ریه[10] نیز نقش بنیادی

1. Statistical Laboratory 2. Iowa State University

۳. این دانشگاه در آن زمان به کالج ایالتی آیووا مشهور بود. ما این مطلب را تا حدی از آن رو ذکر کرده‌ایم که حکایت جالبی دربارهٔ دیدار (و تنها دیدار) فیشر از دانشگاه ایالتی آیووا در دست است. متأسفانه دیدار او همزمان با دوره‌ای بود که دمای هوا به ۱۰۰ درجهٔ فارنهایت (۳۷٫۸ درجهٔ سانتی‌گراد) رسید. این وضعیت برای فیشر چنان تحمل‌ناپذیر بود که تمام مدت دیدار را کف اتاق خود دراز کشید و عزم کرد که هرگز به ایمز (Ames) در آیووا باز نگردد.

4. control chart 5. online 6. *acceptance Sampling*

۷. در حالی که صنعت خودروی ایالات متحده به‌شدت از رقابت بیگانه متضرر شد، صنعت شیمی ایالات متحده شروع به پذیرش روش‌های کنترل فرایند آماری در دههٔ ۱۹۵۰ کرده و یکی از رقابتی‌ترین موردها در جهان باقی مانده است.

8. Deming 9. Taguchi

۱۰. شرکت گروه لیگت (The Liggett Group) نخستین شرکت دخانیات بود که اعتیادآور بودن سیگار و این مسئله را ←

داشته است. جدول زیر برخی از موارد پیدایش و کاربردهای آمار را که در بالا بحث کرده‌ایم به‌طور خلاصه بیان می‌کند.

- پیشرفت‌های اولیه در گردآوری و پردازش داده‌هادر قرن هفدهم: جدول‌های مرگ و میر درانگلستان ۱۶۶۲
- توزیع نرمال. کتله، ۱۸۳۵
- پواسون، قانون اعداد بزرگ
- گالتون ۱۸۷۹. مفاهیم همبستگی و رگرسیون را معرفی کرد. وی توزیع کتله‌گاوسی را به‌کار برد
- پی‌یرسون. بنیادهای آمار نوین را پایه‌ریزی کرد
- فیشر. طرح آزمایش‌ها، اطمینان ۹۵٪
- شوهارت، نمودارهای کنترل کیفیت، حدهای ۳ـسیگما
- آزمایش‌های بیماری فلج اطفال
- ارتباط بین سرطان ریه و سیگارکشیدن

۳.۱ مروری بر مشاورهٔ آماری

پس، مشاورهٔ آماری دقیقاً چیست؟ چرا مورد نیاز است؟ در حالی که مطالعه‌های جمعیت‌شناختی و علمی دو رکن آمار نوین‌اند، این موضوع تا حدی به این معنا منحصر به فرد است که ما همواره داده‌های فرد دیگری را تحلیل می‌کنیم. آمار مطمئناً از پیشرفت‌های سریع در فناوری کامپیوتر بهره برده است و نرم‌افزارهای آماری هم‌اکنون در دسترس مخاطبان بیشتری است. با این حال، پیچیدگی پرسش‌های تحت بررسی در بسیاری از زمینه‌های علمی نیازمند حداقل شکلی از خبرگی در ارتباط با تحلیل داده‌هاست. مسئله این است که اغلب پژوهشگران فرصت کسب این دانش تخصصی همراه با تجربهٔ عملی برای به‌کار بستن مناسب آن را ندارند. لازم است کسی را به این کار بگماریم که فرایند علمی را درک می‌کند و مهارت‌های کتی برای ایفای این نقش مهم را داراست: مشاور آماری. شرحی از مطالبی که در این کتاب گنجانده شده

که موجب مشکلات سلامتی از جمله سرطان ریه می‌شود پذیرفت. یک اطلاعیه برای عموم در ۲۱ مارس ۱۹۹۷ منتشر شد.

فصل ۱ روش علمی. محیط‌های مشاوره
فصل ۲ اهمیت ارتباط شفاهی و نوشتاری
فصل ۳ روش‌شناسی آماری و استفاده از منابع
فصل ۴ همه چیز یک‌جا. یک مثال
قسمت II کسب تجربه: مطالعه‌های موردی

در این کتاب، تأکید ما بر فراهم‌کردن برخی از مهارت‌ها برای آماردانان است که برای یک مشاورِ مؤثربودن به آن‌ها نیاز دارند. هدف بخش اول این فصل توضیح‌دادن اهمیت درک فرایند علمی بوده است. در بخش بعدی محیط‌های گوناگونی را در نظر می‌گیریم که در آن‌ها مشاورهٔ آماری نقش مهمی ایفا می‌کند. برای مثال، شیوه‌ای را که توسط صنعت داروسازی برای ایجاد دارویی جدید به‌کار می‌رود توصیف می‌کنیم؛ از بررسی‌های مربوط به مرحلهٔ ترکیب تا پس از عرضه در بازار. در توصیف این محیط‌های مختلف، بر مسئولیت‌های آماردانان نیز تمرکز می‌کنیم. در کل، می‌توانیم انواع محیط‌های مشاوره‌ای زیر را از هم متمایز کنیم.

۱. آماردان به‌عنوان بخشی از یک گروه تحقیقاتی روی یک پروژهٔ اصلی کار می‌کند و دانش آماری خود را با گروه به مشارکت می‌گذارد. اهداف عبارت‌اند از پیگیری پروژهٔ تحقیقاتی در قالب یک «گروه». این بدان معنی است که آماردان باید زمینهٔ تحقیق را با عمق کافی فراگیرد تا بتواند به‌طور مؤثری در پروژه سهیم باشد. این یک نمونهٔ بارز در صنعت مخابرات و داروسازی است.

۲. آماردان، مشاور تعداد زیادی پروژه، اغلب در مقیاسی کوچک، است یا جایی که آماردان نقش مشورت‌دهندهٔ مهمی را ایفا می‌کند. نمونهٔ بارز آن محیط‌های دانشگاهی یا شرکت‌های مشاورهٔ بازرگانی است.

در هر یک از محیط‌ها، نقش مشاور آماری نیازمند ترکیبی از چهار نوع دانش است:

علمی قادر است به‌سرعت حقایق اساسی زمینهٔ موضوع را که پروژهٔ تحت مطالعه در قلمرو آن قرار می‌گیرد یاد بگیرد.

آماری لازم است دانش و تجربه داشته باشد تا بتواند شیوه‌های مناسب را برای طراحی و تحلیل به‌کار گیرد.

محاسباتی باید سواد کامپیوتری داشته باشد. مهارت کافی برای انجام دادن تحلیل داده‌های واقعی با استفاده از نرم‌افزار را داشته باشد.

ارتباطی آخرین و مهمترین مورد، مشاور آماری باید مهارت‌های ارتباط نوشتاری و گفتاری خوب را در خود به وجود آورد.

در بسیاری از موارد، کار یک آماردان در نهایت انتقال حقایق و ایده‌های ساده است؛ بنابراین، آنچه واقعاً حائز اهمیت است مهارت‌های خوب ارتباطی است ــ هم گفتاری و هم نوشتاری. به خاطر بیاوریم که ارتباط گفتاری شامل توانایی گوش دادن است. همچنین مهارت شمرده بیان کردن ایده‌ها به‌طور شفاهی. مهارت‌های نوشتاری مهم‌اند زیرا محقق باید «گزارش» حاوی نتایج تحلیل را بفهمد و بر آن باور داشته باشد. فصل ۲ به بحث مفصلی از جنبه‌های مختلف ارتباط نوشتاری و گفتاری اختصاص یافته است. خلاصه‌ای از زمینه‌های آماری و محاسباتی مورد نیاز در فصل ۳ در ادامه می‌آید.

البته آزمون واقعی زمانی پیش می‌آید که مشاور آماری لازم است به‌طور واقعی با یک مشتری تعامل داشته باشد. برای اینکه خواننده را یاری کنیم تا درک واقعی‌تری از آنچه مشاورهٔ آماری مستلزم آن است به‌دست آورد، کل فرایند مشاوره در مورد یک پروژهٔ خاص را در فصل ۴ بازسازی کرده‌ایم. از آنجا که بسیاری از موضوعات مطرح در آن فصل به دیگر پروژه‌های آماری عمومیت داده خواهد شد، مجموعه‌ای از پروژه‌ها با سطوح متفاوت دشواری آماری را به شکل مطالعه‌های موردی و تمرین‌های مطالعه‌های موردی ارائه کرده‌ایم.

هدف از این مطالعه‌های موردی آن است که خواننده را در معرض حوزه‌ای گسترده از مباحث و روش‌های تحلیل قرار دهیم که ممکن است ضمن اشتغال به مشاورهٔ آماری برای خواننده پیش آید. اگرچه بر تحلیل آماری در این مطالعه‌های موردی تأکید شده است، سعی کرده‌ایم تا آنجا که ممکن است اطلاعات محتوایی را در اختیار خواننده قرار دهیم به‌طوری که خواننده بتواند واقعیت مباحث مرتبط با هر مطالعهٔ موردی را درک کند. مطالعه‌های موردی، قسمت IIای این کتاب را تشکیل می‌دهند.

کتاب‌های متعدد دیگری در رابطه با مشاورهٔ آماری موجودند که ارزش ذکر دارند. هند و اوریت (۱۹۸۷) مجموعه‌ای از مثال‌ها را ارائه کرده‌اند که مشاورهٔ آماری را در عمل نشان می‌دهند. چت‌فیلد (۱۹۹۵) چشم‌اندازی آماری از نقش یک مشاور به‌عنوان یک مسئله‌حل‌کن فراهم کرده است. در (۲۰۰۰) و کتاب کلاسیک نوشتهٔ بن و ژان (۱۹۸۲) نقش مشاور آماری را به‌عنوان رابطی مؤثر بررسی کرده‌اند. برخی از مقاله‌های مربوط به موضوع مشاورهٔ آماری عبارت‌اند از: تووبدی (۱۹۹۸)، ویل‌فورد و همکاران (۱۹۹۵)، هند (۱۹۹۴)، مین‌دانلد (۱۹۹۲)، جوینر (۱۹۸۲)، گرین‌فیلد (۱۹۷۹)، مارکوارت (۱۹۷۹)، لوری (۱۹۵۸)، و کیمبال (۱۹۵۷).

۴.۱ محیط‌های مشاورهٔ آماری

کارورزی آمار در محیط‌های گوناگونی اتفاق می‌افتد. در باقیماندهٔ این فصل برخی محیط‌های خاص را در نظر می‌گیریم و نقش آماردان را مشخص می‌کنیم. از صنایع داروسازی و مخابرات آغاز می‌کنیم که اولین و دومین بخش‌های (قانونی[1]) درآمدی اقتصاد امریکا هستند. آمارداتان نقش مهمی در هر دوی این محیط‌ها بازی می‌کنند. انواع دیگری از محیط‌ها، شرکت‌های مشاورهٔ بازرگانی و نهادهای دولتی‌اند که کار نظارت را انجام می‌دهند و کارکردهای مختلف جامعه را نظم می‌دهند. سرانجام دانشگاه‌ها اغلب اعمال مشاوره‌ای کوچک یا بزرگ مرتبط با زیست‌آمار، کاربردهای کشاورزی و صنعتی، و نیز پژوهش در حالت عام را انجام می‌دهند.

در بحث دربارهٔ هر یک از این محیط‌ها، نقش آماردان مشاور و تخصص‌هایی را که یک آماردان می‌تواند به این محیط‌ها بیاورد مشخص می‌کنیم. همچنین، با ارائهٔ جزئیاتی دربارهٔ تعامل با دیگر پژوهشگران و نوع دانش پایه‌ای لازم برای آنکه یک آماردان عضوی مؤثر از یک تیم پژوهشی باشد، دربارهٔ معنای عنصری از تیم بودن به بحث می‌پردازیم.

۱.۴.۱ داروسازی

فرایند فراوری داروها

فراوری داروها فرایندی پیچیده و طولانی است که برای یک تک‌دارو با قیمتی که ممکن است سر به ده‌ها میلیون دلار آمریکا بزند شاید ۷ تا ۱۵ سال طول کشد. اگرچه توصیف زیر به فراوری دارو در آمریکا مربوط می‌شود، باید خاطرنشان کنیم که این یک صنعت به‌شدت رقابتی در سطح بین‌المللی است. در واقع، یکی از انتقادها از مدیریت داروی فدرال (FDA، نهادی که بازار داروی آمریکا را تنظیم می‌کند) مدت‌زمانی است که طول می‌کشد تا تأییدیه‌ای برای داروهای جدید به‌دست آید.

برای شرکت‌های اصلی بین‌المللی داروسازی، سه بخش مهم از فرایند فراوری دارو وجود دارد که پشت سر هم اتفاق می‌افتند. خواننده را برای ملاحظهٔ توصیف مشروح فرایند فراوری دارو و آمار زیست‌دارویی به پیس (۱۹۸۸) ارجاع می‌دهیم.

- کشف و تصمیم
- بررسی‌های پیش‌بالینی
- بررسی‌های بالینی

۱. داروهای غیر قانونی به‌عنوان بزرگترین بخش‌های درآمدی تخمین زده می‌شوند.

کشف و تصمیم

این فرایند با کشف یک ترکیب جدید یا کاربرد بالقوهٔ جدیدی از یک ترکیب موجود آغاز می‌شود. در مورد یک ترکیب جدید، آزمون‌های مقدماتی غیرذیروحی[1] و ذیروحی[2] به‌منظور ارزیابی کنشگری ترکیب به‌کار می‌رود.

ذیروحی آزمون‌کردن ترکیب در اندام‌های زنده (یاخته‌ها).

غیر ذیروحی آزمون‌هایی که در محلول‌های آزمایشگاهی (لولهٔ آزمایشگاهی) انجام می‌شوند.

بر پایهٔ نتایج کافی، سپس این تصمیم گرفته می‌شود که آیا دارو فراوری شود یا نه. اگر تصمیم مثبت باشد، در این صورت فرایند فراوری وارد مرحلهٔ بررسی‌های پیش‌بالینی می‌شود.

بررسی‌های پیش‌بالینی

در مرحلهٔ پیش‌بالینی، سم‌شناسی ترکیب روی حیوانات بررسی می‌شود. فرمولبندی مقدماتی فراوری دارو و مطالعه‌های داروشناختی خاص یا جامع در حیوانات نیز در این مرحله صورت می‌گیرد. در پایان این مرحله، شواهد سلامت بالقوه و مؤثربودن دارو از سوی شرکت داروسازی ارزیابی می‌شود. این امر بسیار مهم است، زیرا شواهد باید این اطمینان کافی را در شرکت به وجود آورد که دُز دارو مهلک نخواهد بود و برای انسان قابل تحمل خواهد بود. برای برداشتن گام بعدی، یک شرکت که مرکز آن در ایالات متحده باشد باید پروندهٔ اطلاعیهٔ معافیت داروی جدید تحقیقاتی ادعایی[3] را تشکیل دهد. غرض از این اطلاعیه اجازه‌یافتن شرکت برای انجام‌دادن مطالعه‌های روی موردهای انسانی است. این مرحلهٔ بعدی متضمن چندین نوع آزمایهٔ بالینی است.

مطالعه‌های بالینی

در این وهله، آزمون‌های گسترده‌ای روی حیوانات انجام شده است و از این مطالعه‌ها چنین برداشت شده است که شواهد کافی برای اینکه دارو بر روی موردهای انسانی نافع خواهد بود، وجود دارد. آزمون‌کردن دارو بر روی موردهای انسانی گام بعدی است و سه نوع مطالعه در دست است که پیاپی انجام می‌شوند.

مرحلهٔ I مقصود از مرحلهٔ I آزمایهٔ بالینی، کسب اطلاعات ایمنی آغازین دربارهٔ تأثیر دارو بر انسان‌هاست، از قبیل دامنهٔ دُز قابل قبول و دارو جنبش‌شناسی[4] دارو. این مطالعه‌ها معمولاً

1. in vitro 2. in vivo 3. *Notice of Claimed Investigational New Drug Exemption*
۴. pharmacokinetics: شاخه‌ای از داروشناسی است که به نحوهٔ جذب، گردش و محو دارو در بدن می‌پردازد.م.

با داوطلبان سالم انجام می‌شود. دلیل این امر آن است که تا آنجا که مقدور است از تأثیرات آلودگی دارو و بیماری اجتناب شود. تعداد موردها نوعاً از ۴ تا ۲۰ در هر بررسی متغیر است و جمعاً تا ۱۰۰ مورد در طول آزمایه‌های مرحلهٔ I به‌کار می‌روند.

مرحلهٔ II مرحلهٔ IIای مطالعه‌ها متوجه بیمارانی است که به‌طور بالقوه از داروی جدید منتفع خواهند شد. در اینجا دامنه‌های دُز مؤثر و تأثیرات اولیهٔ دارو بر روی این بیماران ارزیابی می‌شود. معمولاً تا چندصد بیمار بدون پیچیدگی‌های اضافی انتخاب می‌شوند تا از عامل‌های اختلاط جلوگیری شود.

مرحلهٔ III مطالعه‌های مرحلهٔ III ارزیابی‌هایی از سلامت، مؤثربودن، و مقدار دُز بهینه در اختیار ما قرار می‌دهد. این بررسی‌ها بر اساس گروه‌های تیمار و کنترل طراحی می‌شوند که اغلب متشکل از صدها یا حتی هزاران نفر است. بر مبنای نتایج موفقیت‌آمیز حاصل از این بررسی‌ها، شرکت می‌تواند یک NDA یا تقاضانامهٔ داروی جدید[1] تسلیم کند. این تقاضانامه نتایج همهٔ مرحله‌ها را ــ از کشف تا مرحلهٔ IIIــ شامل می‌شود و به‌وسیلهٔ FDA مورد بررسی قرار می‌گیرد.

هیئت بررسی‌کنندهٔ FDA تقاضانامهٔ داروی جدید متشکل از بررسی‌کنندگانی در زمینه‌های زیر است: پزشکی، داروشناسی، زیست‌دارو، شیمی، و آمار. هر بررسی‌کننده ممکن است اطلاعات یا آزمون‌های دیگری را طلب کند که باید پیش از تسلیم یک بررسی کتبی به FDA به‌وسیلهٔ شرکت مرتبط با NDA انجام شود.

فعالیت‌های پسابازار

پس از آنکه NDA به FDA تسلیم شد، فعالیت‌های تبلیغی، تحلیل پژوهشی پیامدها[2]، و مطالعه‌های پیگیرانه برای امتحان‌کردن تأثیرات دارو آغاز می‌شود. این مطالعه‌ها ممکن است به شکل یک آزمایهٔ بالینی باشد و عنوان مطالعه‌های مرحلهٔ IV به آن‌ها اطلاق می‌شود. هدف اصلی از این بررسی‌ها تضمین آن است که همهٔ ادعاهای مطرح‌شده از طرف شرکت دربارهٔ داروی جدید را می‌توان به اصطلاح با «شواهد بالینی» محقق کرد. پیش از آنکه چنان ادعاهایی را بتوان برای مقاصد تبلیغی به‌کار برد نیاز به تأییدیهٔ FDA است. کلیهٔ تأثیرات زیان‌بارگزارش‌شده نیز باید به‌وسیلهٔ شرکت مورد رسیدگی قرار گیرد و در برخی موارد ممکن است که نیاز به جمع‌آوری دارو از بازار باشد.

1. New Drug Application(NDA) 2. outcomes research analysis

مسئولیت‌های آماردان

آماردان به‌وضوح نقش مهمی در سرتاسر فرایند فراوری دارو دارد و ما برخی مسئولیت‌هایی را که در صنعت داروسازی از آماردانان انتظار می‌رود در زیر فهرست می‌کنیم:

- در برنامهٔ فراوری دارو برای بررسی یک دارو شرکت داشته باشند.
- بررسی طرح و موافقت‌نامهٔ فراوری. الگوهای تصادفی‌سازی.
- پاکسازی داده‌ها و فرمت ساختمان پایگاه داده‌ها.
- طرح تحلیل و تدوین برنامه برای تحلیل.
- آماده‌کردن گزارش. تولید جدول‌ها و شکل‌ها.
- یکپارچه‌کردن نتایج مطالعه‌های بالینی، گزارش‌های بی‌ضرربودن و مؤثربودن.
- برقراری ارتباط و دفاعیهٔ NDA برای هیئت بررسی‌کنندهٔ FDA.
- پشتیبانی انتشاراتی و مشورت با دیگر کارکنان شرکت.

۲.۴.۱ مخابرات

در پایان سال ۱۹۸۳، شرکت تلفن و تلگراف امریکایی[1] (که در کلام محبت‌آمیز «مابل[2]» نامیده می‌شود) چندپارچه شد. محصول این چندپارچگی، در بدو امر، غول مخابرات از راه دور AT&T و به وجودآمدن هفت شرکت منطقه‌ای بود که ارائه‌دهندگان خدمات شهری در ایالات متحده، موسوم به «بچه‌بل[3]»ها، بودند.

از آن زمان به بعد تغییرات زیادی صورت گرفته است! به کمک قانون مخابرات که در سال ۱۹۹۶ به تصویب کنگرهٔ ایالات متحده رسید، بچه‌بل‌های رشدکرده حالا با یکدیگر و با AT&T و دیگر شرکت‌های ارائه‌دهندهٔ بین شهری رقابت می‌کنند. بچه‌بل‌ها به صنایع مخابرات بی‌سیم و اینترنتی نیز منشعب شده‌اند و برخی از آن‌ها سرمایه‌گذاری‌های سنگینی در بازارهای بین‌المللی انجام داده‌اند. برعکس، شرکت‌های مخابراتی بین‌المللی نیز وارد بازار ایالات متحده شده‌اند.

بل لَبز[4] اصلی تاریخچه‌ای طولانی در پژوهش‌های آماری دارد و شرکت‌هایی مانند AT&T، لوسنت[5]، و بل‌کور[6]، که از چندپارچه‌شدن مابل به وجود آمده‌اند، قطعاً از این سنت پیروی کرده‌اند.

1. American Telephone & Telegraph 2. Ma Bell 3. Baby Bell 4. Bell Labs(= Bell Laboratories) 5. Lucent 6. Bellcore

این شرکت‌ها تعداد زیادی از آماردانان را، هم در بخش‌های تحقیقات و هم بازرگانی، استخدام می‌کنند. بخش تحقیقات نوعاً به دو بخش تحقیقات پایه و کاربردی تقسیم می‌شود، گرچه تقسیمات سازمانی مانع از این نمی‌شود که افراد در زمینه‌های مختلف بر روی طرح‌های مشترک کار کنند. برای آن‌ها که علاقه به صنعت مخابرات دارند، مقاله‌های کِیتن‌رینگ (۱۹۹۵، ۱۹۹۷) ارزش خواندن دارند.

- پایگاه داده‌های فوق‌العاده بزرگ
- بخش‌های تحقیقات و بازرگانی
- فناوری‌های ارتباطات و اطلاعات

وجه مشخصهٔ منحصر به‌فرد صنعت مخابرات مقادیر بسیار زیاد داده‌های حاصل از بسیاری از پروژه‌هاست که در دسترس آماردانان است. مسائلی در فناوری‌های پردازش اطلاعات، نظیر آزمون‌کردن شبکه، انتقال بسته‌ها و تأخیرها، و خلق نرم‌افزارها، داده‌هایی گیگابایتی تولید می‌کنند که از جهات بسیار باید تحلیل شوند. آماردانان از انعطاف‌پذیری زیادی از لحاظ نحوهٔ تحلیل داده‌ها، مادام که نتایجی تولید شوند، برخوردارند. همچنین فضا برای ایجاد روش‌شناسی جدید، در صورت لزوم، باز است و بسیاری از شغل‌ها به جای اینکه تولیدگرا باشند، پژوهش‌گرا هستند.

آماردانان نوعاً، در قالب تیم‌هایی از مهندسان و متخصصان علوم کامپیوتر، درگیر یک یا چند پروژه‌اند. چنان تیمی مدیر پروژه‌ای دارد که امور اجرایی را می‌چرخاند، یک راهبر فنی دارد که تحقیقات را هدایت می‌کند، و گروهی از اعضای فنی دارد. اغلب پروژه‌ها از راهبران فنی سرچشمه می‌گیرند، اما ممکن است از زمینه‌های دیگری نیز که درخواست کمک دارند نشأت گیرند. امکان آن هست که شخصی مولد پروژه‌ای از خود باشد، اما این اتفاق کمتر رخ می‌دهد. آماردانان نیاز دارند که با مفاهیم و اصطلاحات فناوری ارتباطات آشنا باشند و منظور از شبکه‌ها، سوییچ‌ها، بسته‌ها، و امثال آن‌ها را بفهمند. آماردان نیاز به یادگیری مداوم ابداعات جدید در فناوری خواهد داشت، بنابراین مهم است که به مباحث علمی علاقه‌مند باشد.

بین جوانب فعالیت‌های بازرگانی و تحقیقات وجه تمایزی هست. جانب بازرگانی در سلطهٔ مدیریت تولید، تحقیقات بازار، و بازاریابی است، در حالی که جانب تحقیقاتی عموماً در سلطهٔ فعالیت‌های مرتبط با کاربردهای نرم‌افزاری است که می‌تواند کاملاً متنوع باشد. حقوق و مزایا در مقیاس دیگر صنایع تحقیقاتی نظیر داروسازی است و پاداش‌ها ممکن است بسیار سخاوتمندانه باشند.

۳.۴.۱ بازرگانی

شرکت‌های مشاوره

شرکت‌های کوچک، محیط مشاوره‌ای برای آمارداتان فراهم می‌کنند که دامنهٔ گسترده‌ای از فعالیت‌ها را دربرمی‌گیرد و مشاوران اغلب، به جای صرفاً تمرکز درکاری خاص با نوعی تحلیل، با کلیت فرایند درگیرند. این امر ممکن است تعامل مستقیم با مشتریان از شرکت‌های بیرون، آماده‌کردن گزارش و ارائه‌کرد، و نیز پیشنهادهای بودجه و دیگر وظایف مدیریتی را دربر داشته باشد. هان و هورل (۱۹۹۸)، هورل و همکاران (۱۹۹۳)، و کتاب هیلتون و پارکر (۱۹۹۳) ارزش خواندن دارند.

محرمانگی اطلاعات مشتری، جزء کلیدی مشاورهٔ بازرگانی موفقیت‌آمیز است و برخی زمینه‌های نوعی که شرکت‌های مشاوره‌ای کوچک و مقاطعه‌کاران تحقیقاتی مستقل[1] (IRC) در آن‌ها تخصص دارند عبارت‌اند از:

- تحقیقات بازار
- طرح و تحلیل پیمایشی
- تحلیل مالی
- پروژه‌های داروسازی و مخابراتی
- مدیریت پایگاه داده‌ها
- شاهدان خبره برای دعاوی قضایی

البته، این تنها بخشی از زمینه‌های تحت پوشش مؤسسات کسب و کار کوچک است که در آن‌ها خدمات مشاوره‌ای آماری اثرگذارند و برخی شرکت‌ها ممکن است یا به مقاطعه‌کاری فرعی هم بپردازند، یا اصلاً تخصص آن‌ها در مقاطعه‌کاری‌های فرعی پروژه‌های مشاوره‌ای باشد. به عنوان مثال، یک شرکت ممکن است هیئت علمی یا دانشجویان را از برنامهٔ مشاوره‌ای یک دانشگاه به مقاطعهٔ جزئی بردارد تا در بخش‌های معینی از یک پروژه همکاری کنند.

این نوع محیط‌ها فرصتی برای آمارداتان فراهم می‌آورند تا همهٔ جزئیات کار یک شرکت یا فرایند پروژه‌ها را فراگیرند. همچنین در معرض تعاملات بیرونی با مشتریان قرارگیرند و، به این ترتیب، حوزه‌ای گسترده از فعالیت‌ها را داریم که مشاور می‌تواند در آن‌ها شرکت داشته باشد، در قیاس با یک شرکت بزرگتر که در آن‌ها نقش یک مشاور در یک پروژه ممکن است بیشتر وظیفه‌ای مشخص

1. independent research contractors

باشد. دیگر مزیت‌ها عبارت‌اند از اینکه برنامهٔ کاری مشاور ممکن است انعطاف‌پذیرتر، تنوع کاری بیشتر، و دستمزد ممکن است با سود بیشتری همراه باشد (مثلاً قرارداد مشاور ممکن است شکلی از سهیم‌شدن در سود را به عنوان پاداش دربر داشته باشد).

از سوی دیگر، نوع مهارت‌های مورد نیاز جامع‌ترند. علاوه بر دانش آماری و مهارت‌های کسب وکار، مهارت‌های ارتباطی خوب (شفاهی و کتبی) مورد نیازند. ممکن است لازم باشد که مشاور در تنظیم قرارداد پیشنهادی اولیه کمک کند، بودجه را تدوین کند، و در مورد کیفیت کار توصیه‌های لازم را به مشتریان ارائه یا آن‌ها را متقاعد کند. ارائه‌کردها نقشی مهم در جلب نظر مشتریان و راضی نگه‌داشتن آن‌ها دارد. مشاور ممکن است مسئولیت کامل کردن تحلیل آماری را به‌موقع و در چارچوب بودجه داشته باشد. بدین ترتیب، محدودیت‌های معینی مانند «مهلت‌های انقضا» وجود دارند که اولویت اساسی دارند و باید به آن‌ها عمل شود؛ در غیر این صورت شرکت مشتری خود را از دست خواهد داد. مشهور است که برخی شرکت‌های کوچک مشاوره‌ای برنامه‌های کاری درهم‌برهمی دارند.

مثالی از یک تحلیل پیمایشی بازار

شرحی از مراحل و اقلام کاری که اساس یک برنامهٔ بودجه برای یک پروژهٔ تحقیق پیمایشی بازار را تشکیل می‌دهد در جدول ۱.۱ ارائه شده است. بودجهٔ واقعی برای هر قلم کاری هزینهٔ مشخصی را اختصاص می‌دهد (برای مثال، بر مبنای نرخ ساعتی)، و نیز هزینه‌های بالاسری مشخصی را دربردارد که در اینجا در نظر گرفته نشده‌اند. به‌علاوه، این پروژهٔ تحقیق پیمایشی بازار بر یک طرح پیمایشی مبتنی بود که پیش‌تر برای همان مشتری بکار رفته بود.

در این بررسی خاص، مشتری علاقه‌مند به این ارزیابی بود که مصرف‌کنندگان چه تصوری از کارخانه و محصول از لحاظ هزینه، کیفیت، و ظاهر دارند. اگرچه این نوع بررسی پیش‌تر برای مشتری انجام گرفته بود، ارزش توجه دارد که تعداد قابل توجهی از اقلام کاری در این بودجه مربوط به مدیریت پایگاه داده‌ها، تهیهٔ اسناد، تعاملات مشتریان، و آماده‌سازی گزارش و ارائه‌کرد است. در فصل بعد، این جنبه‌های مهم مشاورهٔ آماری را به‌تفصیل شرح می‌دهیم.

مشاورهای خصوصی

ترجیح می‌دهید که وقتتان در اختیار خودتان باشد؟ برای تغییر ذائقه، کارفرمای خودتان باشید؟ در این صورت به یک مشاور خصوصی تمام‌وقت بدل شوید! اما ... پیش از آنکه از کار روزانهٔ خود استعفا کنید، فهرست کوتاهی از برخی اقلامی که نخست باید از نظر بگذرانید به این شرح است:

- چه سخت‌افزارهای محاسبه‌ای دارید؟ آیا می‌توانید هر چندسال یک‌بار آن‌ها را ارتقا دهید؟

مرحله و اقلام کاری
گردآوری داده‌ها:
پیش‌آزمون و مدیریت گردآوری داده‌ها
مرور نرخ‌های پاسخ
هماهنگی تبادل نمونه/داده
دیدارهایی با مشتری برای رسیدگی به پیشرفت کار
اعتبارسنجی بررسی:
داده‌ها را پالایش، فرمت‌بندی، و برچسب‌دار کنید
نمایهٔ پاسخگو (آمارهای جمعیتی)[1]
جدول‌هایی برای گزارش درست کنید
نمایهٔ برند شرکت / مورد مصرف محصول
امتحان برند و تأثیرات ترتیب در پرسش
خلاصه‌بندی پاسخ‌ها از طریق بخش‌ها و شرکت‌های مورد بررسی
خطاهای استاندارد برای جدول‌ها و نمودارها
نتایج مورد نظر را برای نمایش نموداری در گزارش انتخاب کنید.
ملاقات با مدیر پروژه
تحلیل پیمایشی:
تحلیل مقیاسی، درخت‌های تصمیم برای ارزش/کیفیت
همبستگی‌ها و تحلیل عاملی: بها، ارزش، و پرسش‌ها
همبستگی‌ها و تحلیل عاملی برحسب بخش: مشتری در برابر دیگران
تحلیل عاملی برای پرسش‌های محصول: مشتری در برابر دیگران
رگرسیون لوژستیک: مشتری در برابر رقیبان
ملاقات با مشتری: اعتبارسنجی و نتایج تحلیل را مورد بحث قرار دهید
مدلبندی:
آماده‌کردن داده‌ها برای مدلبندی
مدلبندی برای تعیین رابطهٔ بین متغیرها
مدلبندی معادلهٔ ساختاری
ملاقات با مشتری: مدلبندی را پیش از مکتوب‌کردن مورد بحث قرار دهید
گزارش نتایج:
ملاقات در ارتباط با پروژه: نتایج را مورد بحث قرار دهید و فعالیت‌های مربوط به نوشتن را سازماندهی کنید
درست‌کردن/درج‌کردن جدول‌ها و نمودارها در سند گزارش
ارائه به مشتری: ملاقات برای تمامی روز برنامه‌بندی شود

- چه نرم‌افزارهایی دارید؟ وقتی نسخه‌های جدید وارد بازار می‌شوند، آیا از عهدهٔ ارتقای آن‌ها برمی‌آیید؟
- چه کسی گزارش‌ها، حساب‌ها، و صورت‌حساب‌های شما را آماده می‌کند؟
- از کجاها مشتری خواهید داشت؟

قلم آخری حیاتی‌ترین پرسشی است که باید در نظر داشته باشید، زیرا یک مشاور خصوصی به‌طور بدیهی نیاز به مشتری دارد که از لحاظ مالی دخل و خرج کند. مشکل عمده آن است که پروژه‌های «کوچک» حقیقتاً از لحاظ اقتصادی به صرفه نیستند، زیرا هزینه‌ها بسیار بالا هستند. یعنی، ملاقات با مشتری و تهیه‌کردن اسناد لازم (قرارداد، صورت‌حساب، و گزارش) بیشتر از تحلیل داده‌ها وقت می‌برند. خواننده می‌تواند شرح مفصلی از هزینه‌های بالاسری مربوط به یک پروژهٔ کوچک را در فصل ۴ ملاحظه کند.

بنابراین، مشاوران خصوصی باید قادر باشند که مشتری‌هایی با پروژه‌های «بزرگ» را جلب کنند. معمولاً این کار با ایجاد شبکه‌ای از مشتری‌ها بسیار پیش از تصمیم‌گیری از سوی مشاور برای خصوصی کارکردن دست‌یافتنی است. مشتری‌هایی با پروژه‌های مشاوره‌ای بزرگ به‌روشنی می‌خواهند بدانند که داده‌های گران‌بهای خود را به چه کسی واگذار می‌کنند. این بخش را به بیان نکته‌هایی از هم‌قطارانی که از توفان حوادث مشاورهٔ خصوصی سالم گذشته‌اند، به پایان می‌بریم.

- شبکه، شبکه، شبکه.
  برخی از پروژه‌ها به یک «تیم» از مشاوران با تخصص‌های مختلف نیاز دارند. لازم است که یک شبکهٔ خصوصی فراتر از آمار گسترش یابد.
- زمینهٔ کاری یا تخصص ویژه‌ای را که مشتری به آن نیازمند است بیابید. این امر شاید مستلزم آن باشد که مشاور خصوصی مهارت‌هایی را در زمینه‌های غیر آماری داشته باشد یا در خود به وجود بیاورد.
- شراکت امکان تقسیم کار را به وجود می‌آورد و قدرت مواجهه را افزایش می‌دهد. می‌توان طرح‌های بزرگتری را جذب کرد.
- شاهد خبره. برخی از مشاوران خصوصی در زمینه‌های حقوقی تخصص پیدا می‌کنند که در آن نیاز به حضور شاهد خبره است. برخی از مباحث مطرح در امر شاهد خبره بودن را در ادامه مورد بحث قرار می‌دهیم.

شاهد خبره

در مواقعی شما ممکن است درگیر پروژه‌ای شوید که در آن مباحث حقوقی اهمیت دارند. مثلاً در یک مورد تبعیض، شاکیان ممکن است ادعا کنند که با آن‌ها در مقایسه با همکارانی که وضعیت مشابهی داشتند، ناعادلانه رفتار شده است؛ متهم، البته، سعی خواهد کرد تا استدلالی برخلاف آن مطرح کند. در اینجا نقش آماردان مشاور این است که سعی کند یک ارزیابی عینی از ادعای تبعیض از یک منظر آماری ارائه کند. به این معنا که دانش تخصصی شما در زمینهٔ آمار شما را به «شاهدی خبره» در آن مورد تبدیل می‌کند. برای ملاحظهٔ جزئیات مربوط به آمار و حرفهٔ قضا به کی و زیسل (۱۹۹۷)، فینکل‌استاین و لوین (۱۹۹۰)، و فاینبرگ (۱۹۸۹) مراجعه کنید.

شاهد خبره‌شدن در یک مورد قضایی الزاماتی را در بر دارد. بنابراین، نخست باید تصمیم بگیرید که آیا «پرونده را می‌پذیرید» یا خیر. این تصمیم باید به محض اینکه خود را به‌طور کامل، هم از لحاظ جنبه‌های آماری و هم اخلاقی پرونده، سبک و سنگین کردید، اتخاذ شود. اگر شما مخالفت‌های فکری و اخلاقی با ماهیت پرونده داشته باشید، باید از قبول آن امتناع کنید؛ میزان مؤثربودن شما به‌عنوان یک شاهد خبره محتملاً محدود خواهد شد. اگر به توافق برسید، در این صورت از شما انتظار خواهد رفت که در یک سلسله مراحل رسمی قضایی، نظیر شهادتی با قید سوگند یا محکمه، برای استشهاد حاضر شوید. به خاطر داشته باشید که هر نوع شهادتی که می‌دهید کاملاً در چارچوب قضایی تعبیر و تفسیر خواهد شد. حفظ عینی‌گرایی آماری تحت این شرایط ممکن است دشوار باشد. برخی از جنبه‌های یک شاهد خبره بودن که باید در نظر داشته باشید عبارت‌اند از:

- مدت‌زمان مورد قضایی
- پاسخ‌دادن به پرسش‌ها
- حفظ عینی‌گرایی
- کارکردن با وکیل
- تهیهٔ اسناد مؤثر

مدت تصمیم‌گیری در مورد نتیجهٔ یک مورد قضایی ممکن است زمان‌بر باشد. بنابراین، به خاطر سپردن آن شهادت «دقیق»، که یک سال پیش در یک عمل شهادت با قید سوگند ارائه کرده‌اید، ممکن است دشوار باشد. این امر در صورتی که از این شهادت صرفاً به عنوان یک تاکتیک حقوقی استفاده کرده باشید ممکن است تشویش‌آمیز باشد. بنابراین در مورد نتایج

کلی آماری خود شفاف باشید. این بخش از شهادت شما باید با هر شهادتی که پیش‌تر ارائه کرده‌اید سازگاری داشته باشد.

پاسخ‌ها وقتی شهادت می‌دهید، پاسخ‌های شما باید قابل فهم باشند و «به زبان عوام» ارائه شوند. شما نمی‌توانید از اصطلاحات فنی (بدون اینکه نخست آن‌ها را توصیف کرده باشید) استفاده کنید. به یاد داشته باشید که این یک جلسهٔ تمرین با وکیل مشتری شما نیست؛ هرچه بگویید بی‌درنگ در سوابق ثبت می‌شود. بنابراین، باید هر سؤال را با دقت زیاد پاسخ دهید. بیش از آنکه لازم است نگویید.

عینی‌گرایی قصد شما به‌عنوان یک شاهد خبره ارائهٔ یک ارزیابی عینی بر مبنای جنبه‌های آماری پرونده است. این بدان معنی است که شما نباید اجازه دهید که از شهادت شما چنین برداشت شود که از آن علیتی برمی‌آید. به‌عنوان مثال، نتیجه‌ای معنی‌دار از آزمون دقیق فیشر مستلزم آن نیست که تبعیض رخ داده است. مشکل در اینجا آن است که میل به تعیین علت ممکن است از هر دو سوی دعوی برخیزد. با این وسوسه به هر قیمتی شده مقابله کنید، از شهادت آماری مخدوش شما بعداً علیه خود شما استفاده می‌شود.

وکلا وظیفهٔ شما برنده‌شدن در دعوی نیست. همان‌طور که نمی‌توانید انتظار داشته باشید که وکیل مشتری شما به‌یکباره در آمار خبره شود، خود شما هم نباید سعی کنید که در حقوق خبره شوید. اگر فکر می‌کنید که مثال بسیار خوبی دارید که می‌توانید از آن در طرح دعوی استفاده کنید، نخست دربارهٔ آن با وکیل صحبت کنید. غافلگیری و مثال‌هایی که به حد و حصر آن‌ها فکر نشده در زمینهٔ حقوقی ممکن است به‌سهولت کار دست شما بدهند.

اسناد ممکن است همهٔ گزارش‌ها، پرینت‌های کامپیوتری، یادداشت‌های مربوط به دعوی که در اختیار دارید، و نیز بحث‌های شفاهی شما را (با چه کسانی و دربارهٔ چه چیزی) که به دعوی مربوط باشند، از شما مطالبه کنند. محتاط باشید. این‌ها اطلاعاتی نیستند که همین‌طوری کنار آب‌سردکن با همه در میان بگذارید. به همین نحو، سعی کنید که مستندات خود را به‌گونه‌ای کارآمد نگهداری کنید. لزومی ندارد که پیش‌نویس گزارش خود را نگهداری کنید.

۴.۴.۱ حکومت

حکومت ایالات متحده و نهادهای ایالتی، برخی از مراکز مهم استخدامی برای آماردان‌ها هستند. مقالهٔ راس (۱۹۹۵) برای کسانی که به شغلی در حکومت می‌اندیشند، جالب توجه است. نهادهایی که عده‌ای کثیر از آماردانان را استخدام می‌کنند عبارت‌اند از:

۱. دفتر سرشماری

۲. مدیریت غذا و دارو[2] (FDA)

۳. سازمان حفظ محیط زیست[3] (EPA).

اغلب فکر می‌کنند که دفتر سرشماری، سرشماری جمعیت ایالات متحده را انجام می‌دهد. با این حال، مأموریت دفتر سرشماری بسیار گسترده‌تر از آن است. این مؤسسه بررسی‌هایی در بسیاری از زمینه‌ها از قبیل خرده‌فروشی، رونق خانه‌سازی و دیگر مباحث مربوط به خانه‌سازی، تولید صنعتی، محصولات کشاورزی، ادعاهای بیمه، تجارت خارجی، بازرگانی، حمل و نقل، و موارد متعدد دیگر انجام می‌دهد. آماردانان نقشی مهم در گردآوری و پردازش این داده‌ها دارند.

- طراحی مطالعه.
- گردآوری داده‌ها.
- تحلیل داده‌های آماری و داده‌کاوی.

جنبهٔ مهم دیگری از مأموریت دفتر سرشماری، تحقیق و ایجاد روش‌شناسی جدید است. به‌طور سنتی، این مؤسسه کار خود را بر روش‌شناسی نمونه‌گیری پیمایشی متمرکز کرده است، اما اخیراً علاقه به روش‌های جامع‌تری از قبیل مصورسازی داده‌ها، تحلیل سری‌های زمانی، و داده‌کاوی در آن بروز کرده است.

مدیریت دارو و غذا (FDA) یک نهاد نظارتی است که هدف آن نظارت بر سلامت داروها و غذاهاست. آماردانانی که به علوم پزشکی، زیست‌شناسی، و علوم تغذیه علاقه دارند، شاید شغلی مورد علاقه در FDA به‌دست آورند. در نظارت بر دارو، سه زمینهٔ اصلی وجود دارد که آماردانان نقشی مهم در آن‌ها دارند:

- داروهای انسانی
- داروهای دامی
- ابزارهای پزشکی.

هر یک از این زمینه‌ها با فرایند زیرنظرداشتن داروها سروکار دارد. آماردانان نقشی مهم در پاسخ‌دادن به پرسش‌های اساسی سالم‌بودن و مؤثربودن داروها و ابزارها دارند. آماردانان ممکن است در هیئت‌های ارزیابی شرکت داشته باشند یا گزارش‌هایی برای فرایند ارزیابی داروها یا ابزارها بنویسند.

2. Food and Drug Adminstration 3. Environmental Protection Agency

FDA همراه با دیگر مؤسسات ایالتی یا فدرال، مسئول مراقبت از ذخیره‌های غذایی نیز هستند. در اینجا نقش آمارداتان در تحلیل جریان‌های داده‌هاست که از انجام‌دادن آزمون‌های مداوم غذا و ذخیره‌های آب‌های داخل بطری مطابق با معیارهای معین حاصل می‌شوند. دو زمینهٔ دیگر که FDA در آن‌ها دخالت دارد و آمارداتان در آن‌ها نقش دارند، بررسی و صدور مجوز برای محصولات زیستی و نظارت بر لوازم آرایشی است.

سازمان حفظ محیط زیست (EPA) نیز آمارداتان را در فرایند نظارت در طرح‌هایی مانند پالودن مکان‌های دفن زباله‌های سمی و اعمال مقررات آب و هوای پاک به‌کار می‌گیرد. آمارداتان با گروه‌هایی از شیمیدانان، زیست‌شناسان، یا دانشمندان محیط زیست که مطالعاتی از مکان‌های مورد بحث را در دست دارند، دست‌اندرکارند. این طرح‌ها اغلب در بردارندهٔ پرسش‌هایی‌اند که پاسخ به آن‌ها به کمک روش‌شناسی متعارف دشوار است و آمارداتان ممکن است درگیر به وجودآوردن رویکردهایی جالب و نو برای تحلیل شوند.

زمینه‌های حکومتی دیگری نیز تعداد قابل توجهی از آمارداتان را در گذشته به استخدام درآورده‌اند. سازمان امنیت ملی[1] (NSA)، مؤسسهٔ سلامت ملی[2] (NIH)، ارتش ایالات متحده، ناسا (NASA)، و سیا (CIA) در این زمره قرار دارند. در حالی که ما توجه خود را به مؤسساتی معطوف کرده‌ایم که مختص ایالات متحده‌اند، می‌توان دید که آمار ابزاری مهم در بسیاری از زمینه‌های حکومتی است و مطمئناً این امر در مورد دیگر کشورها نیز درست است. سرانجام، مؤسسات بین‌المللی معروفی نیز از قبیل سازمان همکاری‌ها و توسعهٔ اقتصادی[3] (OECD) هستند که اطلاعات اقتصادی را گردآوری و منتشر می‌کنند.

۵.۴.۱ دانشگاه

برنامه‌های مشاورهٔ آماری دانشگاه می‌تواند منبع مهمی برای صنایع و شرکت‌های مشاوره‌ای کوچک فراهم کند. منافع برقرارکردن تعاملی با بخش خصوصی آشکارا دوطرفه است: هزینه‌های بالاسریِ به‌خدمت‌گرفتن اعضای هیئت علمی و دانشجویان نسبتاً کم است (نیازی به پرداخت مزایا نیست، آن‌ها دانش آماری دارند، و دانشگاه تجهیزات کامپیوتری مورد نیاز را در اختیار دارد). به‌ویژه، دانشجویان آن فرصت بسیار مهم را برای کسب «تجربه»ی عالم واقع به‌دست می‌آورند. اطلاعات دربارهٔ مراکز مشاورهٔ دانشگاهی جاافتاده را می‌توان به‌آسانی از طریق ارتباط اینترنتی با پایگاه وب دانشگاه به‌دست آورد. دیگر مراجع مربوط به مشاوره در محیط‌های دانشگاهی عبارت‌اند از

1. National Security Agency 2. National Institute of Health 3. The Organization for Economic Cooperation and Development

کیرک (۱۹۹۱)، روستاگی و وولف (۱۹۸۲)، و بخش مربوط به آموزش آمار انجمن آمار امریکا[1] (ASA).

سابقهٔ وجود محیط‌های مشاورهٔ آماری دانشگاهی در واقع طولانی است و به‌عنوان مراکز پژوهشی برای کمک به بهبود کشاورزی[2] تأسیس شده‌اند. در حالی که بسیاری از دانشگاه‌های زمین‌اعطایی[3] روابط مستحکم خود با صنعت کشاورزی را حفظ کرده‌اند، محیط‌های مشاورهٔ آماری متنوعی در دانشگاه‌های امروزی موجودند. در برخی موارد، مشاورهٔ آماری ممکن است صرفاً بر مبنای عضویت در دانشگاه باشد: یک مشتری با یک عضو هیئت علمی دانشکده‌ای مناسب تماس می‌گیرد تا کمک آماری دریافت کند. در زیر به‌اختصار وضعیتی را در نظر می‌گیریم که در آن نوعی ساختار رسمی در داخل دانشگاه موجود است که به‌روشنی به‌عنوان یک «مرکز» با هدف مشاورهٔ آماری شناخته می‌شود.

برنامهٔ مشاورهٔ آماری (SCP)

تأسیس یک برنامهٔ مشاورهٔ آماری[4] در یک دانشگاه می‌تواند پشتیبانی درونی باارزشی برای پژوهشگران در طیفی گسترده از موضوعات علمی فراهم آورد. به محض تأسیس، برنامهٔ مشاورهٔ آماری (SCP) فرصتی برای دانشگاه از طریق فعالیت‌های مشاورهٔ بیرونی و جذب منابع مالی پژوهشی از خارج فراهم خواهد آورد. سه هدف عمدهٔ برنامهٔ مشاورهٔ آماری (SCP) عبارت‌اند از:

مرکز مشاوره برای خدمت به نیازهای پژوهشی دانشگاه.

درس مشاوره برای ارائهٔ فرایند تفکر علمی به دانشجویان در نقش مشاورهٔ آماری. این را می‌توان در یک درس رسمی تحصیلات تکمیلی ادغام کرد.

هیئت مشاورهٔ صنعتی برای ایجاد سازوکاری رسمی برای تعامل با جامعهٔ پژوهشگران در صنعت. دانشجویان می‌توانند فرصت کار در یک محیط غیردانشگاهی را از طریق دستیاری به‌دست

1. The American Statistical Association

۲. در واقع، بخش اعظم اصطلاحات مرتبط با طرح آزمایشی نظیر بلوک‌ها، کرت‌ها، و کرت‌های خردشده، مستقلاً به تقسیم‌بندی واقعی قطعه‌زمینی که در طرح به آن نیاز بود، مربوط می‌شوند. بنابراین، «کرت» همان معنی «یک کرت زمین» را داشت.

۳. land grant universities: دانشگاهی است در ایالات متحده که از طرف مجالس ایالتی برای استفاده از مزایای قانون‌های موریل تعیین شده‌اند. دو قانون موریل، مصوبهٔ ۱۸۶۲ و ۱۸۹۰ کنگرهٔ امریکا که توسط جاستین اسمیت موریل (Justin Smith Morrill) به مجلس تقدیم و تصویب شدند، عطیه‌ای بود که از محل آن دانشگاه‌های مربوط موظف بودند به مردم در زمینه‌های کشاورزی، اقتصاد خانه، صنایع مکانیکی، و دیگر حرفه‌هایی که در زمان تصویب، جزء عملی محسوب می‌شدند آموزش دهند.

4. Statistical Consulting Program

آورند و صنعت با داشتن این فرصت که بهترین دانشجویان واجد شرایط را به استخدام درآورده است از این موقعیت سود می‌جوید.

در محیط برنامهٔ مشاورهٔ آماری، یک مشاور اغلب با یک یا چند دانشجو همزمان روی چندین پروژه کار می‌کند. معمولاً زمان اندکی برای آموختن فراتر از حداقل دانش مسئله برای پاسخ‌دادن به پرسش‌های آماری مسئله وجود دارد. با این حال، همان‌طور که در مطالعه‌های موردی ارائه‌شده در قسمت II نشان می‌دهیم، طرح‌های پژوهشی درازمدت نیز پیشنهاد می‌شوند.

۲

ارتباط

توانایی یک آماردان مشاور برای برقراری ارتباط مؤثر بسیار مهم است. البته تعجبی هم ندارد. ارتباط مؤثر مطمئناً مهارتی مطلوب است که ارزش ایجاد آن، هر جا که نیازی به تعامل بین دو طرف وجود داشته باشد، حس می‌شود. با این حال، طی تربیت آماردانان تأکید قابل ملاحظه اغلب بر ایجاد مهارت‌های فنی است و ارتباط به‌عنوان چیزی که می‌توان «بعداً اکتساب کرد» کنار گذاشته می‌شود. برای آماردان مشاور، حق انتخاب «بعداً» وجود ندارد؛ مهارت‌های ارتباطی خوب برای مشاور کارآمد بودن مورد نیاز است.

در این فصل، ما به‌ویژه بر نقش ارتباط تمرکز می‌کنیم و برخی عناصر و مهارت‌های متعارف را که در ارتباط مؤثر دخیل‌اند واکاوی می‌کنیم. این موارد به‌تفصیل مورد بحث قرار می‌گیرند و توصیف کاملی از رویکرد ما و دستورالعمل‌های «چگونه انجام دهیم» ارائه می‌شود. البته به وجودآمدن مهارت‌های ارتباطی خوب مستلزم زمان و تلاش است و دستورالعمل‌های ما بر آن نیستند که جایگزینی برای آن باشند.

با در نظرگرفتن وضعیت زیر مطلب را شروع می‌کنیم. وقتی یک آماردان در پروژه‌ای شرکت دارد، فرایندی در برگیرندهٔ انتقال اطلاعات صورت می‌گیرد. خواه این امر با گروهی از همکاران انجام شود و خواه با یک مشتری، عناصر عام معینی در فرایند ارتباط دخیل‌اند:

۱. تعامل شفاهی با مشتری (مشتریان) که تا زمانی که پیشرفت قابل توجهی در پروژه صورت گیرد ادامه می‌یابد.

۲. آماده‌کردن خلاصه‌های فنی، گزارش‌نویسی، و ارائهٔ یافته‌ها.

برخی مهارت‌ها برای انجام‌دادن مؤثر این وظایف مورد نیاز خواهد بود و ما در دو بخش بعدی آنچه را که در تعامل شفاهی با مشتری دخیل است در نظر می‌گیریم. از آنجا که توجه ما در این فصل بر ارتباط است، مستقیماً به مرحلهٔ گزارش‌نویسی پروژه می‌رویم. در حالی که این امر بر پذیرش این فرض متکی است که تحلیل با استفاده از روش‌های آماری مناسب انجام‌شدنی است، جزئیات مربوط به فنون آماری خاص در فصل آتی پوشش داده می‌شود و برای بحث ما در گزارش‌نویسی ضروری نیست.

همچنین نقش ارائه‌کردهای شفاهی را مورد بحث قرار می‌دهیم و جزئیاتی دربارهٔ نحوهٔ مؤثرکردن این ارائه‌کردها پیش رو می‌گذاریم. اهمیت کیفیت نمودارها برای مقاصد ارائه‌کرد در بخش ۶.۲ مورد بحث قرار می‌گیرد. معرفی کوتاهی از استفاده از نرم‌افزار پاورپوینت برای ارتقای ارائه‌کردها در پایان این بخش انجام شده است.

پیش از ادامه باید تأکید کنیم که هر دستورالعملی که برای ارتباط مؤثر ارائه می‌کنیم مناسب هر وضعیت مشاوره نخواهد بود. با این حال، امید می‌رود که خواننده با داشتن فرصت تبعیت از رویکرد تفصیلی ما از آن سود ببرد و آن را با وضعیت‌های خاص مقتضی تطبیق دهد.

۱.۲ تعامل شفاهی

آماردان در وضعیت‌های گوناگونی باید با یک مشتری یا گروهی از همکاران مستقیماً تعامل کند. هدف اصلی این تعامل مبادلهٔ اطلاعاتی دربارهٔ پروژهٔ مورد نظر است و برای انجام‌دادن مؤثر این کار آماردان نیازمند پدیدآوردن مهارت‌های ارتباطی در زمینه‌های زیر است.

۱. آغاز فرایند ارتباط.

۲. درک و تعریف مسئله.

۳. ارزیابی دانش فنی مشتری یا همکار.

۴. سبک و سنگین‌کردن مباحث و اهداف کلی پروژه.

۵. شناسایی سهم‌های خاص آماردان در پروژه.

در ارتباط با مقاصد این بحث از مشتری (مشتریان) یا همکاران صرفاً به‌عنوان «مشتری» یاد می‌کنیم. در حالی که این کلمه اشارات ضمنی معینی دارد، مهارت‌های ارتباطی که باید به وجود آوریم، چه زمانی که با یک مشتری سروکار داریم چه با گروهی از همکاران، یکسان‌اند. به عبارت دیگر، این مهارت‌ها به نوع خاصی از محیط مشاوره، که ممکن است در آن اصطلاحات بیانی معینی ترجیح داده شود، بستگی ندارد.

آغاز تعامل

پس، ازکجا شروع کنیم؟ چگونه تعامل را آغازکنیم؟ دیر (۲۰۰۰) و بون و ژان (۱۹۶۸) هر دو به‌تفصیل به این مسئلهٔ خاص می‌پردازند و تأکید عمده‌ای بر اهمیت شکل‌گیری نخستین برداشت‌های مثبت دارند. وضع ظاهری ملاقات و رفتار آغازین غیرکلامی ما در مقابل مشتری، اولین برداشت آن‌ها را از ما به وجود خواهد آورد ــ در حالی که هنوز «سلام» هم نکرده‌ایم. برخی از این ایما و اشاره‌ها را در مثال ارائه‌شده در فصل ۴ بررسی می‌کنیم. ادب و مهربانی رایج و احترام به‌طور بدیهی در ایجاد محیطی مثبت در ملاقات‌های مشاوره‌ای ما سهم زیادی دارند. برخی کارهای ساده که می‌توانیم انجام دهیم تا مشتریان هنگام خوشامدگویی به آن‌ها احساس راحتی کنند عبارت‌اند از:

۱. هرکاری را که مشغول انجام‌دادن آن هستیم متوقف کنیم و برای خوشامدگویی از جای خود برخیزیم. شاید لازم باشد که کت مشتری را بگیریم یا نشان دهیم که می‌تواند کیف دستی خود را کجا بگذارد.

۲. چشم در چشم مشتری بیندازید و لبخند بزنید. چنین کاری این پیام را منتقل می‌کند که از دیدن مشتری خوشحالیم و به ما این فرصت را می‌دهد که رفتار عمومی مشتری را ارزیابی کنیم. اگر به نظر دستپاچه می‌رسد، به مشتری کمی فرصت دهید تا خونسردی خود را پیدا کند. پیش از سؤال دربارهٔ پروژهٔ آن‌ها دربارهٔ هوا یا مطلب حاشیه‌ای دیگری صحبت کنید.

پس از آنکه معارفه‌های اولیه را تمام کردیم، مشتری شاید بخواهد داده‌هایی را به ما نشان دهد یا از یک شیوهٔ آماری که می‌خواهد از آن استفاده کند یاد کند. با این حال، لازم است که از ابتدا آغازکنیم. به این معنی که لازم است با زمینهٔ مسئله آغازکنیم زیرا داده‌های بدون زمینه چندان معنایی ندارند.

- چقدر لازم است بدانیم؟
- زیاد پرسش کنید.
- همواره آمادهٔ یادداشت برداشتن باشید.

چقدر لازم است بدانیم؟ آنچه در واقع می‌پرسیم این است که به چقدر اطلاع از مشتری نیاز داریم تا جنبه‌های آماری مسئله را رفع و رجوع کنیم. البته، هنوز هم لازم است که مسئله خوش‌تعریف باشد، بنابراین به استراتژی‌هایی نیازمندیم تا اطلاعات مناسب را بیرون بکشیم. رویکرد مناسب چنین است:

زیاد پرسش کنید: در این مرحلهٔ اول مشاوره اغلب سودمند است که مشتری را وادار کنیم که جلسه را با توصیف پروژه با الفاظ خودش شروع کند. این امر به ما فرصت آموختن دربارهٔ حوزهٔ کار مشتری یا قطع به‌موقع سخنانش را وقتی اصطلاحات ناآشنا یا تخصصی مطرح می‌شوند فراهم می‌سازد. آیا هرگز مقیاس «EQRT» را شنیده‌اید؟ پس سؤال کنید! از ما انتظار نمی‌رود که در کلیهٔ زمینه‌های تحقیقات علمی خبره باشیم. حال جملهٔ آخر را یک‌بار دیگر بخوانید. چرا؟ زیرا نیاز هم نداریم که در حوزهٔ کار مشتری خبره شویم.

وقتی یک سؤال می‌پرسیم باید به آنچه هم مشتری می‌گوید (به‌دقت) گوش دهیم. به خاطر داشته باشید که مشتریان برای مشورت آماری به ما مراجعه می‌کنند. از آن‌ها نمی‌توان انتظار داشت که بدانند برخی اصطلاحات معانی کاملاً خاصی در آمار دارند. گفتن اینکه عاملی «معنی‌دار» است برای ما چیزی کاملاً متفاوت با گفتن آن است که آن عامل «مهم» است. البته عکس این امر نیز به‌سادگی صادق است. مشتری ممکن است بگوید «مهم»، زمانی که در واقع منظورش (مثلاً بر پایهٔ یک بررسی بیشتر) معنی‌دار است. بنابر تجربهٔ ما، مشتریان اغلب گرایش به انجام‌دادن دو کار دارند:

۱. اصطلاحات آماری را به‌طور نامناسب به‌کار ببرند. ما همواره باید منظور مشتری را کنترل مجدد کنیم.
۲. از ذکر متغیرهای مهم نظیر عامل‌های طرح، که در آزمایش به‌کار رفته‌اند، خودداری کنند. این بدان معنی است که ما باید آنچه را هم که نمی‌گویند گوش کنیم.

همواره آمادهٔ یادداشت برداشتن باشید: طبیعتاً در طول جلسهٔ پرسش و پاسخ آمادهٔ یادداشت برداشتن بوده‌ایم. ... فرض نکنید که بعداً توانایی به یادآوردن همهٔ جزئیات دربارهٔ پروژهٔ

مشتری را خواهید داشت. یادداشت برداشتن طی جلسهٔ مشاوره بخشی مهم از فرایند تهیهٔ اسناد (بخش ۳.۲) است و ما بر اهمیت پذیرش این رفتار تأکید داریم.

تعریف مسئله

وظیفهٔ آغازین ما سعی در درک زمینهٔ یک پروژه از منظر مشتری است. این بدان معنی است که لازم است پیش از سعی در تعریف مسئله چیزی دربارهٔ زمینهٔ کار مشتری و اصطلاحات مربوط به آن فرا بگیریم. تدریجاً که با «زمینه»ی مسئله آشناتر می‌شویم و با استفاده از اصطلاحات رایج با مشتری ارتباط برقرار می‌کنیم، متقابلاً می‌توان نیت سؤال‌ها را به سمت جنبه‌های زیر از پروژه متوجه کرد. این اطلاعات در تعریف مسئله و شناسایی موضوعات آماری دخیل در آن، سودمند است.

- زمینهٔ پروژه
- وضعیت پروژه
- اهداف پروژه
- انتظار مشتری

زمینه پروژه‌ها اغلب بر مطالعات پیشین مبتنی‌اند که در این حالت ممکن است روش تحلیل تثبیت‌شده یا پذیرفته‌ای موجود باشد. اگر چنین باشد، به‌دست‌آوردن مرجع مربوطی از مشتری ممکن است به حصول اطمینان از اینکه روش تحلیل تثبیت‌شده معقول و قابل به‌کارگیری در مسئلهٔ مشتری است به ما کمک کند.

وضعیت وضعیت پروژه چیست؟ اگر بررسی در مرحلهٔ پیش‌آزمایشی یا برنامه‌ریزی است، کمک ما ممکن است در تضمین اینکه آزمایش برنامه‌ریزی‌شده داده‌های قابل اعتمادی برای تحلیل بعدی تولید خواهد کرد اهمیت داشته باشد. اگر داده‌ها قبلاً گردآوری شده باشند، لازم است که سؤال‌های خود را به سمت گردآوری جهت ببخشیم. داده‌ها تا چه حد قابل اعتمادند؟ آیا مشتری از نقاط دورافتاده در داده‌ها آگاه است؟ آیا آزمایش مطابق با اصول رایج آزمایشگری آماری یعنی کنترل، تصادفی‌سازی، و تکرار انجام شده است؟ آیا شواهد کافی (اندازهٔ نمونه و ساختار در داده‌ها[1]) برای پشتیبانی از اهداف پروژه موجود است؟

۱. در یک پروژه ما آماده بودیم که تحلیل یک مجموعه‌دادهٔ نسبتاً بزرگ (۱۰۰ مگابایت) را صرفاً برای دریافتن اینکه متغیری اصلی در مطالعهٔ اولیه ثبت نشده بود شروع کنیم. بنابراین، اهداف پروژهٔ ما نمی‌توانست برآورده شود.

اهداف هدف‌ها و فرضیه‌های وابسته به بررسی کدام‌اند؟ آیا هدف‌های مشتری با نتایجی که می‌توان از یک تحلیل آماری به‌دست آورد در توافق‌اند؟ در برخی حالت‌ها، شاید لازم باشد که فرضیه‌های معینی را از نو تدوین کنیم تا تحلیل‌های آماری نتیجه‌گیری‌های معتبری را در اختیار گذارند. همچنین باید مطمئن شویم که مشتری تمایز بین علیت و نتیجه‌گیری‌های مبتنی بر تحلیل‌های آماری را می‌داند.

انتظارات مشتری از ما چه انتظاری دارد؟ ما نه جادوگریم نه مجری پروژه. مسئولیت ما همواره باید متوجه جنبه‌های آماری مسئله باشد؛ این وظیفهٔ مشتری است که اهمیت و انگیزهٔ پروژه را به‌طور شمرده مدون کند.

دانش فنی مشتری

تعریف جنبه‌های آماری مسئلهٔ مشتری برای خودمان، بخش سادهٔ کار است. حال لازم است آن‌ها را برای مشتری تشریح کنیم. در همان حال که مسئله را تعریف می‌کنیم، سودمند (وگاهی کشف‌آمیز) است که دانش مشتری و درک او از جنبه‌های آماری پروژه را معلوم کنیم. به‌عنوان مثال، یک مشتری ممکن است سؤال کند که «به نمونه‌ای به چه بزرگی نیاز دارم؟» چگونه باید پاسخ دهیم؟ برحسب توان یا درستی[1] آزمون‌های آماری یا مفهومی آشناتر مانند حاشیهٔ خطا؟ این امر آشکارا به دانش فنی مشتری بستگی دارد. بنابر تجارب ما، حاشیهٔ خطا اغلب نقطهٔ آغازین سودمندی را در اختیار ما می‌گذارد و از تعبیرهای غلط اصطلاحاتی از قبیل درستی یا دقت به‌دور است.

- مشتری تا چه حد پروژه را می‌فهمد؟
- مشتری چقدر دانش آماری دربارهٔ پروژهٔ خود دارد؟

همان‌گونه که در بالا خاطرنشان کردیم، ممکن است یک روش‌شناسی آماری پایه‌ای یا تثبیت‌شده موجود باشد که در حوزهٔ کار مشتری کاملاً پذیرفته شده باشد. با این حال، دانستن یک شیوهٔ آماری لزوماً به این معنی نیست که مشتری مفاهیم زمینه‌ای شیوهٔ آماری را به‌طور کامل می‌فهمد. بنابراین، بخشی از نقش ما ممکن است یک نقش آموزشی باشد. ما برخی مباحث را در نظر می‌گیریم که شاید لازم باشد به آن‌ها پرداخته شود.

آموزش‌دادن مشتری مشتری برای یک خطابهٔ درسی به ما مراجعه نکرده است؛ توضیحات ما باید در متن پروژهٔ مشتری باشد؛ تعبیری از نتیجه و مقصود یک شیوهٔ آماری، و نه جزئیات ریاضی

۱. وقتی سؤال شد «تا چه حد می‌خواهید نتایج آزمون شما صحت داشته باشند؟» یک مشتری به یکی از مؤلفان گفت که آن‌ها می‌خواهند نتایج ۱۰۰٪ صحت داشته باشند. ما به‌سرعت به موضوع تعیین اندازهٔ نمونهٔ مناسبی بازگشتیم.

را، در اختیار او قرار دهید. به‌عنوان مثال، یک Pـمقدار را می‌توان در قالب «مخاطره» و نه احتمال مبتنی بر نوعی توزیع بیان کرد. صبور باشید، اما از چسبیدن به جزئیاتی که اساسی نیستند خودداری کنید ــ چقدر لازم است بدانیم؟ دربارهٔ مشتری هم صادق است.

سطح پیچیدگی روش‌های آماری به‌کاررفته در تحلیل باید مناسب مسئله باشند و این امر شاید مستلزم آشناکردن مشتری با رویکردهای پیچیده‌تر باشد. با این حال، ما نباید تلاش کنیم که تحلیل آماری را از آنچه واقعاً لازم است پیچیده‌تر کنیم. مشتری باید قادر باشد که نتایج تحلیل را فارغ از سطح پیچیدگی آماری تفسیر کند.

رسمیت‌بخشیدن به مسئله یک موضوع پیچیده‌تر، نیاز بالقوه به رسمیت‌بخشی به ایده‌های مشتری است. آن‌ها ممکن است دربارهٔ کار خود به روشی شهودی‌تر بیندیشند که لازم است پیش از آنکه بتوان تحلیل آماری مسئله را به اجرا درآورد به‌دقت در قالب رسمی درآورده شود. زمانی که صرف رسمیت‌بخشی مسئله می‌شود ارزش این کار را دارد زیرا به مشتری کمک می‌کند که پژوهش را از منظری آماری درک کند. در برخی حالت‌ها، این کار حتی ممکن است به فرمولبندی بهتر اهداف پژوهش کمک کند.

مثال ۱.۲ اثر متقابلی بین PST و PREF به‌واسطهٔ یک GROUP موجود خواهد بود.

در بخش ۱.۹ (بهبود آموزش)، فرضی که در بالا بیان شد بهترین تقریب مشتری از بیان فرض «رسمی» مسئله بود. آنچه مشتری واقعاً می‌خواست بداند این بود که آیا عامل‌های PREF و GROUP تأثیری بر پاسخ PST دارند؛ به این معنی که آیا یک اثر متقابل ناشی از این عامل‌ها وجود دارد یا خیر.

موضوعات و هدف‌های کلی

در این مرحله باید که کانال ارتباطیِ به‌قدر کافی مناسبی را با مشتری برقرار کرده باشیم و حالا می‌توانیم وارد جزئیات مربوط به موضوعات آماری کلی دخیل در پروژه شویم. اقلام زیر باید به زبانی مورد بحث قرار گیرند که هر دو طرف قادر به درک آن باشند.

- هدف‌ها و فرض‌های پروژه،
- روش‌شناسی جاری یا پیشین، در صورت وجود،
- استفادهٔ سنجیده از نتایج پساآزمایشی.

هدف‌ها در برخی مطالعات، هدف‌ها ممکن است صرفاً ماهیت اکتشافی داشته باشند و فرض‌های مناسب باید بعداً فرمولبندی شوند. از سوی دیگر، اگر اهداف مورد توجه خاصی موجود باشند، باید مطمئن شویم که طرح آزمایشی، نتایج از لحاظ آماری معتبری را فراهم خواهد کرد. اگر آزمایش هنوز اجرا نشده است، در این صورت باید موضوعات مرتبط با خود از قبیل اندازهٔ نمونه، تصادفی‌سازی، و کنترل، و نیز موضوعات اجرایی را مورد توجه قرار دهیم. اگر داده‌ها قبلاً گردآوری شده باشند، باید این مسئله را مورد بررسی قرار دهیم که آیا اهداف مطالعه با داده‌های کنونی برآورده خواهد شد یا خیر. ممکن است داده‌های بیشتری مورد نیاز باشند.

روش‌شناسی برای تضمین اینکه شیوهٔ آماری به‌طور مناسب به‌کار بسته شده، توجه به موضوعاتی خاص از قبیل موارد زیر لازم خواهد بود:

- نوع داده‌های هر متغیر چیست؟
- آیا نقاط دورافتاده یا گمشده حضور دارند؟
- آیا محدودیت‌های معینی در فرایند وجود دارند؟

این امر مستلزم تعاملی نزدیک با مشتری در طول یک جلسهٔ مشاوره است. (از سردرگم‌کردن مشتری اجتناب کنید: مقادیر گمشده عموماً میل به، بسیار خوب ... گمشدن دارند!) ما مثالی از این نوع تعامل را در بخش ۵.۳ ارائه می‌کنیم که در آن برخی پرسش‌های خاص رگرسیونی مطرح می‌شوند. همچنین باید روش‌شناسی جاری و پیشین را که در پروژه به‌کار می‌رود با شیوه‌های آماری جاافتاده که ممکن است نوتر یا مناسب‌تر باشند مقایسه کنیم.

پساآزمایش ارزشمند خواهد بود که استفادهٔ سنجیده از نتایج پساتجربی را مورد توجه قرار دهیم. به‌عنوان مثال، آیا نتیجهٔ بررسی به‌طور اساسی به معنی‌دارشدن یک فرض خاص بستگی دارد یا خیر. پیامدهای به‌دست‌آوردن یک نتیجهٔ غیر معنی‌دار چیست؟ اطمینان حاصل کنید که مشتری می‌فهمد که نمی‌توانیم نتیجهٔ تحلیل را صرفاً به این دلیل که هدف‌های آغازین مشتری را تأیید نمی‌کند تغییر دهیم.

سهم‌های مشخص

مرحلهٔ نهایی تعامل شفاهی ما با مشتری عبارت از معین‌کردن سهم مشخص ماست. این امر برای تضمین اینکه هم ما و هم مشتری به‌روشنی نقش‌های مربوط به خود را در پروژه درک می‌کنیم اهمیت دارد. اقلام زیر باید به‌عنوان موارد ضروری مورد توجه قرار گیرند.

مدیریت داده‌ها اگر ما مسئول واقعی اجرای محاسباتیم، لازم است که مشتری داده‌ها را در فرمت مناسب به ما تحویل دهد.

تحلیل داده‌ها وارسی خطا: مشتری باید آگاه باشد که مرحلهٔ آغازین تحلیل ما دربردارندهٔ وارسی داده‌هاست. مشتری باید هر نوع تصحیح را میسر سازد.

تحلیل آماری هم ما و هم مشتری دربارهٔ روشی که در تحلیل داده‌ها بکار خواهد رفت و جزئیات انجام‌دادن محاسبات توافق کرده‌ایم.

گزارش‌نویسی اینکه آیا از ما انتظار می‌رود نمودارها و جدول‌های با کیفیت مورد نظر در ارائه‌کرد یا هر مورد خاص ضروری در گزارشی که نوشته خواهد شد تهیه کنیم یا خیر.

چارچوب زمانی لازم است که چارچوب زمانی واقع‌بینانه‌ای وجود داشته باشد تا به ما اجازه دهد که تحلیل را انجام دهیم و گزارش کتبی برای پروژه را کامل کنیم.

۲.۲ دیگر جنبه‌های تعامل شفاهی

ارتباط تشویقی

در عمل، اجزای مختلف تعامل شفاهی، که هم‌اکنون آن‌ها را مورد بحث قرار دادیم، ممکن است همزمان اجرا شوند. هنر ارتباط تشویقی مستلزم ترکیب‌کردن خلاقانهٔ این اجزا برای کاراکردن تعامل با مشتری است. اداره‌کردنِ خوب یک جلسهٔ مشاوره بر مهارت‌های سازماندهی خوب متکی است. باید آمادگی داشته باشیم که (مؤدبانه) سخن مشتری را در صورت لزوم قطع و او را به موضوعات مربوط بازگردانیم. این کار جلسات مشاورهٔ ما را پربارتر می‌کند و آغاز مرحلهٔ تحلیل را ممکن خواهد ساخت.

تماس آغازین

پیش از قرار آغازین ما با مشتری، ممکن است قبلاً ارتباط‌های غیر مستقیمی برقرارکرده و اطلاعاتی دربارهٔ پروژه کسب کرده باشیم. در این صورت، تمرکز و جهت‌گیری بحث در طول این دیدار آغازین

ممکن است تا حدی از قبل معین باشد. با این حال، مهم است که همهٔ اطلاعات پیشین را در شروع مشاوره تکرار کنیم تا اطمینان حاصل کنیم که توافق دوجانبه‌ای در زمینهٔ ارتباط‌های قبلی حاصل شده است. انجام‌دادن این کار به تنظیم رابطهٔ کاری بین خودمان و مشتری کمک می‌کند.

زمان تصمیم

در برخی حالت‌ها، ممکن است به دلایل گوناگون از شرکت در پروژه خودداری کنیم. به عنوان مثال، ممکن است محدودیت‌هایی داشته باشیم که مانع از تکمیل پروژه در چارچوب زمانی دلخواه مشتری می‌شود. در موارد حقوقی، ممکن است صرفاً به دلیل کار ما با یک مشتری قبلی تضاد منافعی وجود داشته باشد. در این موقعیت‌ها، تصمیم مبنی بر شرکت‌نکردن در پروژه معمولاً باید پیش از جلسهٔ مشاورهٔ آغازین اتخاذ شود.

اگر نیاز به این تصمیم‌گیری حین جلسهٔ مشاوره پیش آید چطور؟ آگاه‌کردن مشتری از تصمیم خود برای شرکت‌نکردن در یک پروژه ممکن است دشوار باشد، اما لازم است طی مراحل آغازین فرایند مشاوره هر زمان که مقدور باشد انجام شود.[۱] بسیار مهم است که مشتری در مورد قصد ما برای شرکت در پروژه به انتظارات واهی دچار نشود. مهم‌تر اینکه مشتری وقت خود را صرف بحث دربارهٔ پروژه با ما کرده است و حالا باید به فکر جایی دیگر باشد. موضوع مهم این است:

دانستن اینکه چه موقع کار را رها کنیم؟

البته، دانستن اینکه چه موقع کار را رها کنیم به این معنی نیست که صرفاً بلند شویم و برویم! در واقع، بخش مهمی از تصمیم ما آگاه‌کردن مشتری از مشکلات پروژهٔ آن‌ها و نصیحت‌ها و توصیه‌هایی است که برای غلبه بر محدودیت‌های بررسی مناسب باشند. ما حتی ممکن است پروژه‌ای «جدید» به وجود آوریم؛ این بار با این ارزش افزوده که می‌توانیم در مرحلهٔ طراحی پروژه هم شرکت کنیم. برخی از دلایلی که بهترین گزینه، سر باززدن از مشاورهٔ یک پروژه است عبارت‌اند از:

- اندازهٔ نمونه برای هر تحلیل پرمعنی، بسیار کوچک یا صرفاً با توجه به منابع محاسباتی فعلی ما بیش از اندازه بزرگ است.
- داده‌ها اریب‌اند یا با بی‌دقتی گردآوری شده‌اند و هیچ فرصتی برای گردآوری داده‌های برنامه‌ریزی‌شدهٔ بیشتر موجود نیست.

۱. تصمیم ما برای رد مشاوره در یک پروژه مطمئناً باید پیش از انجام‌دادن هر تحلیل واقعی گرفته شود؛ ممکن است از لحاظ قانونی یا بنابر قرارداد در صورت شروع تحلیل پروژه مقید به اتمام آن باشیم.

- مشتری ممکن است واقعاً از درک پروژهٔ خود عاجز باشد یا انتظاراتش از نتایج تحلیل غیر واقع‌بینانه باشد.
- ممکن است ما از پروژهٔ مشتری سر در نیاوریم یا خبرگی کمی در نوع تحلیل آماری لازم برای پروژه داشته باشیم.
- نسبت به پروژه ایرادهای فکری یا اخلاقی داریم. این امر اخلاق آماری از این قبیل را نیز دربرمی‌گیرد که مشتری «تقاضا»ی روش تحلیل خاصی را دارد که آشکارا نامناسب است.

اتفاقات منفی

«اوه، ببخشید. می‌توانید آدامس[1] خود را اینجا (سطل زباله) بیندازید.»

چون زمینهٔ این جمله را نداریم، دو سناریو که ممکن است این صحبت طی آن‌ها به‌عمل آمده باشد عبارت‌اند از: ما کمی در فهم ایما و اشاره‌های مشتری کند بوده‌ایم؛ پس از شروع ملاقات مشاوره‌ای، متوجه شده‌ایم که مشتری آدامس می‌جود. در سناریوی اول، اظهارات ما صرفاً پاسخی به یک تقاضای به‌زبان‌نیامده از طرف مشتری بوده که پیش‌درآمد آن پوزش صمیمانه‌ای به‌عنوان یک ژست دوستانه بوده است. در سناریوی دوم، گفتهٔ ما آشکارا منظور متفاوتی داشته است: سعی کرده‌ایم که وضعیتی «منفی» را که طی جلسهٔ مشاوره پیش آمده است رفع و رجوع کنیم.

سناریوی بالا البته تا حدی کم‌مایه است، اما مثالی سودمند برای تشریح برخی از موضوعاتی است که ممکن است لازم باشد مشاور آماری در مواجهه با موقعیت‌های منفی در نظر بگیرد. این موضوعات در زیر به‌طور خلاصه آمده‌اند.

ادراک مشکل چیست؟ هر مشتری خلقیات و ویژگی‌هایی شخصی دارد که ممکن است از نظر ما «آزاردهنده» باشد، اما به‌ندرت می‌توان این را یک «موقعیت» منفی به‌حساب آورد. پرداختن به مسائلی بی‌اهمیت تنها به بدترشدن وضعیت منجر خواهد شد. اگر آدامس‌جویدن یک مشتری مسئلهٔ مهمی برای شما نیست از آن بگذرید.

پیامدها پیامدهای توجه به یک مشکل چیست؟

- با پرده‌دری از وضعیت، ممکن است برداشتی بسیار منفی به او انتقال دهیم. (ممکن است نحوهٔ تقاضای ما به مشتری بربخورد: «این عذرخواهی ناصمیمانه چطور است: 'اوه، من خیلی خیلی متأسفم.'»

۱. ممکن است ما در ایران با ارباب‌رجوعی که موقع صحبت آدامس می‌جود مواجه نشویم، اما عادات ناپسند دیگر چطور؟ بنابراین، باید برای موقعیت‌هایی نظیر این خود را آماده کنیم.ـم

- ممکن است کارمان به جایی بکشد که وقت قابل توجهی را صرف موضوعات غیر مرتبط با پروژهٔ مشتری کنیم (مشتری ما را به خاطر استفاده از رویکردی غیر مستقیم سرزنش می‌کند. «شما می‌توانستید مستقیماً از من بپرسید.» اکنون وضعیت بدتر شده است و ما باید وقت بیشتری صرف کنیم تا مشتری احساس تحقیرشدن کمتری کند).
- ممکن است نتیجه معکوس شود (آدامس وسیله‌ای است که مشتری به دلایل پزشکی به آن نیازمند است. ما تصور کرده‌ایم که آدامسی معمولی است و اکنون با دو مشکل بالا مواجهیم).

موقع‌شناسی چه موقع باید به مسئله بپردازیم؟ معمولاً بهتر است مشکلات ساده را بی‌درنگ عنوان کنیم، اما در برخی از موارد ممکن است بهتر باشد که دربارهٔ آن تأمل کنیم. به این معنا که پرداختن به موضوع را تا پایان جلسهٔ مشاوره به تعویق بیندازید. شاید هم سعی به یافتن راه‌حلی مبتکرانه نمایید تا در ابتدای جلسهٔ بعدی به‌کارگیرید تا مشکل با ترفندی حل شود.

برنده-برنده هدف از پرداختن به یک جوّ منفی چیست؟ البته که هدف دست‌یافتن به یک جوّ مثبت است. مشاور و مشتری هر دو از درک اینکه مشکل چه بوده سود می‌برند و قادرند به سمت موضوعات مهم‌تری حرکت کنند. دِر (۲۰۰۰) از این حالت به‌عنوان یک «وضعیت برنده-برنده» یاد می‌کند و مثال‌های متعددی از جوّهای مثبت و منفی را که مشاور آماری ممکن است با آن‌ها مواجه شود مطرح می‌کند.

ادامه...

شکل‌گیری مهارت‌های ارتباطی خوب یک روند تدریجی است و با اینکه تجربه قطعاً به کمک می‌آید، برای یک مشاور آماری ارزیابی مداوم عملکردش دارای اهمیت است. چگونه ممکن بود وضعیت دشواری که هفتهٔ گذشته با آن روبه‌رو شدیم، به شکل بهتری حل شود؟ چرا دو جلسه طول کشید تا اهداف و تخصیص کار این پروژه روشن شود؟ تفاوت‌های فرهنگی نیز ممکن است بر تعامل ما با مشتری اثر داشته باشد. آشکارا در ارائه‌کرد بالا نه به همهٔ جوانب تعامل شفاهی پرداخته‌ایم و نه همهٔ انواع مختلف وضعیت‌های مشاوره را که یک مشاور انتظار مواجهه با آن‌ها را دارد عنوان کرده‌ایم.

خوشبختانه، منابع فراوان دیگری موجودند که می‌توانند ما را در ایجاد مهارت‌های ارتباطی‌مان یاری کنند! کتاب درسی تألیف دیر (۲۰۰۰) و ویدئوی ضمیمهٔ آن بهواقع ارزش نگاه‌کردن دارند. برخی مقالات مرتبط با این بخش عبارت‌اند از: تویدی (۱۹۹۸)، فینی (۱۹۸۲)، و لاری (۱۹۵۸).

در نهایت، انتشارات تجارت‌محور اغلب اطلاعات سودمندی فراهم می‌کنند که می‌تواند وارد محیط آماری شود. به‌عنوان مثال، همیلتون و پارکر (۱۹۹۳)، و پیتس و هابتن (۱۹۹۸) را ببینید.

۳.۲ چگونه گزارش بنویسیم؟

تحلیل آماری ما در حال حاضر کامل است. یافته‌ها اغلب فرض‌های پروژه را تأیید می‌کنند، و یک ساعت وقت آزاد داریم تا آن «گزارش»ی را که مشتری تقاضا کرده بود ضربتی آماده کنیم. بسیار خوب! به اندازهٔ کافی ساده به نظر می‌رسد، پس بیایید شروع کنیم. به چه چیزی نیاز داریم؟ یک جور مقدمه، یافته‌ها، و نتیجه‌گیری‌های ما. خروجی کامپیوتری که همگی را می‌توان در پیوستی در انتها آورد. کار تمام است! نه کاملاً. طی تماس‌هایمان با محققان برجسته در صنعت، رایج‌ترین شکایتی که از این دانشمندان می‌شنویم این است که فارغ‌التحصیلان آمار ما در نوشتن گزارش‌ها بسیار مشکل دارند. در واقع، مواردی وجود دارند که دانشمندان صنعت که درگیر یک پروژه‌اند به‌طور معمول همکاری یک آماردان تازه‌کار را که مهارت‌های ارتباط نوشتاری خوبی ندارد نمی‌پذیرند. گرچه این کار ممکن است ناعادلانه به نظر برسد، «کار خوب» انجام‌شده در تحلیل، در صورتی که هیچ‌کس نتواند گزارش نوشتاری ما را بفهمد، ارزش چندانی ندارد.

هدف این قسمت فراهم‌کردن دستورالعمل‌های نوشتن گزارش‌هاست. البته این دستورالعمل‌ها «کیفیت» را تضمین نمی‌کنند ــ محتوای نامفهوم در صورت ساختاربندی خوب هم به همان شکل باقی می‌ماند. به‌علاوه، گزارش‌هایی که در محل کار تهیه می‌شوند اغلب لازم است تابع قیود مشخصی باشند که ممکن است با برخی از دستورالعمل‌هایی که ما در اینجا پیشنهاد می‌کنیم متفاوت باشند. با این حال، برخی اصول کلی وجود دارند که ممکن است در فرایند نوشتن یک گزارش کمک کنند.

- کار ما باید به‌خوبی مستندسازی شده باشد.
- ما تابع یک چارچوب زمانی محدود هستیم.
- گزارش باید مختصر اما قابل فهم باشد.

تهیهٔ مستندات هر چه از اهمیت اطمینان‌یافتن از این مطلب بگوییم که همهٔ کارهای ما به‌خوبی مستندسازی شده و یافته‌های ما به‌صورت مکتوب برای افراد مورد نظر قابل درک باشند، کم گفته‌ایم. وقتی سلسله‌گزارش‌هایی دربارهٔ پروژه‌ای خاص می‌نویسیم، در حال ایجاد مدارکی هستیم که سند همکاری ما در پروژه است. مثال‌هایی از نوع مستندسازی استفاده‌شده در یک پروژه در فصل ۴ ارائه می‌شوند.

چارچوب زمانی معمولاً یک سقف زمانی محدود وجود دارد که انتظار می‌رود اسناد و گزارش‌ها در یک قالب قابل ارائه در دسترس باشند. برای به‌انجام‌رساندن کارآمد اهدافمان، باید سعی کنیم چارچوب‌های زمانی واقع‌بینانه‌ای برای تکمیل گزارش‌هایمان معین و آن را دنبال کنیم.

قابل فهم بودن هر گزارشی که می‌نویسیم معطوف به گروه خاصی از خوانندگان است. گزارش باید قابل فهم باشد و باید در سطح مناسبی برای آن گروه نوشته شود. همچنین گزارش باید، به طریقی معقول، مختصر نوشته شود. گنجاندن اطلاعاتی در نوشته که خواننده آن را می‌داند ممکن است لازم نباشد یا در صورت لزوم تنها به‌طور خلاصه ذکر شود. با این حال، نکات مهم گزارش باید با دقت بسیار به‌تفصیل آورده شود.

وانمود نمی‌کنیم که دستورالعمل‌های ما به خواننده یاد خواهد داد که چگونه در یک زمینهٔ کلی بنویسد، اما باید بینشی را نسبت به ویژگی‌های مورد نیاز در نوشتن گزارش برای گروهمان، مافوق‌هایمان، یا دیگر خوانندگان کارمان فراهم کنند. یا در نظرگرفتن این نکات، سبک بسیار دقیقی از نوشتن گزارش را معرفی می‌کنیم.

طرح کلی پروژه

قالب پروژه که در این کتاب دنبال خواهد شد دارای ساختاری است که در زیر نشان داده شده است. در عمل، ممکن است نیاز به ایجاد اصلاحاتی در این ساختار داشته باشیم، اما این نوشته حداقل این شانس را به خواننده می‌دهد که راه مشخصی برای نوشتن گزارش‌ها فراگیرد.

- صفحهٔ عنوان
- مقدمه
- یافته‌ها
- نتیجه‌گیری
- مراجع
- پیوست‌ها



واقعیت لک‌لک[1]

ارائه‌کننده

آمبر. ال. ادینگ
کورل ا. شان

۱ ژانویهٔ ۲۰۰۰

خلاصهٔ اجرایی

این گزارش در بردارندهٔ شواهدی از یک همبستگی قوی بین جمعیت النبرگ و تعداد لک‌لک‌های مشاهده‌شده طی یک دورهٔ هفت‌ساله است. با این حال، طی این مدت مشاهده آزمایش‌هایی بالینی انجام‌شد که برای۹۵٪ همهٔ زاد و ولدهای ثبت‌شده در النبرگ اقامهٔ دلیل می‌کند. برای ۵٪ مابقی، ظن هیچ اثر علّی نمی‌رود.

شکل ۱.۲ مثال صفحهٔ عنوان.

صفحهٔ عنوان

صفحهٔ عنوان بسیار مهم است زیرا آنچه در آن نوشته می‌شود تأثیر فوق‌العاده‌ای در بالابردن شانس اینکه گزارش ما خوانده و فهمیده شود خواهد داشت. اطلاعات موجود در صفحهٔ عنوان باید به خواننده ایدهٔ شفافی از محتوا و نکات مهم پروژه‌مان بدهد. ساختار صفحهٔ عنوان در قالب مثالی در شکل ۱.۲ نشان داده شده است. این ساختار شامل موارد زیر است:

- عنوان پروژه
- مؤلف(ها)
- تاریخ
- خلاصهٔ اجرایی

1. The Stork Reality

عنوان پروژه عنوان، اسم (اسامی)، و تاریخ، اطلاعات مربوط به ارجاع‌دهی گزارش را فراهم می‌کنند. سعی کنید عنوانی آگاهی‌بخش با «جالب» انتخاب کنید: گزارش پایانی پروژه به خواننده چیزی نمی‌گوید! با این حال، هنگام پرداختن به موضوعی که ممکن است به احساسات افراد بربخورد حواستان جمع باشد. در مورد این موضوع با توجه به مثال اسلاید عنوان با پاورپوینت نشان داده‌شده در شکل ۴.۲، بعداً بیشتر توضیح می‌دهیم.

خلاصهٔ اجرایی باید بسیار کوتاه و مرتبط باشد. خلاصهٔ اجرایی دربردارندهٔ شرح مختصری از نتیجه‌گیری ماست. مسئله را توصیف یا در نوع روش‌شناسی بکار رفته بحث نکنید؛ صرفاً نتیجه‌گیری را بیان کنید. به‌عنوان مثال،

> «قیمت سهام IBM در ماه اکتبر نسبت به ماه توامبر در سطح بالاتری قرار داشت. در اکتبر میانگین قیمت ... بود.»

مثال شکل ۱.۲ نشان می‌دهد که خلاصهٔ اجرایی کوتاه و مرتبط است. به یاد داشته باشید که خلاصهٔ اجرایی به خواننده می‌گوید که چه اتفاقی افتاده است. این یک «چکیده» نیست که به خواننده بگوید چه اتفاقی خواهد افتاد. چکیده‌ها در مقالات چاپ‌شده در مجلات پیش از آن‌ها می‌آیند. خطاب گزارش‌های مشاوره‌ای مشخصاً به مشتریمان برمی‌گردد.

مقدمه

هدف مقدمه توصیف پروژه (در حدودی که به ما محول شده)، و ارائهٔ اطلاعات پیش‌زمینه‌ای لازم است. سعی کنید مختصر باشد اما اطلاعات مربوط را حذف نکنید. آمار توصیفی اساسی، نمودارها و خلاصه‌های داده‌ها را می‌توان در اینجا گنجاند. اگر چندین نمودار یا جدول‌های بزرگ داشته باشیم آن‌ها را در پیوست قرار دهید و در متن به آن‌ها ارجاع دهید.

یافته‌ها

نکات یا فرض‌ها را بیان کنید و آن‌ها را اثبات یا رد کنید. نکته به نکته پیش روید و نشان دهید که چگونه آزمون فرض متناظر را انجام داده‌اید و نتایج چگونه قرار است تفسیر شوند. اگر از نرم‌افزاری آماری که یک فایل خروجی به‌طور قابل ملاحظه بزرگ را تولید می‌کند استفاده کرده‌ایم، اطلاعات مربوط را بیرون بکشید و آن‌ها را در قالب جدول در پیوست قرار دهید. صرفاً همهٔ فایل‌های خروجی را در پیوست تلنبار نکنید. چنین کاری از سوی ما غیرحرفه‌ای محسوب می‌شود و وظیفهٔ خواننده نیست که «تشخیص دهد» اشارهٔ ما در متن به کدام بخش خاص از خروجی است.

باید خاطرنشان کنیم که در برخی قالب‌های گزارشی، لازم است که همهٔ فایل‌های خروجی آورده شوند. به‌عنوان مثال، FDA می‌خواهد که گزارش‌های صنعت داروسازی کد منبع به‌کار رفته در تولید فایل خروجی را دربر داشته باشند.

نتیجه‌گیری

بخش نتیجه‌گیری باید صورت گسترش یافته‌ای از خلاصهٔ اجرایی باشد. هنگام نوشتن نتیجه‌گیری به خاطر داشته باشید که باید به شکل یک تفسیر مضمونی نتایج موجود در بخش یافته‌ها باشد. باید از تکیهٔ زیاد بر اصطلاحات فنی مفرط خودداری کنید. به این معنا که استنتاج‌هایی که انجام می‌دهیم باید مبتنی بر حقایق موجود در یافته‌ها باشند و باید توضیح دهد که یافته‌ها با توجه به اهداف و اصطلاحات فنی پروژه چه معنایی دارند. به‌عنوان مثال، در صنعت، زبان هزینه‌های واقعی ($$$$) در گزارش‌های نوشته‌شده برای مدیریت ارشد دارای اهمیت زیادی است. مرتبط‌کردن نتیجه‌گیری‌هایمان با بهره‌وری و کاهش هزینه‌ها می‌تواند به پذیرش اطلاعات کمک شایانی کند.

مراجع

بخش مراجع فهرستی از کتاب‌ها، مقاله‌ها، نرم‌افزارها و دیگر اسنادی است که در متن به آن‌ها ارجاع داده می‌شود. مراجع نامربوطی را که مشخصاً به پروژه مرتبط نیستند در مراجع نیاورید. برخی مراجع مانند اسناد حقوقی یا گزارش‌های فنی را نیز در صورتی که ارتباط خاصی با پروژه دارند می‌توان در پیوست گنجاند. اگر هیچ مرجعی برای ارجاع نداشته باشیم، در این صورت بخش مراجع را به‌طور کامل حذف کنید؛ لازم نیست حتماً آن را داشته باشیم. قاعدهٔ کلی برای ارجاع، نوشتن اسامی مؤلفان است (نام خانوادگی، سپس حرف اول نام کوچک) و به دنبال آن سال انتشار، عنوان اثر، و سپس اطلاعات کلی نشر به‌صورت مقتضی؛ مجله، دوره، صفحات، ناشر، و غیره. کتابشناسی موجود در این کتاب نمونه‌هایی از بسیاری از قالب‌های ارجاع استاندارد فراهم می‌کند.

پیوست

از پیوست باید برای خروجی کامپیوتری، کد برنامهٔ کامپیوتری، نمودارها، جدول‌ها، یا هر سند مرجع دیگری استفاده کرد که نیازی به ظاهرشدن آن در متن نباشد. دوباره، از پیوست به‌عنوان زباله‌دانی استفاده نکنید؛ پیوست باید در قالبی سازمان‌یافته و مناسب برای خواننده ارائه شود. سعی کنید سیستم نمایه‌سازی مناسبی را در پیوست به‌کار بگیرید که بتوان از آن به‌عنوان مرجعی در هر جایی از خروجی که در متن گزارش ذکر شده باشد استفاده کرد. این کار ممکن است مشتمل بر:

برجسته‌سازی و شماره‌گذاری جاهای مربوط در خروجی و ارجاع شماره‌ها در متن، یا به‌کارگیری عنوان‌بندی‌هایی مانند جدول ۲ب و شکل ۴.۳ باشد. متن یا شماره‌ها را با استفاده از یک ماژیک رنگ روشن مشخص کنید (زرد روشن، رنگی مرجع است) که یافتن آن‌ها را آسان کند. اگر میزان مستندات زیاد باشد، می‌توان پیوست را به چندین بخش تفکیک کرد. به‌عنوان مثال، نمودارها را می‌توان در یک بخش، جدول‌ها را در بخشی دیگر، نامه‌هایی مربوط به یک دعوی قضایی را در بخش دیگر، و به همین ترتیب الی آخر قرار داد. اگر تحلیل‌های متعددی انجام شده باشد، شاید نگهداری نتایج در پیوست‌های جداگانه سودمند باشد. قبل از پیوست‌ها، درج فهرست مطالب پیوست‌ها (علاوه بر فهرست مطالب کل گزارش) سودمند است.

۴.۲ دستورالعمل‌های اساسی برای نوشتن

هدف بخش قبل فراهم‌آوردن ساختاری برای نوشتن گزارش بود. حال لازم است ساختار گزارشمان را با جزئیات واقعی تکمیل کنیم. توجه داشته باشید سعدت‌زمان لازم برای تکمیل جزئیات ممکن است گول‌زننده باشد. یک گزارش تا زمانی که همهٔ بخش‌های آن نوشته نشود کامل نیست. با اینکه این مطلب، بیان بدیهیات به نظر می‌رسد، به خاطر داشته باشید که گزارش ما ممکن است تنها بر مبنای آنچه نوشته شده است مورد ارزیابی قرار گیرد. وظیفهٔ خواننده حدس‌زدن آنچه قصد داشتیم بنویسیم، آنچه اشتباهاً حذف کرده‌ایم، یا قصد داشتیم که «بعداً اصلاحش کنیم» نیست. سبک و کیفیت نوشتاری ما نیز سزاوار توجه است و برخی دستورالعمل‌های کلی را که ممکن است در این زمینه سودمند باشد در نظر می‌گیریم. در عالم واقع، خوب‌نوشتن مستلزم تلاش و تجربه است و دستورالعمل‌ها جایگزینی برای این فرایند یادگیری نیست.

از بین مراجع متعدد در رابطه با نوشتن، گوین و سوان (۱۹۹۰) را ذکر می‌کنیم که در بخش مجلهٔ آمار محاسباتی و نموداری (JCGS[1]) در پایگاه ASA به نشانی `www.amstat.org` آمده است.

برخی مفاهیم پایه‌ای

املا خطاهای تایپی حواس‌پرتکن هستند. غلط‌یاب املایی مطلوب ما ممکن است سودمند باشد، اما مصون از خطا نیست. یک فرهنگ لغت می‌تواند ما را در کنترل نحوهٔ نوشتن متداول یک واژهٔ خاص یاری رساند.

روانی جملات ما باید معنی داشته باشند. خواندن جمله با صدای بلند اغلب می‌تواند در شناسایی مشکلات مربوط به نحو و ساخت مفید باشد.

1. Journal of Computational and Graphical Statistics

پاراگراف‌ها تمام جملات موجود در یک پاراگراف باید در یک موضوع اصلی به هم ارتباط داشته باشند. انحراف از موضوع باعث می‌شود تمرکز خواننده به هم بخورد و این ممکن است باعث درست دریافت‌نشدن مقصود ما شود.

بازنگری‌ها چه همزمان با نوشتن به‌طور مستمر بازنگری کنیم، چه صبر کنیم تا پیش‌نویس کاملی قبل از بازنگری به پایان رسد، هیچ گزارش نهایی نباید بدون انجام‌دادن یک وارسی دقیق ارسال شود. یک راهبرد خوب این است که اجازه دهیم شخص دیگری پیش‌نویس ما را بخواند.

مثال۲.۲ نقاط دورافتاده

اصطلاحات آماری اغلب معنای فنی دقیقی دارند که غلط‌یاب املایی آن‌ها را تشخیص نخواهد داد. دو مثال متداول عبارت‌اند از «برآوردگر[estimator]» و «نقاط دورافتاده[outlier]» که برخی از غلط‌یاب‌های املایی ممکن است تلاش کنند که آن‌ها را با برآورد estimate و outliner جایگزین کنند. متأسفانه، اصطلاح اخیر به قدری زیاد ظاهر شده است که ممکن است به نظر برسد که اصطلاح جدیدی برای «Outlier» در آمار داریم.

مثال۳.۲ مال آن‌ها[1]، آنجا[2]

شکل نخستین، مال آن‌ها، تنها به‌عنوان صفتی که نشان‌دهندهٔ مالکیت یک شخص یا یک گروه باشد به‌کار می‌رود، «این ایدهٔ آن‌ها بود.[3]» باید توجه کنیم که «مال آن‌ها» بین شخص مفرد یا جمع تمایز قائل نمی‌شود[4]: آیا این ایدهٔ یک فرد است یا ایدهٔ یک گروه؟ لازم است مؤلف این را مشخص کند. نتیجه می‌شود که آنجا [There] نباید برای نشان‌دادن مالکیت به‌کار رود. کاربرد نادرست این دو کلمه برای خواننده بسیار واضح است.

سبک

بخش مهمی از آماده‌سازی ما این است که بدانیم «چه کسی» گزارش ما را خواهد خواند. ممکن است لازم باشد آن مشتری که در ابتدا پروژه را آورده است گزارش ما را با افراد دیگر در میان بگذارد. بنابراین، سبک نوشتن ما باید با سطح عمومی دانش آماری گروه خوانندگان مورد انتظار تناسب داشته باشد. به‌علاوه، باید به کیفیت و سبک نوشتاری خود فکر کنیم. انتقال اطلاعات فنی

1. their 2. there 3. it was their idea.

۴. البته این مشکل، خاص زبان انگلیسی است. در فارسی ضمیر واحدی برای سوم شخص زن یا مرد داریم که مشابه مشکل یالاست.م

در راستای دیگر رشته‌های علمی کار غیرممکنی نیست و تنها نیازمند توجه و دقت بیشتر به آن چیزی است که می‌نویسیم. مثال مبالغه‌آمیز زیر سبکی نوشتاری است که **هرگز** نباید از آن استفاده کنیم.

> بر مبنای SS نوع III، برای LS مدل رگرسیون برازش‌دادهشدهٔ MLR-IIb، (پیوست II، قسمت ج) که «Xمتغیرهای» AGE88 و HGT-1 را شامل می‌شد، HGT-1 معنادار نیست (توجه کنید: $P = ۰٫۲۳۴۵$). این بدان معناست که HGT-1، مشروط بر AGE88، سهمی در «پیشگویی [PREDICTION]» پاسخ [response]، INC(Y)، ندارد.

تأکید کلمات یا عبارات مهم را می‌توان با تغییردادن فونت‌ها، یا با استفاده از حروف بزرگ، برجسته کرد. با این حال، اگر بیش از اندازه (یا بی‌هماهنگی) به‌کار روند، اثربخشی به‌سرعت از بین می‌رود و خواندن گزارش ما «خسته‌کننده» می‌شود. به‌طور مشابه، استفادهٔ بیش از حد از سرواژه[1]ها غیرضروری است. خواننده نباید مجبور شود برای خواندن گزارش ما آن را رمزگشایی کند!

وضوح وضوح یک جمله مهم‌تر از طول آن است. هر جمله باید سعی در انتقال یک «بسته»ی اطلاعاتی خودبسنده به خواننده داشته باشد. جملات کوتاه خوب‌اند، اما استفادهٔ مداوم از جملات سادهٔ کوتاه باعث می‌شود نوشتهٔ ما «پُردست‌انداز» به نظر برسد و خواننده احساس کند که شرح ما از ماجرا دائماً در حال قطع‌شدن است. به همین نحو، جملات بلند بی‌هدف که پاره‌های اطلاعاتی ناهمگون را به هم اضافه می‌کنند باعث سردرگمی خواننده خواهند شد. آن‌ها از برهم‌افزایی اطلاعات قبل از آنکه جمله تمام شود دست می‌کشند، تا این مرحله تعبیر و تفسیری برای خود می‌کنند و سپس به آخر مطلب می‌روند.

پاراگراف‌ها یک پاراگراف مرکب از یک سلسله از جملات است که همگی باید در موضوع اصلی با هم ارتباط داشته باشند. اگرچه ساختار پاراگراف به موضوع وابسته خواهد بود، باید جریانی طبیعی از اطلاعات بین جملات وجود داشته باشد. به این معنا که لازم است که مطلب اصلی پاراگراف را به شکلی سازمان‌یافته ارائه کنیم. درگزارش‌نویسی، دو نوع از ساختارهای پاراگرافی عبارت‌اند از **نهاده‌ای** و **پیوندیافته**(مثال ۴.۲ را ببینید)

جزئیات فنی اکثر گزارش‌ها باید شامل جزئیات فنی مرتبط با شیوه‌های آماری به‌کار رفته باشند. اگر از سطح دانش خوانندگان مطلع نیستید، جزئیات فنی آماری را به حداقل برسانید. ارجاع

1. acronym

شیوه‌های آماری به یک مرجع مناسب، از نوشتن یک کتاب درسی در مورد شیوه مناسب‌تر است!

مثال۴.۲ سبک‌های پاراگراف

نهاده‌ای: فرضیه را بیان کنید ⇐ آن را ثابت کنید ⇐ دوباره آن را بیان کنید.

این ساختار، روشی مناسب برای نوشتن خلاصهٔ اجرایی است. این ساختار در توصیف نتایج حاصل از یک نوع شیوهٔ آماری خاص، یا برای بیان نتایج مبتنی بر تحلیل رگرسیونی که انجام داده‌ایم نیز سودمند است. سعی کنید بیش از اندازه در یک پاراگراف اطلاعات مرتبط وارد نکنید. به جای مات و مبهوت ساختن خواننده به‌یکباره مطلب اصلی را به خرده‌مطلب‌هایی به‌هم‌مرتبط بشکنید.

پیوندیافته: جملات پیوندی:

[موضوع الف] ⇐ [اطلاع جدید الف]

[موضوع ب = اطلاع جدید الف] ⇐ [اطلاع جدید ب]

منظور از «⇐» این است که کلمات هر جمله، یک «مطلب» مقدماتی را به یک قطعهٔ اطلاعاتی «جدید» ربط می‌دهند. این اطلاع جدید، مطلب مقدماتی جملهٔ بعدی می‌شود که سپس با قطعهٔ اطلاعاتی «جدید» بعدی ربط می‌یابد، و به همین ترتیب الی آخر. این نوع ساختار ممکن است برای پاراگراف‌های آغازین بخش مقدمهٔ گزارشتان مفید باشد که در آن لازم است جزئیات پروژه را توصیف کنیم: اهداف، متغیرها، اندازهٔ نمونه، و غیره. بنابراین، این ساختار هنگامی مفید است که نیاز به توصیف دنباله‌ای از شرایط یا گام‌های تصاعدی یک تحقیق داشته باشیم.

به‌عنوان مثال، ابزارهای تشخیصی مرتبط با یک برازش رگرسیونی را می‌توان به این صورت به یکدیگر ربط داد : نمودار پراکنش ← مانده‌ها. نمودار مانده‌ها در برابر مقادیر برازش‌یافته ← نرمال‌بودن. آزمون‌های نرمال‌بودن ← نتایج معناداری. غیرمعناداری ← نتیجه‌گیری‌ها.

مثال۵.۲ آپوستروف

از آپوستروف «'» برای نشان‌دادن مالکیت استفاده می‌شود ــ داده‌های مشتری[1] ... و برای ادغام ــ $won't$ ،$can't$ ... .

1. the client's data

اگرچه می‌توان با استفاده از شکل بسط‌یافته از ادغام‌ها پرهیز کرد، این باعث می‌شود که نوشتن ما پرزحمت شود زیرا سبک به‌طور غیرضروری، رسمی به نظر می‌آید. این رسمی‌بودن تأکید خاصی را تحمیل می‌کند که ممکن است ما را از نکته‌ای که قصد انتقال آن را به خواننده داریم دور کند. Don't forget to plot the data [۱] به خواننده یادآوری می‌کند که «نمودار داده‌ها را رسم کند» در حالی که Do not forget to plot the data سرزنش دوستانه‌ای برای «فراموش‌کردن» است.

مثال۶.۲ با این حال [*However*]، **بنابراین**[*thus*]، **در نتیجه**[*hence*]

سعی کنید از نوشتن گزارش به سبک برهان ریاضی خودداری کنید. لازم نیست هر جمله‌ای با یکی از این کلمات وصفی آغاز شود. گرچه ممکن است به استفادهٔ مکرر از آن‌ها در بیان فنی عادت کرده باشیم، این کار نباید در گزارش ضروری باشد. حذف استفادهٔ غیر ضروری از این کلمات وصفی ممکن است نیاز به تلاش آگاهانه‌ای از طرف ما باشد، اما بیانِ صرفِ نتیجه اغلب کافی است.

انگلیسی به‌عنوان زبان دوم (*ESL*)

برای مؤلفانی که زبان مادری آن‌ها انگلیسی نیست، ممکن است بررسی موارد فهرست‌شده در بالا آسان نباشد. طبق تجربهٔ ما، دانشجویان در این وضعیت استفادهٔ صحیح از حروف تعریف و نشانه‌گذاری را، در موارد خاص، پرزحمت می‌یابند. اگر این مطلب در مورد شما صادق باشد، یک کتاب راهنمای مرجع در مورد نگارش فراهم کنید؛ ترجیحاً مرجعی که دربردارندهٔ رهنمودهای ESL [۲] باشد. کتاب راهنمای مرجع مؤلفان، نوشتهٔ دایانا هَکر (۱۹۹۸) شامل رهنمودهای خاص ESL است و در کل مرجع مفیدی خواهد بود. اینکه از کسی که زبان مادری‌اش انگلیسی است بخواهید که نوشتهٔ شما را بخواند نیز سودمند خواهد بود، به شرطی که این را به‌عنوان تجربهٔ یادگیری تلقی کنید و نه راهی برای فرار از مسئله. البته، شما هنوز هم یگانه مسئول کیفیت و محتوای گزارش نهایی هستید. سرانجام اینکه نباید فواید نام‌نویسی در یک کلاس نگارش را، چه ESL باشد چه غیر آن، دست‌کم گرفت.

۵.۲ چگونه ارائه‌کردهای [۳] مؤثر انجام دهیم؟

هنر برقراری ارتباط با مخاطبان، آن‌طور که ممکن است برخی وانمود کنند، تنها یک موهبت ذاتی نیست. افرادی هستند که از موهبت ارتباط کلامی برخوردارند، اما برای بقیه، این امر صرفاً یک

۱. رسم نمودار داده‌ها را از یاد نبرید.

2. English as a Second Language 3. presentation

مسئلهٔ آماده‌سازی است. ممکن است قادر نباشیم با سخنرانان بزرگ رقابت کنیم، اما هر کس که به‌خوبی خود را آماده کند قادر خواهد بود ارائه‌کردی مؤثر انجام دهد.

نوع ارائه‌کردی که آماده می‌کنیم به‌وضوح به مخاطبان ما بستگی دارد. اگر صرفاً نتایج تحلیلمان را برای یک یا دو نفر (مشتریمان) ارائه می‌کنیم، در این صورت یک نمایش چندرسانه‌ای پیچیده در اغلب وضعیت‌ها کاملاً نامناسب است (حالت‌هایی که در آن‌ها مشتری مدیر اجرایی ارشد[1] یک شرکت بزرگ باشد به‌وضوح در زمرهٔ آن اغلب وضعیت‌ها نیست). در حالی که برخی از مسائلی که در زیر مطرح می‌کنیم قطعاً در مورد ارائه‌کردهای یک‌به‌یک صادق است، تأکید در این بخش بر ارائه‌کردهای رسمی است؛ یعنی وضعیتی که در آن لازم است برای مشتری ارائه‌کرد به‌خوبی آماده‌شده‌ای از «محصول» (تحلیل داده‌ها) که هزینه‌اش را او داده فراهم کنیم. لازم است که شرکت‌های مشاوره‌ای کوچک همواره ارائه‌کردهای رسمی انجام دهند، و ما باید در هر پروژهٔ مشاوره‌ای بزرگ انتظار انجام‌دادن همین کار را داشته باشیم.

برای انجام‌دادن ارائه‌کردی مؤثر، در ابتدا لازم است جنبه‌های گوناگون دخیل در زبان غیرکلامی و کیفیت صدا را در نظر بگیریم.

- زبان غیرکلامی
- صدا و بیان
- آماده‌شدن برای ارائه

زبان غیرکلامی

زبان غیرکلامی نقش مهمی در ارائه‌کرد دارد. زبان بدن[2] ما ممکن است، از هر دو دید مثبت و منفی، بسیار بیانگر باشد. برخی از چیزهایی که باید سعی کنیم از آن‌ها اجتناب کنیم این است که بیش از حد عصبی به نظر برسیم، بیش از حد این ور و آن ور برویم، یا از چهره به چهره‌شدن با مخاطب اجتناب کنیم و تنها به پردهٔ تصویر نگاه کنیم. به‌وضوح، آراستگی و وضع ظاهر ما باید با شرایط محیطی که ارائه‌کرد در آن اتفاق می‌افتد وفق کند. به همان میزان مهم است که رفتار ما باید شور و شوقی نسبت به موضوع اصلی برانگیزد. زبان غیرکلامی چیزهای زیادی دربارهٔ ما و کیفیت کارمان انتقال می‌دهد.

1. CEO (Chief Executive Officer)

۲. body language ـ «زبان بدن»، شکلی از ارتباط غیرکلامی است که مرکب است از طرز قرارگرفتن (posture)، ژست، وضع قیافه، و حرکات چشم‌ـم.

کیفیت صدا

صدا و لحن بیان، می‌تواند موجب تسهیل در به مقصد رسیدنِ ارائه‌کرد شود. ما باید یاد بگیریم صدایمان را بدون فریادزدن، برای مخاطبان تنظیم کنیم. تمرین می‌تواند ما را در این امر یاری کند. اگر صدای ما بیش از اندازه آرام است، در این صورت لازم است بلندتر و رساتر صحبت‌کردن را تمرین کنیم. لحن صدا نیز اهمیت زیادی دارد؛ هیچ چیز بیهوده‌تر از این نیست که سعی کنیم به سخنرانی‌یی گوش فرا دهیم که با آهنگی یکنواخت بیان می‌شود. با تغییردادن لحن صدایمان می‌توانیم بر مؤلفه‌های خاصی از ارائه‌کردمان تأکید کنیم تا تأکید بیشتری بر نکات مهم داشته باشیم. کتاب دانکل و پرتْهمْ (۱۹۹۳) ممکن است برای این کار مفید باشد.

آماده‌سازی برای ارائه

به محض تمام کردن تحلیل و گردآوری همهٔ اطلاعاتی که در گزارش آمده یا خواهد آمد، باید برای ارائه‌کردی احتمالی آماده شویم. گام اول این است که سعی کنیم مدت ارائه‌کردمان را مشخص کنیم و ساختار را بر طبق آن آماده سازیم. معمول است که ارائه‌کنندگان بی‌تجربه با تلاش در زیاد سخن‌گفتن یا تهیهٔ تعداد زیادی اسلاید، بیش از اندازه ارائهٔ خود را به درازا می‌کشانند. سعی کنید به میزانی معقول از نظر مطالب و زمان‌بندی با احتیاط عمل کنید. «کمتر» سخن‌گفتن اغلب مؤثرتر از این است که سعی کنیم همه چیز را بیان کنیم.

- ساختار و زمان‌بندی.
- چندرسانه‌ای، پاور پوینت، اُورهد پروژکتور.
- آماده‌سازی اسلاید.
- عینی‌بودن.
- برگه‌های توضیحات.
- تمرین ارائه‌کرد.
- آمادهٔ جواب به پرسش‌های ممکن شوید.
- با یک همکار مشورت کنید.
- وقت ارائه‌کرد را تنظیم کنید!

ساختار

ساختار ارائه‌کرد ممکن است به ساختار گزارش شبیه باشد. با شرح خلاصه‌ای از آنچه قرار است به آن بپردازیم شروع کنید و سپس با مقدمه‌ای کلی دربارهٔ مسئلهٔ تحت بررسی، ادامه دهید. سپس با توصیف مشروح‌تری ادامه دهید و با نتیجه‌گیری خاتمه دهید. مجبور نیستیم همهٔ مطالب گزارش را در ارائه‌کرد بیاوریم و این فرصت را نیز داریم که بیشتر ابتکار به خرج دهیم. گاه در صورت ناآشنایی مخاطبان با برخی از جوانب مبحث ممکن است استفاده از مقدمه در ایجاد انگیزه در آن‌ها مفید باشد.

انتخاب رسانه

مؤلفهٔ دیگری که می‌توان در این مرحله در نظر گرفت رسانه‌ای است که قصد استفاده از آن را داریم. این رسانه ممکن است از طاق سیاه و سفید ساده برای پروژکتور گرفته تا پیچیده‌ترین نمایش چندرسانه‌ای متفاوت باشد. زمان آماده‌سازی متناسب با پیچیدگی رسانه است، اما یک ارائه‌کرد چندرسانه‌ای پر زرق و برق که خوب تهیه شده باشد می‌تواند خیلی مؤثر باشد. از سویی، رسانه جایگزینی برای محتوا نیست و نباید مطالب ارائه‌کرد را تحت‌الشعاع قرار دهد.

یک راه در تسهیل ارائه‌کردمان استفاده از رنگ است. یک ارائه‌کرد نوعی متوسط مستلزم استفاده از رنگ و ظرافت کافی گرافیکی است که استفاده از یک نرم‌افزار ارائه‌کرد مانند پاورپوینت را ضروری می‌سازد.

آماده‌سازی اسلایدها

اکنون زمان آماده‌کردن اسلایدهاست. این کار باید با دقت انجام گیرد. تأکید بر محتوا دارای اهمیت است و نه شکل ظاهر. از جلوه‌های ویژه و «بنجل‌نمودار»[1]ها استفادهٔ نابه‌جا نکنید. این واقعیت که این ابزارها در نرم‌افزار ارائه‌کرد موجودند به این معنا نیست که ما باید از همهٔ آن‌ها در هر اسلاید استفاده کنیم. برخی از نکات دیگری که لازم است هنگام آماده‌سازی اسلایدها به خاطر داشته باشیم عبارت‌اند از:

۱. تعداد جدول‌ها در حداقل ممکن باشد. استفاده از نمودارها و شکل‌ها در انتقال اطلاعات به‌مراتب مؤثرتر است.

۲. حاشیه‌نویسی‌ها باید در حداقل ممکن باشند؛ ممکن است موجب حواس‌پرتی بیننده شود.

۱. junk-chart، اصطلاحی برای اجزای تصویری غیرلازم و گیج‌کننده در نمودارها و جدول‌ها.م.

۳. انتخاب رنگ مهم است. هرگز از رنگ زرد استفاده نکنید (این رنگ برای مخاطبان قابل دیدن نخواهد بود) و از برخی ترکیبات مانند قرمز روی آبی سیر بپرهیزید.

۴. در آخر، اطمینان یابید که تعداد مناسبی از اسلایدها را تهیه کرده‌ایم؛ نه خیلی زیاد، نه خیلی کم.

اگر ارائه‌کرد ما بسیار پیچیده یا نیازمند مقایسه‌های دقیقی از اجزایی از اطلاعات باشد (به‌عنوان مثال، جدول‌ها یا نمودارها)، باید برگه‌های توضیحات برای پخش بین حضار را نیز در نظر داشته باشیم. اگر نیاز به تأکید بر برخی نکات داریم، در این صورت از اسلایدها و یک پروژکتور اضافه استفاده می‌کنیم.

عینی‌بودن

همیشه به یاد داشته باشید که، در جایگاه یک آماردان و یک دانشمند، متعهد به تأمین ارائه‌کردی عینی و صادقانه از اطلاعات هستیم. از بسیاری از ابزارهایی که برای ارائه‌کرد موجودند می‌توان برای سوق‌دادن شواهد به سمت نتیجه‌گیری مشخصی استفاده کرد که توجیه‌پذیر نیست. مهم‌تر اینکه، ما نباید از گفتن اینکه نسبت به یک نتیجه مردد یا نامطمئنیم ترس داشته باشیم. بگذارید مخاطب «مسئله» را دریابد. بر اساس تجربه، اغلب، مخاطبان را بهترین مسئله‌حل‌کن یافته‌ایم.

تمرین

به محض اینکه جزئیات اسلایدها تکمیل شده باشند باید ارائه‌کرد را در محیطی واقع‌بینانه تمرین کنیم. برای ارائه‌کرد زمان بگیرید؛ ارائه‌کردهایی که به وقت اضافه کشیده می‌شوند به‌سرعت توجه مخاطبان را از دست می‌دهند. برای پاسخ‌دادن به سؤال‌ها در طول ارائه‌کرد آمادگی داشته باشید. لازم است این مدت‌زمان در وقت کلی که برای استفاده از آن در ارائه‌کرد برنامه‌ریزی کرده‌ایم منظور شود. انعطاف‌پذیری اسلایدها در اینجا راهبردی مفید است زیرا، پس از اتمام ارائه‌کرد، نخواهیم دانست که چه تعداد یا چه نوع سؤالاتی ممکن است از ما پرسیده شود. این صرفاً بدان معناست که ما باید اسلایدهای مشخصی داشته باشیم که بتوان آن‌ها را بسته به نیاز به اینکه ارائه‌کرد را «تند» و «کند» کنیم نادیده بگیریم یا با میزان متغیری از جزئیات مورد بحث قرار دهیم. این امر نیازمند تمرین است، اما از اینکه یک ارائه‌کرد «شنایزده» به نظر برسد جلوگیری می‌کند:

- هیچ سؤالی پرسیده نمی‌شود و ما کار را در نصف مدت‌زمان تخصیص داده‌شده به ارائه‌کردمان تمام کرده‌ایم؛ در مخاطبان این احساس به وجود می‌آید که ارائه‌کرد ما دارای کمبود محتواست.

- سؤالات زیادی پرسیده می‌شوند و ما باید به انتهای ارائه‌کرد بپریم. زمان‌بندی ما ضعیف تلقی می‌شود و مخاطبان نمی‌دانند که کدام موضوع را از دست داده‌اند.

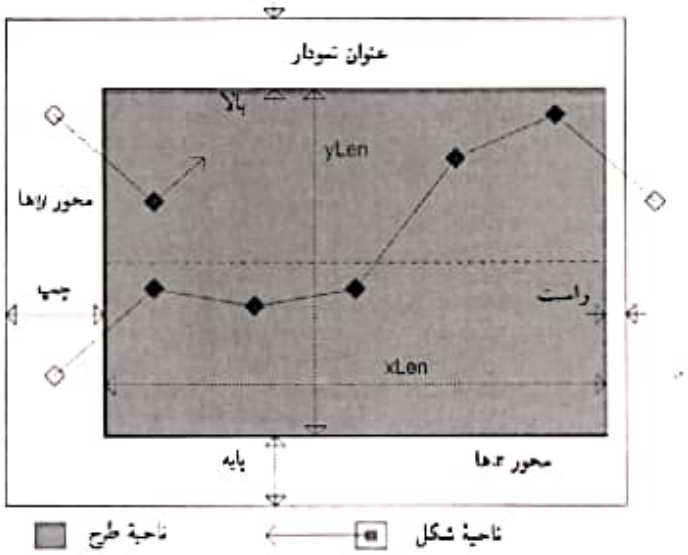
سؤالاتی که در طول ارائه‌کرد مطرح می‌شوند اغلب کوتاه و سریع هستند و مستلزم شفاف‌سازی ساده‌ای از یک اصطلاح، مفهوم یا مدخلی در یک شکل یا جدول‌اند. با این حال، در پایان ارائه‌کرد، ما باید برای پاسخگویی به پرسش‌های جدی‌تری آماده باشیم. ما باید فهرستی از پرسش‌هایی را که ممکن است انتظار داشته باشیم تهیه و پاسخ‌ها را از پیش آماده کنیم. پاسخ‌های ما باید معطوف به همهٔ مخاطبان باشد (نه صرفاً به کسی که سؤال را پرسیده است)، و لازم است زیر و بمی صدا در سطحی باشد که هر شخصی بتواند آن را درک کند. سرانجام، این کار ارزش آن را دارد که ارائه را در برابر اشخاص دیگری انجام دهیم که می‌توانند به ما در تصحیح اشتباهات آشکار و صیقل‌دادن جزئیات ارائه‌کرد یاری رسانند.

۶.۲ اهمیت نمودارهای باکیفیت

زیبایی‌شناسی:	استفادهٔ کارا از ناحیهٔ طرح.
حاشیه‌نویسی:	برچسب‌ها، فهرست علائم و اختصارات، عنوان و زیرعنوان‌ها.
کنتراست‌ها:	مقیاس‌های خاکستری، رنگ، خطوط و نمادها.
ادراک:	انتقال نموداری اطلاعات.

بسیاری از شیوه‌های آماری مبتنی بر ابزارهای تشخیصی نموداری و فنون دیداری‌سازی داده‌ها در جهت اهداف تحلیل هستند. نمودارهای پراکنش، بافت‌نگارها، و نمودارهای میله‌ای‌مثال‌هایی از نمایش‌های گرافیکی ساده اما بسیار مؤثر هستند که می‌توان آن‌ها را در انتقال یافته‌های تلاش‌های مشاوره‌ایمان به‌کار گرفت. در این بخش، توجه خود را به برخی از اصول مهم طراحی گرافیکی که در ایجاد نمودارهای با کیفیت برای مقاصد ارائه‌کرد دخیل‌اند معطوف می‌کنیم. این اصول در مستطیل بالا نشان داده شده‌اند.

نخستین چیزی که باید خاطرنشان کرد این است که تولید گرافیک‌های باکیفیت ارائه‌کرد مستلزم صرف زمان و تلاش است. حتی با یک بستهٔ نرم‌افزاری آماری خوب مانند S-PLUS، اغلب لازم خواهد بود نمایش «پیش‌فرض» در جهت اهداف ارائه‌کرد اصلاح شود. یک فایدهٔ S-PLUS این است که به کاربر اجازه می‌دهد نظارت دقیقی بر مؤلفه‌هایی که یک نمایش گرافیکی را درست می‌کنند اعمال نماید.



شکل ۲.۲ طرح خلاصهٔ ابعادی ناحیهٔ طرح و حاشیه‌ها.

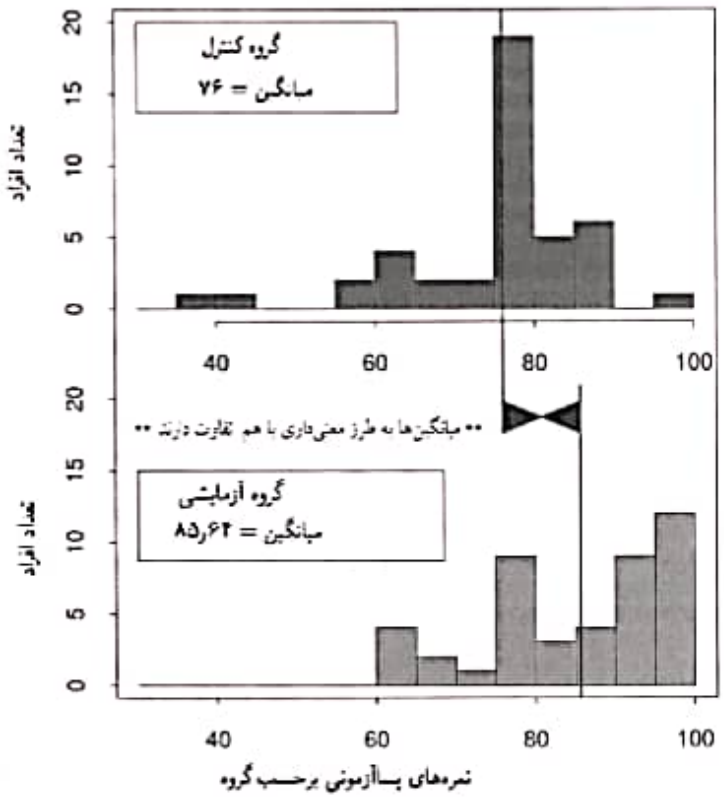
ناحیهٔ طرح

از اصطلاح شکل برای توصیف یک نمایش نموداری کامل استفاده می‌شود که متشکل از یک ناحیهٔ طرح است که با چهار حاشیه (بالا، پایین، چپ، راست)، همان‌گونه که در طرح اجمالی سری زمانی شکل ۲.۲ نشان داده‌شده، احاطه شده است. از حاشیه‌ها برای عناوین، بر چسب‌های محورها، علائم نشانه و برچسب‌ها، و دیگر حاشیه‌نویسی‌ها استفاده می‌شود. حاشیهٔ سمت راست معمولاً دربردارندهٔ هیچ نوشته‌ای نیست و می‌توان آن را به منظور افزایش سطح ناحیهٔ طرح به حداقل رساند. هر جا که عملی باشد، فهرست علائم و اختصارات را نیز می‌توان به‌طور راهبردی درون ناحیهٔ طرح قرار داد (شکل ۳.۲ را ببینید) تا از کاسته‌شدن سطح موجود برای نمایش نموداری جلوگیری کرد.

طرح نموداری

مثال موجود در شکل ۳.۲ استفاده از بافت‌نگار را در نمایش نتایج یک آزمون tی استیودنت دونمونه‌ای مبتنی بر یک بررسی برای مقایسهٔ اثر واردکردن ترجیح‌های سبک یادگیری (گروه آزمایشی) در برابر روش‌های آموزشی سنتی نشان می‌دهد.[۱] از این مثال برای بحث در مورد اصول طراحی گرافیکی مذکور استفاده می‌کنیم.

۱. جزئیات این مطالعه در فصل ۴ ارائه شده است.



شکل ۳.۲ مثالی از تحلیل مقایسه‌ای با استفاده از نمودارها.

زیبایی‌شناسی هدف یک شکل انتقال اطلاعات به‌صورت دیداری است. مقصود از بالابردن کیفیت زیبایی‌شناختی یک شکل ــ«خوب» به نظر رسیدن آن ــ این است که توجه را به نمودار نمایش‌داده‌شده معطوف کنیم. این امر مستلزم استفادهٔ مؤثر از ناحیهٔ طرح و فضاهای حاشیه‌ای است که اغلب نیازمند یک رویکرد آزمون و خطاست. برخی نکات کلی که باید به خاطر سپرده شوند عبارت‌اند از:

۱. نمودار نمایش‌داده‌شده باید ابعاد عمودی و افقی ناحیهٔ طرح را بپوشاند.

۲. باید فضایی کافی در حاشیه‌ها برای نوشتن توضیح‌ها با اندازهٔ قلم بزرگتر در نظر گرفته شود. اندازهٔ قلم پیش‌فرض به‌کارگرفته‌شده برای برچسب محورها اغلب بیش از حد کوچک است که باعث می‌شود خواندن آن برای مخاطبان در جلسهٔ ارائه غیرممکن باشد.

۳. کاستن از پهنای فضای حاشیه‌ای سمت راست (شکل ۳.۲ را ببینید) دارای این عوارض جانبی است که شکل پرت از مرکز به نظر می‌رسد. ترکیب مناسب همهٔ اجزا بخش مهمی از بهبود یک شکل است.

۴. عنوان و برچسب‌های نمودارها متون مهمی هستند. نوشتهٔ اضافی موجب بیشتر برهم‌ریخته به‌نظررسیدن طرح است و بنابراین باید با احتیاط از آن استفاده کرد.

۵. رنگ یا مقیاس‌های خاکستری، انواع خط، نشانه‌های نقطه‌ای، اندازهٔ قلم‌ها، و قرارگیری فهرست علائم و اختصارات را باید امتحان کرد. کنتراست دیناری در یاری‌رساندن به بیننده در تشخیص گروه‌بندی‌ها و دیگر اثرها اهمیت زیادی دارد.

۶. لازم نیست که هر محیط پیش‌فرض ارائه‌شده در بستهٔ نرم‌افزاری را تغییر دهیم. اگر محیط پیش‌فرض قابل قبول است، آن را تغییر ندهید!

حاشیه‌نویسی از حاشیه‌ها برای عناوین، برچسب محورها، علامت‌های نشانه و برچسب‌ها، و دیگر حاشیه‌نویسی‌ها استفاده می‌شود. در شکل ۳.۲، عنوان به جای برچسب محور X نیز به‌کار برده شده است.

۱. برچسب‌های محورها را ساده (اما آگاهی‌بخش) در نظر بگیرید و از علائم نشانه‌ای «دلپسند» استفاده کنید. برای مقیاس‌های کمی، واحد اندازه‌گیری نیز باید در برچسب محور قرار گیرد.

۲. فهرست علائم و اختصارات برهم‌ریختگی دیداری را افزایش می‌دهد. بنابراین، اطلاعات فهرست علائم را به‌دقت انتخاب کنید. در شکل ۳.۲، بافت‌نگارها یک نمایندهٔ بصری برای پراکندگی‌های گروهی و اندازهٔ نمونه‌ها هستند بنابراین فقط میانگین هر گروه در نظر گرفته شده است.

۳. حاشیه‌نویسی‌های دیگری مانند نوشته‌ای نشان‌دهندهٔ نتیجه‌ای معنادار در شکل ۳.۲ ممکن است مفید یا ضروری باشد. به‌عنوان مثال، لازم است خط مبنای به‌کار رفته برای نمایش درصدهای نسبی مشخص شود تا بیننده ارزیابی بامعنایی از نتایج داشته باشد.

کنتراست‌ها اگر بخواهید از رنگ برای تمایز خطوط یا نقاط در ارائه‌کرد استفاده کنید از رنگ زرد استفاده نکنید ــ ممکن است قابل دیدن نباشند.[۱]

۱. یکی از مؤلفان [این کتاب] این شهرت مایه شرمساری را در سابقه دارد که اسلایدی که در اختیار استاد راهنمای رسالهٔ دکترای خود قرار داده بود که در آن سه سری زمانی بر روی هم رسم شده بود. موقع ارائه تنها دو تا از آنها در معرض دید قرار گرفتند.

۱. رنگ برای بیننده طیف وسیعی از کنتراست‌ها را فراهم می‌کند و نیز می‌تواند ابعاد فوق‌العاده‌ای را به نمایش گرافیکی بیفزاید. به‌عنوان مثال، در نیوتن (۱۹۹۳) کدبندی با استفاده از رنگ به‌عنوان یک فن برهم‌افزایی برای نمایش سری‌های زمانی بسیار طویل به‌کار گرفته شده است. همچنین پیچیدگی بیشتری را از نظر تفسیر و زیبایی‌شناختی وارد می‌کند (خوانندهٔ علاقه‌مند را به کار گسترده‌ای که توسط کار (۱۹۹۸) دربارهٔ روش‌های دیداری‌سازی داده‌ها با استفاده از رنگ انجام گرفته ارجاع می‌دهیم).

۲. انتخاب جدول رنگ هنگام تبدیل نمودارهای رنگی به دستگاه مقیاس خاکستری دارای اهمیت است. دستگاه‌های کپی سیاه و سفید اغلب ممکن است نسخهٔ فاقد اطلاعی از یک نمودار رنگی ارائه کنند. به‌دست‌آوردن یک مقیاس خاکستری معقول برای نمودارهای ارائه‌کرد ممکن است به‌طور شگفت‌آوری دشوار باشد. در شکل ۳.۲ «رنگ» بافت‌نگار بالایی همچنان بیش از حد تیره است.

۳. اندازهٔ پیش‌فرض ارائه‌شده در بسیاری از بسته‌های نرم‌افزاری برای نقاط در یک نمودار پراکنش معمولاً برای مقاصد ارائه‌کرد نامناسب است. لازم است به جای این‌ها نشانه‌های بزرگتری همچون لوزی‌های مات در شکل ۲.۲ قرار دهیم.

۴. تعداد خطوط کنتراست‌سازی که می‌توان از آن‌ها به‌طور مؤثری در نمایش‌های مقیاس خاکستری استفاده کرد کاملاً محدود است. بهتر است انواع خطوط سادهٔ خط‌چین و نقطه‌چین را برای خطوط میانهٔ درونی افقی یا عمودی اختصاص دهیم؛ خطوط ضخیمهٔ نامنقطع که با نشانه‌های مختلف علامت‌گذاری شده‌اند راحت‌تر از ترکیب‌هایی از انواع خطوط نقطه-خط از هم تمیز داده می‌شوند. همان‌طور که در شکل ۳.۲ نشان داده شده است پهنای خط (ضخامت) گاه برای خطوط محورها یا برای تأکید افزایش داده می‌شود که در آن میانگین گروه‌ها توسط خط کمی ضخیم‌تر به هم وصل شده‌اند.

درک مطلب مهمترین جزء طراحی گرافیکی اطمینان‌یافتن از این است که نمایش، اطلاعات قابل درکی را منتقل می‌کند. به‌عنوان مثال، نمودار سری زمانی موجود در شکل ۲.۲ درحقیقت نشان می‌دهد که چگونه فن «پیچاندن» یک سری زمانی بر خودش انجام می‌شود ــ دو نقطه‌ای که دو طرف لبهٔ راست از ناحیهٔ طرح را به هم متصل می‌کنند در موقعیت یکسانی نسبت به لبهٔ چپ ناحیهٔ طرح دوباره رسم می‌شوند. این فن دیداری‌سازی داده‌ها توسط مک دوگال و کوک (۱۹۹۴الف،ب) در یک محیط پویای گرافیکی به‌کار گرفته شده است (گرافیک‌های با اثر متقابل در فصل ۳ بیشتر مورد بحث قرار گرفته‌اند). اگرچه مفهوم این فن

پویانمایی بی‌درنگ[1] سرراست است، برای کامل‌کردن شکل ۲.۲ به‌وضوح نیاز به توضیحاتی اضافی بود تا این انتقال اطلاعات انجام شود.

انتقال اطلاعات از طریق گرافیک‌های باکیفیت ارائه‌کرد کار ساده‌ای نیست؛ اما مهارتی است که یک مشاور آماردان لازم است آن را در خود بپرورد. کلیولند (۱۹۸۵، ۱۹۹۳) و سری کتاب‌های تألیف تافت (۱۹۸۳، ۱۹۹۷، ۱۹۹۹) دو مجموعه از مراجعی هستند که دربارهٔ گرافیک آماری و ساختن نمودارهای علمی به‌طور کلی، ارزش خواندن دارند. ما این بخش را با ذکر نکاتی بر مبنای تجارب خود به پایان می‌بریم:

۱. حواستان به شنوندگان باشد ــ نمایش‌های گرافیکی ساده، مانند بافت‌نگار، می‌توانند به همان اندازهٔ نمایش‌های پیچیده مؤثر باشند.

۲. توانایی یک مشاور آماری در تشخیص گرافیک گول‌زننده یا گمراه‌کننده نیز دارای اهمیت است. مشتریان ناخودآگاه گاهی نمایش‌های گرافیکی گمراه‌کننده‌ای درست می‌کنند.

۳. مشتری را تشویق کنید که از یک نمودار میله‌ای به جای یک نمودار کلوچه‌ای استفاده کند. این کار برای اهداف مقایسه‌ای، که نکتهٔ اصلی نمایش است، بسیار مؤثر است.

۴. نقاط موجود در یک نمودار سری زمانی باید همواره به هم وصل باشند. روندها، فصلی‌بودن، نقاط دورافتاده و دیگر خصوصیات مورد نظر را هنگامی که نقاط، به هم وصل باشند بسیار آسان‌تر می‌توان شناسایی کرد.

۵. ناحیهٔ طرح معمولاً با شکل مستطیلی یک صفحه کاغذ در حالت پرتره‌ای[2] یا منظره‌ای[3] در تطابق قرار می‌گیرند. در نتیجه، مقیاس مطلق محورهای X و Y در یک نمودار پراکنش از شکل طبیعی به هم می‌ریزد که ممکن است در برخی وضعیت‌ها گمراه‌کننده باشد. به‌عنوان مثال، یک دایره ممکن است به شکل یک بیضی به نظر بیاید. یک ناحیهٔ طرح مربعی روشی ساده برای تطابق مقیاس‌های مطلق است.

۶. تفسیر ابزارهای تشخیصی گرافیکی پیچیده به تجربه نیاز دارد. همبستگی‌نگارها، نمودارهای Q-Q، و نمودارهای موزاییکی نمونه‌هایی از ابزارهای تشخیصی آگاهی‌بخش هستند که ممکن است تفسیرشان برای بسیاری از مشتری‌ها دشوار باشد زیرا سطح لازم دانش آماری و تجربهٔ آماری لازم را ندارند.

---

1. real-time

۲. portrait: صفحهٔ کاغذ عمودی

۳. landscape: صفحهٔ کاغذ افقی

استفاده از پاورپوینت برای ارائه‌کردها

پاورپوینت ابزاری عالی برای ارائه‌کردهاست و شاید مهم‌تر از آن «قالب»هایی را فراهم می‌کند که تولید یک ارائه‌کرد پایه‌ای خوب را آسان می‌کند. البته، نرم‌افزارهای ارائه‌کرد دیگری را می‌توان برای این مقصود به‌کار برد و انتخاب پاورپوینت از سوی ما عمدتاً به خاطر سهولت به‌کارگیری آن در سطح پیش‌فرض و در دسترس‌بودن گستردهٔ آن است. پاورپوینت دنباله‌ای از اسلایدها را نشان خواهد داد که فرد ارائه‌دهنده هر اسلاید را مطابق با سرعت ارائه‌کرد پیش می‌آورد. در ضمن، می‌توانیم جلوه‌های ویژه مانند زمان‌بندی خودکار، صدا، و پویانمایی را اضافه کنیم. جلوه‌های ویژه قطعاً می‌توانند باعث بهبود شوند و در افزایش شوق‌انگیزی ارائه‌کرد کمک می‌کنند،[1] اما ممکن است آن‌ها بیشتر عامل مزاحمت نیز باشند و توجه را از نکات مهم منحرف سازند. همانند بسیاری از وضعیت‌ها، اعتدال معمولاً برنده است.

آماده‌سازی برای ارائه‌کرد با پاورپوینت از همان گام‌هایی که در بخش قبل مورد بحث قرار گرفت تبعیت می‌کند:

- اسلایدها را به‌دقت تهیه کنید ــ نه خیلی زیاد، نه خیلی کم.
- از به‌کارگیری بسیار زیاد «بنجل‌نمودار»ها بپرهیزید.
- در صورتی که لازم باشد نکات مشخصی را در ارائه برجسته‌سازی کنید، برگه‌هایی حاوی توضیحات مربوط به ارائه‌کرد تهیه کنید یا از اسلایدهای اضافه روی یک پروژکتور استفاده کنید.
- مهم‌تر از همه، تمرین کنید! ارائه‌کرد را تحت آن شرایطی که ارائه در آن انجام می‌شود زمان‌بندی کنید.

تفاوت اساسی این است که این یک ارائهٔ الکترونیکی است که به این معناست که به سخت‌افزار کامپیوتر، کابل، و نرم‌افزار نیاز خواهیم داشت. رایج‌ترین وضعیت آن است که لپ‌تاپ به یک واحد نمایش‌دهندهٔ الکترونیکی که در بالای یک پروژکتور قرار می‌گیرد متصل شده باشد. پیش از ارائه، لازم است کارهای زیر را انجام دهیم:

۱. دریابیم که باید چه تجهیزاتی برای ارائه‌کرد تهیه کنیم.
۲. اطمینان حاصل کنید که لپ‌تاپ با واحد نمایش‌دهندهٔ الکترونیکی سازگار است و کابل‌های اتصال درست را در اختیار دارید.
۳. یاد بگیرید که چگونه واحد نمایش‌دهنده را همراه با لپ‌تاپ به‌کار بیندازید. ممکن است نیاز باشد که تنظیمات کنترلی مشخصی مانند Video Mirroring[2] روشن باشند.

---

۱. ارائه‌کردهایی که بیش از حد فنی باشند ممکن است سبب به خواب‌رفتن مخاطب شوند. افزودن صدا یا پویانمایی به جلب علاقهٔ مخاطبان کمک خواهد کرد حتی اگر موضوع کاملاً فنی باشد.

۲. وقتی کامپیوتر به بیش از یک واحد نمایش یا یک تلویزیون یا پروژکتور متصل باشد، می‌توان تصویر صفحهٔ نمایش اصلی را در واحد نمایش دیگر نیز عیناً نشان داد. این کار Video Mirroring نامیده می‌شود.

۴. طرز استفاده از لپ‌تاپ را بدانید! صرف‌کردن چند دقیقه برای یافتن پاورپوینت یا فایل‌های ارائه‌کردمان شروع خوبی نخواهد بود.

۵. همیشه، همیشه، یک نسخهٔ پشتیبان از ارائه‌کرد همراه داشته باشید. امکان از کارافتادن کامپیوتر هست و ممکن است اسلایدهای پاورپوینت به‌دقت تهیه‌شدهٔ ما را پاک کند.

دستورهای زیر ما را قادر می‌سازد تا ارائه‌کردی پایه‌ای به وجود آوریم. برای قوت‌بخشیدن به ارائه‌کرد می‌توانیم بسیاری از گزینه‌های موجود در پاورپوینت و Wizard tool را بررسی کنیم. نمونه‌هایی از اسلایدهای درست‌شده با پاورپوینت در شکل ۴.۲ و ۵.۲ نشان داده شده‌اند. پیش از ادامه، دوباره بر نکتهٔ مهم زیر تأکید می‌کنیم:

نرم‌افزار ارائه‌کرد، جایگزینی برای محتوا نیست.

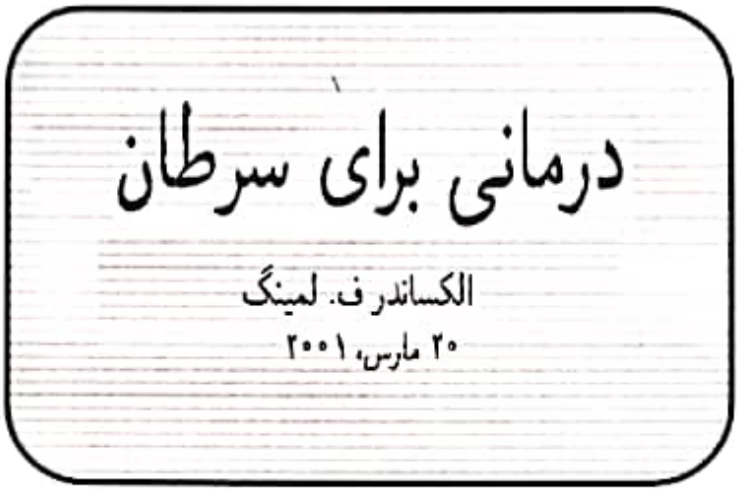
آیا عنوان «درمانی برای سرطان» مورد استفاده در شکل ۴.۲ مناسب است؟ برای افرادی که تحت درمان سرطان بوده‌اند، این عنوان احتمالاً رنج‌آور و قطعاً بی‌احترامی است. در مواجهه با مباحث حساس، سعی نکنید که خیلی بانمک یا تحریک‌آمیز باشد. ممکن است ما قصد سوئی نداشته باشیم، اما دیگر کار از کار گذشته است: آن اسلاید قبلاً با کم‌توجهی ما به حساسیت‌های مخاطبان بر شهرت ما تأثیر سوء گذاشته است.

روی پاورپوینت کلیک کنید این کار، شما را به منویی خواهد برد که می‌توانید Autocontent Wizard را اجرا یا یک قالب را انتخاب کنید. توصیه می‌کنیم که یک قالب ساده را انتخاب کنید. در مثال‌های اسلاید شکل‌های ۴.۲ و ۵.۲ از قالب خالی[1] استفاده شده است. یک رنگ زمینهٔ روشن از منوی color که با انتخاب آیتم Background از منوی Format به آن دست یافته‌ایم اضافه کرده‌ایم.

اسلاید عنوان پاورپوینت، اسلاید عنوان را روی صفحهٔ نمایش می‌آورد و اکنون می‌توانید عنوان ارائه‌کرد، اسم خود، و تاریخ را پر کنید. شکل ۴.۲ مثال ساده‌ای را نشان می‌دهد.

اسلاید جدید سپس بر روی گزینهٔ New Slide از پنجرهٔ Common Tasks کلیک کنید. این کار، پنجره‌ای را باز می‌کند که طرح‌بندی‌های خودکار[2] اسلاید را نشان می‌دهد. آن موردی را انتخاب کنید که به بهترین صورت با هدف مورد نظر اسلاید متناسب است (به‌عنوان مثال، متن تنها، متن با تصاویر، و غیره). اگر طرح‌بندی متفاوتی را می‌خواهید، همیشه

1. blank 2. autolayouts



شکل ۴.۲ مثالی از اسلاید عنوان با استفاده از پاورپوینت.

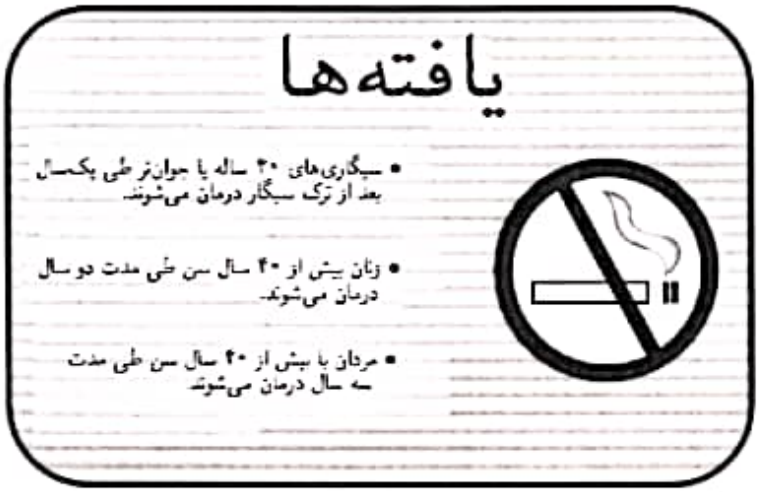
می‌توانید بعداً آن را تغییر دهید. اکنون می‌توانید متن اسلاید دوم را وارد کنید، تصاویر یا اشکالی را در صورت اجرایی‌بودن اضافه کنید.

تصاویر شما می‌توانید از مسیر Picture+From files در منوی Insert یک تصویر را بر اسلاید پاورپوینت قرار دهید. شکل ۵.۲ مثالی از این دست را نشان می‌دهد. پاورپوینت گستره‌ای از قالب‌های گرافیکی از جمله GIF، JPEG، EPS، و PNG را می‌پذیرد.

گرافیک‌های آماری برای واردکردن یک تصویر از نرم‌افزار S-PLUS-PC آن را به‌صورت یک فایل GIF ذخیره کنید. در S-PLUS تحت یونیکس، از درایور گرافیکی PostScript با گزینهٔ onefile=F استفاده کنید، بنابراین، این کار هر تصویر را بر فایل متفاوتی در قالب EPS ذخیره خواهد کرد. در SAS می‌توانید آیتم export از منوی file را انتخاب کنید و قالب GIF را برگزینید. هر یک از این قالب‌های فایل‌های تصویری را می‌توان توسط پاورپوینت خواند.

مرور برای اجرای ارائه‌کرد View Show را از منوی Slide Show انتخاب کنید. برای رفتن به اسلاید بعدی روی retutn کلیک کنید.

روشی جایگزین برای تولید اسلایدهایی برای ارائه‌کرد ایجاد فایل‌های HTML است. یک روش ساده برای تولید فایل‌های ارائه‌کرد HTML، استفاده از حروفچین Netscape است که یک



شکل ۵.۲ مثالی در اسلاید یافته‌ها با استفاده از پاورپوینت.

ویراستار صفحهٔ وب است که استفاده از آن آسان است و در ویرایش‌های آخر Netscape می‌آید. Netscape به‌طور رایگان برای اکثر سازمان‌های غیرانتفاعی موجود است و بسیاری از شرکت‌ها دارای مجوز سایت هستند. حروفچین Netscape کم و بیش مانند یک ویراستار عمل می‌کند. می‌توانیم نوع قلم، ظاهر، اندازه، و رنگ را تغییر دهیم و تصاویری را در چندین قالب به متن اضافه کنیم. همچنین می‌توانیم سرپیوند[1]هایی به دیگر صفحات وب، هم بر روی دیسک محلی هم بر روی اینترنت، اضافه کنیم. میکروسافت‌آفیس، که دربردارندهٔ پاورپوینت، وُرد، و اِکسل است، مورد استفادهٔ اغلب محیط‌های کسب و کار هستند و این برنامه‌های کاربردی از قالب HTML برای ارائه‌کردها پشتیبانی می‌کنند.

چه از پاورپوینت و چه از حروفچین وب استفاده کنیم، ارائه‌کرد نیازمند میزان معقولی زمان برای آمادگی و تمرین است. این یک کار پنج‌دقیقه‌ای نیست.

1. hyperlink

۳

جنبه‌های روش‌شناختی

هدف این فصل، مروری بر برخی فنون استاندارد است که عموماً در مشاورهٔ آماری به‌کار می‌روند. در پایبندی به حال و هوای این کتاب، تأکید در عرضهٔ مطالب از منظر مشتری است و نه جزئیات فنی روش‌ها. پیوست ج مقداری جزئیات فنی و جدول‌ها در اختیار می‌گذارد، اما شرح جامع‌تری از این روش‌ها را می‌توان در جای دیگری یافت. مراجع مربوط به روش‌های مورد بحث را ارائه می‌کنیم و مروری بر نرم‌افزارهای آماری در انتهای این فصل انجام شده است.

البته، پیش از آنکه بتوانیم هر شیوهٔ آماری را به‌کار بگیریم به داده‌هایی نیاز داریم. از این رو، با نگاهی بر روش‌هایی برای گردآوری داده‌ها آغاز می‌کنیم.

۱.۳ گردآوری داده‌ها

روش‌های متعددی وجود دارند که می‌توان در گردآوری داده‌ها از آن‌ها استفاده کرد؛ اما به‌منظور استخراج نتایج معتبر، گردآوری داده‌های «خوب» دارای اهمیت است. متأسفانه، بسیاری از مشتریان

بعد از گردآوری داده‌ها در پی مشورت‌گرفتن از یک مشاور آماری می‌روند. البته، این امر عمدتاً تقصیر ماست. در مقام آماردان، اغلب این تقصیر متوجه ماست که تکیه بر فنون داریم و بر اهمیت نقش مهمی که مشاور آماری می‌تواند طی مرحلهٔ طراحی یک تحقیق آماری داشته باشد تأکید نمی‌کنیم. با این حال، یک مشتری آشکارا حق دارد بداند که مزایای مشورت‌کردن با یک مشاور طی مرحلهٔ برنامه‌ریزی چیستند. چت‌فیلد (۱۹۹۵) سه مؤلفهٔ یک مطالعهٔ برنامه‌ریزی‌شده را که در آن‌ها نظر یک مشاور آماری می‌تواند به‌طور ویژه‌ای مفید باشد فهرست کرده است:

- تعیین متغیرهای مهم.
- مشخص‌کردن اهداف مطالعه.
- فرمولبندی مسئلهٔ آماری.

اهداف یک پروژه باید به‌خوبی فرمولبندی شوند تا نتایج، تفسیر معناداری داشته باشند. این امر مستلزم مشخص‌کردن اهداف «صحیح» برای پروژهٔ مشتری است که همیشه هم کار آسانی نیست. مشاور باید دقت کند که کارش به دادن پاسخ صحیح به سؤال نادرست مشتری نکشد. این را گاهی «خطای نوع III» یا خطای نوع سوم می‌نامند (کیمبال، ۱۹۵۷). ما به این موضوعات در فصل قبل در قالب تعامل کلامی با مشتری طی یک جلسهٔ مشاوره پرداخته‌ایم و اکنون توجه خود را به خود فرایند گردآوری داده‌ها معطوف می‌کنیم.

روش‌های گردآوری داده‌ها

فرایند گردآوری داده‌ها و اهداف یک مطالعه ممکن است همواره در مرزبندی سنتی بین آزمایش طراحی‌شده در برابر یک مطالعهٔ مشاهده‌ای قرار نگیرد. در حالی که قطعاً سودمند است که بین مفاهیم پایه‌ای مرتبط با این روش‌های گردآوری داده‌ها تمایز قائل شویم، اما تا حدودی این گرایش بین آماردانان هست که به نتیجه‌گیری‌های مبتنی بر مطالعه‌های مشاهده‌ای سوءظن داشته باشند. این موضوع بی‌علت نیست. اثرهای تیماری، در یک آزمایشِ به‌خوبی طراحی‌شده، دارای تفسیر روشنی هستند در حالی که مطالعه‌های مشاهده‌ای اغلب حاوی پاسخ‌های ذهنی هستند که ممکن است کمّی‌کردن و تفسیر آن‌ها دشوار باشد. از این رو، اثرهای جالبی که از یک مطالعهٔ مشاهده‌ای پدیدار شده‌اند ممکن است لزوماً واقعی نباشند.

ما به‌اختصار، انواع زیر از روش‌های گردآوری داده‌ها را که مشاور آماری عموماً با آن‌ها مواجه خواهد شد در نظر می‌گیریم.

- مطالعه‌های مشاهده‌ای
- بررسی‌های نمونه‌ای
- مطالعه‌های طولی
- آزمایه‌های بالینی
- آزمایش‌های طراحی‌شده

مطالعه‌های مشاهده‌ای

در یک مطالعهٔ مشاهده‌ای، محقق عموماً نقشی انفعالی دارد و صرفاً افراد هم‌گروه خاصی را مورد مشاهده قرار می‌دهد و پاسخ‌ها را گردآوری می‌کند. در حالی که واژهٔ «انفعالی» ممکن است اولین صفتی نباشد که هنگام مواجههٔ ما با یک بازاریاب از راه دور[1] به ذهن می‌رسد، فرایند ترغیب به خرید عملاً نتیجهٔ یک مطالعهٔ مشاهده‌ای است: تحلیل نمایه‌ای مبتنی بر پایگاه داده‌های موجود مرکب از اطلاعات مرتبط با عادات هزینه‌کردهای ما، اطلاعات جمعیت‌شناختی[2]، کاربرد اینترنت، و غیره. هدف تحلیل نمایه‌ای شناسایی منبع بزرگی از افراد بالقوه‌ای است که واکنش نشان داده‌اند، زیرا حتی درصد کمی از موارد مثبت ممکن است سودآور باشد. برای جزئیات بیشتر در مورد برنامه‌ریزی و تحلیل مطالعه‌های مشاهده‌ای کوکران (۱۹۸۳) را ببینید.

دستور کار

همان‌طور که از مثال بازاریابی از راه دور برمی‌آید، ممکن است فرایند گردآوری و تحلیل داده‌های مطرح در مطالعه‌های مشاهده‌ای کاملاً گسترده باشد. در مطالعه‌های بزرگتر، اغلب داشتن یک دستور کار[3] کتبی ضروری است. این بدان علت است که بخش اعظم (اگر نگوییم همهٔ) فرایند گردآوری داده‌ها توسط افرادی اجرا خواهد شد که آموزش آماری ندیده‌اند. به‌عنوان مثال، در یک بررسی نمونه‌ای، «مصاحبه‌کننده» می‌خواهد بداند که وقتی پاسخ‌دهندهٔ مورد نظر در دسترس نباشد چه باید بکند ــ آیا لازم است فرد مصاحبه‌کننده بعداً مراجعه کند؟ اگر چنین باشد کی و چند مرتبه؟ بنابراین، دستور کار کتبی باید به‌تفصیل دستورالعمل‌هایی دربارهٔ هر یک از جنبه‌های فرایند گردآوری داده‌ها را در اختیار بگذارد. هدف دستور کار این است که خطاهای غیر تصادفی را کنترل کند.

1. telemarketer

۲. demographics، ویژگی‌های جوامع انسانی و زیرجامعه‌ها، بویژه وقتی برای شناسایی بازارهای مصرف به‌کار گرفته می‌شود ــ م.

3. protocol

بررسی‌های نمونه‌ای

بررسی‌ها و نظرخواهی‌های نمونه‌ای به‌صورت روزافزونی متداول شده‌اند و به‌طور گسترده‌ای در تحقیقات بازار به‌کار می‌روند. گرچه مفهوم و اهمیت انتخاب یک نمونهٔ تصادفی عموماً به‌خوبی دریافته شده است، نرخ پاسخ در بسیاری از بررسی‌ها ممکن است بسیار ناچیز باشد. در برخی موارد، یافته‌ها ممکن است کیفیت پایینی نیز داشته باشند. بنابراین، یک فرمت طراحی خوب برای کمک به حداکثرکردن نرخ پاسخ و بهبودبخشیدن کیفیت داده‌های گردآوری‌شده مهم است. بررسی‌های نمونه‌ایِ خوب‌طراحی‌شده روشی مقرون به صرفه برای برآورد درست مشخصه‌های جامعه فراهم می‌کنند. در اینجا، برخی از حیطه‌های اصلی مورد توجه در بررسی‌های نمونه‌ای را با شروع از یادآوری زیر خلاصه می‌کنیم.

درستی یا یک نمونهٔ تنها تضمین نمی‌شود.

این بدان معناست که، صرف‌نظر از اینکه یک بررسی نمونه‌ای چقدر خوب طراحی شده و خوب اجرا شده باشد، ما همواره ناچار به تبعیت از نمایه‌ای هستیم که نمونهٔ (تنهای) گردآوری‌شده حاوی آن‌هاست. کوکران (۱۹۷۷) کتاب درسی کلاسیکی دربارهٔ نظریهٔ آماری مرتبط با روش‌های نمونه‌گیری است. دیلمان (۱۹۷۸)، تامپسون (۱۹۹۲)، و فینک (۱۹۹۵) کتاب‌های دیگری هستند که به روش‌های نمونه‌گیری پرداخته‌اند.

فرمت طرح فرمت طرح یک بررسی بسیار مهم است زیرا می‌تواند تأثیر بسیاری برکیفیت داده‌های گردآوری‌شده بگذارد. سؤال‌ها باید ساده و روشن باشند تا موجب تعابیر و تفاسیر مختلف نشوند. پاسخ‌دهنده را راحت بگذارید تا به هر سؤال پاسخ بدهد یا پاسخ ندهد. بی‌پاسخی‌ها باید از رسته‌ای مانند «شامل حال نیست» قابل تشخیص باشند. شاید نکتهٔ مهم‌تر این باشد که سعی نکنید بیش از اندازه سؤال بپرسید.

روش نمونه‌گیری هدف از یک نمونه استنباط چیزی دربارهٔ یک جامعه است و برای تضمین اعتبار آماری نتایج به نمونه‌ای تصادفی نیاز داریم. می‌توان نشان داد که نمونه‌گیری تصادفی ساده (SRS) راهبرد نمونه‌گیری بهینه‌ای است (کوکران، ۱۹۷۷)، اما معمولاً کاربرد مستقیم آن غیر عملی است. روش نمونه‌گیری باید با ویژگی‌های جامعهٔ مورد نظر سازگاری داشته باشد، اما در ضمن باید بالاسری‌های عملی همچون زمان و هزینه را نیز برآورده کند. برخی از طرح‌های نمونه‌گیری استاندارد در زیر ذکر شده‌اند.

نمونه‌گیری تصادفی ساده نمونه‌گیری تصادفی ساده معادل با استخراج اسامی از یک کلاه بدون جایگذاری است. به این معنا که هر واحد در یک جامعهٔ مورد نظر، شانس برابری

برای انتخاب‌شدن داشته باشد. در بسیاری از طرح‌های بررسی نمونه‌ای، نمونه‌گیری تصادفی ساده عمدتاً درون زیرگروه‌های مشخص‌شده در جامعه به‌کار می‌رود.

طبقه‌بندی برای اجتناب از بیش‌نمونه‌گیری از گروه (طبقه)هایی که در یک جامعه غالبند، نمونه‌گیری تصادفی طبقه‌بندی‌شده را می‌توان بر مبنای سهم‌های نسبی طبقه‌های موجود در جامعه به‌کار برد. سهم‌های طبقه‌ای جامعه باید از قبل معلوم باشد (یا از یک بررسی مقدماتی برآورد شوند) و نمونه‌گیری تصادفی ساده برای انتخاب یک نمونه از درون هر طبقه به‌کار می‌رود.

خوشه‌ای نمونه‌گیری تصادفی از بین تمام جامعه یا طبقه اغلب غیر عملی است. در عوض، نخست خوشه‌هایی همگن مانند شهرها مشخص می‌شوند و نمونه‌ای تصادفی (با استفاده از نمونه‌گیری تصادفی ساده) از این خوشه‌ها برای تولید نمونهٔ نهایی انتخاب می‌شود.

چندمرحله‌ای در جوامع بزرگ و متنوع، ممکن است چندین مرحله خوشه‌بندی و طبقه‌بندی برای تولید نمونه‌ای مقرون به صرفه مورد نیاز باشد.

چندفازی بررسی‌های مقدماتی، پیگیری‌ها، و طرح‌های نمونه‌گیری دوگانه را می‌توان برای بهینه یا تعدیل‌کردن اندازهٔ نمونهٔ کل به‌کار برد. نمونه‌گیری دوگانه اغلب در وضعیت‌هایی به‌کار می‌رود که در آن‌ها فقط لازم باشد که کمیت مورد نظر (پارامتر جامعه) را با دقت مشخصی برآورد کنیم.

اندازهٔ نمونه قصد از یک نمونه برآوردکردن برخی از پارامترهای مورد نظر مربوط به جامعه است. دقت برآورد نمونه‌ای (آماره) از طریق واریانس نمونه‌ای آن به اندازهٔ نمونه بستگی دارد. در عمل، هزینه‌های واقعی در گردآوری داده‌های بررسی دخیل‌اند که با اندازهٔ نمونه افزایش می‌یابند. معمولاً ایجاد توازنی بین هزینه و دقت مورد نیاز ضروری است.

اجرا نکتهٔ کلیدی در به‌دست‌آوردن نمونه‌ای معتبر، انتخاب تصادفی است و برای اجتناب از اریبی انتخاب، از اعداد تصادفی استفاده می‌شود. نتایج اریب هنگامی حاصل می‌شوند که نمایهٔ نمونه‌ای با نمایهٔ جامعهٔ مورد نظر متفاوت باشد. بی‌پاسخی و کم‌پوششی، دو منبع رایج اریبی هستند که می‌توان از طریق بررسی‌های بعدی آن‌ها را به میزان جزئی کاهش داد. با این حال، حتی جمله‌بندی یک سؤال ممکن است تأثیر قابل ملاحظه‌ای بر پاسخ داشته باشد ــ سؤالات راهنما اغلب برای کمک به ایجاد محبوبیت برای مناصب سیاسی و محصولات بازار به‌کار می‌روند.

مطالعه‌های طولی

نوع خاص‌تری از مطالعهٔ مشاهده‌ای در گردآوری داده‌های طولی مطرح است. در این وضعیت، هر واحد در هم‌گروه نمونه‌ای در چندین موقع در طول یک دورهٔ زمانی اندازه‌گیری می‌شود. کتاب‌های تألیف پلی‌ویس (۱۹۸۵) و سلوین (۱۹۹۶) به تحلیل آماری داده‌های طولی می‌پردازند. هدف این نوع مطالعه ارزیابی تأثیر یک متغیر تبیینی مانند X بر متغیر پاسخ مانند Y تحت شرایط کنترل شده است. یک کاربرد نوعی می‌تواند چنین باشد:

مثال۱.۳ عوارض جانبی یک داروی خاص چیست؟

عوارض جانبی مرتبط با یک دارو ممکن است تنها با استفادهٔ طولانی‌مدت از آن بروزکند و ممکن است به دُز دارو نیز بستگی داشته باشد. در این حالت، یک گروه به دقت انتخاب‌شده از افراد ممکن است در طول یک دورهٔ زمانی تحت پیگیری قرار گیرند و اثر دُز (X) بر عوارض جانبی متعدد (Yها) مشاهده شود.

اگر مقادیر ثبت‌شدهٔ X و Y از داده‌های پیشینه‌ای به‌دست آمده باشند، در این صورت مطالعهٔ طولی را گذشته‌نگر می‌نامند. مطالعه‌های گذشته‌نگر آشکارا به کیفیت داده‌های پیشینه‌ای بستگی دارند، اما این مزیت را دارند که می‌توان مسئلهٔ مورد نظر را فوراً مورد تحلیل قرار داد. وقتی قرار است Y در طی مطالعه تولید شود، آن را آینده‌نگر می‌نامند. مثال‌هایی در زیر آورده شده‌اند.

گذشته‌نگر متغیر پاسخ Y از روی داده‌های پیشینه‌ای معلوم است و توجه معطوف به اثر متغیر تبیینی X است: ۵۰ فرد سیگاری و ۵۰ غیر سیگاری به‌صورت گذشته‌نگر بر اساس سن در زمان مرگ رده‌بندی می‌شوند.

آینده‌نگر سطوح متغیر تبیینی X در آغاز مطالعه ثابت گرفته می‌شوند و توجه به مقادیر آیندهٔ متغیر پاسخ Y است: ۵۰ سیگاری و ۵۰ غیر سیگاری برای یک مطالعهٔ طولی مربوط به وقوع سرطان ریه انتخاب می‌شوند.

آزمایه‌های بالینی

آزمایه‌های بالینی در فصل ۱ در قالب اصطلاحات فراوری و آزمون‌کردن داروها توسط صنعت داروسازی مورد بحث قرار گرفتند. وقتی یک مشتری یا آماردان به یک آزمایهٔ بالینی اشاره می‌کند، معمولاً منظور او آزمون تمام و کمال انجام‌شده در مطالعه‌های فاز III است. فلیس (۱۹۸۶)، پیکاک (۱۹۸۳)، مینرت (۱۹۸۶)، و فریدمن و همکاران (۱۹۸۵) کتاب‌هایی هستند که مشخصاً

به آزمایه‌های بالینی می‌پردازند. مؤلفه‌های عمدهٔ یک طرح آزمایه‌ای بالینی عبارت‌اند از:

- گروه کنترل
- دستور کار کتبی
- تصادفی‌سازی

کنترل گروه کنترل مرکب از بیمارانی است که درمان متعارف یا یک داروننما را دریافت می‌کنند. یک داروننما شکل «دروغین» آن دارویی است که به بیماران در گروه تیمار تجویز شده است و دربردارندهٔ مواد کاملاً بی‌اثر است و معیاری موسوم به «اثر داروننما» مرتبط با تأثیر روانی خوردن دارو در اختیار محقق می‌گذارد. ممکن است چندین گروه تیمار برای ارزیابی اثربخشی یک داروی واقعی با دُزهای مختلف وجود داشته باشند.

دستور کار آزمایه‌های بالینی فاز III ممکن است نوعاً دربردارندهٔ هزاران بیمار باشد که در پایگاه‌های آزمایش متعددی پراکنده‌اند. در برخی از مطالعه‌ها، این پایگاه‌ها ممکن است حتی در کشورهای مختلف باشند. بنابراین، مدیریت آزمایه‌های بالینی کاملاً پیچیده است و همان‌گونه که در بالا خاطرنشان کردیم، فرایند گردآوری داده‌ها عمدتاً به دست پرستاران، پزشکان، و دیگر کارکنانی انجام خواهد شد که هیچ‌گونه مهارت آماری رسمی ندارند. در چنین وضعیتی، یک دستور کار کتبی ضروری است زیرا امکان تخلف از دستور کار تقریباً به‌طور یقین وجود دارد. به عنوان مثال، یک بیمار ممکن است به خاطر مسمومیت غذایی موقتاً مجبور به قطع مصرف دارو باشد یا ممکن است یک بیمار از قواعد مصرف دارو تبعیت نکند. به همین نحو، تخلف ممکن است ناشی از این باشد که کارکنان اجرایی به‌درستی از شیوه‌های استاندارد پیروی نمی‌کنند.

دستور کار کتبی برای یک آزمایهٔ بالینی تمام اطلاعات مرتبط با مطالعه را به‌طور مستند ثبت می‌کند. این دستور دربردارندهٔ اقلامی همچون موارد زیر است:

- مقصود از مطالعه و الزامات واجد شرایط بودن برای اینکه یک بیمار در مطالعه گنجانده شود.
- اندازهٔ نمونه و روش تخصیص بیماران در گروه‌های تیمار و کنترل.
- برنامهٔ زمانی درمان و شیوه‌های ارزیابی پاسخ بیمار. آموزش لازم برای به‌کارگیری برخی ابزارهای ارزیابی.

- شیوه‌های تصمیم‌گیری در مواجهه با تخلف از دستورکار.

اطلاعات خام از مطالعه معمولاً در یک فرم گردآوری می‌شود که توسط شخصی که پاسخ بیمار را ارزیابی می‌کند پر می‌شود. بعضی فرم‌ها ممکن است دربردارندهٔ چندین صفحه اطلاعات باشند که همگی باید در یک پایگاه دادهٔ الکترونیکی وارد شوند. یک فرمت طراحی مناسب برای کمک به تضمین کیفیت داده‌های به‌دست آمده در طول فرایند گردآوری داده‌ها به مطلوبیت کار کمک می‌کند.

تصادفی‌سازی تخصیص تصادفی بیماران به سطوح تیماری به‌خصوص از دیدگاه آماری حائز اهمیت است زیرا این کار تنها روش عینی برای مقابله با منابع بالقوهٔ اریبی است. با این حال، تصادفی‌سازی به خودیِ خود برای حذف منابع غیر تصادفی اریبی از جمله اینکه پزشکی که «می‌داند» یک بیمار تحت چه تیماری است، کافی نیست. پزشک ممکن است مراقبت بسیار خوبی در خصوص همهٔ بیماران مورد مطالعه به‌عمل آورد، اما آشکارا انتظار هیچ‌گونه اثر تیماری در بیماری که دارونما دریافت می‌کند ندارد. اجرای یک برنامهٔ دوسرکور[1]، که در آن هم بیمار و هم پزشک هیچ‌یک اطلاعی از اینکه بیمار چه تیماری دریافت می‌کند ندارند، می‌تواند برای غلبه بر این مشکل به‌کار گرفته شود.

آزمایه‌های دارویی تصادفی‌ساخته برای ارزیابی درمان‌های جدید به‌صورت اطمینان‌بخشی بسیار مؤثر از کار درآمده‌اند، اما در عین حال برخی موارد اخلاقی نیز پیش آورده‌اند. آیا واقعاً عادلانه است که اجازه دهیم بیماران گروه کنترل درمان‌نشده باقی بمانند؟ این پرسش، در صورتی که بیمار دچار بیماری یا ویروسی است که اگر بدون درمان رها شود ممکن است کشنده باشد، جدیت بیشتری پیدا می‌کند. پاسخ آماری این است که یک گروه کنترل برای مقاصد مقایسه‌ای مورد نیاز است . پرسش اخلاقی آشکارا پاسخ ساده‌ای ندارد.

آزمایش‌های طراحی‌شده

آخرین روش گردآوری داده‌ها که در نظر می‌گیریم مربوط به مطالعه‌های مبتنی بر طرح‌های آزمایشی است. در این پژوهش‌ها، اثر یک یا چند تیمار بر متغیر پاسخ، مورد نظر است. در بخش ۷.۳ دوباره به طراحی آماری آزمایش‌ها بازمی‌گردیم و برخی طرح‌های خاص را که استفاده از آن‌ها در عمل متداول است در نظر می‌گیریم. کاربرد عملی طراحی آزمایشی در کتاب باکس و همکاران (۱۹۷۸) مورد تأکید قرار گرفته است (این کتاب باید در کتابخانهٔ هر مشاور آماری موجود باشد). اصول اساسی طراحی آماری آزمایش‌ها در زیر فهرست می‌شود.

1. double-blind

- کنترل
- تصادفی‌سازی
- تکرار

کنترل اهمیت گروه کنترل و تخصیص تصادفی‌ساختهٔ تیمارها به بیماران در آزمایه‌های بالینی در بالا مورد بحث قرار گرفت. مقایسهٔ چندین تیمار ساده‌ترین کاربرد اصل کنترل است. به‌صورت کلی‌تر، این مفهوم به مسئلهٔ حذف منابع خطای سیستماتیک (اریبی) و کنترل تأثیرات متغیرهای دیگر بر پاسخ اشاره دارد. به‌عنوان مثال، بلوک‌بندی عوامل، گروه‌بندی طبیعی مشاهدات را به‌کار می‌گیرد. با کنترل‌کردن تغییرپذیری بین بلوک‌ها در یک طرح آزمایشی، مقایسهٔ دقیق‌تری از میانگین‌های تیماری امکان‌پذیر است.

تصادفی‌سازی مزیت اصلی آزمایش‌های به‌خوبی‌طراحی‌شده در مطالعه‌های مشاهده‌ای این است که آزمایش‌ها می‌توانند شواهد خوبی برای علیت‌ها فراهم کنند. با کنترل‌کردن دیگر عامل‌ها و حذف اریبی، تفاوت‌ها در پاسخ مشاهده‌شده را می‌توان به اثر تیمار نسبت داد. با زدودن داوری ذهنی محقق، و تخصیص آزمودنی‌ها به گروه‌های تیمار، مستقل از هر مشخصهٔ واحدهای آزمایشی، تصادفی‌سازی برای ایجاد گروه‌های قابل مقایسه، تنها متکی بر شانس است. یک طرح کاملاً تصادفی‌شده لزوماً عملی یا حتی مطلوب نیست (به‌عنوان مثال، بلوک‌بندی). در این وضعیت، به نوعی از تصادفی‌سازی مقید نیاز داریم.

تکرار این یک مسئله در آزمایه‌های بالینی نیست، اما همهٔ شاخه‌های علوم آن بودجه‌ای را که از شرکت‌های دارویی انتظار می‌رود ندارند. در واقع، قصد بسیاری از طرح‌های آزمایشی این است که برای محقق روشی کارآمد به‌منظور تحلیل همزمان چندین اثر تیماری فراهم کنند. به این معنا که بتوان تعداد کل ترکیب‌های تیماری یا تعداد اجراهای مورد نیاز را در حداقل نگاه داشت (ما توصیف رسمی‌تری از چنین طُرح‌هایی را به قسمت‌های بعدی این فصل وامی‌گذاریم). با این حال، تکرار حداقل برخی از اجراها به‌منظور ارزیابی خطای آزمایشی مورد نیاز است. مانند همیشه، توازنی بین اندازهٔ نمونه (تعداد اجراها) و دقت (حساسیت آزمایش در تشخیص تفاوت‌های تیماری) که محقق باید با آن مواجه شود وجود دارد. نقش یک مشاور آماری این است که رهنمودی بر انتخاب سازش مناسبی بین این دو گروه ارائه کند.

گاهی، مشاور ممکن است با مسئلهٔ شبه‌تکرار (هارلبرت، ۱۹۸۴) روبه‌رو شود. به‌منظور

مطالعهٔ رشد مجدد پس از آتش‌سوزی جنگل، تعداد نهال‌های جدید در هر متر مربع در یک مکان خاص شمرده می‌شود. اگر تعداد شمرده‌شده از بلوک‌های ۱۰ متر مربعی به‌دست آمده باشد، آیا ۱۰ تکرار داشته‌ایم؟ چت‌فیلد (۱۹۹۵) مثالی مشابه ارائه می‌کند که مربوط به یک مطالعهٔ تجزیهٔ چندین کیسه برگ در یک آبگیر است.

آزمایش‌های طراحی‌شده (؟)

خیر، این یک شیوهٔ شبه‌تکرار از نوع عنوان قبلی نیست! در عمل، بسیاری از آزمایش‌های «طراحی شده» نمی‌توانند نتایج مفیدی عاید کنند زیرا اهداف مطالعه را برآورده نمی‌کنند. آزمودنی‌ها، تیمارها، یا اجراهای آزمایش ممکن است تکرار واقع‌بینانهٔ شرایطی که محقق واقعاً خواهان مطالعهٔ آن است، نباشد. فقدان واقع‌گرایی ممکن است حتی ناشی از خود آزمایش باشد ــ اگر آزمودنی‌ها بدانند «این صرفاً یک آزمایش است»، آیا پاسخ‌های آن‌ها نمایندهٔ نمایهٔ کلی جامعه است؟

به موضوعات مطرح در طراحی واقعی یک آزمایش، اغلب طی آموزش یک آماردان کمتر اعتنا می‌شود و تأکید به‌مراتب بیشتری بر جزئیات فنی مرتبط با طرح‌های متعادل و متعامد به‌عمل می‌آید. برای مشاور آماری، نتایج آزمایشی بیشتر محتمل است که از طرح‌های نامتعادل و نامتعامد آمده باشند تا غیر آن. این امر معمولاً، حداقل از لحاظ محاسباتی، مشکلی ایجاد نمی‌کند، زیرا بسته‌های نرم‌افزاری آماری خوبی موجودند.

شایان ذکر است که نگرانی عمدهٔ ما دربارهٔ داده‌های گردآوری‌شده از یک آزمایش طراحی‌شده مربوط به کیفیت طرح است. مثال‌های متعددی از آزمایش‌های سرهم‌بندی‌شده و نتایج گمراه‌کننده مبتنی بر تحلیل طرح نادرست وجود دارند (اندرسن (۱۹۹۰) کتابی کامل دربارهٔ آزمایش‌های معیوب است). نخستین وظیفهٔ یک مشاور آماری این است که به دقت تعیین کند که داده‌ها چگونه گردآوری شده‌اند، مگر اینکه خودمان داده را گردآوری کرده باشیم. گرچه این امر آشکارا شامل حال هر یک از روش‌های گردآوری داده‌ها که در این بخش مورد بحث قرار داده‌ایم می‌شود، در آزمایش‌های طراحی‌شده اهمیتی ویژه دارد.

۲.۳ پردازش داده‌ها

اکنون که پایگاه داده‌ها گردآوری شده و به فرمتی الکترونیکی تبدیل شده است، لازم خواهد بود که با برنامهٔ نرم‌افزار آماری خوانده شود. در اینجا فرض می‌کنیم که انتقال داده‌ها به سیستم ما با موفقیت تکمیل شده و اکنون پایگاه داده‌ها برای پروژه در فرمت مناسب بر سیستم ما قرار دارد.

همواره یک نسخهٔ پشتیبان از پایگاه داده‌های اصلی تهیه کنید.

(به جزئیات مربوط به چنین موضوعاتی بعداً در این فصل پرداخته می‌شود). نخستین مرحلهٔ تحلیل آماریمان با پردازش داده‌ها آغاز می‌شود. بنابراین، نخستین وظیفهٔ ما خواندن داده‌ها با برنامهٔ نرم‌افزاری آماریمان و یک ارزیابی اولیه از کیفیت داده‌هاست.

کیفیت داده‌ها

سه مرحلهٔ عمدهٔ ارزیابی کیفیت داده‌های اولیه عبارت‌اند از:

- پشتیبان ــ یک نسخهٔ پشتیبان از داده‌های اصلی تهیه کنید.
- فرمت ــ گدها و نوع داده‌ها برای خواندن داده‌ها.
- مقادیر ــ دراید‌های داده‌های مرتبط با متغیرها.

پشتیبان نگه‌داشتن نسخه‌ای از داده‌های اصلی منظورهای زیادی را برآورده می‌کند. این کار، اقدام احتیاطی متعارفی در وضعیتی است که فایل داده‌های کاری خراب یا پاک شود ــ این اتفاق ممکن است در هر لحظه رخ دهد، اغلب بدون هشدار، و ممکن است برای جلوگیری از آن کاملاً ناتوان باشیم (به‌عنوان مثال، یک درایو معیوب دیسک را پاک می‌کند، به اشتباه * rm را بر یک دستگاه UNIX تایپ می‌کنیم و نظایر آن). این کار به ما اجازه می‌دهد که با مشکلات مربوط به داده‌ها در فرمتی که مشتری با آن آشناست ارتباط برقرار کنیم. سطر ۵۳ در فایل داده‌ای (مرتب‌شدهٔ) کاری ما ممکن است با نسخهٔ اصلی آن کاملاً متفاوت باشد. به همین نحو، این اجازه را به ما می‌دهد که دراید‌های مشخصی را در پایگاه داده‌ای کاری خود با دراید‌های نسخهٔ اصلی مقایسه کنیم. این کار می‌تواند به‌خصوص هنگامی که سعی در رفع خطاهای ورودی داریم سودمند باشد.

فرمت بسته‌های نرم‌افزاری آماری همچون SAS و S-PLUS به‌منظور خواندن صحیح یک فایل داده‌ای نیازمند دستورات فرمتی هستند. به‌عنوان مثال، متغیرهای کاراکتری لازم است به‌درستی به SAS شناسانده شوند و در S-PLUS جدول‌های داده‌ها مستطیلی در نظر گرفته می‌شوند. اغلب زمانی که داده‌های اصلی در بردارندهٔ مقادیر گمشده‌ای باشند که با فضاهای خالی (مانند یک خانهٔ خالی در یک صفحهٔ گستردهٔ اکسل) نمایش داده می‌شوند، مشکلاتی بروز می‌کند. از آنجا که پیش‌فرض این است که از فضاهای خالی چشم‌پوشی شود، دراید‌های دادهٔ باقیمانده به متغیری اشتباه تخصیص داده می‌شوند. در برخی موارد شاید ویرایش جنبه‌های معینی از فایل داده‌ها (نسخهٔ کاری ما) ساده‌تر از سعی در به‌کارگیری دستوات فرمتی پیچیده باشد. به

روالی دیگر، ممکن است خواندن فایل داده‌ها در یک فرمت پروکسی (مثلاً به‌عنوان متغیرهای کاراکتری)، سپس خواندن خروجی داده‌ها در فرمت مطلوب امکان‌پذیر باشد. البته، فراهم‌کردن دستورالعمل‌های روشن برای مشتریان در مورد نوع فرمتی که انتظار داریم که فایل داده‌ها مطابق آن تهیه شوند، قطعاً می‌تواند در کاهش مدت‌زمانی که صرف این فعالیت می‌کنیم، کمک کند.

مقادیر به محض اینکه خواندن صحیح داده‌ها عملی شد، گام بعدی، وارسی کیفیت آماری داده‌هاست. این کار، یافتن خطاهای بدیهی یا بالقوه در خود مقادیر داده‌ها را دربرمی‌گیرد. آماره‌های خلاصه و جدول‌های فراوانی راهی سودمند برای مجزاکردن و ارزیابی خطاهای بالقوه فراهم می‌کنند. مثال‌هایی چند در زیر ارائه می‌شوند.

باید بر این نکته تأکید کنیم که ارزیابی کیفیت داده‌ها شامل حالتی که در آن مشتریان مسئولیت انجام‌دادن تحلیل آماری را یا استفاده از نرم‌افزار خودشان به عهده گرفته‌اند نیز می‌شود. در چنین وضعیتی باید به‌دقت آماره‌های خلاصهٔ مرتبط با پردازش داده‌های مشتری را بازبینی، و انواع خطاهای بالقوه یا مشکلاتی را که در زیر مورد بحث قرار می‌دهیم وارسی کنیم. این کار باید پیش از کمک به مشتری در تفسیر نتایج انجام گیرد ــ به یاد داشته باشید که استنباط‌های مبتنی بر داده‌های نامعتبر معنایی ندارند. بیان مختصرتری برای این سناریو این است که «زباله جواب زباله».

خطای بالقوه	علت ممکن
برخی از رده‌های رسته‌بندی‌شده دربردارندهٔ مشاهدات کمی هستند	خطای ورود داده‌ها خطای تولیدشده در پایگاه داده‌ها
یک متغیر رسته‌بندی‌شده شامل سطوح زیادتری است در مقایسه با آنچه باید باشد	تغییرات در کدبندی به‌کاررفته انتقال ستونی طی read-in
متغیرهایی که دربردارندهٔ نسبت بالایی از مقادیر گمشده‌اند	ساختار اشتباه به‌کاررفته در read-in انتقال ستونی در طول read-in خطای برنامه‌نویسی
مقادیر مرتبط با متغیرهای تعریف‌شده توسط کاربر هیچ مفهومی ندارند	خطای ورود داده‌ها خطای برنامه‌نویسی عمل غیرقانونی انجام‌شده
مقادیری که خارج دامنهٔ مورد انتظار یا مجاز یک متغیر قرار می‌گیرند	نقطهٔ دورافتاده: خطای ورود داده‌ها خطای read-in یا برنامه‌نویسی داده‌های با کیفیت ضعیف

۱. رده‌هایی از یک متغیر رسته‌ای که شامل مشاهدات بسیار کمی باشند.
این کار ممکن است صرفاً یک خطای تایپی باشد که منجر به ایجاد رده‌ای ناموجود برای یک متغیر رسته‌ای می‌شود.

۲. تغییرات در کدبندی به‌کار رفته برای رده‌های یک متغیر رسته‌ای.
متغیرهای رسته‌ای که رده‌هایشان به‌صورت عددی کدبندی شده‌اند بیشتر در معرض خطاهای ناشی از کدبندی ناصحیح قرار دارند. گرچه واردکردن مقادیر عددی برای متغیرهای رسته‌ای آسان‌تر است، تشخیص خطاهای سیستماتیک ممکن است دشوار و وقت‌گیر باشد. در مواردی که متغیری سطوح بیشتری از آنچه باید داشته باشد دارد، یک بررسی دقیق برای اطمینان‌یافتن از اینکه داده‌ها به‌درستی خوانده شده‌اند یا نه ضروری است (به‌عنوان مثال، ۳۰٫۰ منتج به آن می‌شود که ۰ به‌طور ناصحیح به متغیر بعدی تخصیص یابد).

۳. مقادیر داده‌ای که خارج دامنهٔ مورد انتظار یا مجاز قرار دارند.
یک وارسی سریع آماره‌های خلاصه (مینیمم، ماکسیمم، میانگین، انحراف استاندارد) می‌تواند خطاهای بالقوه را برای متغیرهای پیوسته مشخص کند. خطاهای بدیهی معمولاً به‌عنوان نقاط دورافتادهٔ فاحش آشکار می‌شوند. یک انحراف استاندارد بزرگ را باید به‌عنوان مورد مشکوک تلقی کرد، به‌خصوص زمانی که به‌لحاظ کمّی مشابه میانگین باشد. متغیرهای ID و شمارشی را می‌توان از طریق جدول‌های فراوانی مورد بررسی قرار داد (باید فراوانی هر ID برابر با ۱ باشد).

۴. متغیرهایی که دربردارندهٔ نسبت بالایی از مقادیر گمشده‌اند.
این حالت ممکن است در صورتی اتفاق بیفتد که یک تغییر مکان ستونی ناشی از یک فاصلهٔ خالی اتفاق افتاده باشد زیرا در صورتی که متغیر بعدی پیوسته اعلام شود، مقادیر کاراکتری گمشده تلقی می‌شوند.

۵. مقادیر داده‌ای مرتبط با متغیرهای تعریف‌شده توسط کاربر.
اگر مشتری از پیش داده‌ها را پردازش کرده باشد و متغیرهایی را که از کمیت‌های دیگر محاسبه شده‌اند در اختیار ما بگذارد، ما باید با استفاده از نرم‌افزار خود همین کمیت را محاسبه کنیم و مواردی را که تطبیق نمی‌کنند وارسی کنیم.

مقادیر گمشده و خطاها

چندین مسئلهٔ مهم زمانی بروز می‌کنند که پایگاه داده‌ها دربردارندهٔ مقادیر گمشده یا خطاها باشد. در برخی از موارد، تأثیر بر یک تحلیل آماری ممکن است قابل ملاحظه باشد و صرفاً چشم‌پوشی از

مقادیر گمشده به سبک کتاب‌های درسی توصیه نمی‌شود! به همین نحو، در مورد خطاهایی که پس از پردازش داده‌ها بر جای مانده‌اند باید تصمیمی اتخاذ شود. باز هم، صرفاً تخصیص یک مقدار گمشده به هر خطای شناسایی‌شده ممکن است بهترین رویکرد نباشد. مطالب زیر فهرستی از برخی از موقعیت‌هایی را در اختیار می‌گذارند که در آن‌ها خطاها و مقادیر گمشده به‌طور نوعی اتفاق می‌افتند.

گردآوری داده‌ها

- بدخواندن یک اندازه‌گیری
- بدثبت‌کردن یک اندازه‌گیری (۲۰ به جای ۰٫۲)
- برآوردکردن یک اندازه‌گیری
- لوازم معیوب به‌کاررفته در مطالعه
- ابزارها یا کارکنان متفاوت به‌کارگرفته‌شده در مطالعه
- بریدن یا گردکردن اریب اندازه‌گیری‌ها

واردکردن داده‌ها

- بدتایپ‌کردن مقادیر هنگام تبدیل به فرمت الکترونیکی
- نمایش اشتباه مقادیر داده‌ها
- تکرار مقادیر در ستون‌ها یا سطرها
- ردیف‌بندی نادرست مقادیر در ستون‌ها یا سطرها
- تخصیص نادرست مقادیر به‌عنوان گمشده
- تخصیص یک مقدار وقتی مشاهده گمشده باشد

مسئلهٔ اصلی که در رابطه با خطاها و مقادیر گمشده بروز می‌کند آن است که با از دست‌دادن اطلاعات مواجه می‌شویم. در قالب اصطلاحات آماری، این مطلب متناظر با از دست‌دادن درجه‌های آزادی و در نتیجه، استنباط کمتر قابل اطمینان است. بنابراین، گام اول کار با مشتری و تلاش برای بازیابی این داده‌ها در صورت عملی‌بودن آن است. خواستن تعریف دقیق از مشتری دربارهٔ اینکه برخی متغیرها چگونه ثبت و وارد شده‌اند اغلب نقطهٔ خوبی برای شروع کار بازیابی است.

> ممکن است شگفت‌زده شویم که چگونه دستورالعمل‌های «صریح ما» دربارهٔ واردکردن داده‌ها، کدبندی، و فرمت‌ها کاملاً آن‌گونه که می‌پنداشتیم روشن نبوده‌اند. مشتری را سرزنش نکنید. آن‌ها نهایت تلاش خود را برای انجام‌دادن آنچه ما خواسته‌ایم کرده‌اند ــاکنون نوبت ماست که از تجربه درس بگیریم.[۱]

به محض اینکه کار بازیابی از منبع داده‌های اصلی به‌طور مؤثری به اتمام رسید، لازم است که در مورد خطاهای باقیمانده و مقادیر گمشده نیز تصمیماتی اتخاذ شوند. در این مرحله باید تأکید کنیم که تنها خطاهای «بالقوه» عملاً باقی می‌مانند زیرا مقدار صحیح را نمی‌توان از منبع اصلی بازیابی کرد. دشواری در اینجا این است که تحلیلگر گرایشی به اتخاذ تصمیمات موردی همچون تلقی همهٔ خطاهای «بالقوه»ی باقیمانده به‌عنوان گمشده، یا حذف کل یک مشاهدهٔ چندمتغیره حتی اگر تنها یک مؤلفهٔ آن گمشده باشد دارد. از سویی دیگر، وقتی یک مقدار آشکارا یک خطا است (به عنوان مثال، خارج دامنهٔ مجاز قرار دارد)، اصلاح یک خطا یا مقدار دادهٔ گمشده آشکارا یک تصمیم شخصی است و اریبی (یا بدتر از آن) را وارد تحلیل می‌کند. این مورد، رویهٔ متداول جایگزینی یک مقدار گمشده با میانگین یا میانهٔ داده‌های مرتبط با متغیر را نیز شامل می‌شود. رویکردی عینی‌تر برای جانهی مقادیر گمشده در لیتل و رابین (۱۹۸۷) توصیف شده است.

۳.۳ موضوعات آماری

با اتمام مرحلهٔ پردازش داده‌های پروژه، محقق اکنون آمادهٔ انجام‌دادن «تحلیل آماری» داده‌هاست. حتی در پروژه‌های نسبتاً ساده، این کار به‌طور نوعی مرکب از مراحل زیر است:

۱. محاسبهٔ خلاصه‌های عددی و نمایش‌های گرافیکی

۲. انجام‌دادن یک «آزمون» آماری بر روی داده‌ها

۳. تفسیر یافته‌ها و نتیجه‌گیری‌ها

مورد آخر معمولاً بیشترین دشواری را برای مشتری‌ها ایجاد می‌کند زیرا این کار دربردارندهٔ مفهوم استنباط آماری است. مبانی نظری زمینه‌ای روش‌شناسی آماری گسترده‌اند و در حیطهٔ تخصص مشاور آماری جای دارند ــآن‌ها را همان‌جا نگه دارید! مشتری برای یاری‌جستن نزد ما آمده است، نه برای گرفتن یک درس احتمال و نظریهٔ استنباط. البته، موضوع استنباط آماری مهم است و لازم است به آن پرداخته شود. برخی از این موضوعات را با توجه به حوزه‌های استنباطی زیر در نظر می‌گیریم.

۱. راهی برای پرهیز از این نوع سوءتفاهم‌ها آن است که واقعاً با مشتری بنشینیم و چند مورد را به‌صورت مثال وارد کنیم. مشتری پس از آن می‌تواند منظور واقعی ما را دربارهٔ کدبندی دودویی و نمیزدادن مقادیر گمشده دریابد.

- برآورد
- آزمون‌های فرضی
- اندازهٔ نمونه و توان

برآورد

یک آمارهٔ خلاصه مانند میانگین نمونه‌ای $\overline{X}$ مثالی از یک **برآورد نقطه‌ای** است که می‌توان از آن برای استنباط چیزی دربارهٔ میانگین نامعلوم جامعه (μ) که نمونه از آن استخراج شده بود استفاده کرد. در حالی که این گزاره ممکن است ظاهر بی‌آزاری داشته باشد، اما در صدر حجم قابل ملاحظه‌ای از نظریهٔ برآورد جای دارد. «استنباط» دربارهٔ پارامتر جامعه (μ) مبتنی بر ویژگی‌های **برآوردگر** میانگین نمونه‌ای است: کمیت متغیر تصادفی نظری پیش از آزمایش تعریف شده است... و تا همین حالا هم، احتمالاً مشتری را سر درگم کرده‌ایم! آنچه مشتری به واقع لازم است بداند این است که چه نوع روش برآوردی برای پروژه مناسب است. مثال‌های زیر برای تشریح این نکته بکار می‌روند.

مثال۲.۳ متوسط سن [AGE] یک بیمار

پایگاه‌های داده‌ها تاریخ تولد یک بیمار را در اختیار می‌گذارند، اما نه لزوماً متغیر سن [AGE] مربوط را. می‌توان از تفاوت بین تاریخ جاری و تاریخ تولد بیمار برای محاسبهٔ یک صورت بریده‌شده از سن [AGE] برحسب سال‌های کامل استفاده کرد.

مشکل استفاده از یک صورت بریده‌شده از سن [AGE] این است که زمانی که سن [AGE] متوسط محاسبه می‌شود، چگونه باید مقداری مانند ۶۸٫۷۵ سال را تفسیر کرد؟ در حالی که به «نظر» می‌رسد که سن متوسط ۳ ماه کمتر از ۶۹ سال است، پیونددادن «ماه» با جزء کسری سال (۷۵٫۰) گمراه‌کننده است زیرا این مقدار از یک اندازهٔ سن [AGE] بریده‌شده به‌دست آمده است که سن بیمار را برحسب ماه منعکس نمی‌کند. این بدان معناست که آشکارا سن متوسط، سن بیماران هم‌گروه را کم برآورد می‌کند زیرا جزء کسری سال (ماه‌ها + روزها) مربوط به یک فرد در نظر گرفته نشده است.

مثال۳.۳ فشار خون طبیعی

فشار خون «طبیعی» یک شخص اغلب به‌صورت یک دامنه بیان می‌شود.

در این وضعیت، یک **برآورد بازه‌ای** اندازهٔ مناسب‌تری برای فشار خون طبیعی برای یک

شخص است. آشکار است که یک برآوردگر نقطه‌ای تنها مانند مقدار میانگین، در صورتی که فشار خون شخص بالاتر از این مقدار «نرمال» میانگین باشد، اطلاعات مورد نیاز شخص را به همراه ندارد.

تبصره‌ها

۱. دو مثال سادهٔ بالا تنها انواع مسائل برآورد نیستند، که مشاور با آن‌ها مواجه خواهد شد. مثال جالبی در تیستد(۱۹۸۸) داده شده است که در آن از الگوریتم E-M (دمپستر و همکاران، ۱۹۷۷) برای برآورد نسبت مرتبط با یک جامعهٔ آمیخته استفاده شده است. این داده‌ها به مجموعه‌داده‌های «Widows» معروف است که در آن لازم است برای شمارش «صفر»، نسبت آمیختگی توزیع‌های پواسون و دوجمله‌ای برآورد شود: بیوه‌های بدون اطفال تحت تکفل.

۲. در برخی از موارد، نظریهٔ مجانبی نرمال که اغلب برای استخراج خطاهای استاندارد به‌کارگرفته می‌شود، ممکن است برقرار نباشد. در این مورد، می‌توان از روش‌های بازنمونه‌گیری (بعداً مورد بحث قرار خواهد گرفت) برای تولید خطاهای استاندارد و بازه‌های اطمینان استفاده کرد. ممکن است لازم باشد که کمی مشتری را متقاعد کنیم که این کار، «مجاز» است و شاید مهم‌تر اینکه، لازم است مشتری درک کند که روش‌های بازنمونه‌گیری نمونه را از آنچه هست بزرگتر نمی‌کنند.

آزمون‌های فرض

هدف یک شیوهٔ استنباط آماری آن است که تحلیلگر را قادر به نتیجه‌گیری دربارهٔ مسئلهٔ تحت بررسی بر مبنای شواهد آزمایشی کند. در کاربردهای سنتی استنباط آماری، تحلیل یک مسئله دربردارندهٔ چندین مؤلفه است. نخست لازم است مسئله در قالب اصطلاحات ریاضی توصیف شود، که کار شناسایی ماهیت تصادفی را انجام می‌دهد. مقداری ساده‌سازی از ماهیت دقیق تصادفی‌بودن به‌منظور مهارپذیر بودن تحلیل‌ها اغلب ضروری است.

با تبدیل‌کردن مسئله به یک «فرایند تصادفی»، گام بعدی انتخاب یک شیوهٔ آماری مناسب برای ارزیابی شواهد آزمایشی است. در اینجا، انتخاب شیوه هم به زمینهٔ مسئله و هم به نوع شواهدی که قرار است ارزیابی شوند بستگی دارد. تطبیق این اطلاعات با یک مدل احتمالاتی مناسب، شیوه‌ای در اختیار تحلیلگر می‌گذارد که می‌توان از آن برای اندازه‌گیری میزان محتمل‌بودن هر برآمد آزمایش استفاده کرد. با این حال، به‌منظور محاسبهٔ احتمال درست یک برآمد واقعی، لازم است مدل نشان‌دهندهٔ وضعیت «صحیح» فرایند تصادفی باشد. آشکار است که اگر این را می‌دانستیم، نیاز به استنباط آماری نمی‌بود. بنابراین، می‌توان راهبرد زیر را به‌کار گرفت.

۱. تحلیلگر دربارهٔ وضعیت درست فرایند، «فرض»ی را در نظر می‌گیرد و سپس احتمال به‌دست‌آوردن این برآمد گزارش‌شدهٔ آزمایش را تحت این فرض محاسبه می‌کند.

۲. اگر این احتمال محاسبه‌شده بیش از حد نامحتمل به نظر برسد، منطقی خواهد بود که استنباط کنیم که فرض بیان شده نادرست بوده است.

۳. در این حالت، تحلیلگر نتیجه خواهد گرفت که به نظر نمی‌رسد شواهد، از فرض در نظر گرفته‌شده حمایت کند.

این «نتیجه‌گیری» و فرض مربوط به زمینهٔ مسئله بستگی دارند. به بیان مختصرتر، مراحل مطرح در این نوع شیوهٔ استنباط آماری عبارت‌اند از:

مفروضات شیوه‌های آماری نیازمند این هستند که مفروضاتی معین دربارهٔ «فرایند تصادفی» تحت مطالعه قید شود. این مفروضات ما را قادر می‌سازد که آن مدل احتمالی را استخراج کنیم که فرایند تصادفی را به‌لحاظ ریاضی توصیف می‌کند. پس از آن، شواهد آزمایشی بر مبنای این مدل تحلیل می‌شود.

فرض در رویکرد آزمون فرضی به استنباط آماری، فرض بر این است که شواهد زمانی که آزمایش تحت شرایط خاصی انجام شده باشد به‌دست آمده‌اند. به این فرض، فرض صفر (H_0) اطلاق می‌شود. تحت H_0، احتمال به‌دست‌آوردن نتیجه‌ای دستکم به همان اندازه «نامحتمل» به‌قدر برآمد مشاهده‌شده از آزمایش از روی مدل احتمالاتی محاسبه می‌شود. احتمال محاسبه‌شده Pـمقدار نامیده می‌شود و به شکل فرض مقابل (H_1) از آزمون فرض بستگی خواهد داشت. با این حال، اگر Pـمقدار کمتر از ۵٪ باشد، آنگاه ممکن است H_0 به نفع H_1 رد شود. در این حالت، نتیجهٔ آزمون فرض را از لحاظ آماری معنادار می‌نامند.

نتیجه‌گیری‌ها توجه به این نکته که یک نتیجهٔ معنادار از یک آزمون فرض از لحاظ آماری دلالت بر علیت نمی‌کند حائز اهمیت است. در واقع، Pـمقدار معیاری برای احتمال ارتکاب خطای نوع I (ردکردن H_0 زمانی که درست باشد) فراهم می‌کند. ممکن است لازم باشد، به‌منظور اثبات علیت، آزمایش را دفعات بسیاری تکرار کنیم. در عمل، معمولاً فرصت این را که آزمایش را دوباره به‌طور کامل اجرا کنیم نداریم و لازم است بر مبنای نتایج موجود نتیجه‌گیری کنیم. هنگام نتیجه‌گیری بر مبنای یک شیوهٔ استنباط آماری دقت قابل ملاحظه‌ای لازم است تا این اطمینان حاصل شود که ادعاهای بی‌پایه به‌طور ضمنی با ماهیت آماری تحلیل پیوند داده نشده‌اند.

حاشیه‌ای بر استنباط

شاید بهتر باشد استنباط آماری را به‌عنوان یک فلسفه توصیف کنیم زیرا مباحثهٔ قابل توجهی دربارهٔ چارچوب نظری‌یی که رویکرد فراوانی‌گرایی بالا به استنباط بر مبنای آن بنا نهاده شده است وجود دارد. در واقع، مسائلی فلسفی در پایه‌ای‌ترین مفاهیمی که می‌توانیم احتمال‌ها را تعیین کنیم و استنباط‌هایی انجام دهیم مطرح است. این کتاب آشکارا جای پرداختن به این موضوع نیست، اما مهم است توجه کنیم که رویکردهای دیگری به استنباط آماری وجود دارند (برای یک گزارش مقایسه‌ای از رویکردهای متفاوت به استنباط آماری خواننده را به بارنلت (۱۹۸۲) ارجاع می‌دهیم). به‌ویژه، استنباط بیزی اطلاعات «پیشین» را که ممکن است در مورد داده‌ها داشته باشیم (معمولاً برحسب یک توزیع احتمال بیان می‌شود) با داده‌های نمونه ترکیب می‌کند که محاسبهٔ اطلاعات «پسین» را امکان‌پذیر می‌سازد.

در این کتاب، دیدگاهی واقع‌گرایانه در پیش می‌گیریم که برای بسیاری از اهداف عملی، رویکرد آزمون معناداری به استنباط آماری قابل عمل است (البته به شرطی که به‌درستی انجام شود). با این حال، موقعیت‌های بسیاری وجود دارند که رویکرد بیزی در آن‌ها به‌طور شهودی خوشایندتر است. به‌عنوان مثال، در تحقیقات بازار اعتقاد پیشین ما دربارهٔ جامعهٔ مشخصی برای یک محصول جدید باید در برنامه‌ریزی آزمون بازار گنجانده شود. بنابراین، مشاوران آماری لازم است در رویکردشان به استنباط آماری انعطاف‌پذیر باشند.

تبصره‌ها

۱. آن تأکید آماری که ما (تعمداً) در خلاصه‌کردن مراحل یک آزمون فرض به‌کار گرفته‌ایم احتمالاً بیشتر از آنکه به عمیق‌ترشدن درک مشتری از استنباط آماری بینجامد، حاکی از احساس خودبینی ماست. گرچه ممکن است مشتریان کاملاً با «اجراکردن» یک آزمون فرض آشنا باشند، اکثراً این نوع از خلاصه‌کردن را سخت می‌یابند.

۲. بنابراین، چه رویکردی باید اتخاذ کنیم؟ اگر مشتری قبلاً با این نوع شیوهٔ آزمون آماری آشنا باشد، در این صورت می‌توانیم کار را با تأکید بر اهمیت استفاده از شواهد برای وارسی این مطلب آغاز کنیم که آیا مفروضات آزمون پیش از نتیجه‌گیری بر مبنای آن‌ها برقرارند یا خیر. این کار را می‌توان با آزمودن معنای زمینه‌بنیان[1] فرضِ بیان‌شده و، سرانجام، با اطمینان‌یافتن از این امر ادامه داد که مشتری محدودیت‌های یک نتیجه‌گیری آماری را درک می‌کند.

۳. تبصرهٔ بالا شامل حال مشتریانی نیز هست که با آزمون‌های آماری آشنا نیستند، اما این کار را سودمند یافته‌ایم که این بحث را با مثالی ساده از نوع زیر پیش ببریم.

1. contextual

غریبه‌ای ادعا می‌کند که در ۱۰ بار پرتاب سکه ۱۰ شیر به‌دست آورده است. اگر فرض شود که سکه سالم است، در این صورت احتمال به‌دست‌آوردن این برآمد صرفاً بر اساس تصادف کمتر از ۱ در ۱۰۰۰ است. از آنجا که این مقدار کمتر از استاندارد مبنیم ۵٪ است، نتیجه می‌گیریم که این امر ناشی از چیز دیگری بجز شانس است. از این رو، ممکن است ما غریبه را به دروغگویی متهم کنیم. در حالی که این «نتیجه‌گیری» ممکن است کاملاً منطقی به نظر برسد، فرض کنید این فرصت به ما داده شده که خود ما آزمایش را با پرتاب‌کردن سکهٔ مورد بحث تکرار کنیم. اگر چنین باشد که ۹ شیر در ۱۰ پرتاب به‌دست آوریم، در این صورت نتیجه‌گیری ما با این مفهوم آماری که ناشی از چیزی بجز شانس بوده است مطابقت دارد. این بدان معناست که ما فرض سالم‌بودن سکه را رد خواهیم کرد. با این حال، این بیان تلویحی که غریبه دروغ می‌گوید آشکارا نامعتبر است.

۴. مشتریان اغلب اهدافشان را در قالب فرض‌های جهت‌دار بیان می‌کنند. وقتی این را به نمادهای آماری برگردانیم، فرض جهت‌دار نشان‌دهندهٔ H_1 است. مسئله این است که Pـمقدار اندازه‌ای از احتمال برآمد آزمایشی را نسبت به فرض صفر H_0 به‌دست می‌دهد. نکتهٔ کلیدی این است که می‌توان تصمیمی دربارهٔ فرض جهت‌دار مشتری بر مبنای Pـمقدار مرتبط با آزمون آماری اتخاذ کرد. اینجاست که مشتری معمولاً برای نتیجه‌گیری نیاز به کمک دارد.

۵. اغلب آموزنده است که آن نتیجه‌گیری را که با یک نتیجهٔ آزمون غیر معنادار مرتبط است در نظر بگیریم. در برخی زمینه‌های علمی، نقطهٔ برش ۵٪ تغییرناپذیر دانسته می‌شود (حتی نتیجه‌ای مانند ۰٫۵۲ نادیده گرفته می‌شود!) این مایهٔ تأسف است، اما باید حداقل تلاش کنیم مشتریان را دربارهٔ استدلال نادرست اتخاذکردن یک نقطهٔ برش سخت‌گیرانه متقاعد کنیم ــ شاید آن‌ها بتوانند تغییردادن نگرش همقطاران خود را آغاز کنند. یک راه برای انجام‌دادن این کار، در نظرگرفتن توان آزمون است. ممکن است لازم باشد که کار در آغاز به‌طور غیر مستقیم انجام گیرد (با شروع از ارتباط بین اندازهٔ نمونه و «قابلیت اعتماد» یک آزمون)، اما فراهم‌کردن نوعی جدید از «Pـمقدار» (توان) برای مشتری و بیان اینکه عدد مزبور چه معنایی دارد می‌تواند سودمند باشد. موضوع اندازهٔ نمونه و توان را بعداً مورد بحث قرار می‌دهیم.

اندازهٔ نمونه و توان

تعیین اندازهٔ نمونهٔ مناسب در مطالعه‌های مشاهده‌ای و آزمایش‌های برنامه‌ریزی شده دغدغه‌ای مهم است. این امر مستقیماً به هزینهٔ انجام‌دادن مطالعه مربوط می‌شود و بر کیفیت استنباط آماری

مرتبط با نتایج، تأثیر می‌گذارد. روش‌های تعیین اندازهٔ نمونه به بهینه‌سازی میزان اطلاعاتی می‌پردازد که لازم است با توجه به کاربردی خاص گردآوری شوند. برخی از این مثال‌ها عبارت‌اند از:

مطالعه‌های طولی مطالعه باید چه مدت انجام گیرد و هر چند وقت یکبار باید داده‌ها را گردآوری کرد؟

بررسی‌های نمونه‌ای در بررسی‌های پیچیده، لازم است که تابع‌های هزینه را نسبت به کیفیت بهینه کرد تا حداکثر بازگشت اطلاعات تضمین شود.

آزمایه‌های بالینی لازم است تأثیر یک داروی جدید برخی مقررات دولتی مشخص را برآورده کند. به چه تعدادی از بیماران در هر سطح تیماری و برای گروه کنترل نیاز است؟

تضمین کیفیت لازم است که برنامهٔ زمانی بازرسی و طرح‌های نمونه‌گیری با دستورکارهای کیفی سازگاری داشته باشند.

طرح‌های آزمایشی در یک طرح خاص، چه تعدادی از اقلام در هر خانه مورد نیاز است تا تفاوت مشخصی در اثرها را آشکار کنند؟

در حالت کلی، یک اندازهٔ نمونهٔ مناسب را می‌توان با مشخص‌کردن یک سطح دقت یا با ماکسیمم‌کردن یک تابع هدف مشخص به‌دست آورد. روش‌های صریحی برای بسیاری از مسائل استاندارد وجود دارند و جدول‌های جامعی بر مبنای تحلیل توان موجودند: کرامر و تایمن (۱۹۸۷)، کوهن (۱۹۸۸)، و دسو و راگاوارائو (۱۹۹۰) سه تکنگاشت اساسی هستند. برای مروری بر روش‌های تعیین اندازهٔ نمونه، از جمله روش‌های بیزی، ادکاک (۱۹۹۷) را ببینید. تحلیل توان نیز در تعدادی از بسته‌های نرم‌افزاری آماری گنجانده شده است.

با این حال، در عمل، تعیین نمونه‌ای مناسب برای مشتری ممکن است بسیار دشوار باشد. به‌عنوان مثال، تعریفِ توان $= 1 - \beta$ را در نظر بگیرید که در آن $\beta = P[\text{II}$ [خطای نوع $]$.

> توان احتمال به‌دست‌آوردن نتیجه‌ای معنادار و تابعی از اندازهٔ نمونهٔ n، اندازهٔ تأثیر δ، انحراف استاندارد σ، و سطح معناداری α است. توان به شما می‌گوید که چقدر محتمل است که آزمایش شما تفاوت مشخص σ را در سطح معناداری مفروض α آشکار کند.

بنابراین، در تعیین اندازهٔ نمونه n، مشخص‌کردن هر یک از کمیت‌های دیگر دخیل است. مقادیر مرسوم برای سطح معناداری و توان به‌ترتیب ۵٪ و ۸۰٪ هستند. در بسیاری از آزمون‌های استاندارد، اغلب می‌توان σ را برای ارائهٔ اندازه‌های اثر استانداردشده‌ای تحلیل برد. در غیر این صورت، برای به‌حساب‌آوردن تغییرپذیری مرتبط با برآوردکردن σ اِعمال اصلاحاتی در محاسبات اندازهٔ نمونه ضروری است. مسئلهٔ اصلی موقع تلاش برای مشخص‌کردن یک اندازهٔ اثر منطقی رخ می‌دهد

زیرا این اندازه، درجهٔ اعتقاد محقق را به نادرست‌بودن فرض صفر (H_0) نشان می‌دهد. این یک تصمیم شخصی است و مقیاس‌های اندازه‌گیری متفاوت آن را پیچیده‌تر هم می‌کند زیرا اندازهٔ اثر به نوع آزمون مورد استفاده بستگی دارد.

در علوم رفتاری، جدول‌هایی (به‌عنوان مثال، کوهن (۱۹۸۸، ۱۹۹۲) را ببینید) به‌منظور تعیین اندازهٔ نمونه اغلب بر رده‌بندی اندازهٔ اثرها به‌صورت کوچک، متوسط، و بزرگ برای برخی آزمون‌ها مبتنی هستند. در مهندسی و کاربردهای تضمین کیفیت، اندازه‌های نمونهٔ معقول اغلب به‌خوبی جاافتاده‌اند و از منحنی مشخصهٔ عملکرد (OC) برای ارزیابی β در برابر اندازه‌های اثر استانداردشده استفاده می‌شود. منحنی (OC) صرفاً نمودار β در برابر δ بمازای اندازهٔ نمونهٔ مفروض n است. یک نمودار OC مرکب از چندین منحنی روی هم در چندین اندازهٔ نمونه‌ای متفاوت است. برای آزمایش‌ها و طرح‌های ساده، این جدول‌ها و نمودارهای OC معمولاً برای مقاصد تعیین اندازهٔ نمونه مناسب‌اند. بحث را با چند توضیح دیگر به پایان می‌رسانیم.

۱. توضیح‌دادن مفاهیم مطرح در تحلیل توان به مشتری ممکن است بسیار دشوار باشد. اتخاذ یک رویکرد بازهٔ اطمینانی، ممکن است مؤثرتر باشد. به‌عنوان مثال، حاشیهٔ خطای محافظه‌کارانهٔ $۲\sqrt{p(۱-p)/n} \leq ۱/\sqrt{n}$ که در نظرسنجی‌ها بکار می‌رود راهی ساده برای آشناکردن مشتری با موضوع اندازهٔ نمونه در برابر کیفیت است.

۲. محاسبات توان پیش از آزمایش انجام می‌گیرند. آن‌ها توضیح نمی‌دهند که چرا یک نتیجهٔ پساآزمایشی معنادار نیست.

۳. یک اندازهٔ نمونهٔ مشخص تضمین نمی‌کند که آزمایش نتیجه‌ای معتبر به وجود خواهد آورد. معمولاً مفروضاتی زمینه‌ای در شیوهٔ آزمون وجود دارند که ممکن است در آزمایش برآورده نشوند. تأکید این امر به مشتری مهم است.

۴. برای اندازه‌های نمونهٔ بزرگ، انتظار می‌رود که نتایج معنادار باشند. اینکه این‌ها همیشه معنادارند یا نه روشن نیست. اندازه‌های نمونهٔ بزرگتر همچنین بیشتر در معرض خطاهای پردازش داده‌ها و مسائل پیاده‌سازی هستند، و ممکن است حاوی اطلاعاتی با کیفیت پایین باشند.

۵. برای آزمایش‌های برنامه‌ریزی‌شده، اندازه‌های نمونه با فرض طرح‌های متعادل تعیین می‌شوند: اندازه‌های خانه‌ها یا گروه‌های برابر. مقادیر گمشده ممکن است به‌طرز نامناسبی بر توان مرتبط با یک طرح نامتعادل تأثیر بگذارد. لازم است مشتری را از پیامدهای بیش از حد نامتعادل‌شدن آزمایش طراحی‌شده آگاه کنیم.

۶. در آزمایش‌های چندمتغیره، توان و اندازهٔ نمونه به تعداد متغیرهای پاسخی که قرار است اندازه‌گیری شوند نیز بستگی خواهند داشت.

۴.۳ روش‌های آماری مورد استفاده در مشاوره

در دو بخش زیر، برخی از روش‌های آماری متداول در محیط‌های مشاوره را مرور می‌کنیم. در بخش بعد، توجه خود را بر روش‌های «استاندارد» معطوف می‌کنیم. این روش‌ها دربردارندهٔ روش‌های توصیفی هستند که طی مرحلهٔ اکتشافی تحلیل به‌کار گرفته می‌شوند و نیز شامل شیوه‌های آماری پایه‌ای (همان‌گونه که در جدول زیر فهرست شده است) می‌شوند. برای شیوه‌های تخصصی‌تر، ارائه‌کرد ما از مطالب لزوماً مختصر و مفید است زیرا شرح جامع این روش‌ها را می‌توان در جای دیگر یافت. شاید از این مهم‌تر، جزئیاتی دربارهٔ کاربرد برخی از این روش‌ها در زمینهٔ مطالعه‌های موردی ارائه‌شده در قسمت II فراهم شده است. ارجاعاتی به دیگر کتاب‌های درسی و مطالعهٔ موردی مربوط انجام می‌شود. فهرست موجز روش‌های آماری در جدول زیر خلاصه شده است.

روش‌های استاندارد

- تحلیل اکتشافی داده‌ها
- جدول‌های پیشایندی
- آزمون t
- تحلیل واریانس
- رگرسیون

روش‌های کلی

- مدل‌های خطی کلی
- تحلیل چندمتغیره
- تحلیل سری‌های زمانی
- تحلیل داده‌های رسته‌ای
- شیوه‌های خاص

۵.۳ روش‌های استاندارد

در توصیف روش‌ها و فنون آماری مطرح در این بخش، سعی ما بر این است که بر اهمیت دخالت‌دادن درک مشتری از هدف، جزئیات، و تفسیر هر روش تأکید کنیم. این بدان معناست که این‌ها روش‌هایی هستند که ممکن است لازم باشد که دربارهٔ آن‌ها به مشتری آموزش دهیم. آماردان در درک این روش‌های مقدماتی مشکلی ندارد، اما در مورد مشتریان قطعاً چنین است! این همان جایی است که مشاور آماری وارد عمل می‌شود.

تحلیل اکتشافی داده‌ها (*EDA*)

تحلیل اکتشافی داده‌ها یک اصطلاح فنی نسبتاً بی‌چارچوب است به این معنا که طیف گسترده‌ای از روش‌های آماری وجود دارند که می‌توان آن‌ها را با به‌اصطلاح تحلیل اکتشافی داده‌ها مرتبط دانست. برای مقاصد ما، EDA دربردارندهٔ هرگونه نمایش گرافیکی، خلاصهٔ عددی، یا شیوهٔ آماری است که برای بررسی ویژگی‌های توزیعی و ساختاری یک مجموعهٔ داده‌ها طی مراحل اولیهٔ تحلیل به‌کار می‌رود. بنابراین، EDA نمایندهٔ تحلیلی «غیر رسمی» از داده‌ها در تلاشی برای آشکارساختن جنبه‌های نامعمول یا جالب است.

آمار توصیفی

به محاسبهٔ آماره‌های خلاصهٔ عددی و ساختن نمودارهای ساده آمار توصیفی اطلاق می‌شود. برای مقاصد گزارش‌نویسی، یک تحلیل توصیفی خلاصه‌ای همه‌جانبه از فرد به فرد متغیرها در پایگاه داده‌ها فراهم می‌کند. خلاصه‌های دیگر در درون زیرگروه‌های مهم نیز ممکن است سودمند باشند. از منظر مشاور، این ابزارهای تشخیصی توصیفی قبلاً نقشی بنیادی طی مرحلهٔ پردازش داده‌های پروژه ایفا کرده‌اند. جنبه‌های مربوط مانند کیفیت داده‌های برخی متغیرها ممکن است در جهت‌بخشی تحلیل پروژه سودمند باشد.

شکل: جنبه‌های توزیعی داده‌ها
مکان: موقعیت داده‌ها (مرکز، مقادیر کرانگین)
پراکنش: اندازه‌های تغییرپذیری در داده‌ها

برای این سه ویژگی توصیفی فهرست‌شده در بالا، بافت‌نگار، میانگین نمونه‌ای، و انحراف استاندارد به‌یقین بیشترین استفاده را در خلاصه‌کردن متغیرهای فردی دارند. با این حال، تشخیص اینکه این آماره‌ها چه زمانی ممکن است گمراه‌کننده باشند حائز اهمیت است.

نمایش‌های گرافیکی

اهمیت نمودارهای کیفی در بخش ۶.۲ مورد بحث قرار گرفت. در اینجا توجه خود را بر جنبه‌های آماری مرتبط با انتخاب نمایش‌های گرافیکی معطوف می‌کنیم که در EDA سودمندند. به‌عنوان مثال، گرچه نمودار کلوچه‌ای اغلب برای نمایش اطلاعات توزیعی به‌کار می‌رود، نمودار میله‌ای ابزاری مؤثرتر در دیداری‌سازی داده‌هاست.

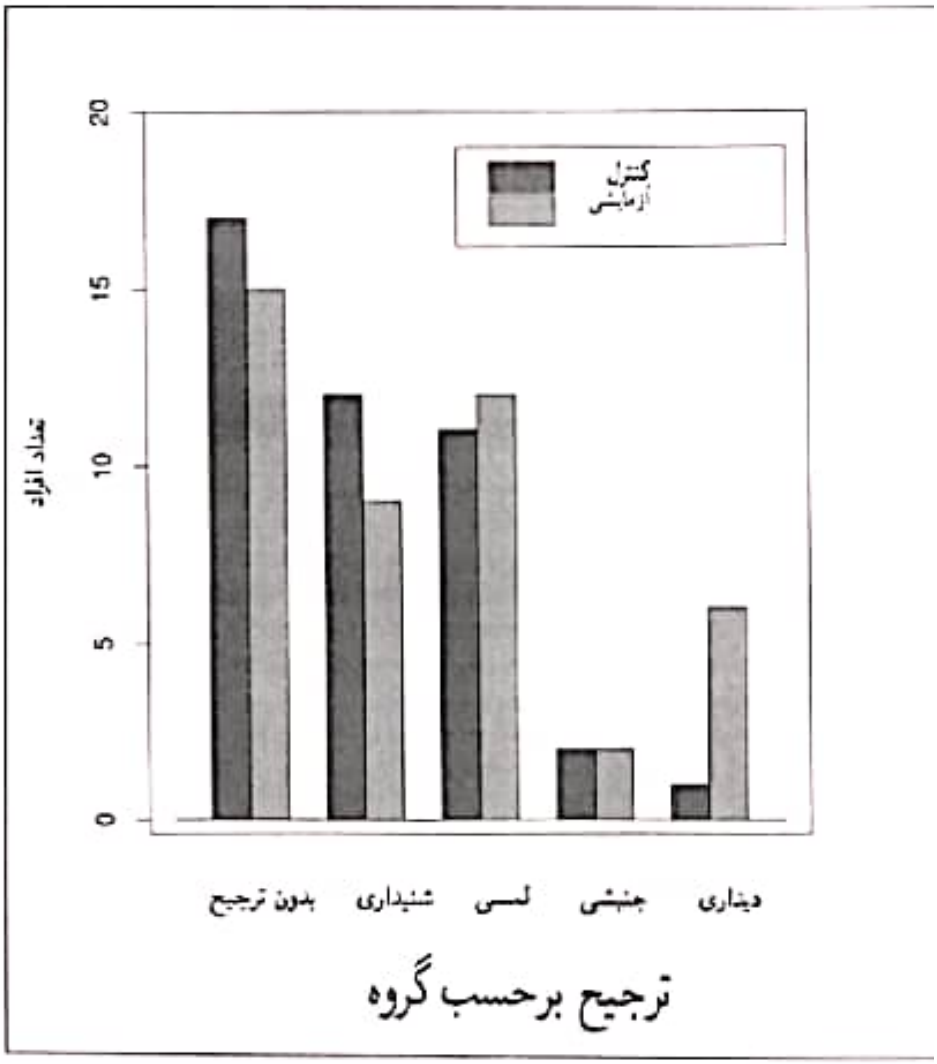
توزیع:	بافت‌نگار، نمودار میله‌ای، نمودار ساقه‌ای، نمودار جعبه‌ای
Y در برابر X:	نمودار پراکنش، نمودار زمان، نمودار موزاییکی
چندمتغیره:	نمودار خم تراز، نمودار داربستی، گرافیک‌های پویا
تشخیصی:	نمودار Q-Q، دوره‌نگار، مانده‌ها در برابر برازش‌یافته

توزیع بافت‌نگار، نمودار میله‌ای، نمودار ساقه‌ای، نمودار جعبه‌ای

بافت‌نگار و نمودار میله‌ای به‌طور گسترده‌ای برای نمایش توزیع فراوانی داده‌های کمّی و رسته‌ای به‌کار می‌روند. برای متغیرهای کمّی، خصیصه‌هایی همچون تقارن، چولگی، چندمدی‌بودن، گروه‌بندی‌ها، و نقاط دورافتاده مورد توجه‌اند. دو فن مؤثر دیگر برای بررسی این خصیصه‌های مربوط به «شکل»، نمودارهای ساقه‌ای و جعبه‌ای هستند. یک بافت‌نگار در شکل ۳.۲ فصل قبل (بخش ۶.۲) نشان داده شده است. مثال‌هایی از انواع دیگر نمایش در شکل‌های ۱.۳ تا ۳.۳ نشان داده شده است.

مقصود از بافت‌نگار یا نمودار ساقه‌ای[1] فراهم‌کردن یک نمایش معرف از توزیع زمینه‌ای است که داده‌های نمونه‌ای از آن تولید شده‌اند. بنابراین، انتخاب پهنای بازه‌ای که قرار است در نمایش بافت‌نگار به‌کار رود، و تعداد خطوط استفاده شده در نمایش نمودار ساقه‌ای موضوع مهمی است. این امر حاکی از مصالحه‌ای بین دقت و اندازهٔ نمونه (n) است:

۱. این نمودار در اصل نمودار «ساقه و برگ» نامیده می‌شود (توکی، ۱۹۷۰)؛ این عنوان از تفکیک معنادارترین ارقام یک مقدار داده‌ای به مؤلفهٔ «ساقه» (ارقام اصلی) و «برگ‌های» در پی دارندهٔ ارقام باقیمانده از تمام مقادیر داده‌ای که ساقهٔ مشترک دارند گرفته شده است.

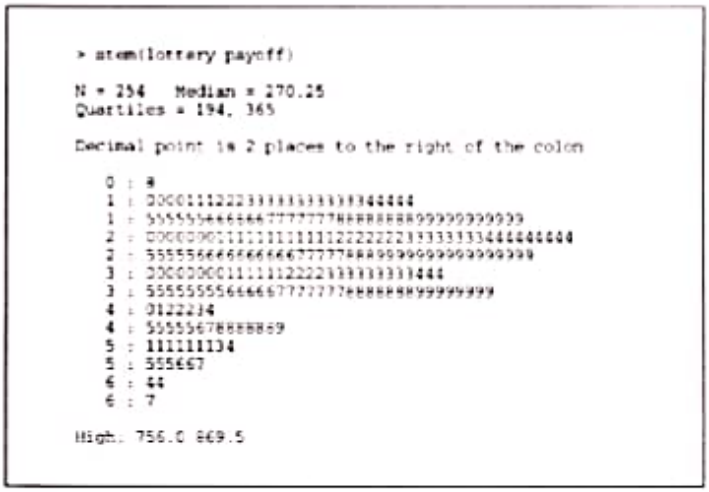


شکل ۱.۳ نمودار میله‌ای دوبخشی.

اگر تعداد بازه‌ها بسیار کم باشد، خصیصه‌های توزیعی از دست خواهند رفت؛ اگر تعداد بسیار زیاد باشد، خصیصه‌های غیر واقعی توزیع از طریق خود نمایش وارد کار خواهند شد.

برای نمونه‌ای به اندازهٔ n، ملاک‌های زیر سودمند واقع شده‌اند.

قرایند گاوسی $h_n = ۳٫۴۹\hat{\sigma}n^{-1/3}$ توسط اسکات (۱۹۷۹) به‌عنوان انتخابی بر مبنای داده‌ها برای پهنای بازه مطرح شد تا در نمایش بافت‌نگار داده‌هایی به‌کار رود که



شکل ۲.۳ نمودار ساقه‌ای: داده‌های بخت‌آزمایی (منبع S-PLUS).

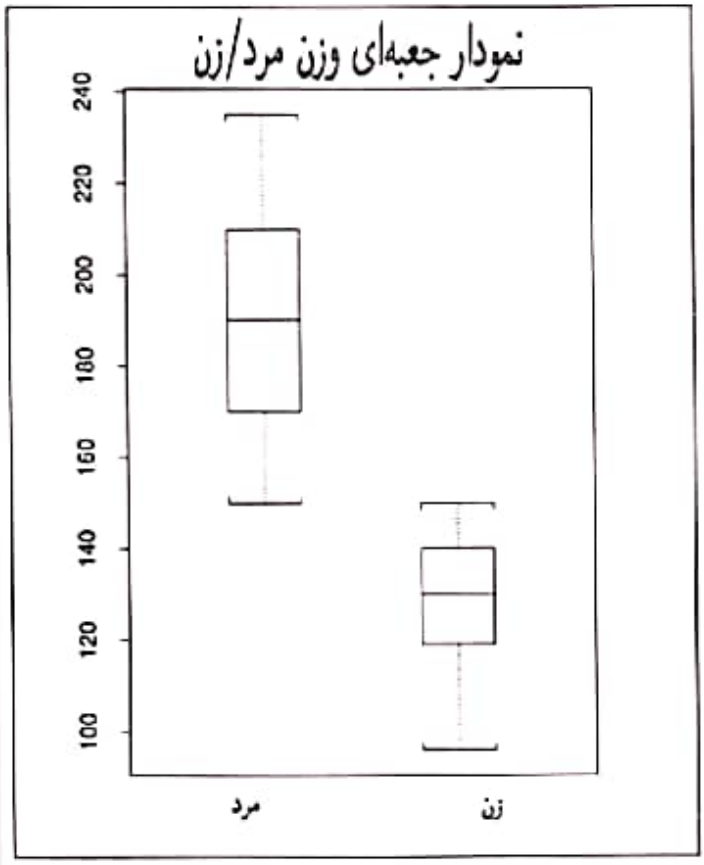
فرض می‌شود توزیع آن‌ها نرمال است. در اینجا، s برآوردی از σ است (مانند انحراف استاندارد نمونه). قاعده‌ای که توسط فریدمن و دیاکونیس (۱۹۸۱) مطرح شد $h_n = 2(IQR)n^{-1/3}$ است که در آن IQR دامنهٔ میان‌چارکی نمونه است.

نمودار ساقه‌ای فرض کنید $L_{\max}$ نشان‌دهندهٔ ماکسیمم تعداد خطوط مورد استفاده در نمایش نمودار ساقه‌ای باشد. در هاگلین و همکاران (۱۹۸۳)، $L_{\max} = \lfloor 10 \log_{10} n \rfloor$ که در آن $\lfloor x \rfloor$ نشان‌دهندهٔ جزء صحیح x است، به عنوان یک حد بالای معقول برای تعداد خطوط مورد استفاده در نمایش نمودار ساقه‌ای پیشنهاد می‌شود. این نتیجه منسوب به دیکسون و کرونمال (۱۹۶۵) است که از آن در بافت‌نگار استفاده کردند. این نتیجه برای نمونه‌های با اندازه‌های کوچک تا متوسط ($n \leq 300$) به‌خوبی کار می‌کند. برای نمونه‌های با اندازهٔ بزرگتر، فایدهٔ نگاه‌داشتن مقادیر واقعی داده‌ها کمتر می‌شود.

تبصره‌ها

۱. اطمینان بیابید که مشتریان می‌دانند که «چگونه» یک بافت‌نگار را تفسیر کنند. ممکن است لازم باشد که مفهوم تقارن توضیح داده شود.

۲. ساختن نمودارهای میله‌ای ممکن است گاهی گمراه‌کننده باشد، به‌ویژه اگر چندین



شکل ۳.۳ نمودارهای جعبه‌ای موازی.

متغیر بر روی یک نمودار نمایش داده شوند. به‌عنوان مثال، اگر تعداد در هر گروه مبتنی بر اندازه‌های نمونهٔ مختلف باشد باید از فراوانی‌های نسبی استفاده کرد.

۳. یک اِشکال نمودارهای ساقه‌ای و میله‌ای این است که ممکن است برای برخی از مشتریان یا زمینه‌ای که مشتریان به‌طور معمول نتایج خود را ارائه می‌کنند، نامأنوس باشند. اگر ساختن این نمودارها در مقایسه با تفسیر آن‌ها مستلزم توضیح بیشتری باشد، از یک بافت‌نگار کمک بگیرید.

۴. تفاوت عمدهٔ نمودارهای ساقه‌ای و بافت‌نگارها این است که اثرهای ناشی از خود

داده‌ها مانند گردکردن یا مکرر بودن را می‌توان با استفاده از نمودار ساقه‌ای آشکار کرد. بافت‌نگار به این انواع اختلالات حساس است: یک جابه‌جایی کوچک در کرانه‌های رده‌ها (نقاط انفصال) ممکن است تغییرات بزرگی در توزیع فراوانی ایجاد کند.

۵. نمودار جعبه‌ای نمایش دیداری فشرده‌ای از داده‌ها بر مبنای چارک‌ها و کرانگین‌ها به‌دست می‌دهد: $\{\min, Q_1, M, Q_3, \max\}$، که در آن M میانه است (مکان را در ادامه ببینید). نمودار جعبه‌ای به‌ویژه برای مقایسهٔ توزیع‌های چندین متغیر وابسته یا وارسی توزیع یک متغیر فردی در زیرگروه‌ها سودمند است.

Y در برابر X نمودار پراکنش، نمودار زمانی، نمودار موزاییکی

اصطلاح «نمودار پراکنش» به وضعیتی اشاره دارد که در آن رابطهٔ بین دو متغیر کمّی، که برای راحتی با X و Y نشان داده می‌شوند، به‌صورت نقاطی بر روی صفحهٔ دکارتی نمایش داده می‌شود. قوت و ماهیت پیوند (خطی یا غیر خطی) همراه با نقاط مؤثر[1] در صفحهٔ دوبعدی مدنظر است. در زمینهٔ رگرسیونی، Y به‌عنوان متغیر پاسخ تلقی می‌شود و کیفیت پیشگویی متغیر ورودی یا تبیینی X مورد نظر است. مثال‌هایی در شکل ۴.۳ نشان داده شده‌اند.

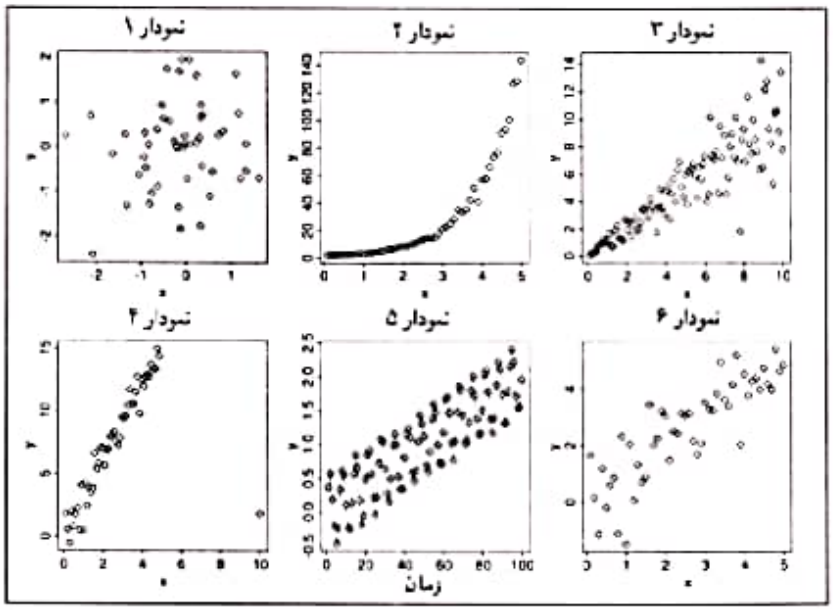
یک مورد خاص آن است که X نشان‌دهندهٔ یک شاخص زمانی باشد. در این وضعیت، مقادیر Y معمولاً به هم متصل می‌شوند تا یک نمودار پیوستهٔ «سری زمانی» تشکیل دهند. در تحلیل سری‌های زمانی، دوره‌ها، روندها، و نقاط دورافتاده مورد نظر هستند (شکل ۴.۷، مطالعهٔ موردی ۴.۷ را ببینید).

تبصره‌ها

۱. در حالی که ممکن است تعریف یک نمودار پراکنش تا حدی زائد به نظر برسد، اما از آن به‌منظور خاطرنشان‌کردن محدودیت بنیادی در توانایی ما برای «دیداری‌سازی» داده‌ها استفاده می‌شود: ما ذاتاً محدود به دید دو بعدی هستیم. به همین دلیل، مهم است که با احتیاط به یک نمودار پراکنش توجه کنیم. فقدان رابطه‌ای مشهود بین X و Y لزوماً مستلزم آن نیست که هیچ پیوندی بین این متغیرها وجود ندارد.

۲. حضور فصلی‌بودن، دوره‌ها، و روندها را معمولاً می‌توان از روی نمودار زمانی تشخیص داد. اگر این اثرهای ساختاری حذف شوند (به‌عنوان مثال، با تفاضل‌گیری یا برازش یک مدل سری زمانی)، در این صورت توجه به این معطوف می‌شود که آیا در

1. influential



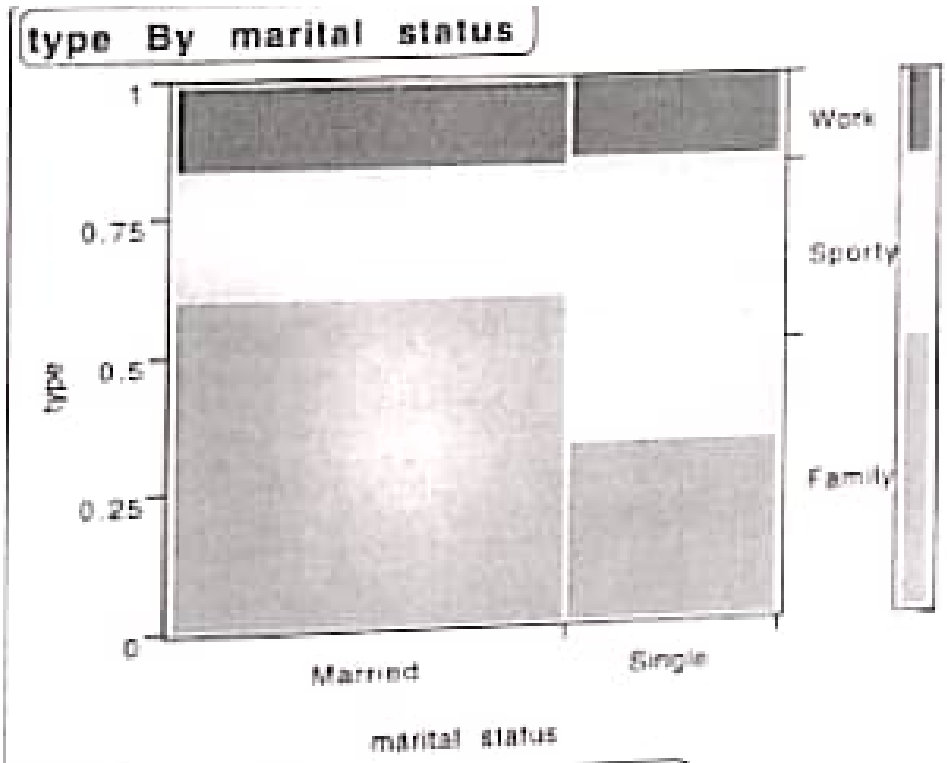
شکل ۴.۳ نمودارهای پراکنش.

سری «مانده‌ها»، مثلاً $r(t)$، همبستگی پیاپی وجود دارد یا خیر. یک نمودار پراکنش $Y = r(t)$ در برابر $X = r(t-1)$ امکان وارسی سادهٔ خودهمبستگی را فراهم می‌کند.

۳. وقتی X و Y هر دو متغیرهای رسته‌ای باشند، یک نمودار موزاییکی، نمایشی گرافیکی از فراوانی‌های نسبی را در یک جدول فراوانی دوطرفه فراهم می‌کند. این مطلب با جزئیات بیشتری در بخش زیر دربارهٔ جدول‌بندی آمده است.

چندمتغیره نمودار خم تراز، نمودار داربستی، گرافیک‌های پویا

همان‌گونه که در بالا خاطرنشان شد، فنون دیداری‌سازی داده‌ها به فضای دوبعدی محدود می‌شوند. البته، این امر مانع از بررسی روابط چندبعدی نشده است و نمودارهای خم تراز، فنی مؤثر برای آزمودن همزمان سه متغیر کمّی فراهم می‌کنند. نمودار خم تراز مشابه ارتفاع کناره‌ها بر روی یک نقشهٔ توپوگرافیکی است. نمودارهای پرسپکتیو، نمودارهای سه‌بعدی، و بافت‌نگارهای سه‌بعدی را نیز می‌توان بکار برد. اما به‌دست‌آوردن یک چشم‌انداز «خوب»



Crosstabs

	marital status		
Count Total % Row % Col %	Married	Single	
Family	119 39.27 76.77 60.71	36 11.88 23.23 33.64	155 51.16
Sporty	45 14.85 45.00 22.96	55 18.15 55.00 51.40	100 33.00
Work	32 10.56 66.67 16.33	16 5.28 33.33 14.95	48 15.84
	196 64.69	107 35.31	303

Tests

Source	DF	-LogLikelihood	RSquare (U)
Model	2	13.38260	0.0441
Error	299	289.81208	
C Total	301	303.19548	
Total Count	303		

Test	ChiSquare	Prob>ChiSq
Likelihood Ratio	26.766	<.0001
Pearson	26.963	<.0001

شکل ۵.۳ نمودار موزاییکی: داده‌های نظ...

از داده‌ها ممکن است به دلیل روش‌های ترسیم مات[1] بکار رفته در این فنون دشوار باشد.

نمودارهای دار بستی (مت‌سافت (۱۹۹۷)) تلاشی در تدارک یک سیستم گرافیکی «خودکار» برای نمایش داده‌های چندمتغیره در فضای دوبعدی است. «نمودار» اولیه ممکن است هر یک از نمودارهای بحث‌شده در بالا باشد ــ نمودار میله‌ای، بافت‌نگار، نمودار جعبه‌ای، نمودار پراکنش، نمودار خم تراز، نمودار سه‌بعدی ــ که بعداً مشروط بر سطوح وابسته به همهٔ ترکیب‌های بقیهٔ متغیرها بارپرداری می‌شوند. مثالی از نمودار دار بستی را می‌توان در مطالعهٔ موردی ۲.۷، شکل ۱.۷ دید.

فن گرافیکی چندمتغیره‌ای دیگر آن است که نمودارها را «پویانمایی» کنیم که به آن گرافیک‌های پویا اطلاق می‌شود. نمودارهای چرخان[2] اغلب در نرم‌افزارهای آماری گنجانده شده‌اند و این امکان را برای کاربر فراهم می‌کنند که نمایی سه‌بعدی از داده‌ها را به تصویر در آورد. XGobi (سوین و همکاران، ۱۹۹۱) به کاربر توانایی اجرای «تور» همراه با راهنمای EDA از داده‌های چندمتغیره را می‌دهد.

تبصره‌ها

۱. نمودارهای جعبه‌ای موازی و نمودارهای پراکنش جفتی دو مثال از نمودارهای گرافیکی ساده برای داده‌های چندمتغیره هستند. گرچه برای نمایش همزمان $k \geq 3$، متغیر فنون مبتکرانهٔ بسیاری خلق شده است، خوب است به خاطر داشته باشیم که هر نمای دوبعدی از داده‌های چندمتغیره هرگز تمام و کمال نخواهد بود.

۲. نمودارهای گرافیکی چندمتغیره تا چه اندازه برای مشاور اهمیت دارند؟ اغلب نمودارهای تشخیصی، از قبیل آن‌ها که در زیر ارائه می‌شوند، تا حد زیادی کفایت می‌کنند. شاید پرسشی بهتر این باشد: «آیا می‌توانیم نتایج را از نمایش‌های چندمتغیرهٔ پیچیده‌تر تفسیر کنیم؟ تجربه نشان می‌دهد که توانایی «نگاه» به داده‌ها همواره بر امکان بالقوهٔ دچار اشتباه‌شدن بوسیلهٔ ابزارهای تشخیصی عددی غلبه خواهد کرد.»

ابزارهای تشخیصی نمودار Q_Q، نمودار مانده در مقابل برازش‌یافته، دوره‌نگار

از نمودار Q_Q اغلب برای ارزیابی نرمال‌بودن مانده‌های مرتبط با یک نمودار مانده در مقابل برازش‌یافته استفاده می‌شود. دوره‌نگار در تحلیل سری زمانی برای ارزیابی وجود بسامدی معنی‌دار یا چرخه در داده‌ها بکار گرفته می‌شود. هر سه نمودار، مثال‌هایی رایج در نمودارهای تشخیصی‌اند که به‌طور گسترده در تحلیل‌های آماری بکار می‌روند.

1. opaque 2. spinning plots

تبصره‌ها

۱. نمودار Q-Q مرکب است از چندک‌هایی از یک توزیع احتمال $f(x)$ که در مقابل داده‌های مرتب‌شده رسم می‌شود. اگر داده‌ها در واقع امر توسط $f(x)$ تولید شده باشند، آنگاه نمودار پراکنش باید رفتاری راست‌خط از خود نشان دهد. در عمل، توزیع مورد نظر اغلب توزیع نرمال است و به $n > ۳۰$ مشاهده برای فراهم‌آوردن یک ابزار تشخیصی مورد اعتماد نیاز داریم. برای ملاحظهٔ مثال‌هایی از نمودارهای Q-Q شکل ۶.۳ را ببینید.

۲. نمودار مانده در مقابل برازش‌یافته نتایج حاصل از برازاندن مدلی به داده‌ها را نمایش می‌دهد. اگر مدل صحیح باشد، آنگاه مانده‌ها نباید الگوهایی غیر تصادفی را به نمایش گذارند. این نمودار به‌طور گسترده در بسیاری از روش‌های آماری به‌کار می‌رود و مثال‌های ارائه‌شده در شکل ۷.۳ به روش‌شناسی رگرسیون که بعداً مورد بحث قرار می‌گیرد مربوط می‌شود.

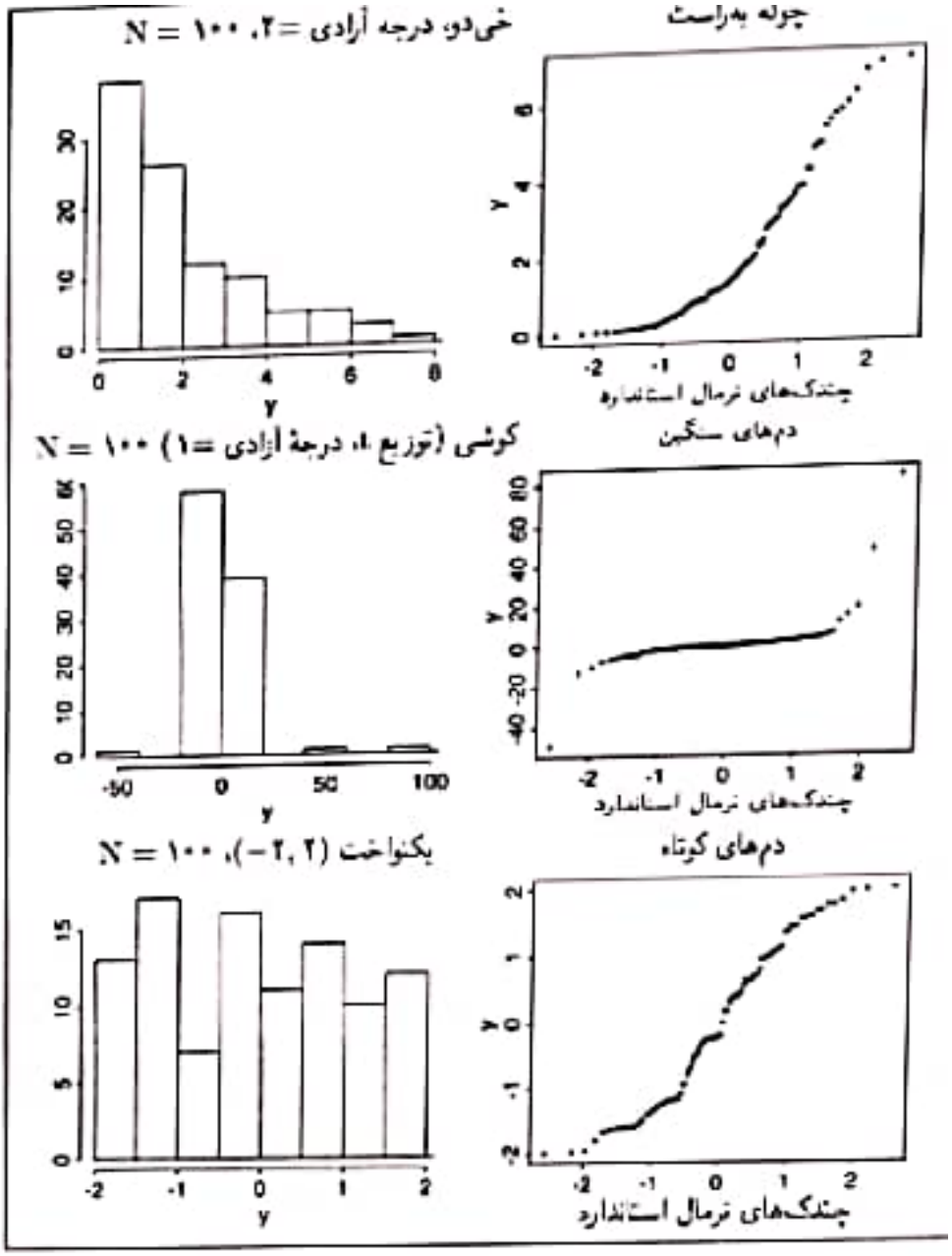
۳. دوره‌نگار اساساً یک افراز ANOVAی یک‌طرفه از تغییرات کل به مؤلفه‌های بسامد است که به کمک تجزیهٔ فوریهٔ سری زمانی $y(t)$ تعریف شده است. اگر یک بسامد خاص، ω_j، غالب باشد، آنگاه یک «تیک» بزرگ $I(\omega_j)$ در نمودار دوره‌نگار $Y = I(\omega_j)$ در مقابل $X = \omega_j$ ظاهر خواهد شد.

ابزارهای تشخیصی عددی

جدول‌بندی: جدول‌های فراوانی و خلاصه
مکان: موقعیت داده‌ها: مرکز، چندک‌ها
پراکندگی: اندازه‌های تغییرپذیری
همبستگی: اندازه‌های پیوند بین X و Y

جدول‌بندی جدول‌های فراوانی و خلاصه

یک جدول فراوانی یک‌طرفه صرفاً فهرستی از تعداد مشاهدات ــ فراوانی یا شمار ــ مرتبط با سطوح یک متغیر رسته‌ای است. معمولاً لازم است که داده‌های کمّی در بازه‌های ناهمپوش گروه‌بندی (یا رسته‌بندی) شوند تا در تحلیل جدول فراوانی سودمند باشند. یک جدول

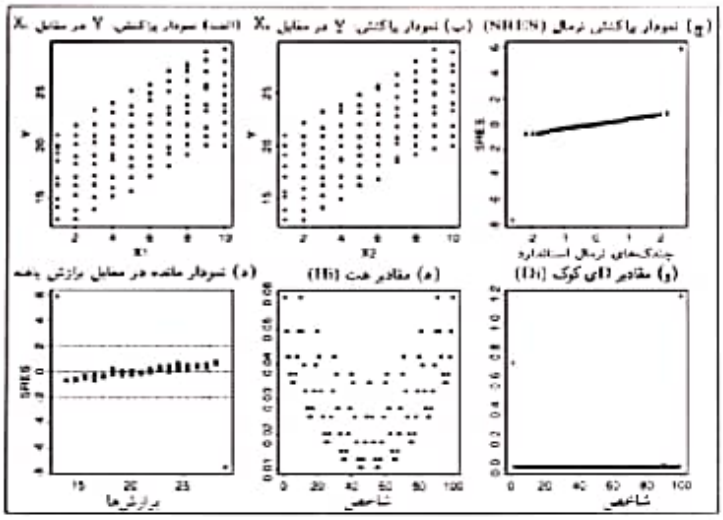


شکل ۶.۳ نمودارهای Q-Q.

شایندی یک جدول فراوانی دوطرفه است که شمار «خانه»ی مرتبط با هر ترکیبی از طوح دو متغیر را فهرست می‌کند. برای مقاصد EDA منظورکردن فراوانی نسبی یا درصد ی هر سطح یا خانه در جدول فراوانی سودمند است. در یک جدول پیشایندی، می‌توان رصدهای خانه‌ای را نیز به شرط شمارهای «حاشیه‌ای» سطری و ستونی محاسبه کرد.

ال جدول پیشایندی ارائه‌شده در شکل ۵.۳ با نرم‌افزار آماری JMP (SAS 2000)[1]

نشان داده‌شده با استفاده از ویرایش ۳ تولید شد. ارجاع مربوط به ویرایش ۴ است.



شکل ۷.۳ ابزارهای تشخیصی رگرسیونی.

تولید شده بود و دربردارندهٔ نمودار موزاییکی است که نمایشی گرافیکی از فراوانی‌های نسبی مرتبط با خانه‌های جدول پیشایندی است. سطح «آجرهای مستطیلی‌شکل» درون مربع واحد متناظر با فراوانی‌های نسبی خانه‌هاست. استفاده از رنگ (مقیاس خاکستری) برای متمایزکردن سطوح یک متغیر، ظاهری موزاییکی به نمودار می‌دهد که وجه تسمیهٔ آن است.

جدول‌های فراوانی چندطرفه را می‌توان به‌عنوان دنباله‌ای از جدول‌های دوطرفه به شرط ترکیب‌های سطح بقیهٔ متغیرها نمایش داد. در عمل، **جدول‌های خلاصه** هنگامی که متغیرهای متعددی مطرح باشند مؤثرترند. قالب پایه‌ای یک جدول خلاصه، یک جدول دوطرفه است که توسط سطوح عامل‌های سطری و ستونی، مثلاً A و B، تولید می‌شود. سپس سطوح هر عامل اضافی در درون سطوح A و B «آشیانی» می‌شوند و آشیانی‌شدن عامل سطری بیشتر صورت می‌گیرد. این امر تا حدودی ناشی از ابعاد یک صفحهٔ کاغذ است. همچنین خواندن یک جدول، هنگامی که سطرها بیشترین میزان آشیانی‌شدن را داشته باشند، آسان‌تر است. درایه‌های متن یک جدول خلاصه ممکن است شمارها، میانگین‌ها، یا هر آمارهٔ توصیفی دیگری باشد که مورد نظر مشتری است.

تبصره‌ها

۱. یک جدول باید آگاهی‌بخش باشد! برخی از مثال‌هایی که چنین نیست عبارت‌اند از:

- تنها درصدها ارائه شده‌اند. اندازهٔ نمونهٔ مورد استفاده برای محاسبهٔ این درصدها همواره باید داده شوند.
- درصدها با میزان دقت ناکافی ارائه می‌شوند.
- نسبت بزرگی از داده‌ها (مثلاً ۹۰٪) دلالت بر یک سطح دارند. اگر سطوح دیگری مورد نظر باشند باید در جدول جداگانه‌ای ارائه شوند.
- اندازهٔ نمونه کوچکتر از آن است که مؤید یک تحلیل فراوانی معتبر باشد.

۲. در صورتی که تعداد زیادی سطح مطرح باشد، ممکن است با جدول‌های تنک روبه‌رو شویم. نسبت بزرگی از خانه‌ها شمارهای صفر یا یک دارند. برای مقاصد تحلیل، رسته‌بندی‌کردن دوبارهٔ یک متغیر از طریق ترکیبکردن برخی سطوح ضروری خواهد بود. این تصمیم‌ها باید به همراه مشتری اتخاذ شوند.

۳. گاه جدول‌های فراوانی با استفاده از رسته‌های همپوش ساخته می‌شوند، به این معنی که یک فقره ممکن است به بیش از یک سطح تعلق داشته باشد. این امر آشکارا تفسیر و تحلیل آماری جدول را تحت تأثیر قرار می‌دهد و هنگامی که مشتری یک جدول فراوانی در اختیار ما می‌گذارد وارسی اینکه آیا این وضعیت وجود دارد یا خیر دارای اهمیت است.

۴. تعداد جدول‌های تولیدشده در یک تحلیل فراوانی ممکن است قابل ملاحظه باشد. به‌عنوان مثال، یک بررسی شامل ۱۰۰ سؤال (متغیر)، در صورتی که همهٔ اثرهای متقابل دومتغیره در نظر گرفته شوند، $\binom{100}{2} = 4950$ جدول تولید خواهد کرد. به همین نحو، تعداد جدول‌های شرطی دوطرفهٔ تولیدشده در یک تحلیل فراوانی چندعاملی به‌طور هندسی افزایش می‌یابد. انتخاب داهیانهٔ متغیرها برای استفاده در جدول‌بندی متقاطع و تحلیل فراوانی چندطرفه حائز اهمیت است.

هدف اصلی، جدول‌بندی اطلاعات خلاصه به سبکی موجز و در عین حال واضح است. گرچه آرایش پایه‌ای یک جدول، یک طرح دوطرفه است، اما به‌کارگیری گروه‌بندی‌های فرعی درون سطرها و / یا ستون‌ها اجازهٔ نمایش روابط پیچیده‌تری را می‌دهد. با این حال، قالب یک جدول دارای اهمیت است و معمولاً مستلزم انتخاب داهیانهٔ متغیرهای سطری و ستونی و زیرگروه‌بندی‌ها است. برخی از نکاتی که لازم است هنگام ساختن یک جدول خلاصه در نظر گرفته شوند عبارت‌اند از:

آماره‌ها خانه‌های یک جدول به‌طور طبیعی متشکل از آماره‌های خلاصهٔ ساده‌ای همچون فراوانی‌ها، درصدها، میانگین‌ها، مجموع‌ها، و چندک‌ها هستند.

خطاهای استاندارد این‌ها اغلب با برآورد متناظر، به شکل SE± یا (SE) ترکیب می‌شوند. روشن‌ساختن اینکه از انحراف استاندارد استفاده می‌شود یا خطای استاندارد، حائز اهمیت است.

Pـمقدارها معمولاً فقط P ـمقدارهای معنادار مورد نظر هستند و اغلب فهرست‌کردن برخی آستانه‌ها (مثلاً، ۰٫۰۵ و ۰٫۰۱) در زیرنویس جدول کفایت می‌کند.

اندازهٔ جدول یک جدول خلاصه باید تنها در یک صفحه جای گیرد. از آنجا که جدول به لحاظ شکل مستطیلی است، فضای فیزیکی کمتری برای ستون‌ها در مقایسه با سطرها (در اسلوب پرتره) وجود دارد. معمولاً لازم است که گروه‌بندی فرعی نخست درون متغیر سطری انجام گیرد.

تعداد خانه‌ها تعداد خانه‌ها در یک جدول با حاصل‌ضرب همهٔ سطوح گروه‌بندی مورد استفاده تعیین می‌شود. در صورت لزوم، یک جدول را می‌توان به چندین جدول کوچکتر، هر یک «به شرطه سطوح یک متغیر خاص، افراز کرد.

منظره‌ای جدول‌های نشان داده‌شده در اسلوب منظره‌ای (در جهت عرض صفحه) اجازه می‌دهد که ستون‌های بیشتری منظور شوند. از آنجا که این کار مستلزم چرخاندن گزارش برای دیدن جدول، سپس بازگشتن برای خواندن متن مربوط است، باید از حالت منظره‌ای تا جایی که امکان دارد پرهیز کرد.

مکان گرایش به مرکز، چندک‌ها، اندازه‌های مقاوم

میانگین نمونه‌ای $\overline{X}$ و میانهٔ M، متداول‌ترین آماره‌های مورد استفاده برای توصیف مقدار «متوسط» فرایند تحت مطالعه هستند. برای مقاصد اکتشافی، آماره‌های ترتیبی راهی مناسب برای بررسی ویژگی‌های توزیعی یک فرایند فراهم می‌کنند. به‌عنوان مثال، چارک‌های یک مجموعه از داده‌ها، Q_1 و Q_3، به‌ترتیب، اندازه‌های مکانی ۲۵٪ و ۷۵٪ هستند (یعنی، ۲۵٪ از مقادیر داده‌ها زیر Q_1 قرار دارند). برخلاف میانه، هیچ راه محاسبهٔ «استانداردی برای Q_1 و Q_3 وجود ندارد. بنابراین، تفاوت‌های جزئی بین بسته‌های نرم‌افزاری آماری وجود دارد.

تبصره‌ها

۱. تفاوت بین M و $\overline{X}$ به‌سادگی با در نظرگرفتن مجموعه‌داده‌های زیر، $\{۱۸۰, ۳۳, ۳۱, ۲۹, ۲۷\}$، نشان داده می‌شود که برای آن $\overline{X} = ۶۰$ و $M = ۳۱$. اگر این اعداد نشان‌دهندهٔ حقوق‌ها در یک شرکت کوچک باشند که در پی استخدام کارکنان جدید است، یک حقوق «متوسط» ۶۰ هزار دلار [در سال] بسیار جذاب به نظر خواهد رسید! واضح است که یک اندازهٔ مکانی مانند $\overline{X}$ به‌تنهایی ممکن است کاملاً گمراه‌کننده باشد ــ باید همواره یک اندازهٔ تغییرپذیری ارائه شود.

۲. **خلاصهٔ ۵ـعددی** نمودار جعبه‌ای یک نمایش گرافیکی از خلاصهٔ ۵ـعددی است که گردایه‌ای از پنج آمارهٔ مکانی زیر است: $\{\min, Q_۱, M, Q_۳, \max\}$. یک مؤلفهٔ مهم در ساختمان یک نمودار جعبه‌ای، تعیین نقاط دورافتادهٔ بالقوه است که مبتنی بر ملاک‌های زیر است.

نقاط دورافتاده هر X_i، $X_i < C_L$ یا $X_i > C_U$ به‌عنوان نقطهٔ دورافتادهٔ بالقوه‌ای در نظر گرفته می‌شود، که در آن[۱]،

$$C_L = Q_۱ - ۱٫۵ * (Q_۳ - Q_۱)$$

و

$$C_U = Q_۳ + ۱٫۵ * (Q_۳ - Q_۱).$$

۳. همان‌گونه که در مثال بالا نشان داده شد، میانگین نمونه‌ای نسبت به نقاط دورافتاده مقاوم نیست. گرچه مقاوم‌بودن در مقابل خطاهای فاحش از یک منظر کاربردی مطلوبیت آشکاری دارد، کارآیی آماری[۲] یک برآوردگر نیز دارای اهمیت است. این امر به مطالعهٔ برآوردگرهای استوار منجر شده است و یک مثال ساده میانگین پیراسته است:

میانگین‌های پیراسته یک میانگین پیراسته ۵٪ از کوچکترین مقادیر داده‌ها و ۵٪ از بزرگترین مقادیر داده‌ها را کنار می‌گذارد؛ یعنی، در مجموع ۱۰٪ از مجموعه‌داده‌ها کنار گذاشته می‌شود. سپس، میانگین نمونه بر اساس ۹۰٪ باقیماندهٔ داده‌ها محاسبه می‌شود. در هوگلین و همکاران (۱۹۸۳)، نشان داده می‌شود که میانگین پیراستهٔ ۲۵ درصدی کارایی معقولی را حفظ می‌کند.

---

۱. ثابت ۱٫۵ انتخاب متداولی است؛ از ۲ نیز استفاده می‌شود.

۲. این مفهوم در واقع اندازه‌ای نسبی است که دقت دو برآوردگر را مقایسه می‌کند. به‌عنوان مثال، میانهٔ M تقریباً به ۶۴٪ دادهٔ بیشتر نیاز دارد تا قادر باشد «مرکز» یک توزیع نرمال (μ) را با همان دقت $\overline{X}$ اندازه بگیرد.

۴. از منظر مشتری، آرایهٔ بی‌پایانِ اندازه‌های مختلف برای مکانی یکسان ممکن است تا حدودی اضطراب‌آور باشد. تأکید بر هدف اکتشافی این اندازه‌های مختلف دارای اهمیت است. در واقعیت امر، پرسش این نیست که «باید از کدام‌یک استفاده کنیم؟» بلکه این است که «آیا داده‌ها مفروضات مورد نیاز برای تحلیل مورد نظر را تأیید می‌کنند؟» اگر این‌طور نیست، حداقل می‌توانیم نشان دهیم که چرا.

پراکندگی انحراف استاندارد، خطای استاندارد، دامنه، IQR

یک برآورد مکانی صرفاً نقطه‌ای مرجع برای مجموعه‌داده‌ها فراهم می‌کند؛ این برآورد، اطلاعی دربارهٔ «تغییرپذیری» مقادیر داده‌ها به‌دست نمی‌دهد. دو اندازهٔ متداول پراکندگی، انحراف استاندارد S و دامنهٔ R هستند. از دامنهٔ میان‌چارکی $IQR = Q_3 - Q_1$ می‌توان برای فراهم‌کردن یک اندازهٔ پراکندگی مقاوم استفاده کرد.

تبصره‌ها

۱. **انحراف استاندارد** گرچه به اندازهٔ $\overline{X}$ متداول است، یکی از «غیرشهودی‌ترین» آماره‌های خلاصه نیز هست. تعریف S از طریق واریانس نمونه‌ای S^2 قطعاً کمکی به موضوع در این رابطه نمی‌کند. مسئلهٔ اصلی این است که برخی مشتریان سعی می‌کنند بفهمند که S دقیقاً چه چیزی را اندازه می‌گیرد. «آیا مقدار S من خوب یا بد است؟» ممکن است پرسشی احمقانه به نظر برسد، اما طبق تجربهٔ ما این آن چیزی نیست که مشتری واقعاً می‌پرسد. اغلب، مسئلهٔ واقعی دربارهٔ اعتبار تحلیل است که انتظار می‌رفته ربطی به S داشته باشد. معرفی مفهوم خطای استاندارد برحسب حاشیهٔ خطا ممکن است در اینجا سودمند باشد.

۲. **خطای استاندارد** آماره‌هایی که در بالا تعریف شدند همگی اندازه‌ای از تغییرپذیری داده‌های «خام» را به‌دست می‌دهند. برای مقاصد تحلیل، نیازمند اندازه‌ای از تغییرپذیری نمونه‌گیری مرتبط با یک برآورد (مانند $\overline{X}$) هستیم که از روی مجموعه‌داده‌ها محاسبه می‌شود. مسئله در اینجا این است که ما تنها «یک» برآورد داریم! این اندازهٔ پراکندگی، خطای استاندارد برآورد نامیده می‌شود و معمولاً مبتنی بر واریانس نظری برآوردگر است. به‌عنوان مثال، خطای استاندارد میانگین نمونه‌ای $\sqrt{\sigma^2/n}$ است که در آن σ^2 واریانس نامعلوم فرایند است. بنابراین، $S/\sqrt{n}$ برآوردی از خطای استاندارد $\overline{X}$ به‌دست می‌دهد. مجدداً، ممکن است این توصیف از خطای استاندارد برای مشتری خیلی معنا

نداشته نباشد (به‌عنوان مثال، ممکن است مشتری در حال دیدن خروجی رگرسیون باشد). در این حالت، ممکن است که لازم باشد صرفاً خطای استاندارد را بیان کنیم و توضیح دهیم که، برحسب تکرار آزمایش مشتری به دفعات زیاد، چه چیزی را نشان می‌دهد.

۳. **بازهٔ اطمینان** (CI) برآورد مکانی و پراکندگی را ترکیب می‌کند تا بازه‌ای مانند $[a, b]$ به‌دست دهد که نشان‌دهندهٔ دامنه‌ای از مقادیر ممکن برای پارامتر (نامعلوم) فرایند مورد نظر است. اغلب، حدود CI به‌گونه‌ای انتخاب می‌شوند که بازه با شانس ۹۵٪، پارامتر فرایند را در بر داشته باشد.

به‌عنوان مثال، بازهٔ اطمینان محافظه‌کار ۹۵٪، که با رابطهٔ $\bar{p} \pm n^{-1/2}$ داده می‌شود، اغلب برای گزارش نتایج نظرسنجی به‌کار می‌رود. در اینجا، $\bar{p}$ نسبت پاسخ‌های آری در یک نظرسنجی پیمایشی آری/خیر به اندازهٔ n است و $n^{-1/2}$ حاشیهٔ خطا است که معمولاً به‌عنوان مثال به‌صورت نقاط ۳± درصدی بیان می‌شود.

۴. **بازنمونه‌گیری** روش‌هایی همچون جک‌نایف[1] و بوت‌استرپ[2] فنونی محاسباتی هستند که می‌توان از آن‌ها برای برآورد خطاهای استاندارد استفاده کرد. روش جک‌نایف این امکان را می‌دهد که n برآورد مختلف برای مثلاً $\overline{X}$ از روی یک مجموعه‌داده، هر بار با کنارگذاشتن یکی از داده‌ها، بسازیم. بنابراین، برآوردی از خطای استاندارد $\overline{X}$ را می‌توان بر مبنای فرمول انحراف استاندارد معمولی محاسبه کرد که در مورد n مقدار مختلف حاصل برای میانگین نمونه‌ای اعمال می‌شود. روش بوت‌استرپ از مجموعه‌داده‌ها با جایگذاری بازنمونه‌گیری می‌کند و هر بار برآورد را می‌سازد. این روش از لحاظ محاسباتی زمان‌برتر است و اغلب از چندین هزار بازنمونه‌گیری استفاده می‌کند.

۵. ضرایب **چولگی** و **کشیدگی** دو اندازه‌ای هستند که می‌توان آن‌ها را برحسب ویژگی‌های مرتبط با شکل توزیع تفسیر کرد. با این حال، این دو نسبت به نقاط دورافتاده بسیار حساسند و ممکن است اطلاعاتی کاملاً گمراه‌کننده به‌دست دهند.

روشی بسیار معتبرتر در تعیین چولگی و وزن «دُم» (کشیدگی) این است که یک نمایش گرافیکی از داده‌ها مانند بافت‌نگار یا نمودار Q-Q را وارسی کنیم.

همبستگی اندازه‌های پیوند بین X و Y

ضریب همبستگی r، یک اندازهٔ بی‌بعد از پیوند **خطی** بین دو متغیر کمّی است. تمیزدادن خطی‌بودن دارای اهمیت است زیرا یک مقدار r نزدیک به صفر نشان می‌دهد که هیچ

1. jacknife 2. bootstrap

همبستگی معناداری بین X و Y وجود ندارد ــ این لزوماً بدان معنا نیست که هیچ پیوندی بین این متغیرها وجود ندارد زیرا پیوند ممکن است غیر خطی باشد. افزون بر آن، r صرفاً وابستگی جفتی را اندازه می‌گیرد؛ r نمی‌تواند وابستگی‌هایی را که ممکن است بین چندین متغیر وجود داشته باشد آشکار سازد.

اینجا دادن هشداری ضروری است، ضریب همبستگی پی‌یرسون (گشتاور ضربی) به‌صورت

$$r = \frac{\sum_{i=1}^{n}(X_i - \overline{X})(Y_i - \overline{Y})}{\sqrt{\sum_{i=1}^{n}(X_i - \overline{X})^2 \sum_{i=1}^{n}(Y_i - \overline{Y})^2}}$$

تعریف می‌شود. می‌توان نشان داد که $-1 \leq r \leq 1$ و پیوند خطی کامل زمانی رخ می‌دهد که $r = \pm 1$. مفهوم همبستگی را می‌توان در مورد متغیرهای شامل داده‌های ترتیبی همچون آن‌هایی که از داده‌های پیمایشی حاصل می‌شوند نیز بکار برد. در این وضعیت، اندازهٔ پیوند خطی، بین «رتبه»های دو متغیر است و اغلب از ضریب همبستگی رتبه‌ای اسپیرمن استفاده می‌شود. شکل سادهٔ آن به‌صورت

$$r = 1 - \frac{6\sum_{i=1}^{n} d_i^2}{n(n^2 - 1)}$$

تعریف می‌شود که در آن، رتبهٔ(Y_i) − رتبهٔ$(X_i) = d_i$. وقتی در رتبه‌های هر یک از گروه‌ها «هم‌رتبگی» وجود داشته باشد، در این صورت می‌توان از فرمول ضربی پی‌یرسون استفاده کرد و به جای X_i و Y_i رتبهٔ (X_i) و رتبهٔ (Y_i) را قرار داد. به روش دیگر، شکل سادهٔ بالا را می‌توان تعدیل کرد به‌طوری که هم‌رتبه‌ها نیز به‌حساب آورده شوند[1].

تبصره‌ها

۱. ضریب همبستگی پی‌یرسون (گشتاور حاصل‌ضربی) اندازه‌ای از جهت و قوت پیوند خطی بین دو متغیر کمّی به‌دست می‌دهد. وقتی متغیرها ترتیبی باشند، باید از ضریب همبستگی اسپیرمن استفاده شود زیرا این ضریب مبتنی بر رتبه‌های هر متغیر است. این کار از مشکل به‌طور کمّی بامعنا فرض‌کردن فواصل بین‌نقطه‌ای واقعی یک مقیاس ترتیبی اجتناب می‌کند.

۲. توجه داشته باشید که مشتریان گاه ممکن است از اصطلاح «همبستگی» به معنای محاوره‌ای‌تری استفاده کنند که دلالت بر علیت دارد. این امر ممکن است منجر

۱. پرس و همکاران (۱۹۹۲، صفحهٔ ۶۴۰) را ببینید.

به سوءتفاهمی دربارهٔ این شود که یافته‌های یک تحلیل همبستگی (آماری) واقعاً چه چیزی را نشان می‌دهد.

۳. در عمل، مطالعه‌های مشاهده‌ای ممکن است مقادیر همبستگی پایین ۳، ۰٫۳ > |r|، را حاصل کنند. معناداری آماری r را می‌توان تحت این فرض ارزیابی کرد که هیچ پیوندی (خطی) بین X و Y وجود ندارد. تحت این فرض، خطای استاندارد r تقریباً $n^{-1/2}$ است و بنابراین از $|r| > 2n^{-1/2}$ چنین برداشت می‌شود که یک همبستگی معنادار وجود دارد.

۴. اگر n بزرگ باشد، در این صورت به‌دست‌آوردن یک همبستگی «معنادار» با مقادیر پایین r امکان‌پذیر است. در این موقعیت، تفسیر رگرسیونی r^2 سودمند است: تغییرپذیری موجود در Y که توسط X تبیین می‌شود برابر با r^2 است. بنابراین، ۰٫۳ > |r| دلالت بر این دارد که X کمتر از ۱۰٪ از تغییرپذیری در Y را تبیین می‌کند.

۵. یک اندازهٔ خاص که اغلب در رابطه با داده‌های آزمون پیمایشی مورد نیاز واقع می‌شود آمارهٔ آلفای کرونباخ (کرونباخ، ۱۹۵۱) است. از این آماره برای برآورد قابلیت اطمینان ابزار پیمایش استفاده می‌شود. فرض بر این است که مقادیر بالاتر این ضریب دلالت بر مناسب‌بودن قابلیت اطمینان دارد.[۱]

تبدیلات

- تقارن
- پراکندگی در برابر سطح
- تحلیل مانده‌ها

مجموعه‌داده‌هایی که واقعاً برای تحلیل به‌کار می‌روند اغلب چندین گام با مقادیر داده‌های اصلی یا خام فاصله دارند. فرایند پاکیزه‌سازی داده‌ها، مجموعه داده‌های اصلی را با انجام‌دادن اعمالی چون حذف خطاها، تعدیل نقاط دورافتاده، برآورد مقادیر گمشده، کدبندی متغیرهای رسته‌ای، و استانداردکردن

۱. ابزارهای پیمایشی اغلب شامل دو صورت از سؤالی واحد هستند (که سؤالات دیگری بین آن‌ها قرار داده می‌شوند). بنابراین، اگر بسیاری از پاسخ‌دهندگان با این سؤال «موافق» باشند که «آیا دربارهٔ x < x > احساس اطمینان می‌کنید؟» اما با این سؤال هم موافق باشند که «آیا x < x > موجب از دست‌رفتن اطمینان شما می‌شود؟»، در این صورت قابلیت اطمینان ابزار پیمایشی سؤال‌برانگیز خواهد بود.

متغیرها «مبدّل می‌کند». اعمال ساده‌ای مانند جمع، تفاضل، یا نسبت نیز ممکن است برای ایجاد متغیرهای جدید به‌کار روند و، در برخی کاربردها، متغیرها به‌طور خودکار به شیوه‌ای «پذیرفته‌شده» تبدیل می‌یابند. به‌عنوان مثال، عایدی‌های سهام به‌صورت $R_t = \log Y_t - \log Y_{t-1}$ تعریف می‌شود، که در آن Y_t قیمت سهام است. پس از کامل‌شدن مرحلهٔ پردازش داده‌ها، ممکن است تبدیل‌های دیگری برای مقاصد آماری مناسب باشند. برخی از دلایل متداول برای تبدیل‌کردن عبارت‌اند از:

- توزیع یک متغیر شدیداً چوله است.
- یک اثر پراکندگی در برابر سطح در داخل دسته‌ها وجود دارد.
- مانده‌های یک مدل برازش‌داده‌شده یک الگوی سیستماتیک را نشان می‌دهند.
- داده‌ها مفروضات یک روش آماری را برآورده نمی‌کنند.

دشواری عمده‌ای که در این موقعیت‌ها ظاهر می‌شود، بروز غیرخطی‌بودن است که ممکن است پیچیدگی تحلیل آماری را به میزان قابل توجهی افزایش دهد. با بکارگیری یک تبدیل غیرخطی بر داده‌ها، می‌توانیم این مشکلات را برطرف و ساختاری را ایجاد کنیم که تکیه بر تحلیل غیر پیچیده‌ای دارد و تفسیر ساده‌تری از نتایج را میسر می‌سازد.

تبدیل‌های توانی ردهٔ تبدیل‌های توانی ساده‌ترین و پرکاربردترین فن برای بیان مجدد مجموعه‌داده‌هایی است که مرکب از مقادیر مثبت باشند. این تبدیل‌ها به شکل زیرند: $T_p(x) = \text{sign}(p)x^p$ به‌ازای $p \neq 0$ و $T_p(x) = \log x$ به‌ازای $p = 0$، که در آن از $\text{sign}(p) = \pm 1$ برای حفظ ترتیب داده‌ها در مجموعه‌داده‌ها استفاده می‌شود. صورت کمی اختصاصی‌تر، ردهٔ **باکس-کاکس** از تبدیل‌های توانی است که در زیر تعریف می‌شود. این تبدیل، $\log_e$ را به‌عنوان حالتی خاص[1] شامل می‌شود که برای مقایسهٔ ویژگی‌های تبدیل‌های مختلف سودمند است.

$$Y_i(p) \equiv T_p^*(X_i) = \begin{cases} \frac{1}{p}(X_i^p - 1) & , p \neq 0 \\ \log_e(X_i) & , p = 0 \end{cases}$$

در عمل، از یک توان «مناسب» برای پرهیز از مشکلات تفسیری استفاده می‌شود. تبدیل‌هایی که بتوان یک توجیه زمینه‌ای به آن‌ها بخشید جذابیت بیشتری نسبت به توانی مانند $p = 0.43$ برای مشتری خواهند داشت. جدول ۱.۳ برخی از توان‌های متداول برای تبدیل

۱. می‌توان نشان داد که وقتی $p \to 0$، $\log_e x \to (x^p - 1)/p$.

جدول ۱.۳ تبدیل‌ها برای متقارن‌سازی

مجموعه داده‌های اصلی			
تبدیل	p	شکل	کاربرد
توان دوم (X)	۲	چوله به چپ	آستانهٔ بالایی
همانی	۱	متقارن	بدون تغییر
ریشهٔ دوم	۰٫۵	چوله به راست	اندازهٔ مساحت
ریشهٔ سوم	$\frac{1}{3}$	چوله به راست	اندازهٔ حجم
لگاریتمی	۰	چوله به راست	رشد جمعیت
ریشهٔ معکوس	۰٫۵−	چوله به راست	سطح معکوس
معکوس (X)	۱−	چوله به راست	نقاط دورافتادهٔ کرانگین

یک مجموعه‌داده را فهرست می‌کند. مقصود از یک تبدیل تلاش برای تحمیل تقارن به توزیع داده‌های تبدیل‌یافته است.

تقارن کاربردی مهم از تبدیل‌های توانی ایجاد تقارن در توزیع یک متغیر است. این کار دارای ویژگی مطلوب ایجاد یک «مرکز» طبیعی برای مجموعه‌داده‌ها است که درک برآوردهای مکانی را به‌لحاظ مفهومی ساده‌تر می‌سازد. کارکردن با برآوردهای مکانی بر مبنای توزیع‌های متقارن از مزیت‌های نظری نیز برخوردار است. وجه اشتقاق تبدیل باکس‌ـکاکس، از شیوهٔ آن‌ها برای انتخاب p تحت این فرض است که داده‌های تبدیل‌یافته دارای توزیع نرمال هستند (باکس و کاکس، ۱۹۶۴). در حال و هوای EDA، روشی منسوب به هینکلی (۱۹۷۷) را توصیف می‌کنیم که اجرای آن ساده‌تر است و می‌توان آن را به شکلی مقاوم درآورد.

تبدیل هینکلی برای تقارن هینکلی (۱۹۷۷) روشی برای انتخاب تبدیلی مناسب برای تقارن نمونه‌ای از داده‌ها، مبتنی بر این فرض منطقی ساده بنیان نهاد که نمونه‌های تصادفی حاصل از یک توزیع «متقارن»، اتحاد میانگین=میانه را بازتاب می‌دهند. در نتیجه، یک اندازهٔ نسبی از تقارن در یک نمونه به‌صورت

$$D_p = \frac{\{Y_i(p)\}\text{میانگین} - \{Y_i(p)\}\text{میانهٔ}}{\{Y_i(p)\}\text{پراکندگی}}$$

است که در آن $Y_i(p)$ مقدار تبدیل‌یافتهٔ X_i است و $D_p = 0$ دلالت بر تقارن دارد.

دو انتخاب برای آمارهٔ «پراکندگی» S و IQR هستند. از یک میانگین پیراسته نیز می‌توان برای رسیدن به مقاوم‌بودن استفاده کرد.

نمودار D_p این ابزار تشخیصی عبارت است از رسم نمودار D_p در برابر p که در آن p روی محور xها قرار دارد و حل معادلهٔ $D_p = 0$ از طریق درون‌یابی.

پراکندگی در برابر سطح برای K دسته از داده‌های مربوط، یک نمودار پراکندگی در برابر سطح، از طریق یک تبدیل توانی متداول، روشی برای پایدارسازی پراکندگی در داخل دسته‌ها فراهم می‌کند. این روش ارتباط نزدیکی با مسئلهٔ کلاسیک پایدارسازی واریانس دارد که در آن $\sigma_Y^2 = g(\mu_X)$ و ما در جستجوی تبدیلی مانند $Y = \psi(X)$ هستیم به‌قسمی که σ_Y^2 تقریباً مستقل از μ_Y باشد. در ابزار تشخیصی زیر، که توسط توکی (۱۹۷۷) معرفی شد، فرض می‌شود که وابستگی بین IQR و میانه به شکل $IQR_X = kM_X^b$ است که در آن k و b ثابت‌اند. هوگلین و همکاران (۱۹۸۳) نشان داده‌اند که تابع $\psi(X) = X^{1-b}$، IQR را پایدار می‌سازد و b متناظر با شیب نمودار زیر است.

نمودار پراکندگی در برابر سطح نمودار $\log(IQR_j)$ در برابر $\log(M_j)$ را برای دسته‌های $j = 1, 2, \ldots, K$، رسم کنید که در آن $\log(M_j)$ روی محور x قرار دارد. سپس می‌توان یک توان مناسب را با در نظرگرفتن $p = 1 - b$ به‌دست آورد که در آن b شیب خط مستقیم برازش‌داده‌شده به این نقاط است.

تبصره‌ها

۱. گاهی به‌کارگیری یک تبدیل برای مقاصد تحلیل با مقاومت شدیدی روبه‌رو می‌شود.

- مجموعه‌داده‌ها از قبل برای تحلیل «زمینه‌چینی» شده است ــ چه لزومی به پیچیده‌ترکردن تحلیل وجود دارد؟
- تبدیل موجب دشوارترشدن تفسیر می‌شود.

ترغیب‌کردن یک مشتری شکاک یا سمج به لزوم تبدیل داده‌ها ممکن است واقعاً دشوار باشد و در برخی موارد، حتی زمانی که مزیت‌های تبدیل به آن‌ها نشان داده شود، ممکن است باز هم مجاب نشوند. بنابر تجارب ما، توضیح‌دادن یافته‌ها بر مبنای داده‌های تبدیل‌یافته ممکن است آسان‌تر از تلاش برای دفاع از دلایل لزوم یک تبدیل باشد.

۲. افزایش پراکندگی با افزایش سطح در داده‌های دسته‌ای، اغلب با چولگی راست همراه

است. در این وضعیت، تأثیر تعمت غیرمترقبهٔ به‌کارگیری یک تبدیل این است که هر دو متبع غیر خطی‌بودن برطرف می‌شوند. با این وجود، وضعیت‌هایی وجود دارند که در آن‌ها تبدیل‌ها همیشه کارساز نخواهند بود:

- یک ساختار داده‌ها، که جنبه‌هایی مانند چندمدی‌بودن یا چولگی هم راست و هم چپ را نمایش می‌دهد، از یک تبدیل توانی سودی نخواهد برد.
- نمی‌توان یک تبدیل توانی را (مستقیماً) بر مجموعه‌ای از داده‌ها اعمال کرد که شامل مقادیر منفی است. در برخی حالت‌ها، افزودن یک مقدار ثابت یا به‌کارگیری تبدیل‌های جداگانه در مورد مقادیر منفی و مثبت «شاید» مناسب باشد.

۳. در بسیاری از روش‌های آماری فرض بر این است که می‌توان از مدل جمعی

$$\text{داده} = \text{مدل} + \text{خطا}$$

برای تحلیل یک ساختار داده‌ای استفاده کرد. تبدیل داده‌ها گاه می‌تواند، با حذف اثرهای متقابل بین مدل و خطا و پایدارسازی واریانس خطا، به ایجاد ساختار جمعی کمک کند. به‌عنوان مثال:

- نمودار پراکنش Y در برابر X ــ به‌کارگیری یک تبدیل توانی بر یک یا هر دو متغیر ممکن است امکان برازش یک مدل خط راست را فراهم سازد که تحلیل و استفاده از آن برای مقاصد پیشگویی ساده‌تر است.
- ANOVA ــ فرض برابری واریانس‌ها در یک مدل تحلیل واریانس، متناظر با ارزیابی نمودار پراکندگی در برابر سطح است.

جدول‌های پیشایندی

روش‌شناسی جدول‌های پیشایندی معطوف به تحلیل پیوند بین سطوح دو متغیر رسته‌ای است. روش تحلیل استاندارد، انجام‌دادن یک آزمون استقلال با محاسبهٔ آمارهٔ خی‌دوی پی‌یرسون یا، در حالت نمونه‌های کوچک، آزمون دقیق فیشر است. هر دو آزمون در زیر به‌تفصیل مورد بحث قرار می‌گیرند. پانل Crosstabs در شکل ۵.۳ مثالی از یک جدول پیشایندی و پانل Tests دربردارندهٔ نتایج آزمون خی‌دوی پی‌یرسون است. مطالعهٔ موردی ۱.۶ و تمرین مطالعهٔ موردی ۵.۹، هر دو مثال‌هایی از کاربردهایی هستند که دربردارندهٔ جدول‌بندی فراوانی بین دو متغیر رسته‌ای‌اند.

جدول‌های پیشایندی را می‌توان از انواع مختلفی از مطالعه‌ها تولید کرد و ماهیت یک مطالعهٔ خاص را باید در نظر گرفت. به‌عنوان مثال، در مطالعه‌های طولی (بخش ۱.۳)، تفسیر پیوند بین دو

متغیر رسته‌ای X (سطرها) و Y (ستون‌ها) به این بستگی خواهد داشت که آیا مطالعه، گذشته‌نگر است یا آینده‌نگر. تغییرپذیری دیگری نیز موسوم به مطالعه‌های مقطعی وجود دارد. تفاوت‌های بین این سه نوع مطالعه در زیر نشان داده شده است:

گذشته‌نگر مجموع‌های ستونی برای Y ثابت‌اند و توزیع سطوح X، مورد توجه است. به‌عنوان مثال، ۵۰ سیگاری و ۵۰ غیرسیگاری بر مبنای جنسیت به‌طور گذشته‌نگر رده‌بندی می‌شوند.

آینده‌نگر مجموع‌های سطری برای X ثابت‌اند و توزیع سطوح Y، مورد توجه است. به‌عنوان مثال، ۵۰ سیگاری و ۵۰ غیرسیگاری برای مطالعهٔ طولی وقوع یا عدم وقوع سرطان ریه انتخاب می‌شوند.

مقطعی اندازهٔ نمونهٔ کل ثابت است و توزیع توأم X و Y مورد توجه است. فرض بر این است که X و Y «به‌طور معناداری» به هم مربوط‌اند.

آزمون خی‌دو

آمارهٔ خی‌دوی پی‌یرسون را می‌توان در بسیاری از کتاب‌های درسی مقدماتی یافت (به‌عنوان مثال، مور و مه‌کیب (۱۹۹۸) را ببینید) که آزمونی را برای آزمودن فرض صفر فقدان پیوند دو متغیر (استقلال) فراهم می‌کند. با این حال، از آزمون خی‌دو صرفاً باید وقتی استفاده کرد که اندازهٔ نمونه به قدر کافی بزرگ باشد. در عمل، معیار متداول این است که نباید در بیش از ۲۰٪ از خانه‌ها تعداد مورد انتظار کمتر از ۵ باشد. گرچه جزئیات زیر برای بسیاری از خوانندگان آشنا خواهد بود، برای کامل‌بودن مطالب، فرایند استنباط را با توجه به این آمارهٔ آزمون خاص تشریح می‌کنیم.

جزئیات آزمون خی‌دو

برای مقاصد نمادگذاری، فرض کنید A و B نشان‌دهندهٔ دو عامل رسته‌ای با n_a و n_b سطح (رده) باشند. جدول پیشایندی مرتبط با نمونه‌ای از مشاهدات که برحسب سطوح این عامل‌ها به‌صورت متقاطع رده‌بندی شده‌اند، شامل $n_a \times n_b$ خانه است که هر خانه n_{ij} مشاهده دارد. آمارهٔ خی‌دوی پی‌یرسون به‌صورت

$$W = \sum_{i=1}^{n_a} \sum_{j=1}^{n_b} \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

محاسبه می‌شود که در آن $O_{ij} = n_{ij}$ تعداد (فراوانی) مشاهده‌شدهٔ خانه و E_{ij} تعداد مورد انتظار خانه است.

تحت این فرض که A و B مستقل‌اند، $E_{ij} = n_{i\bullet}n_{\bullet j}/n_{ij}$، که در آن $n_{i\bullet}$ و $n_{\bullet j}$ مجموع‌های حاشیه‌ای برای سطر A_i و ستون B_j هستند. فرض استقلال معادل با حکم احتمالاتی $P[A_i \cap B_j] = P[A_i]P[B_j]$ است که می‌توان آن را به‌صورت فرض زیر بیان کرد:

$H_\circ$: هیچ وابستگی بین A و B وجود ندارد.

برای نمونه‌های بزرگ، می‌توان نشان داد که تحت $H_\circ$ (درست‌بودن فرض)، W دارای توزیع تقریبی خی‌دو با $(n_a - ۱)(n_b - ۱)$ درجه‌آزادی است. اغلب می‌توان این را به‌اختصار به‌صورت $W \sim \chi^۲_{(n_a-۱)(n_b-۱)}$ نوشت.

از آنجا که هر تفاوتی بین O_{ij} و E_{ij} آمارهٔ آزمون را افزایش خواهد داد، از مقادیر بزرگ W چنین برمی‌آید که شواهد علیه $H_\circ$ است. بنابراین، یک نتیجهٔ معنادار ($P = P[\chi^۲_{(n_a-۱)(n_b-۱)} > W] < ۰٫۰۵$ـمقدار) این نتیجه‌گیری را که پیوندی بین عامل‌های A و B وجود دارد تأیید می‌کند. همان‌طور که در بالا ذکر شد، از آزمون خی‌دو صرفاً باید وقتی استفاده کرد که اندازهٔ نمونه به قدر کافی بزرگ باشد. برای اجتناب از داشتن تعداد زیادی E_{ij}، با $E_{ij} < ۵$، ممکن است ترکیب کردن سطوح لازم باشد.

آزمون دقیق فیشر

آزمون دقیق فیشر در اصل برای حالت ۲×۲ به‌کار گرفته شد که می‌توان نشان داد، تحت فرض استقلال، توزیع دقیق نمونه‌ای به شرط مجموع‌های حاشیه‌ای ثابت، توزیع فوق هندسی است (پیوست ج، جدول ج.۱ را ببینید). بنابراین، مزیت آزمون دقیق فیشر این است که حتی وقتی اندازهٔ نمونه کوچک باشد می‌توان از آن استفاده کرد.

برای جدول‌های ۲×۲، Pـمقدار برای آزمون دقیق فیشر بر اساس احتمال فوق هندسی مرتبط با فراوانی خانهٔ (۱،۱) تعیین می‌شود. از آنجا که مجموع‌های حاشیه‌ای ثابت‌اند، تغییرات در فراوانی خانه‌های جدول ۲×۲ی اصلی را می‌توان به کمک نسبت بخت‌ها که با رابطهٔ

$$\theta = \frac{n_{۱۱}n_{۲۲}}{n_{۱۲}n_{۲۱}}$$

ارائه می‌شود، مرتب کرد. بنابراین، می‌توان Pـمقدارهای یک‌طرفه و دوطرفه را که متناظر با فرض‌های مقابل مرتبط با فرض معادل با استقلال، یعنی $H'_\circ : \theta = ۱$، هستند محاسبه کرد. Pـمقدار دوطرفه را می‌توان به‌صورت مجموع احتمال‌های فوق هندسی همهٔ جدول‌هایی (تغییرات) تعریف کرد که وقوع آن‌ها محتمل‌تر از جدول مشاهده‌شده نیست.

با این حال، برای نمونه‌های کوچک، Pـمقدارها به‌شدت گسسته خواهند بود و آزمودن در یک سطح معناداری ثابت معمولاً امکان‌پذیر نیست.[۱] البته، به‌لحاظ شهودی، همواره این پرسش وجود دارد که آیا تحلیل آماری یک نمونهٔ کوچک لزوماً معنادار است یا خیر. معمولاً محاسبهٔ توان‌ها می‌تواند اندازه‌ای از کیفیت یک آزمون فراهم کند، اما این کار برای آزمون دقیق فیشر تا حدودی پیچیده‌تر است (به‌عنوان مثال، گیل و گارت (۱۹۷۳) را ببینید). به جای این کار، می‌توان بازه‌های اطمینان برای نسبت بخت‌ها، θ، را محاسبه کرد. بستهٔ نرم‌افزاری StatXact بازه‌های اطمینان دقیقی برای θ فراهم می‌کند.

آزمون دقیق فیشر برای استقلال را می‌توان به هر جدول پیشایندی $n_0 \times n_1$ تعمیم داد، گرچه در برخی موارد با هزینهٔ محاسباتی نسبتاً قابل توجه. این آزمون از نوع آزمون‌هایی نیست که بر روی جدول‌های پیشایندی متعددی با اندازه‌های دلخواه صرفاً برای اهداف اکتشافی اجراشدنی باشد. آزمون دقیق فیشر در اگرستی (۱۹۹۰) به‌تفصیل مورد بحث قرار گرفته است.

آزمون‌های دیگر

آزمون‌ها و اندازه‌های پیوند دیگر بسیاری وجود دارند که می‌توان آن‌ها را برای جدول‌های پیشایندی به‌کار برد. برای خلاصه‌ای مفصل‌تر، خواننده به بحثی که در بخش PROC FREQ از راهنمای SAS ارائه‌شده ارجاع داده می‌شود (SAS 1990).

تفسیر معناداری

موضوع مهم دیگر در تحلیل جدول‌های پیشایندی تفسیر نتایج در زمانی است که Pـ مقدارهای محاسبه‌شده از هر یک از آزمون‌های بالا معنادارند. در این موارد، در جدول به جستجوی الگوهای صریحی از پیوند بین متغیرها می‌پردازیم که یافته را توضیح می‌دهند، اما کاملاً امکان‌پذیر است که چنین الگوهایی واضح نباشند یا اینکه تفاسیر متناقضی موجود باشند. برای جدول‌های ۲ × ۲ تفسیر آسان است زیرا الگوهای جالب تنها آن‌هایی هستند که، در خانه‌های قطری یا غیر قطری، فراوانی‌های بیشتری از فراوانی‌های مورد انتظار داشته باشند. اما برای جدول‌های بزرگتر ممکن است تفسیر رابطه در برخی موارد دشوار باشد، حتی هنگامی که آزمون رابطه‌ای معنادار را نشان می‌دهد. در مطالعهٔ موردی ۱.۶ در رابطه با تحلیل جدول‌های ۱.۶ و ۳.۶ به این نکته بازمی‌گردیم.

۱. منظور ما از «به‌شدت‌گسسته‌بودن» این است که تنها تعداد کمی از Pـمقدارهای ممکن واقعاً رخ می‌دهند. در نتیجه، مقایسهٔ Pـمقدار با ۰٫۰۵ تا حدودی گمراه‌کننده است زیرا ممکن است این کران بالای اندازهٔ نمونهٔ داده‌شده دست‌یافتنی نباشد. می‌توان از آزمون‌های تصادفی استفاده کرد، اما در این صورت لازم است (مشتری را) توجیه کنیم که چرا برای تصمیم‌گیری دربارهٔ معناداری نتیجه، یک سکه (اریب) انداختیم.

تبصره‌ها

۱. تحلیل جدول‌های پیشایندی را شاید با جزئیاتی بیشتر از آنچه مقدمتاً لازم است ارائه کرده‌ایم زیرا این نوع از تحلیل اغلب (!) در موارد تشخیصی بکار می‌رود. همانگونه که از مطالعهٔ موردی ۱.۶ برمی‌آید، تلاش برای تبیین کاربرد و تفسیر آزمون خی‌دوی پی‌یرسون یا آزمون دقیق فیشر به یک وکیل دعاوی که طرف مقابل ما است ممکن است تجربه‌ای نسبتاً ناامیدکننده باشد.

۲. دیگر مشتریان در مقایسه با وکیل دعاوی طرف مقابل ما حرف‌شنوتر خواهند بود، اما همچنان بر سر جزئیات مشابهی مجادله می‌کنند. درایه‌های `% row` و `% col` در پانل **Crosstab** شکل ۵.۳ راهی ساده برای نشان‌دادن فرض استقلال فراهم می‌کنند (اگر H_0 پذیرفته شود، تقریباً یکسان خواهند بود؛ در غیر این صورت تفاوت‌هایی بدیهی وجود خواهد داشت).

۳. اگر شواهدی بر یک پیوند معنادار وجود داشته باشد، ممکن است لازم باشد که مشتری درایه‌های `% row` و `% col` و `% total` را به‌دقت مرور کند. توانایی تعیین معناداری از طریق P-مقدار خود به خود به این معنا نیست که مشتری قادر به تفسیر پیوند خواهد بود. این امر، به‌ویژه وقتی مهم است که یک متغیر دربردارندهٔ رسته‌ای همچون «سایر موارد» باشد. در این حالت، پیوند معنادار ممکن است تفسیر معناداری نداشته باشد.

آزمون t

علاوه بر آزمون‌های tی تک‌نمونه‌ای و دونمونه‌ای، از آزمون t به‌طور گسترده‌ای در بسیاری از شیوه‌های آماری دیگر همچون رگرسیون و تحلیل واریانس استفاده (و سوءاستفاده) می‌شود. یک آزمون tی جفتی نیز وجود دارد که، همانگونه که در مثال زیر نشان داده شده، می‌توان از آن در آزمایش‌های قبل‌ـوـبعد، استفاده کرد.

مثال۴.۳ آزمایش دوسرکور

در یک مطالعهٔ دارونما در برابر داروی فعال، پاسخ هر بیمار به دو دارو در زمان‌هایی جداگانه اندازه‌گیری می‌شود. ترتیب داروها تصادفی شده است و فرض می‌شود که، در هر دو زمان، شرایط مشابهی[1] حاکم باشد. برای اجتناب از اریبی «مصاحبه»، باید نه بیمار و نه فردی که دارو را به بیمار می‌دهد از اینکه چه دارویی داده می‌شود مطلع نباشند. به این آزمایش یک اجرای دوسرکور یک آزمایش قبل ـوـ بعد اطلاق می‌شود.

۱. فرض «شرایط مشابه» نشان‌دهندهٔ این است که هیچ اثر به‌جامانده یا منتقل‌شده‌ای از نخستین دارو وجود ندارد.

آزمون tی جفتی از یک آزمون tی جفتی می‌توان برای تحلیل تفاوت‌های بین دو پاسخ از هر بیمار استفاده کرد. از آنجا که این طرح جفتی شرایط اولیهٔ یک بیمار (پاسخ طبیعی به شرایط مورد نظر) را به‌حساب می‌آورد، میانگین تفاوت ($\overline{D} = \overline{X}_{\text{بعد}} - \overline{X}_{\text{قبل}}$) به تغییرپذیری درون‌گروهی بزرگ، که اغلب در این گونه آزمایش‌ها وجود دارد، حساسیت کمتری خواهد داشت.

آزمون tی دونمونه‌ای یک اجرای ساده‌تر این است که از دو گروه مستقل استفاده شود. با تصادفی‌سازی، هر بیمار فقط یکی از داروها را دریافت می‌کند و می‌توان از یک آزمون tی دونمونه‌ای برای ارزیابی معناداری تفاوت بین دو میانگین پاسخ استفاده کرد. با این حال، اگر پاسخ‌های درون هر گروه (مطابق معمول) به‌شدت متغیر باشند، این امر بر خطای آزمایشی غلبه می‌کند و تفاوت معنادار بین میانگین‌های گروهی را می‌پوشاند.

در فصل بعد، ارائهٔ مفصلی از یافته‌های حاصل از آزمون tی دونمونه‌ای به مشتری را دنبال می‌کنیم (جلسهٔ مشاوره). همچنین از آزمون tی دونمونه‌ای برای تحلیل مقدماتی در مطالعهٔ موردی ۳.۶ استفاده می‌شود و پیوست ج دربردارندهٔ جدول‌هایی از برخی شیوه‌های استاندارد آزمون به‌انضمام آزمون tی تک‌نمونه‌ای و دونمونه‌ای است. بنابراین، ما بر روی برخی از ویژگی‌های آماری آزمون tی دونمونه‌ای تمرکز می‌کنیم. نخست صورت‌های دوطرفه و یک‌طرفهٔ فرض مقابل برای این آزمون را در نظر می‌گیریم.

آزمون دوطرفه ترجمهٔ آماری آزمون tی دوطرفهٔ تفاوت بین میانگین‌های دو گروه مستقل مانند A و Bی مشتری به‌صورت

$$H_0 : \mu_A = \mu_B \quad \text{در برابر} \quad H_1 : \mu_A \neq \mu_B$$

است. این معمولاً آزمونی از وضع موجود است: هیچ چیز تغییر نکرده در برابر اینکه چیزی تغییر کرده. Pـمقدار این آزمون $P[T > |t_0|]$ است که در آن $t_0 = (\overline{x}_A - \overline{x}_B)/se(\overline{x}_A - \overline{x}_B)$ مقدار مشاهده‌شدهٔ آمارهٔ آزمون است. قریباً دربارهٔ عبارت خطای معیار $se(\overline{x}_A - \overline{x}_B)$ مختصراً بحث می‌کنیم.

آزمون یک‌طرفه یک آزمون یک‌طرفه اغلب به‌صورت

$$H_0 : \mu_A = \mu_B \quad \text{در برابر} \quad \begin{cases} H_1 : & \mu_A > \mu_B \quad \text{(راست‌دم)} \\ H_1 : & \mu_A < \mu_B \quad \text{(چپ‌دم)} \end{cases}$$

بیان می‌شود. مسئلهٔ اصلی این است که آیا آزمون‌کردن صرفاً در یک جهت قابل توجیه است یا خیر. دربارهٔ جهت آزمون نمی‌توان پس از آزمایش تصمیم‌گیری کرد. شکل H_0

نشان‌دهندهٔ نقطه‌ای است که در آن خطای نوع I آزمون بیشترین مقدار را خواهد داشت. آمارهٔ آزمون مانند t_0 بالا است، اما Pـمقدار برای آزمون راست‌دم $P[T > t_0]$ و برای آزمون چپ‌دم $P[T < t_0]$ است.

خطای استاندارد آزمون tی دونمونه‌ای

اینجا جایی است که مشتری به احتمال زیاد سردرگم می‌شود. چرا دو آزمون t وجود دارد؟ آیا یکی از آن‌ها چیز دیگری را آزمون می‌کند؟ چگونه می‌شود درجه‌های آزادی «کسری» داشت؟ اینجا، لازم است مشاور مراقب باشد که بیش از حد به مطالب فنی متوسل نشود زیرا یافته‌ای معنادار از هر یک از این دو آزمون تفسیر یکسانی از لحاظ فراهم‌کردن شواهدی مبنی بر اینکه میانگین‌های دو فرایند متفاوت به نظر می‌رسند دارد.

یک حالت فرض می‌کند که دو نمونه از فرایندهای مستقل (نرمال)، که واریانس‌های برابر دارند، می‌آیند ($\sigma_A^2 = \sigma_B^2$). در این حالت، از برآورد واریانس نمونه‌ای ادغام‌شده استفاده می‌شود (ن. ک. پیوست ج، جدول ج.۶) و، تحت H_0، $T \sim t_{n_A+n_B-2}$. در صورت دوم، واریانس‌های فرایندهای نرمال می‌توانند نابرابر باشند، اما آمارهٔ آزمون تحت H_0 دارای توزیع t نیست. آزمون اسمیت‌سترت‌ویت[1]، توزیع فرض صفر را با t_ν تقریب می‌زند که در آن پارامتر درجه‌آزادی از روی میانگین موزون واریانس‌های نمونه‌ای (ن. ک. جدول ج.۶) محاسبه می‌شود.

موضوع اصلی این است که آیا فرض برابری واریانس‌ها معقول است یا خیر. انحراف استانداردهای نمونه‌ای شواهدی تجربی در اختیار می‌گذارند، اما یک دلیل بنیادین مربوط به زمینهٔ بحث را باید ارجحیت داد (به‌عنوان مثال، اندازه‌گیری عملکرد مردان و زنان در یک امتحان علوم: ممکن است تفاوتی جنسیتی وجود داشته باشد، اما این فرض که تغییرپذیری در عملکرد برای هر دو گروه یکسان است معقول به نظر می‌رسد). این موضوع مجدداً در فصل بعد مورد بحث قرار می‌گیرد.

تبصره‌ها

۱. البته، پیش از تفسیر یا اجرای یک آزمون t، باید وارسی کنیم که آیا شرایط آزمون به‌وسیلهٔ داده‌ها تأیید می‌شوند یا خیر. امتحان‌کردن چیزهایی ساده از قبیل موارد زیر را فراموش نکنید: آیا داده‌ها تشکیل نمونه‌ای تصادفی می‌دهند؟ آیا دو نمونه مستقل‌اند؟

۲. آزمون tی دونمونه‌ای نسبت به دورشدن از فرض‌های توزیعی، استوار است. مور و مک‌کیب (۱۹۹۳) توصیفی عالی از این ویژگی استواری فراهم کرده‌اند. شبیه‌سازی‌های آن‌ها نشان

1. Smith-Satterthwaite

می‌دهد که آزمون tی دونمونه‌ای، حتی وقتی که توزیع‌ها نرمال نیستند، استنباط قابل اعتمادی را میسر می‌سازد.

۳. در برخی وضعیت‌ها اجرای یک آزمون t «ابلهانه» است. به‌عنوان مثال، اندازهٔ نمونه آن‌قدر بزرگ است (مثلاً $n > ۱۰۰۰$) که می‌توان نتایج معنی‌داری را انتظار داشت. با این حال، «معنی‌داری» این یافته‌ها، شاید هم چندان افادهٔ معنی نکند. مثالی دیگر زمانی است که یکی از نمونه‌ها بسیار کوچک باشد. در این حالت، توان آزمون tی دونمونه‌ای شاید به قدری ضعیف باشد که نتیجهٔ قابل اعتمادی از آن حاصل نشود.

تحلیل واریانس (ANOVA)

برای مقایسهٔ میانگین‌های چندین گروه، شیوهٔ تحلیل واریانس (ANOVA)، روشی برای ارزیابی تفاوت‌های معنی‌دار بین (برخی از) میانگین‌ها را فراهم می‌آورد. این یک فن عام است که فیشر آن را برای افرازکردن تغییرپذیری کلی یک متغیر پاسخ به مؤلفه‌های وابسته به یک یا چند عامل و با خطای تصادفی ابداع کرد. نوشتگان موجود دربارهٔ فن ANOVA بسیار گسترده است و ارائه‌کرد ما تنها قادر است که به برخی از مباحث مرتبط به این روش‌شناسی مختصراً اشاره کند. مونتگومری (۱۹۹۷) و باکس و همکاران (۱۹۷۸) بحث جامعی از این روش‌شناسی را همراه با مثال‌های مرتبط ارائه می‌کنند.

ANOVAی یک‌طرفه

در ساده‌ترین حالت با یک عامل تک یا تیمار با دو سطح یا بیشتر، شیوهٔ ANOVAی یک‌طرفه تعمیمی از آزمون tی دونمونه‌ای است. به این معنی که، به جای مقایسهٔ I، $I > ۲$، میانگین سطوح تیمار به‌صورت «دو به دو»، شیوهٔ ANOVA فرض صفر وجودنداشتن هیچ اثر تیمار، یعنی $H_۰ : \mu_۱ = \mu_۲ = \cdots = \mu_I$ را مستقیماً آزمون می‌کند. مقایسهٔ (همزمان) میانگین‌های سطوح (گروهی) تیمار با مقایسهٔ تغییرپذیری بین میانگین‌های گروهی با تغییرپذیری (خطای) درون گروه‌ها انجام می‌شود. بنابراین، نتیجه‌ای معنی‌دار از یک ANOVAی یک‌طرفه صرفاً شواهدی از تفاوت بین میانگین‌های گروهی فراهم می‌کند؛ این نتیجه مشخص نمی‌کند که میانگین‌های گروهی «چگونه» با هم تفاوت دارند.

نتایج حاصل از به‌کارگیری شیوهٔ ANOVA را می‌توان به شکل یک «جدول ANOVA» ارائه کرد. برای یک طرح یک‌طرفهٔ متعادل، با I سطح تیمار و K مشاهده در هر سطح، جدول ANOVA اثر تیمار، مثلاً A، را برحسب مجموع توان‌های دوم تجزیهٔ تغییرات پاسخ (Y) خلاصه می‌کند: SST=SSA+SSE، که در آن $\text{SST} = (n-۱)s_y^۲$ و $n = IK$. بنابراین، جملات SSA و SSE متناظر با تجزیهٔ SST به‌ترتیب به تغییرات «بین‌گروهی» و «درون‌گروهی» هستند.

منبع	DF	SS	MS=SS/DF	مقدار F
تیمار A	$I-1$	SSA	MSA	$F_0 = \frac{MSA}{MSE}$
خطا	$I(K-1)$	SSE	MSE	
کل	$n-1$	SST		

مقدار F تعریف‌شده به‌وسیلهٔ F_0، آمارهٔ آزمونی برای ارزیابی اثر تیمار در اختیار می‌گذارد. برای فراهم‌کردن یک مقایسهٔ بامعنی، F_0 با استفاده از میانگین مجموع توان‌های دوم ساخته می‌شود (MS=SS/DF) که در آن DF نشان‌دهندهٔ پارامتر درجه‌آزادی است. بنابراین، MSA و MSE به‌ترتیب نمایندهٔ تغییرپذیری «متوسط» مرتبط با مؤلفه‌های تیمار و خطا هستند. اگر بتوان فرض‌های نرمال‌بودن و ثابت‌بودن واریانس را در سطوح تیمار در نظر گرفت، آنگاه تحت فرض صفر H_0: فقدان اثر تیمار، $F_0 \sim \mathcal{F}_{I-1,I(K-1)}$. هر اختلافی بین میانگین‌های سطوح مشاهده‌شدهٔ تیمار سهمی مثبت بر SSA دارد (و بنابراین از SSE خواهد کاست زیرا SST ثابت است)، در نتیجه مقدار بزرگ F_0 شواهدی علیه H_0 در اختیار خواهد گذاشت.

Pـمقدار برای آزمون ANOVAی H_0 برابر با $P[\mathcal{F}_{I-1,I(K-1)} > F_0]$ است که در اغلب بسته‌های نرم‌افزاری آماری در خروجی جدول ANOVA منظور شده است. بنابراین، اگر مدل ANOVAی یک‌طرفه برای داده‌ها مناسب باشد و H_0 رد شود، می‌توان یک تحلیل پساآزمایشی را برای بررسی رابطهٔ بین میانگین‌های سطوح تیمار اجرا کرد. به این موضوعات در شیوهٔ ANOVAی دوطرفه که در زیر بحث می‌شود پرداخته می‌شود.

ANOVAی دوطرفه

ANOVAی دوطرفه وقتی به‌کار می‌رود که هدف مطالعه، تشخص همزمان در تأثیر دو عامل، مثلاً A و B، باشد. این کار تحلیل را برای آماردان جالب‌تر می‌کند زیرا ممکن است یک اثر متقابل (که با AB نشان داده می‌شود) بین دو عامل وجود داشته باشد؛ به این معنی که اثر عامل A بر پاسخ ممکن است در همهٔ سطوح عامل B یکسان نباشد. اگر این امر رخ دهد، آنگاه مقایسه‌ها بین میانگین‌های سطوح تیمار یک عامل (مثلاً A) ممکن است به دلیل اثر متقابل AB پوشیده بماند که از این مطلب (به‌طور نادرست) چنین برداشت می‌شود که عامل A هیچ اثری بر پاسخ ندارد. بنابراین، نخست باید تغییرپذیری وابسته به اثر متقابل AB را امتحان کرد.

در حالی که این موضوعات ممکن است علاقهٔ آماردان را برانگیزد، برای برخی مشتریان، ANOVAی دوطرفه ممکن است بیشتر شبیه بازکردن جعبهٔ پاندورا[۱] باشد. اندازه‌گیری‌های

۱. Pandora's Box (پاندورا، در افسانه‌های یونانی، نخستین زن فانی است که از روی کنجکاوی جعبه‌ای را باز و ←

مکرر، عامل‌های بلوک‌بندی، طرح‌های متقاطع و آشیانی، اثرهای متقابل، فرض‌های نوع III، تحلیل اثرهای آمیخته... . حالا دیگر خیلی دیر است. آن مشتری چراغ آن روز آماردان را روشن کرده است. بنابر تجارب ما، آشناکردن مشتری با روش‌شناسی و مفاهیم زمینه‌ای ANOVAی دوطرفه ممکن است گاهی فرایندی دشوار باشد. صبور باشید. صرف وقت و تلاش برای روشن‌کردن این مفاهیم دشوار برای مشتری بسیار ارزشمند است. و سود آن برای مشتریان ما؟ ... خوب، آن‌ها بودند که سر جعبه را باز کرده‌اند، پس بگذارید نگاهی به درون آن بیندازیم.

جعبهٔ پاندورا

گذر به ANOVAی دوطرفه اغلب دلالت بر مطالعاتی دارد که در آن‌ها یکی از عامل‌ها تنها دو سطح دارد (مثلاً کنترل در مقابل تیمار). مشتری ممکن است صرفاً بخواهد بداند که آیا «تفاوت آماری معنی‌داری» بین دو سطح متغیر نسبت به متغیر پاسخ وجود دارد یا خیر. ... و، «اوه، بله. می‌خواهیم این را نیز بدانیم که آیا این مطلب برای مردان و زنان، هر دو، درست است» (ای وای! مشتری همین حالا در جعبهٔ پاندورا را باز کرد).

آزمون‌های *t*ی چندگانه مشکل انجام‌دادن آزمون‌های *t*ی چندگانه آن است که خطای نوع I آزمایشی (کل) کنترل نمی‌شود؛ به این معنی که شانس یک خطای نوع I تکی (که به‌غلط اختلاف معنی‌داری بین دو سطح یک عامل تیمار را آشکار می‌کند) برای هر آزمون *t*ی مقایسه‌ای α است (نوعاً ۰٫۰۵). اما شانس ارتکاب حداقل یک خطای نوع I مقایسه‌ای با آزمون‌های *t*ی چندگانه ممکن است بسیار بزرگتر باشد. برای k مقایسه، کران بالایی برای خطای نوع I مقایسه‌ای، $1 - (1 - \alpha)^k$ است. چون این مقدار به ۱ همگراست، نتیجه می‌شود که نباید همزمان از آزمون‌های *t*ی چندگانه در ارزیابی معنی‌داری عامل‌های ANOVA استفاده کرد. به جای آن، فرض ANOVAی H_0: «هیچ اثر تیمار وجود ندارد» را باید همواره پیش از اجرای هر گونه تحلیل مقایسه‌ای پسآزمایشی آزمون کرد.

در سناریوی جعبهٔ پاندورای ما، این مطلب اساساً بدین معنی است که ممکن است تا ۲۵٪ شانس[1] آشکارسازیِ به‌غلط اختلافی معنی‌دار بین برخی ترکیب‌های مذکر/مؤنث و سطوح تیمار/کنترل وجود داشته باشد. این بدان معنی است که این آزمون‌های *t* شواهدی قابل

قمتهٔ اعمال شر انسان را به جهان وارد کردیم.)

۱. در واقع، این مقدار $1 - ۰٫۹۵^۴ = ۰٫۲۶۴۹$ است، اما مشتری احتمالاً انجام‌دادن دو آزمون *t*ی اثر متقابل را در نظر نداشته است. نکتهٔ اصلی این است که این مطلب به ما کمک می‌کند تا بر این مشکل که در صورت کنترل‌نکردن خطای مقایسه‌ای پیش می‌آید تأکید کنیم.

اعتماد از اثری معنی‌دار ناشی از جنسیت یا عامل کنترل/تیمار در اختیار ما **نمی‌گذارند**. بنابراین ضروری است که تحلیل ANOVAی دوطرفه را اجرا کنیم.

هر بار یک عامل این رویکرد جاذبهٔ شهودی خاصی دارد: برای بررسی اثر یک عامل بر پاسخ، صرفاً عامل دیگر را ثابت نگه دارید. با این حال، مشکل رویکرد هر بار یک عامل آن است که این روش نه‌تنها از طرح عاملی دوعاملی متناظر کارایی کمتری دارد، بلکه نتایج آن ممکن است هنگامی که اثر متقابل وجود دارد گمراه‌کننده باشد. این مطلب در مثال زیر تشریح شده است.

مثال۵.۳ هر بار یک عامل در مقابل طرح عاملی ۲ × ۲.

آن‌گونه که در جدول‌های مثال ۵.۳ تشریح شده است، مزیت طرح ۲ × ۲ آن است که برآوردهای اثرهای اصلی برای A و B درست به همان اندازهٔ مقدارهای به‌دست آمده از هر بار یک عامل دقیقند، اما تنها چهار مشاهده برای این کار مورد نیاز است. مهم‌تر از آن، این نتیجه‌گیری که امکان دارد ترکیب A_2B_2ی مشاهده‌نشده در طرح هر بار یک عامل حتی پاسخ بزرگتری تولید کند، ممکن است در صورت وجود اثر متقابل کاملاً نادرست باشد.

هر بار یک عامل

	B_1		B_2	
A_1	۸	۱۲	۳۸	۴۶
A_2	۴۷	۵۳		

۲ × ۲ عاملی

	B_1	B_2
A_1	۱۰	۴۰
A_2	۵۰	۵

تصادفی‌سازی مثال بالا ممکن است کمی با طرح و تدبیر جلوه کند از آن رو که احتمالاً ترکیب A_2B_2 به هر صورت اجرا می‌شد (حتی اگر تنها به این علت بود که ببینیم پاسخ‌ها چه بودند). در نتیجه، این آزمایش امکان برآوردکردن اثر متقابل را نیز می‌دهد زیرا ما اکنون دو تکرار از یک ۲ × ۲ی عاملی را داریم ... . یا آیا چنین است؟ این مطلب بستگی به این دارد که آیا تصادفی‌سازی به‌درستی در آزمایش به‌کارگرفته شده است یا خیر. برای اینکه بتوان تحلیل آماری یک آزمایش عاملی دوعاملی را معتبر دانست، لازم است که همهٔ ترکیب‌های سطوح تیمار به ترتیبی تصادفی اجرا شوند. بنابراین تعیین اینکه مشاهدات دقیقاً «چگونه» گردآوری شده‌اند مهم است.

برای تشریح بهتر مسئله، فرض کنید که عامل B دارای k، $k > ۲$، سطح است. در رویکرد هر بار یک عامل کاملاً امکان‌پذیر است که عامل دوسطحی A را (مثلاً در A_1) ثابت گرفت و پاسخ را در هر سطح عامل B مشاهده کرد. سپس همین شیوه با ثابت نگه‌داشتن سطح A_2 اجرا می‌شود. حتی اگر سطوح B به ترتیبی تصادفی اجرا شوند،

نتایج با یک طرح عاملی توافق نخواهند داشت زیرا تصادفی‌سازی به «تیمارهای درون بلوک‌ها» محدود شده است؛ بدین معنی که A عملاً به یک عامل بلوک‌بندی تبدیل می‌شده است و این امر بر چگونگی عمل در تحلیل آماری اثر می‌گذارد.

طرح با اثرهای ثابت

جدول ANOVA برای یک طرح عاملی با **اثرهای ثابت** دوطرفهٔ متعادل، با I سطح تیمار برای عامل A، J سطح تیمار برای عامل B، و K تکرار، در زیر ارائه شده است. این فرض که هر دو عامل تیمار «ثابت‌اند» بدین معنی است که استنباط‌های برگرفته از این تحلیل تنها قابل اعمال به سطوح خاصی از A و B هستند که محقق برگزیده است. به روالی مشابه با حالت یک‌طرفه، جدول ANOVA اثرهای اصلی A و B، و اثر متقابل AB را، بر اساس تجزیهٔ مجموع توان‌های دوم تغییرپذیری متغیر پاسخ، خلاصه می‌کند.

منبع	DF	SS	MS	مقدار F
تیمار A	$I-1$	SSA	MSA	$F_A = \frac{\text{MSA}}{\text{MSE}}$
تیمار B	$J-1$	SSB	MSB	$F_B = \frac{\text{MSB}}{\text{MSE}}$
اثر متقابل AB	$(I-1)(J-1)$	SSAB	MSAB	$F_{AB} = \frac{\text{MSAB}}{\text{MSE}}$
خطا	$IJ(K-1)$	SSE	MSE	
کل	$n-1$	SST		

مدل ANOVAی دوطرفه را می‌توان به‌صورت

$$y_{ijk} = \mu_{ij} + \epsilon_{ijk} \qquad (۱.۳)$$

نوشت که در آن μ_{ij} نشان‌دهندهٔ اثر ترکیبی سطوح iام و jام عامل A و B بر پاسخ y_{ijk} است، و به‌طور مستقل $\epsilon_{ijk} \sim N(۰, \sigma^۲)$. آمارهٔ آزمون برای یک اثر تکی را می‌توان با استفاده از پارامتری‌سازی $\mu_{ij} = \mu + \alpha_i + \beta_j + \gamma_{ij}$ ساخت که در آن μ میانگین پاسخ کل، α_i اثر سطح iام عامل A، β_j اثر سطح jام عامل B، و γ_{ij} اثر متقابل بین α_i و β_j است. اگر فرض‌های زمینه‌ای این مدل به‌وسیلهٔ داده‌ها برآورده شوند، آنگاه توزیع صفر مقدار F متناظر برای هر اثر عبارت است از $\mathcal{F}_{\text{DF(effect)},\text{DF(error)}}$.

مانند حالت یک‌طرفه، یک مقدار F بزرگ، شواهدی بر ضد فرض صفر برای اثر متناظر فراهم می‌کند. با این حال، همیشه باید نخست آزمون اثر متقابل AB را بررسی کرد. دلیل این امر آن است

که آزمودن H_A یا H_B در صورتی که فرض «H_{AB}: اثر متقابلی وجود ندارد» رد شود فایده‌ای ندارد، زیرا اختلاف بین هر دو سطح یک اثر اصلی، یک اثر متقابل «متوسط» را نیز در بردارد. به عبارت دیگر، صرف‌نظر از نتایج آزمون برای H_A و H_B، در صورت وجود یک اثر متقابل معنی‌دار، «هر دو» عامل تیماری مهم‌اند (در غیر این صورت چگونه ممکن است اثر متقابل وجود داشته باشد؟).

مباحث تشخیصی

پیش از اتخاذ هر گونه نتیجه‌ای بر مبنای جدول ANOVA، لازم است که کفایت مدل را وارسی کنیم. روش‌های EDA را می‌توان برای امتحان‌کردن داده‌های پاسخ مستقیماً بکار گرفت، اما آنچه در واقع به آن نیاز داریم «مانده‌ها»ی حاصل از مدل ANOVAی برازش‌یافته است تا امتحان‌های تشخیصی را برای کفایت مدل به اجرا گذاریم. از (۱.۳) نتیجه می‌شود که می‌توان مانده‌ها را به‌صورت $e_{ijk} = y_{ijk} - \overline{y}_{ij}$ محاسبه کرد که در آن $\overline{y}_{ij}$ میانگین روی K پاسخ در خانهٔ i، jـام است.

تحلیل مانده‌ها نقشی بنیادی در وارسی کفایت مدل دارد. اگر مدل ANOVA برای داده‌ها مناسب باشد، مانده‌ها نباید هیچ ساختار بدیهی را نشان دهند. برای مقاصد استنباطی فرض می‌شود که خطاها دارای توزیع نرمال با واریانس ثابت‌اند. مباحث تشخیصی زیر برای وارسی اینکه آیا این جنبه‌ها در نمونهٔ مانده‌ها منعکس شده‌اند یا خیر بکار می‌روند.

نرمال‌بودن نمودار QـQ را می‌توان برای ارزیابی نموداری انحراف از نرمال‌بودن بکار برد. آمارهٔ شاپیرو[1]ـ ویلک[2] یک آزمون آماری رسمی برای نرمال‌بودن در اختیار می‌گذارد. برای آزمایش‌های با اندازه‌های کوچک تا متوسط (یعنی تعداد کل اجراهای $n < ۳۰$)، این ابزارهای تشخیصی را در ارتباط با شواهد غیر نرمال‌بودن باید با احتیاط تفسیر کرد.

واریانس نمودار مانده‌ها در مقابل برازش‌یافته را می‌توان برای ارزیابی نموداری اینکه آیا تغییرپذیری درون‌نمونه‌ای ثابت است یا خیر بکار برد. رسم‌کردن نمودار مانده‌ها نسبت به سطوح یک عامل را می‌توان برای ارزیابی اینکه آیا تغییرپذیری در راستای گروه‌ها همگن به نظر می‌رسد یا خیر، بکار برد. توجه کنید که استانداردکردن مانده‌ها پیش از رسم نمودار، مقایسهٔ مقادیر آن‌ها را با ± ۲ آسان‌تر می‌کند (ن. ک. شکل ۷.۳). همچنین آزمون‌های رسمی «همگنی واریانس‌ها در ANOVA» موجودند که مانند آزمون بارتلت می‌توان از آن‌ها استفاده کرد. ما خواننده را به اندرسون و مک‌لین (۱۹۷۴) برای مروری بر آزمون‌های برابری واریانس ارجاع می‌دهیم.

1. Shapiro 2. Wilk

تحلیل پساآزمایشی

نخست باید مقدار F مرتبط با اثر متقابل AB را امتحان کرد. اگر اثر متقابل معنی‌دار باشد، مقایسه‌های بین سطوح یک اثر اصلی ممکن است به دلیل اثر متقابل AB پوشیده بمانند. یک نمودار اثر متقابل، نمایش نموداری ساده‌ای از میانگین پاسخ است که در برابر سطوح یک عامل، مثلاً A، در سطحی مفروض از عامل B رسم می‌شود. میانگین‌ها به‌وسیلهٔ خطوط به هم وصل می‌شوند تا «نمایه‌ای از پاسخ متوسط در امتداد سطوح A ایجاد شود. این کار برای هر سطح عامل B تکرار می‌شود و نمایه‌ها در یک نمودار روی هم قرار داده می‌شوند. اگر نمایه‌ها به‌طرز آشکاری متفاوت باشند، از آن چنین برداشت می‌شود که اثر متقابل در بین است. اگر همهٔ نمایه‌ها شکل یکسانی داشته باشند، وجود هیچ اثر متقابلی از آن برداشت نمی‌شود.

وقتی اثر متقابل معنی‌دار نباشد، اختلاف‌های بین سطوح از یک اثر اصلی (معنی‌دار) را می‌توان با استفاده از روش‌های مقایسهٔ چندگانهٔ گوناگونی بررسی کرد:

- کمترین اختلاف معنی‌دار[1] (LSD) فیشر متناظر با استفاده از آزمون‌های tی جفتی است. وقتی تعداد سطح‌ها افزایش می‌یابد (لازم است تعداد بیشتری مقایسهٔ جفتی انجام شود)، خطای نوع I آزمایش ممکن است با استفاده از روش LSD بزرگ شود.
- آزمون دامنهٔ چندگانهٔ دانکن یکی از رایج‌ترین روش‌های مقایسهٔ چندگانه است زیرا توان خوبی دارد؛ بدین معنی که در آشکارسازی اختلاف‌های بین میانگین‌ها، وقتی واقعاً اختلاف‌هایی وجود داشته باشند، بسیار کاراست. این آزمون به‌وسیلهٔ دانکن (۱۹۵۵) ابداع شد.
- توکی (۱۹۵۳)[2] آزمون دامنهٔ استیودنتیده (HSD) را مطرح کرد که یک مقدار بحرانی را برای همهٔ مقایسه‌های جفتی (بالاترین اختلاف معنی‌دار[3]) مطرح کرد. این آزمون از روش دانکن محافظه‌کارتر است.

برآورد پارامترها

در حالت کلی، تمرکز اصلی مباحثهٔ بین یک مشتری و مشاور آماری بر تفسیر درست جدول ANOVAی دوطرفه و نتیجه‌گیری مبتنی بر تحلیل پساآزمایشی خواهد بود. افزون بر آن، از اشارهٔ

1. least significant difference

۲. این ارجاع اشاره به یادداشت‌های منتشرنشده‌ای با عنوان «مسئلهٔ مقایسه‌های چندگانه» دارد. با این حال، آزمون دامنهٔ استیودنتیده به‌تفصیل در شفه (۱۹۵۹) تشریح می‌شود.

3. highest significant difference

مستقیم به‌صورت بازپارامتریدهٔ مدل ANOVAی (۱.۳) می‌توان با استفاده از شبه‌کد

$$Y = \mu + A + B + AB + \text{خطا}$$

اجتناب کرد. با این حال، آموزنده است که مسئلهٔ برآوردکردن پارامترهای یک مدل ANOVAی دوطرفه را در نظر بگیریم زیرا این کار به موضوع «تابع‌های برآوردپذیر» و انواع مختلف مجموع‌های توان‌های دومی که می‌توان از یک تحلیل ANOVA تولید کرد منجر می‌شود.

همان گونه که در بالا خاطرنشان کردیم، مدل ANOVAی دوطرفهٔ (۱.۳) را می‌توان به‌صورت

$$y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + \epsilon_{ijk}, \quad \epsilon_{ijk} \sim N(0, \sigma^2) \qquad (۲.۳)$$

بازپارامتریده کرد که در آن، مانند قبل، μ میانگین پاسخ کل است، α_i اثر سطح iام عامل A است، β_j اثر سطح jام عامل B است، γ_{ij} اثر متقابل بین α_i و β_j است، و ϵ_{ijk} مؤلفهٔ خطای تصادفی است. با در دست‌داشتن داده‌های کافی (یعنی، $n = IJK > 1 + I + J + IJ$)، نتیجه می‌شود که ما «باید» قادر باشیم که پارامترهای مدل راکه در (۲.۳) ظاهر می‌شوند برآورد کنیم. علاوه بر ساختن آزمون‌های فرض که در جدول ANOVA ظاهر می‌شوند، برآوردهای پارامتری باید به ما این امکان را بدهند که آزمون‌های فرض پس‌آزمایشی مورد نظر را نیز بسازیم. به عنوان مثال، α_1 =کنترل در برابر همهٔ سطوح دیگر تیمار $(i > 1, \alpha_i)$.

مسئله آن است که این مدل ذاتاً بیش‌پارامتریده است و صرف‌نظر از اینکه چه مقدار داده در دسترس باشد، پارامترهای مطرح در (۲.۳) به‌طور خطی وابسته‌اند؛ به این معنی که معادلات معمولی (نرمال) وابسته به برآوردکردن این پارامترها زیر بار جوابی یکتا نمی‌روند مگر اینکه برخی قیود شناسایی‌پذیری را بر آن‌ها اعمال کنیم. رویکرد استاندارد، اعمال‌کردن قید صفربودن مجموع است که مستلزم آن است که برآوردهای پارامترها در

$$\sum_i \hat{\alpha}_i = \sum_j \hat{\beta}_j = 0, \quad \text{و} \quad \sum_i \hat{\gamma}_{ij} = \sum_j \hat{\gamma}_{ij} = 0 \qquad (۳.۳)$$

صدق کنند. تحت قیود مذکور در (۳.۳)، برآوردهای پارامتری حاصل که از معادلات نرمال به‌دست آمده‌اند عبارت‌اند از

$$\hat{\mu} = \overline{y}_{\cdot}, \quad \hat{\alpha}_i = \overline{y}_{i\cdot} - \overline{y}_{\cdot}, \quad hat\beta_j = \overline{y}_{\cdot j} - \overline{y}_{\cdot}, \quad \text{و} \quad \hat{\gamma}_{ij} = \overline{y}_{ij} - \overline{y}_{i\cdot} - \overline{y}_{\cdot j} - \overline{y}$$

که در آن $\overline{y}_{\cdot}$ نشان‌دهندهٔ میانگین y_{ijk} روی اندیس(های) ناموجود است.

در حالی که این جواب برای معادلات نرمال به‌طور شهودی پذیرفتنی است، اغلب بسته‌های نرم‌افزاری آماری برآوردهای پارامتری را به این روال ارائه نمی‌کنند. به جای آن، تبدیل‌هایی از پارامترهای اصلی به‌کار گرفته می‌شوند و تفاوت‌هایی در روش تبدیل مورد استفاده در هر بستهٔ نرم‌افزاری وجود دارد. برای ساده‌کردن بحث، وضعیتی را در نظر می‌گیریم که در آن عامل A سه سطح دارد. جدول زیر تبدیل‌های مختلف به‌کار رفته به‌وسیلهٔ JMP، S-PLUS، و SAS را نشان می‌دهد.

		SAS			JMP		S-PLUS	
پارامتر کدگذاری‌شده		A_1	A_1	A_2	A_1	A_2	A_1	A_2
پارامتر اصلی	a_1	۱	۰	۰	۱	۰	−۱	−۱
	a_2	۱	۱	۰	−۰	۱	۱	−۱
	a_3	۰	۰	۱	−۱	−۱	۰	۲

صفر بودن مجموع JMP پارامتر سطح آخر را برابر با مجموع دو سطح دیگر با علامت منفی قرار می‌دهد. این کار قید صفرشدن مجموع را برای اثر A برآورده می‌کند. با این حال، برآوردها (و آزمون‌های) پارامتری مندرج در «گزارش پارامتری JMP» $\hat{a}_1$ و $\hat{a}_2$ نیستند. در واقع، تفسیر درست پارامتر برچسب‌گذاری‌شده با "A[1]"[1] این است که $a_1 - (a_1 + a_2 + a_3)/3$.

مقایسه‌های مقید هلمرت تبدیل‌های پیش‌فرض به‌کاررفته به‌وسیلهٔ S-PLUS، مقایسه‌های مقید هلمرت هستند. این مقایسه‌های مقید معادل با مقایسهٔ سطح تام A با «متوسط» $(i-1)$ سطح قبلی است. S-PLUS تبدیل صفر بودن مجموع را به عنوان گزینه‌ای در تابع تحلیل واریانس خود "aov()" در اختیار می‌گذارد و نیز اجازهٔ به‌کارگیری مقایسه‌های مقید تعریف‌شده به‌وسیلهٔ کاربر در این تابع را می‌دهد.

تابع‌های برآوردپذیر SAS پارامترها را مستقیماً تبدیل نمی‌کند. به جای آن، SAS با استفاده از معکوس‌های تعمیم‌یافته[2] به جستجوی تابع‌های برآوردپذیر می‌پردازد. در مثال بالا، جواب انتخاب‌شده چنین از کار درمی‌آید که A_3 برابر صفر گذاشته می‌شود. این روش از لحاظ محاسباتی پرهزینه‌تر است اما به SAS اجازه می‌دهد که چهار نوع تابع برآوردپذیر تولید کند که می‌توان از آن‌ها برای ساختن چهار نوع مجموع توان‌های دوم برای آزمودن فرض ANOVA استفاده کرد.

۱. ویرایش ۴. در JMP 3.x، این پارامتر حتی به‌صورت گمراه‌کننده‌تر به‌صورت "A[1-3]" برچسب‌گذاری‌شده که این ظاهر را دارد که که آزمون متناظر، آزمون a_1 در مقابل a_3 بوده است.

۲. خواننده را برای جزئیات خاص به SAS (۱۹۹۰؛ فصل ۴) ارجاع می‌دهیم.

انواع SSها

مجموع‌های توان‌های دوم اثرها در ستون SS جدول ANOVAی بالا را می‌توان به راه‌های مختلف محاسبه کرد. در SAS، چهار نوع مجموع توان‌های دوم در دسترس‌اند که به‌طور فرضی SSهای نوع I، نوع II، نوع III، و نوع IV نامیده می‌شوند. در S-PLUS (ویرایش ۶ برای UNIX)، تنها نوع I و III را در دست داریم. اختلاف زمینه‌ای بین این مجموع توان‌های دوم در ساختن «تابع‌های برآوردپذیر» برای آزمون‌کردن فرض صفر «H_0: هیچ اثری وجود ندارد» است. به‌طور مختصر اختلاف‌ها را برای دو مدل ANOVA شرح می‌دهیم. برای جزئیات بیشتر خواننده را به فصل ۹ کتاب راهنمای SAS/STAT (SAS 1990) ارجاع می‌دهیم.

نوع I SS برای هر اثر به‌طور دنباله‌ای محاسبه می‌شود؛ به این معنی که «ترتیب»ی که اثرها طبق آن‌ها به مدل (۲.۳) افزوده می‌شوند اهمیت دارد. از آنجا که عامل A در ابتدا افزوده شده است، SSA نمایش‌دهندهٔ تغییر در مجموع توان‌های دوم ناشی از اثر A به‌تنهایی است: مدل صفر $y_{ijk} = \mu + \epsilon_{ijk}$ در برابر $y_{ijk} = \mu + \alpha_i + \epsilon_{ijk}$. با این حال، SSB نمایش‌دهندهٔ تغییر پس از احتساب قبلی اثر عامل A است؛ یعنی $y_{ijk} = \mu + \alpha_i + \epsilon_{ijk}$ در برابر $y_{ijk} = \mu + \alpha_i + \beta_j + \epsilon_{ijk}$. SSAB آخرین تغییر ناشی از اثر AB است که مدل قبلی را با (۲.۳) مقایسه می‌کند.
بنابراین، فرض‌های نوع I اثر «نموّی» هر جملهٔ جدید را در مدل، مشروط بر بودن همهٔ جمله‌های قبلی در مدل، آزمون می‌کنند. این مطلب، برای مدل‌های به‌طور خالص آشیانی و مدل‌های رگرسیونی چندجمله‌ای مناسب خواهد بود.

نوع II برای طرح‌های متعادل، مدل‌های اثرهای اصلی، و مدل‌های رگرسیونی خالص، فرض‌های نوع II مناسب‌اند. SSهای نوع II، III، و IV مجموع توان‌های دوم جزئی نامیده می‌شوند زیرا اندازه‌ای از سهم هر اثر، مثلاً عامل A، را پس از به‌حساب‌آوردن همهٔ اثرهای دیگری که شامل A نیستند در اختیار می‌گذارند. بنابراین، SS نوع II برای عامل A مدل‌های $y_{ijk} = \mu + \beta_j + \epsilon_{ijk}$ را در برابر $y_{ijk} = \mu + \alpha_i + \beta_j + \epsilon_{ijk}$ مقایسه می‌کند (جملهٔ اثر متقابل AB منظور نشده است زیرا این مدل، اثر A را دربردارد). به‌استثنای مدل‌هایی که در بالا فهرست شده‌اند، SSهای نوع III بر SSهای نوع II ترجیح داده می‌شوند.

نوع III، IV SSهای نوع III به سبکی مشابه با SSهای نوع II ساخته می‌شوند، بجز اینکه فرض‌ها بر تابع‌های برآوردپذیری استوارند که متعامدند. این بدان معنی است که فرض‌های نوع III برای A بر تابع‌های برآوردپذیری مبتنی هستند که بر تابع‌های AB عمودند؛ یعنی

اینکه فرض‌های نوع III زمانی که طرح نامتعادل یا شامل خانه‌های گمشده است (هیچ مشاهده‌ای برای ترکیب‌های خاصی از سطوح عامل وجود ندارد)، همچنان مناسب‌اند. فرض‌های نوع IV با فرض‌های نوع II متناظرند اما نوع IV را می‌توان وقتی به‌کار برد که خانه‌های گمشده‌ای وجود داشته باشند. تفاوت بین SS نوع III و IV در نحوهٔ انتخاب تابع‌های برآوردپذیر است. نوع IV سعی می‌کند که این تابع‌ها را «متعادل» کند در حالی که نوع III آن‌ها را متعامد می‌کند.

اثرهای تصادفی

آزمون‌های F نشان داده‌شده در جدول ANOVAی دوطرفهٔ بالا بر این فرض مبتنی هستند که هر دو عامل تیماری ثابت‌اند. این بدان معنی است که استنباط‌های منتج از این تحلیل، مختص به سطوح خاصی از A و B خواهند بود که پژوهشگر انتخاب کرده است. با این حال، وقتی که سطوح یک عامل به‌تصادف از دامنه‌ای از مقادیر ممکن انتخاب شود، استنباط‌های منتج دربارهٔ اثر آن عامل در مورد همهٔ سطوح قابل اعمال خواهد بود. از آنجا که اثر این عامل بستگی به این خواهد داشت که کدام سطح‌ها انتخاب شوند، به آن اثر **تصادفی** اطلاق می‌شود.

گرچه ساختمان اساسی جدول ANOVA بدون تغییر باقی می‌ماند، لازم است که آزمون‌های F را برای به‌حساب‌آوردن تغییرپذیری بیشترِ مطرح‌شده توسط یک اثر تصادفی اصلاح کنیم. میانگین توان‌های دوم مورد انتظار برای ساختن آماره‌های آزمون مناسب در مسائل تحلیل واریانس به‌کار گرفته می‌شوند. برای مدل ANOVAی دوطرفه، آزمون‌های F مناسب عبارت‌اند از:

هر دو تصادفی به جای MSE از MSAB در مخرج هر دو کسر F_A و F_B استفاده می‌شود.

مدل‌های آمیخته اگر A ثابت و B تصادفی باشد، آنگاه اثر متقابل AB نیز تصادفی است. با این حال، قیود متفاوتی را می‌توان بر γ_{ij} در (۲.۳) اعمال کرد که بر نحوهٔ اصلاح F_A و F_B تأثیر می‌گذارد. این مطلب با جزئیات بیشتر در مطالعهٔ موردی ۱.۸ مورد بحث قرار گرفته است که در آن آزمون‌های F مناسب ارائه می‌شوند.

تبصره‌ها

۱. پَه!... آیا واقعاً انتظار داریم که مشتری تمام جزئیاتی را که دربارهٔ ANOVA مطرح کردیم بفهمد؟

نه، البته نه همهٔ جزئیات را، اما مهم است قادر به توضیح این مطلب باشیم که چرا از فرض‌های نوع III، یک مدل اثر تصادفی، یا آزمون دامنهٔ چندگانهٔ دانکن استفاده می‌کنیم.

بدیهی است که مشتری لازم نیست جزئیات فنی این مباحث تشخیصی را بداند، اما باید بداند که چرا از آن‌ها استفاده می‌شود.

نکتهٔ اصلی که در اینجا در صدد طرح آن بوده‌ایم این است که ANOVAی دوطرفه یک روش تحلیل «غیر بدیهی» است، هم از لحاظ آماری و هم مفهومی. با این حال، به‌طور گسترده در حوزه‌های مطالعاتی بسیار متفاوتی به‌کار گرفته می‌شود (حتی اگر مشتری از این امر بی‌خبر باشد)، و نقش مشاور آن است که رهنمون مشتری در طراحی و تحلیل باشد. همان‌گونه که قبلاً بیان کردیم، این کار ممکن است فرایند ساده‌ای نباشد. صبور باشید. مشتریان ما ناتوان درگذر به ANOVAی دوطرفه نیستند؛ تنها نیاز به کمی راهنمایی دارند.

۲. توصیف ما از شیوهٔ ANOVA آشکارا وابسته به طرح آزمایشی و عامل‌های مطرح در بررسی است. به‌عنوان مثال، یکی از عامل‌ها در ANOVAی دوطرفه ممکن است یک عامل «بلوک‌بندی» باشد. در بخش ۷.۳، برخی طرح‌های بلوکی خاص را به‌تفصیل مورد بررسی قرار می‌دهیم. همان‌گونه که قبلاً ذکر کرده‌ایم (بخش ۱.۳)، یکی از موضوعات اصلی هنگام پرداختن به مطالعهٔ پساآزمایشی مشتری، درک دقیق این مطلب است که داده‌ها «چگونه» گردآوری شده‌اند.

۳. در عمل، آزمایش‌ها اغلب به داده‌های «نامتعادل» می‌انجامند: یعنی، تعداد مشاهدات در هر خانه (سطوح A در ANOVAی یک‌طرفه، یا AB در ANOVAی دوطرفه). بجز وقتی که طرح بسیار نامتعادل باشد، این امر مشکلی از لحاظ تحلیل ایجاد نمی‌کند.

۴. روش ANOVA را می‌توان به‌طور بدیهی برای امتحان‌کردن اثر بیش از دو عامل گسترش داد. با این حال، تعداد ترکیب‌های سطوح عامل‌ها ممکن است به‌سرعت رشد یابند و باید طرح‌های آزمایشی خاص‌تری را به جای آن‌ها به‌کار برد. در برخی آزمایش‌ها، سطوح یک عامل ممکن است درون هر سطح عامل دیگر آشیانی باشد. دوباره، طرح‌های آزمایشی خاص‌تر را در بخش ۷.۳ مورد بررسی قرار می‌دهیم.

۵. در توصیف بالا از تعریف‌های صریح جمله‌های مجموع‌های توان‌های دوم اجتناب کردیم. در حالی که مطمئناً توضیح‌دادن این فرمول‌ها به یک مشتری چندان دشوار نیست، پرسش بدیهی مربوط و نامربوط‌بودن پیش می‌آید: نتیجهٔ نمایش داده‌شده در جدول ANOVA یک عدد خواهد بود، نه یک فرمول! تمرینی پر حاصل‌تر توضیح «مفهوم» شیوهٔ ANOVA مثلاً برحسب نمودارهای جعبه‌ای موازی خواهد بود. این کار مشتری را به سمت مفهوم مهم وارسی کفایت مدل پیش از پذیرش پیامدهای مبتنی بر نتایج جدول ANOVA رهنمون خواهد شد.

۶. پارامتر درجه‌های آزادی (DF) پیش‌تر در توصیف‌های ما از جدول‌های پیشایندی و آزمون‌های t مطرح شده‌اند. جدا از آزمون t ای دونمونه‌ای (تقریبی) که در آن‌ها واریانس‌ها برابر فرض نمی‌شوند، مشتری‌هایی که نیاز به کمک در ارتباط با این روش‌ها دارند، نسبت به موضوع DF بی‌تفاوت‌اند. با این حال، در مورد ANOVA همواره چنین نیست وگاهی پارامتر DF مشتری‌ها را «دستپاچه می‌کند». در این وضعیت، معرفی مفهوم مشاهدات به عنوان اجزای «اطلاعات» ممکن است سودمند باشد؛ به این معنی که لازم است توضیح دهیم که یک اندازهٔ نمونهٔ «کارا» در رابطه با مدلبندی داده‌ها چه معنایی دارد.

۷. جنبه‌های بسیار دیگری از ANOVAی دوطرفه هم هستند که ما به آن‌ها نپرداخته‌ایم (از قبیل وضعیتی که در آن یک عامل، سطوح ترتیبی دارد: از چندجمله‌ای‌های متعامد استفاده کنید). با این حال، حدی بر میزان اطلاعاتی که باید ارائه کنیم وجود دارد و تأکید ما در این فصل بر ارائهٔ اطلاعات دربارهٔ روش‌شناسی آماری از دیدگاه یک مشاور آماری است که سعی دارد مثلاً ANOVAی دوطرفه را به یک مشتری توضیح دهد. بنابراین، توصیف خود از شیوهٔ ANOVA را با مطلب زیر به پایان می‌بریم:

> به مطالعهٔ مشتریان احترام بگذارید. مسئله‌های ما تابعی از تلاش‌های پژوهشی آن‌ها است.

رگرسیون

تحلیل رگرسیونی دربارهٔ مسئلهٔ برازش‌دادن یک مدل خطی برای توصیف رابطهٔ بین یک متغیر پاسخ کمّی Y و یک یا چند متغیر تبیینی[1] است. راهنمای ما از وجود چنین رابطه‌ای ممکن است دانش زمینه‌ای دربارهٔ متغیرهای مطرح یا انگیزهٔ ما برای آن، شواهد تجربی حاصل از مطالعه‌های قبلی باشد. در عمل، تحلیل رگرسیونی اغلب برای داده‌های مشاهده‌ای به‌کار می‌رود که در آن چنین دانشی ممکن است در بهترین حالت سطحی باشد و احتمال ندارد که رابطهٔ بین متغیرهای پاسخ و تبیینی (در صورت وجود) دقیقاً خطی باشد. همچنین این تمایل در مشتریان وجود دارد که بخواهند تا حد ممکن متغیرهای تبیینی بسیاری را در مدل رگرسیونی بگنجانند ــ«صرفاً به این دلیل که ممکن است مهم باشند.» علی‌رغم امکان بالقوهٔ سوءاستفاده از آن، تحلیل رگرسیونی ابزاری نیرومند است و می‌تواند بینشی سودمند دربارهٔ فرایند و متغیرهای تحت مطالعه در اختیار گذارد.

کتاب‌های درسی بسیاری دربارهٔ رگرسیون به ANOVA نیز می‌پردازند. سیبر (۱۹۷۷) نظریهٔ استاندارد رگرسیون را دربردارد وکتاب درسی دریپر و اسمیت (۱۹۸۱) تقریباً مترادف با اصطلاح

۱. اصطلاحات مترادف عبارت‌اند از: رگرسور، پیشگو، ورودی، و متغیرهای مستقل.

رگرسیون است. به مباحث تشخیصی رگرسیونی به‌تفصیل در بسلی و همکاران (۱۹۸۰) پرداخته می‌شود.

رگرسیون خطی چندگانه

مدل رگرسیون خطی چندگانه به شکل

$$Y = \beta_0 + \beta_1 X_1 + \cdots + \beta_k X_k + \epsilon$$

است که در آن ϵ مؤلفهٔ خطای تصادفی است. برای مقاصد استنباطی، ϵ را معمولاً مستقل و با توزیع نرمال با میانگین صفر و واریانس ثابت σ^2 در نظر می‌گیرند. یک «مشاهده»ی تک، مرکب از k مقدار تبیینی همراه با مقدار پاسخ مشاهده‌شدهٔ

$$(x_{1i}, x_{2i}, \ldots, x_{ki}, y_i)$$

است. خطی بودنی که بر اثر این مدل تحمیل می‌شود، کمتر از آنچه در بدو امر به نظر می‌رسد، مقیدکننده است. تبدیل‌های سادهٔ متغیر پاسخ را می‌توان به‌کار برد و متغیرهای تبیینی را می‌توان به‌صورت تابع‌هایی از دیگر متغیرها تعریف کرد. این مطلب، به‌عنوان مثال، دربردارندهٔ متغیرهای نشانگر یا ظاهری برای پیشگوهای رسته‌ای، جمله‌های چندجمله‌ای ($X_j = X^j$) در مدل‌های خمیدهٔ خطی، و جمله‌های اثر متقابل (مثلاً $X_3 = X_1 * X_2$) در مدل‌های رویهٔ پاسخ است. در ساده‌ترین حالت وقتی که تنها یک متغیر تبیینی مطرح باشد ($k = 1$)، به این مدل، رگرسیون خطی ساده[1] (SLR) اطلاق می‌شود.

تحلیل مقدماتی

پیش از فرایند برازش واقعی، چندین موضوع مهم وجود دارند که لازم است در ابتدا برای اطمینان از به‌کارگیری مناسب روش‌شناسی رگرسیونی مورد توجه قرار گیرند. این امر مستلزم تعامل نزدیک با مشتری طی یک جلسهٔ مشاوره به هنگامی است که مباحث کلی و هدف‌های پروژه مورد بحث قرار می‌گیرند. پرسش‌های ویژهٔ رگرسیونی زیر، توصیف‌کنندهٔ نوع مهارت‌های ارتباطی زبانی است که به‌طور خلاصه در بخش ۱.۲ آورده شده‌اند. با این حال، پیش از در نظرگرفتن هر یک از این پرسش‌ها یک قاعدهٔ بنیادی در تحلیل رگرسیونی وجود دارد:

نمودار داده‌ها را رسم کنید.

1. simple linear regression

- هدف از مدل رگرسیونی چیست؟
- آیا واقعاً همهٔ متغیرهای تبیینی ضروری‌اند؟
- آیا مشکلی با مقادیر گمشده داریم؟
- چگونه باید با داده‌های دورافتاده رفتار کنیم؟
- آیا هیچ‌یک از متغیرها ترتیبی یا کیفی هستند؟
- آیا در این فرایند قیدهای خاصی موجودند؟
- آیا یک اثر سری زمانی بالقوه وجود دارد؟

نمودارها همیشه، همیشه به نمودارهای داده‌ها نگاه کنید. نمودارهای پراکنش Y در مقابل هر یک از X_jها ممکن است اطلاعات فوق‌العاده باارزشی دربارهٔ قوت هر یک از متغیرهای تبیینی به‌عنوان پیشگو در اختیار گذارند و جنبه‌های مهمی نظیر داده‌های دورافتاده، غیر خطی‌بودن، و خوشه‌بندی را آشکار سازند. نمودارهای جفتی (X_j در مقابل X_k) نیز ممکن است برای آشکارسازی زائدبودن‌های بالقوه، به‌صورت همبستگی شدید بین دو متغیر تبیینی، سودمند باشد. البته، فروریختن داده‌های چندمتغیره به دو بعد، تصویر «کامل»ی از نمونهٔ رگرسیونی فراهم نخواهد کرد و گاهی ممکن است گمراه‌کننده باشد. بنابراین، هدف این نمودارهای پراکنش، اکتشافی است ساز آن‌ها نمی‌توان به‌عنوان وسیله‌ای برای تصمیم‌گیری دربارهٔ اینکه آیا باید متغیرهای تبیینی را حذف کرد یا خیر استفاده کرد.

هدف اگر قرار است که مدل را برای مقاصد پیشگویی به‌کار بریم، در این صورت لازم است که کیفیت خوبی داشته باشد تا پیشگویی‌های قابل اعتمادی در اختیار گذارد. به‌طور بدیهی، یک پایگاه داده‌ای به‌قدر کافی بزرگ و دانش زمینه‌ای دربارهٔ فرایند در این حالت مطلوب خواهد بود.

تعداد پیشگوها بیش‌برازش[1] یک مدل رگرسیونی، اثر نامطلوب تورم واریانس وابسته به برآوردهای پارامتر را دارد. مسائل چندخطی‌بودن نیز ممکن است بر اثر وابستگی‌های خطی بین متغیرهای تبیینی مطرح شوند. توجیه‌کردن نیاز به هر یک از متغیرهای تبیینی در بدو امر می‌تواند تمرینی ارزشمند باشد.

۱. برای بعضی مشتریان، این مطلب ممکن است خلاف (متناقض با) شهود به نظر برسد: آیا R^2 افزایش نیافت؟ توضیح‌دادن مسئله با اضافه‌کردن برداری تصادفی یا استفاده از یک مدل اشباع‌شده ($p = n$، که در آن p تعداد متغیرهای تصادفی تبیینی است) در مثالی ساده ممکن است سودمند باشد.

مقادیر گمشده شیوه‌های نرم‌افزاری آماری به‌طور عادی همهٔ مشاهدات شامل مقادیر گمشده را پیش از شروع محاسبه‌های خواسته‌شده حذف می‌کنند. بنابراین، لازم است که متغیرهای کمکی مشتمل بر بخش بزرگی از مقادیر گمشده، از مدل رگرسیونی حذف شوند (یا مجدداً کدبندی شوند، مشروط بر اینکه این کار معنی داشته باشد).

داده‌های دورافتاده نقاط تأثیرگذار مانند نقاط دورافتاده و نقاط با ضریب نفوذ بالا ممکن است اثری قابل ملاحظه بر مدل برازش‌یافته داشته باشند وگاهی نتایجی به‌دست می‌دهند که در متن فرایند مورد عمل، فاقد معنی هستند. آشکارسازی اولیهٔ نقاط دورافتادهٔ بالقوه از روی آماره‌های خلاصهٔ تکی و نمودارهای پراکنش جفتی ممکن است مقدور باشد، اما ممکن است نقاط تأثیرگذار دیگری نیز ضمن تحلیل آشکار شوند. هر جا که مناسب باشد، باید دربارهٔ این نقاط، مورد به مورد، با مشتری تصمیماتی اتخاذ شود. این بدان معنی است که باید مشتری را از این امر آگاه کرد که ممکن است، پیش از آنکه تحلیل نهایی کامل شود، تعامل بیشتری لازم باشد.

متغیرهای رسته‌ای گاهی از درنظرگرفتن داده‌های کیفی (مثلاً جنسیت) در مدل رگرسیونی صرف‌نظر می‌شود، یا به‌طور نامناسب آن‌ها را به مقیاسی ترتیبی تبدیل می‌کنند. به همین نحو، یک متغیر آشکارا کمّی نظیر سن ممکن است در واقع نمایش‌دهندهٔ یک متغیر گروه‌بندی باشد زیرا تنها برخی از مقادیر ثبت شده‌اند. این وضعیت بر چگونگی پیش‌بردن تحلیل رگرسیونی اثر می‌گذارد و ضروری است توجه خاصی به متغیرهای‌تبیینی‌دربرگیرندهٔ انواع‌داده‌های ترتیبی یا کیفی‌معطوف شود.
البته، در صورتی که متغیر پاسخ، رسته‌ای باشد، لازم است که یک مدل خطی تعمیم‌یافته به‌کارگرفته شود. این مدل به‌اختصار در بخش **روش‌های خاص** زیر مورد بحث قرارگرفته است.

قیدها اثر رگرسیونی وابسته به یک متغیر تبیینی ممکن است روی برخی دامنه‌ها متفاوت باشد. به‌عنوان مثال، دما روی مصرف برق خانگی، در دامنه‌های دمای پایین، همبستگی منفی، اما روی دامنه‌های دمای بالا، همبستگی مثبت دارد. لازم است که نقاط تغییر، آستانه‌ها، و دیگر قیدهای فیزیکی در تحلیل رگرسیونی منظور شوند.

همبستگی پیاپی مشاهده‌هایی که در طول زمان گردآوری می‌شوند، ممکن است دستخوش همبستگی پیاپی باشند. این امر ممکن است به‌طرز نامناسبی واریانس خطا را متورم کند و به برآوردهای رگرسیونی ناکارایی منجر شود. استفسار از مشتری دربارهٔ فرایند گردآوری داده‌ها

به‌طوری که یک اثر سری زمانی بالقوه به‌طور غیر قابل انتظاری پیش نیاید ارزشمند است. در این موقعیت می‌توان از صورت‌های بهبودیافتهٔ شیوهٔ رگرسیونی مانند برآورد دو مرحله‌ای بر مبنای مدل‌های اتورگرسیونی استفاده کرد. این مطلب در سیبر و ویلد (۱۹۸۹) بحث شده است. در مقابل، ممکن است تحلیل سری زمانی با استفاده از پاسخ با تأخیر و متغیرهای تبیینی مورد نیاز باشد. به این مطلب، مدلبندی تابع انتقال اطلاق می‌شود که در براکول و دیویس (۱۹۹۱) بحث می‌شود.

تحلیل رگرسیونی

برازاندن یک مدل رگرسیونی نوعاً فرایندی تکراری است که بین ارزیابی کیفیت مدل برازش‌یافته و بهبودسازی‌های مدل در رفت و برگشت است. با این حال، تأکید بر مطلب زیر ارزشمند است:

در شیوهٔ برازش افراط نکنید.

به تعدادی نامحدود، مدل رگرسیونی وجود دارد که می‌توان آن‌ها را امتحان کرد و تقریباً به همان تعداد امتحان تشخیصی را می‌توان انجام داد. در عمل، بهترین مدل اغلب آن است که ساده‌ترین تعبیر را دارد.

رایج‌ترین روش برای برازش‌دادن مدل‌های رگرسیونی، روش کمترین توان‌های دوم است که می‌توان آن را به‌صورت بسیار کارا محاسبه کرد. از روی برآوردهای کمترین توان‌های دوم پارامترهای $\hat{\beta}_j$، $j = 0, \ldots, k$، مقدارهای برازش‌داده‌شدهٔ $\hat{y}_i = \sum_{j=0}^{k} \hat{\beta}_j x_{ij}$ (که در آن $x_{0i} \equiv 1$) و مانده‌های $e_i = y_i - \hat{y}_i$ را به‌ازای $i = 1, \ldots, n$، به‌دست می‌آوریم. این کمیت‌ها ما را قادر می‌سازند که جدول ANOVAی رایج و برآوردهای پارامترها را، به‌صورتی که در زیر نشان داده شده است بسازیم.

تحلیل واریانس

منبع	DF	SS	MS=SS/DF	مقدار F
رگرسیون	k	SSR	MSR	$F_0 = \frac{MSR}{MSE}$
خطا	$n-k-1$	SSE	MSE	
کل	$n-1$	SST		

برآورد پارامتر			
متغیر	برآورد	انحراف استاندارد	مقدار T
عرض از مبدأ	$\hat{\beta}_0$	$\text{se}(\hat{\beta}_0)$	$t_0 = \hat{\beta}_0/\text{se}(\hat{\beta}_0)$
X_1	$\hat{\beta}_1$	$\text{se}(\hat{\beta}_1)$	$t_0 = \hat{\beta}_1/\text{se}(\hat{\beta}_1)$
⋮	⋮	⋮	⋮
X_k	$\hat{\beta}_k$	$\text{se}(\hat{\beta}_k)$	$t_0 = \hat{\beta}_k/\text{se}(\hat{\beta}_k)$

در اینجا، مقدار F، آمارهٔ آزمونی برای ارزیابی اینکه آیا یک «اثر» کلی رگرسیونی وجود دارد یا خیر، فراهم می‌کند. اگر بتوان فرض نرمال بودن و ثابت‌بودن واریانس را در نظر گرفت، آنگاه تحت فرض صفر H_0 : هیچ اثر رگرسیونی از هیچ‌یک از متغیرهای تبیینی وجود ندارد، خواهیم داشت $F_0 \sim \mathcal{F}_{k,n-k-1}$. مقدار T، آمارهٔ[1] آزمون t برای فرض صفر $H_0 : \beta_j = 0$ و به این معنی است که هیچ اثر رگرسیونی ناشی از متغیر تبیینی X_j وجود ندارد.

مباحث تشخیصی مانده‌ای

مرحلهٔ بعد، انجام‌دادن وارسی‌های تشخیصی روی مدل رگرسیونی برازش یافته است. اگر مدل رگرسیونی و فرض‌ها برای فرایند تحت بررسی مناسب باشند، آنگاه مانده‌های $e_i = y_i - \hat{y}_i$ باید ویژگی‌های اعمال‌شده بر ϵ_i را منعکس کنند. بنابراین، تحلیل مانده‌ای نقش مهمی در مباحث تشخیصی رگرسیونی دارد. زمینه‌های اصلی مورد نظر عبارت‌اند از:

- کفایت مدل: نمودار مانده‌ها در مقابل مقدارهای برازش‌یافته.
- ارزیابی نرمال‌بودن: نمودار Q-Q، آزمون شاپیرو-ویلک.
- آشکارسازی نقاط تأثیرگذار: روش‌های تشخیصی حذفی.

نمودار مانده‌ها در مقابل برازش‌یافته این نمودار نباید هیچ الگوی غیر تصادفی بدیهی را نشان دهد. اغلب e_iها را به‌صورت $r_i = e_i/s_e$ استاندارد می‌کنند که در آن $s_e^2 = MSE$. مقادیر بزرگ r_i را می‌توان با ± 2 مقایسه کرد.

۱. درایه‌های Std Error متناظر با ریشه‌های دوم عناصر قطری $\text{MSE} * (X'X)^{-1}$ هستند که در آن X ماتریس رگرسور $n \times (k+1)$ است.

نمودار Q-Q برای اندازه‌های نمونهٔ متوسط (مثلاً، $n > ۳۰$) یک نمودار Q-Qی مانده‌های e_i را می‌توان برای ارزیابی فرض نرمال‌بودن خطاها به‌کار برد: $\varepsilon \sim N(۰, \sigma^۲)$. مانده‌های مرتب‌شده باید یک الگوی خط مستقیم بر یک نمودار Q-Q را به نمایش بگذارند. همان‌طور که در وارسی‌های تشخیصی مانده‌ها برای یک مدل ANOVA بیان کردیم، آمارهٔ شاپیرو-ویلک یک آزمون رسمی آماری برای نرمال‌بودن را در اختیار می‌گذارد.

روش‌های تشخیصی حذفی روش‌های تشخیصی به‌اصطلاح «یکی را حذف کن[1]» متعددی نظیر مانده‌های استیودنتیده، Dی کوک[2]، DFFITS، DFBETAS وجود دارند که اطلاعاتی دربارهٔ تأثیر یک مشاهدهٔ خاص در اختیار تحلیلگر می‌گذارند. این کار با حذف iامین مشاهده و برازش مجدد مدل رگرسیونی به $n - ۱$ مشاهدهٔ باقیمانده انجام می‌شود. اگر آن نقطهٔ خاص «تأثیرگذار» باشد، آنگاه رگرسیون برازش‌داده‌شده به $n - ۱$ مشاهده باید متفاوت از برازش رگرسیونی اولیه باشد. تعریف‌های این ابزارهای تشخیصی را می‌توان در اغلب کتاب‌های درسی رگرسیون یافت. به‌عنوان مثال، نگاه کنید به بلی و همکاران (۱۹۸۰). در عمل، D (فاصله)ی کوک شاخص سودمندی از مشاهداتی که ممکن است تأثیرگذار باشند، در اختیار می‌گذارد، زیرا نشان‌دهندهٔ مصالحه‌ای بین مانده‌های بزرگ و دادهٔ پرنفوذ است. بنابراین، اگر پاسخ، دادهٔ دورافتادهٔ آشکاری نسبت به مقادیر Y باشد، اما نزدیک «مرکز» دامنهٔ X رخ دهد، این مشاهده ممکن است کمترین «تأثیر» را داشته باشد. به همین نحو، یک مشاهده ممکن است نسبت به دامنهٔ X دورافتاده باشد، اما پاسخ Y به رگرسیون برازش‌یافته (ابرصفحه) نزدیک است. در هر دو حالت، هیچ‌یک از مشاهده‌ها مقدار Dی کوک بزرگی نخواهد داشت که از آن چنین برداشت می‌شود که مشاهده تأثیرگذار نیست. نموداری از Dی کوک در شکل ۷.۳ نشان داده شده که آشکارا اولین و آخرین مشاهده‌ها را بسیار تأثیرگذار شناسایی می‌کند.

انتخاب «بهترین» رگرسیون

متغیرهای تبیینی را گاهی متغیرهای مستقل می‌نامند که گمراه‌کننده است زیرا Xها اغلب از لحاظ اطلاعات پیشگویانه دربارهٔ Y همپوشانی دارند. مهم‌تر اینکه برخی متغیرهای تبیینی ممکن است زمانی که با سایر متغیرهای تبیینی آمیخته شوند سهم چندانی نداشته باشند. از لحاظ آماری، اجتناب از استفاده از اطلاعات زائد و تلاش برای یافتن کوچکترین زیرمجموعه از متغیرهای مستقلی که همگی در پیشگویی متغیر پاسخ سهم داشته باشند قابل توجیه است. به این مطلب، اصل اصاک اطلاق می‌شود. روش‌های تشخیصی زیر برای انتخاب یک مدل رگرسیونی ممسک، سودمند است.

1. delete one 2. Cook's D

Tـمقدارها آماره‌های مقدار T مشخص می‌کنند که آیا X_j سهم قابل توجهی در مدل رگرسیونی دارد یا خیر. مشکل این است که آماره‌های t بر مدل رگرسیونی کامل[1] استوارند؛ بدین معنی که اهمیت «هر» X_j بر مبنای اینکه همه متغیرهای تبیینی دیگر در مدل حضور دارند سنجیده می‌شود. بنابراین، حذف‌کردن همهٔ X_jهای غیر مهم به‌یکباره، ممکن است گمراه‌کننده باشد، زیرا آماره‌های t مبین وابستگی‌های چندگانه بین متغیرهای تبیینی نیستند.

R^2ی تعدیل‌یافته کمیت $R^2 = \text{SSR}/\text{SST}$ اندازه‌ای از کیفیت برازش را برای یک مدل رگرسیونی فراهم می‌کند. با این حال، این اندازه را همواره می‌توان صرفاً با افزودن متغیر تبیینی دیگری افزایش داد. آمارهٔ R^2 «تعدیل یافته»، که به‌صورت $R^2_{\text{adj}} = 1-(n-1)*\text{SSE}/(n-k)*\text{SST}$ تعریف می‌شود، اندازه‌ای از کیفیت برازش نسبت به پیچیدگی مدل رگرسیونی فراهم می‌آورد. بنابراین، R^2_{adj} را می‌توان برای مقایسهٔ مدل‌های برازش‌یافته با زیرمجموعه‌های متفاوتی از متغیرهای تبیینی به‌کار گرفت.

شیوه‌های گام به گام یک شیوهٔ گام به گام، یک روش برازش رگرسیونی تکراری است که در اغلب بسته‌های نرم‌افزاری آماری موجود است. این شیوه در هر تکرار یک متغیر تبیینی را، بسته به اهمیتی که آن متغیر در مدل رگرسیونی جاری دارد، اضافه یا حذف می‌کند. این کار تا آنجا ادامه پیدا می‌کند که هیچ تغییر دیگری رخ ندهد یا همان X_j اضافه و سپس بلافاصله حذف شود.

تبصره‌ها

۱. «و، منظور شما دقیقاً از الگوی غیر تصادفی چیست؟» مشتری می‌پرسد (آیا «الگوی تصادفی»، یک ترکیب تضادآمیز نیست؟).

آنچه می‌خواهیم در واقع بگوییم این است که هیچ رابطهٔ بدیهی‌ای در پراکنش عمودی مانده‌ها، وقتی که آن‌ها را در برابر مقدار برازش‌یافته یا هر متغیر تبیینی رسم می‌کنیم، حضور نداشته باشند. انتقال این مفهوم ساده نیست و طبق تجربیات ما، رسم‌کردن تعدادی الگوی «غیر تصادفی» معمولاً به روشن‌شدن موضوع کمک می‌کند.

۲. شیوهٔ گام به گام صرفاً الگوریتمی است که سطوح معنی‌داری معینی را برای هر ورود و خروج به‌کار می‌برد. این شیوه تضمین نمی‌کند مدل زیرمجموعه‌ای نهایی لزوماً معنی‌دارترین مدل موجود است. دیگر معیارهای اطلاع همچون آمارهٔ C_pی مالوز[2] و AIC را نیز می‌توان در انتخاب زیرمجموعه به‌کار گرفت.

---

1. FULL 2. Mallows

۶.۳ روش‌های عمومی

روش‌هایی که در بخش قبل ارائه شد، ممکن است برای کامل‌کردن تحلیل آماری مورد نیاز در برخی پروژه‌ها، اما مطمئناً نه برای همهٔ آن‌ها، کافی باشند. اغلب، مشاور آماری نیاز به به‌کارگیری فن تخصصی‌تری برای اجرای تحلیلی مناسب دارد و در برخی موارد ممکن است نخستین باری باشد که آن روش به گوش مشتری خورده باشد. در این وضعیت، بار مسئولیت آشکارا بر دوش مشاور منتقل می‌شود تا توصیفی مناسب از این روش‌شناسی را بدون مستغرق‌کردن مشتری در جزئیات فنی ارائه دهد.

از آنجا که هر کس که آموزش آمار دیده است با روش‌هایی که در زیر بحث می‌کنیم آشناست، ارائه‌کرد خود را به‌ناچار مختصر و مفید کرده‌ایم زیرا شرحی عام از این روش‌ها را می‌توان در هر جای دیگری یافت. به‌علاوه، جزئیات مربوط به کاربرد روشی خاص در چارچوب یک مطالعهٔ موردی در قسمت II ارائه می‌شود. ارجاعاتی به دیگر کتاب‌ها و مطالعهٔ موردی مربوط را در اینجا تدارک دیده‌ایم.

گرچه ارائهٔ ما از این روش‌ها مختصر است، هدف ما فراهم‌کردن پایهٔ یک «توصیف مکفی» از روش‌شناسی عمومی است. البته انتظار رسیدن به این هدف کاملاً غیر واقعی است، اما مشتری منتظر توضیح ماست... .

آزمون‌های ناپارامتری

آزمون‌های ناپارامتری، بدیل‌هایی آزادتوزیع در مقابل روش‌های سنتی تحلیل در اختیار می‌گذارند، در حالی که در روش‌هایی از قبیل آزمون t و ANOVAی یک‌طرفه فرض می‌شود که نمونه دارای یک توزیع نرمال زمینه‌ای است، بسیاری از آزمون‌های ناپارامتری مبتنی بر تحلیل رتبه‌های وابسته به داده‌های نمونه‌ای هستند. بنابراین، فرض‌های توزیعی صریح برای نمونه در آزمون‌های ناپارامتری ضرورت ندارد.

شیوه‌های ناپارامتری به‌ویژه برای تحلیل داده‌های ترتیبی نظیر آن‌ها که در مطالعه‌های تحقیق بازار تولید می‌شوند، سودمندند. پاسخ‌های به شکل «هرگز ــ به‌ندرت ــ گاهی ... » به‌طور عددی به‌صورت مثلاً ۱ ــ ۲ ــ ۳ ــ ... کدبندی می‌شوند، اما در موقع به‌کاربردن آزمون t به‌طور صریح فرض می‌شود که فاصله‌های بین این عددها دارای معنی‌اند. این فرض ممکن است واقع‌بینانه نباشد و یک آزمون رتبه‌ای، که تنها از ترتیب ذاتی پاسخ‌ها استفاده می‌کند، مناسب‌تر است. با این حال، در کار با داده‌های متریک محدودیت‌هایی مربوط به شیوه‌های ناپارامتری به دلیل از دست‌رفتن اطلاعات توزیعی موجود است. برخی آزمون‌ها و شیوه‌های ناپارامتری معروف عبارت‌اند از:

همبستگی اسپیرمن اندازه‌ای از پیوند خطی بین رتبه‌های دو متغیر ترتیبی در اختیار می‌گذارد. این همبستگی همان تعبیر و ویژگی‌های ضریب همبستگی را داراست، اما از بامعنی تلقی‌کردن فاصله‌های بین‌نقطه‌ای مقیاس‌های ترتیبی اجتناب می‌کند.

آزمون علامت جایگزینی برای آزمون tی تک‌نمونه‌ای. این آزمون صرفاً تعداد مقدارهای بالای مبانهٔ فرضی را در نظر می‌گیرد و توزیع دوجمله‌ای را برای آزمودن $H_0 : p = ۱/۲$ به‌کار می‌برد. این آزمون بسیار محافظه‌کار است و ممکن است به خطای گردکردن حساس باشد.

آزمون رتبهٔ علامت‌دار ویلکاکسون بدیلی برای آزمون tی تک‌نمونه‌ای. فرض می‌کند که توزیع زمینه‌ای نمونه متقارن است.

آزمون U‌ی من‌ـویتنی بدیلی برای آزمون tی دونمونه‌ای. به آن آزمون مجموع رتبه‌های ویلکاکسون نیز گفته می‌شود. ضروری است که به هم‌رتبگی در رتبه‌های داده‌ها به‌دقت پرداخته شود.

آزمون کروسکال‌ـوالیس بدیلی برای ANOVAی یک‌طرفه.

اطلاعات بیشتر دربارهٔ آزمون‌ها و شیوه‌های ناپارامتری را می‌توان در اسپرنت (۱۹۹۳) وکنوور (۱۹۸۰) یافت.

تبصره‌ها

۱. احتمال دارد آزمون‌های مبتنی بر میانگین‌ها و واریانس‌ها برای یک مشتری، شهودی‌تر از آزمون‌های مبتنی بر رتبه‌ها باشند. تفسیر و مشخص‌کردن فرض ناپارامتری برای مشتری، خود ممکن است مشکل باشد.

۲. از به‌کارگیری آزمون‌های ناپارامتری صرفاً برای اجتناب از یک نتیجهٔ «معنی‌دار» مانند موارد دعاوی حقوقی حذر کنید.

۳. آزمون‌های ناپارامتری ممکن است با نمونه‌های کوچک سودمند نباشند و این دقیقاً زمانی مطرح است که محتملاً فرض نرمال‌بودن در معرض بحث و جدل است. شاید لازم باشد روش‌های بازنمونه‌گیری را در نظر بگیریم.

۴. برآورد چگالی هسته‌ای و روش‌های kـنزدیک‌ترین همسایه را می‌توان در شیوه‌هایی از قبیل تحلیل ممیزی و خوشه‌ای به‌کار گرفت و بدیلی بر فرض نرمال چندمتغیره‌بودن تدارک دید.

مدل‌های خطی کلی

مدل‌های رگرسیون و ANOVA حالت‌هایی خاص از مدل خطی کلی‌اند. مثال‌هایی دیگر از مدل خطی کلی عبارت‌اند از:

مدل‌های رویهٔ پاسخ در آزمایش‌های عاملی، علاقهٔ ما اغلب معطوف به یافتن پاسخ بهینه نسبت به عامل‌های طرح کنی است. وقتی اثر متقابل وجود داشته باشد، مدل‌های رویهٔ پاسخ اثرهای درجهٔ دوم را در رگرسیون وارد می‌کنند.

تحلیل رویهٔ پاسخ در کاربردهایی متنوع از قبیل مهندسی شیمی، کشاورزی، و علم تغذیه مورد استفاده قرار می‌گیرد که در آن‌ها «می‌دانیم» که نتیجهٔ ناشی از یک متغیر پاسخ به برخی متغیرهای ورودی وابسته است. این بدان معنی است که روش‌شناسی رویهٔ پاسخ، یک شیوهٔ دنباله‌ای است که هر آزمایش جدید نتایج حاصل از قبلی‌ها را مورد بهره‌برداری قرار می‌دهد. از آنجا که این نوع رویکرد تکراری در عمل گران است، طرح‌های تخصصی‌تری در آزمایش‌های رویهٔ پاسخ بکار گرفته می‌شوند.

هدف نهایی روش‌های رویهٔ پاسخ، تعیین شرایط عمل بهینهٔ سیستم تحت مطالعه است. تعریف «بهینه» در مطالعهٔ موردی ۲.۷ دنبال شده است. باکس و همکاران (۱۹۷۸) کاربرد روش‌شناسی رویهٔ پاسخ را در یک مثال مربوط به شیمی در نظر می‌گیرند و ارجاعاتی به چندین بررسی دیگر در زمینه‌های متنوعی ارائه می‌کنند. برای ارائهٔکرد جامع‌تری از روش‌شناسی رویهٔ پاسخ نگاه کنید به مه‌یرز و مونتگومری (۱۹۹۵).

تحلیل کوواریانس در برخی وضعیت‌های آزمایشی، ممکن است اطلاعات کنی بیشتری به شکل به‌اصطلاح متغیرهای همراه در دسترس باشد. برای داخل‌کردن این اطلاعات «رگرسیونی» در تحلیل طرح واریانس، می‌توان مدل تحلیل کوواریانس را تحت برخی فرض‌ها به‌کار گرفت. فرض اضافی اصلی آن است که متغیرهای همراه تحت تأثیر عامل‌های «تیمار» در طرح ANOVA نباشند. به‌عنوان مثال، یک آزمایش بالینی بر سن بیمار تأثیر نخواهد داشت، اما تغییرات وزن ممکن است به سهولت رخ دهند. اگر قرار بود که وزن را به‌عنوان یک متغیر کمکی منظور کنیم، لازم می‌بود که به وزن پیش از آزمایش بیمار اشاره کنیم. کتاب‌های زیادی دربارهٔ رگرسیون و طرح، تحلیل کوواریانس را نیز مورد بحث قرار می‌دهند. در اینجا، کوکران (۱۹۵۷) را ذکر می‌کنیم که در ویراست بیومتریکس[1] درآمده و تقریباً به‌طور کامل به تحلیل کوواریانس اختصاص داده شده، است.

MANOVA شیوهٔ ANOVA در قالب فنی عمومی برای افراز تغییرپذیری کلی یک متغیر

1. Biometrics

پاسخ تک به مؤلفه‌هایی مرتبط با اثرهای تیمار شرح داده شد. با این حال، در بسیاری از آزمایش‌ها ممکن است چندین مشخصهٔ پاسخ مورد توجه پژوهشگر باشند و با بکارگیری همان شیوهٔ ANOVA برای هر متغیر پاسخ به‌طور مجزا همبستگی‌های متغیر پاسخ منظور نخواهند شد. برای آزمودن همزمان اثرهای تیمار روی بیش از یک متغیر پاسخ، به تحلیل واریانس چندمتغیره (MANOVA) نیاز داریم. دو حالت خاص مورد توجه عبارت‌اند از:

اندازه‌های مکرر وضعیتی را در نظر بگیرید که در آن دو گروه از مردم ــ یک گروه کنترل و یک گروه تیمارــ در دوره‌ای از زمان بررسی می‌شوند که طی آن مشخصهٔ خاصی (Y) در سه بازهٔ منظم ثبت می‌شود. چون Y برای هر مورد چندین بار اندازه‌گیری می‌شود، این اندازه‌ها همبسته خواهند بود. افزون بر آن، پاسخ موردهای مختلف نیز ممکن است به‌شدت متغیر باشند که این امر، تشخیص اختلاف معنی‌دار بین دو گروه را دشوار می‌کند. برای به‌حساب‌آوردن همبستگی در پاسخ می‌توان یک مدل ANOVAی اندازه‌گیری‌های مکرر چندمتغیره را بکار گرفت. اوریت (۱۹۹۵) مروری عملی بر تحلیل اندازه‌گیری‌های مکرر را ارائه می‌کند. برای ملاحظهٔ کتابی در باب این موضوع نگاه کنید به کرادر و هند (۱۹۹۰).

تحلیل نمایه‌ای آزمایش‌های روان‌شناسی اغلب متضمن اجرای یک سری از آزمون‌ها در مورد گروه‌های خاصی از مردم و ایجاد یک «نمایه»ی پاسخ است که از یک گروه به گروه دیگر قابل مقایسه‌اند. این آزمون‌ها معمولاً برحسب یک مقیاس مشترک استاندارد شده‌اند و نمودار «سری زمانی» حاصل از متصل‌کردن نقاط (j, Y_j)، $j = ۱, ۲, \ldots, k$، که در آن k تعداد مقیاس‌ها (آزمون‌ها) است، نمایهٔ پاسخ نامیده می‌شود. برآوردی از یک نمایهٔ گروهی مبتنی بر نمره‌های متوسط درون گروه است.

اگر تنها دو گروه، بدون عامل‌های اضافی، مطرح باشد، آنگاه آزمون‌های مختلفی نظیر توازی‌گرایی (نمایه‌ها هم‌شکل‌اند) را می‌توان بر آمارهٔ $T^۲$ی هتلینگ استوار کرد. با وجود بیش از دو گروه یا وجود عامل‌های طرح اضافی، تحلیل نمایه‌ای منجر به شیوهٔ MANOVA می‌شود. تحلیل نمایه‌ای در سپر (۱۹۸۴) مورد بحث قرار می‌گیرد.

تبصره‌ها

۱. اینجا باید نکته‌ای را محض احتیاط اضافه کنیم. خروجی حاصل از شیوهٔ MANOVA پیچیده است و، بنابر تجارب ما، مشتریان اغلب در درک نتایج دشواری‌های زیادی دارند. در حالی که ممکن است وسوسه شویم که تقاضای صادقانهٔ مشتری را که «آیا نمی‌توانیم

برای هر پاسخ صرفاً یک ANOVA اجرا کنیم؟» برآورده کنیم، اما این نتایج جایگزین نتایج MANOVA نخواهند شد. همان‌طور که در آغاز این بخش خاطرنشان کردیم، این وظیفهٔ مشاور است که مشتری را در صورت لزوم با شیوه‌های جدید آشنا کند.

۲. در برخی موارد که تنها دو متغیر پاسخ داریم، «ممکن است» اهداف پروژه با آزمودن تفاضل بین پاسخ‌ها در مدل ANOVAی تک‌متغیره تحقق پیدا کنند. دوباره، تحلیل MANOVA نیز باید انجام شود و نتایج حاصل از دو شیوه با هم مقایسه شوند. اگر شانس با ما همراه باشد، شاید لازم باشد که تنها نتایج حاصل از ANOVA را ارائه کنیم، در غیر این صورت ... .

۳. ارائهٔ نتایج حاصل از تحلیل MANOVA به مشتری جرأت و شاید تردستی می‌خواهد. به‌عنوان مثال، اگر بپرسند «اثر پیلایی[1] چیست؟» چگونه باید پاسخ دهیم؟ یک رویکرد این است که صرفاً بگوییم که یک آزمون F از قبیل آن‌هایی است که در جدول‌های ANOVA دیده می‌شوند ــ که البته دقیقاً درست نیست، اما ما به جای اینکه در جزئیات فنی دربرگیرندهٔ نسبت‌های دترمینان‌های ماتریس‌ها گیر بیفتیم، بر تفسیر نتایج تأکید کرده‌ایم. به یاد داشته باشیم که دروغ داریم، دروغ شاخدار داریم، و ... آمار را!

روش‌های چندمتغیره

روش‌های چندمتغیره‌ای که در اینجا ارائه می‌شوند معطوف به نوعی فروکاهی بُعد (متغیرها) یا گروه‌بندی همگن (مشاهدات) یک ماتریس داده‌های $n \times p$، مثلاً X، مرکب از n مشاهده است که در آن p مشخصه (متغیر) برای هر مشاهده اندازه‌گیری می‌شوند. سپر (۱۹۸۴) مرجع عمومی خوبی است که هم مباحث نظری و هم عملی مطرح در تحلیل چندمتغیره را مورد بحث قرار می‌دهد. موئیرهد (۱۹۸۲) صرفاً نظری است، اما مشتمل بر چندین جدول است که مشاور آماری شاید آن‌ها را سودمند ببیند. دیگر متون مربوط عبارت‌اند از: گناندسیکان (۱۹۹۷)، ماردیا و همکاران (۱۹۷۹)، و رنچر (۱۹۹۵). ترکیبی از چندین روش چندمتغیره در مثال داده‌کاوی مطالعهٔ موردی ۸.۴ مورد بررسی قرار می‌گیرد.

تحلیل مؤلفه‌های اصلی مؤلفه‌های اصلی، ویژه‌بردارهای مرتبط با تجزیهٔ متعامد ماتریس واریانس (یا همبستگی) نمونه‌ای X هستند. بنابراین، ویژه‌مقدارهای متناظر، اندازه‌ای از آن بخش واریانس کل (اثر ماتریس واریانس) را به‌دست می‌دهند که توسط هر یک از مؤلفه‌های اصلی توضیح داده می‌شوند. بنابراین، در صورتی که k، $k < p$، ویژه‌مقدار بزرگتر، سهم عمده‌ای از واریانس کل را به خود اختصاص دهند، می‌توان تعداد متغیرهای «مؤثر» در

1. Pillai

ماتریس داده‌های X را کاهش داد. این، هدف تحلیل مؤلفه‌های اصلی[1](PCA) است. توجه کنید که این متغیرهای «مؤثر» آن مؤلفه‌های اصلی[2]اند که ترکیب‌های خطی متغیرهای اولیه‌اند که ضرایب آن‌ها به‌وسیلهٔ k ویژه‌بردار معین می‌شود. استفاده از ماتریس همبستگی مستلزم آن است که هر یک از p متغیر در X اهمیتی برابر با سایرین دارد و بسته به اینکه از کدام ماتریس استفاده شود ممکن است نتایج مختلفی عاید شود. مثال زیر نیز نشان می‌دهد که بزرگترین ویژه‌مقدارها لزوماً «جالب»ترین سیماهای آماری X را در اختیار نمی‌گذارند.

مثال ۶.۳ درآمدهای مردان و زنان برحسب تعداد ماه‌های سابقهٔ کاری

بزرگترین ویژه‌مقدار معمولاً متناظر با تغییرپذیری کلی درآمد [income] خواهد بود. در این حالت، کوچکترین ویژه‌مقدار متناظر با تغییرپذیری وابسته به هر تفاضل بین درآمدهای ناشی از جنسیت[gender] خواهد بود. بنابراین، تصویرکردن داده‌ها بر روی دومین ویژه‌بردار نشان خواهد داد که آیا جنسیت تأثیری بر درآمد دارد یا خیر (که آشکارا مورد توجه بیشتری است).

تحلیل عاملی تحلیل عاملی[3] (FA) اساساً با مشاهدات X این‌گونه رفتار می‌کند که گویی در رگرسیونی به شکل $x = \mu + \Gamma\eta + \epsilon$ صدق می‌کنند، اما در آن همه چیز در سمت راست مجهول است. فرض می‌شود که بردار η مرکب از k، $k < p$، «عامل» (مشترک) با «سربارهایی» است که به‌وسیلهٔ عنصرهای ماتریس Γ ی $p \times k$ داده شده‌اند. با وجود ماهیت نامعین مدل، آماردانان در مورد نتایج حاصل از تحلیل عاملی، به دلیل اینکه هر دوران متعامد η نیز در مدل صدق خواهد کرد، بسیار بااحتیاط عمل می‌کنند. به همین نحو، k را باید تعیین کرد و حتی اگر مدل درست باشد، انتخاب ناصحیح k ممکن است نتایج ضعیفی عاید کند. بنابراین، نتایج حاصل از تحلیل عاملی را باید با احتیاط تفسیر کرد. با این حال، تحلیل عاملی به‌طور کاملاً گسترده در تحقیقات بازار به‌کار می‌رود و ما نتایج یک تحلیل عاملی را در مطالعهٔ موردی ۲.۸ ارائه می‌کنیم.

باید خاطرنشان کنیم که تفاوت بین تحلیل عاملی و تحلیل مؤلفه‌های اصلی اغلب موجب سردرگمی می‌شود زیرا روش مؤلفه‌های اصلی اغلب برای برازش یک مدل تحلیل عاملی به‌کارگرفته می‌شود... گیج شدید؟ سپیر (۱۹۸۴، صفحهٔ ۲۱۵) مقایسهٔ روشنی بین روش‌های تحلیل عاملی و تحلیل مؤلفه‌های اصلی انجام می‌دهد.

1. principal components analysis

۲. این‌ها را متغیرهای پنهان نیز می‌نامند.

3. factor analysis

تحلیل ممیزی تحلیل ممیزی با مسئلهٔ تخصیص یک مشاهده به یک «گروه» خاص (از بین تعدادی ثابت از گروه‌ها) بر مبنای p مشخصهٔ اندازه‌گیری‌شده سروکار دارد. وقتی گروه‌ها همپوشانی دارند، برای اجرای بهینهٔ عمل تخصیص لازم است که تابع ممیزی کمترین نرخ بدرده‌بندی‌کردن را حاصل کند. با این حال، این امر بستگی به این دارد که چه اندازه دربارهٔ توزیع مشاهدات در هر گروه اطلاع داریم (یا می‌توانیم فرض کنیم که داریم) و اینکه هر مشاهده چقدر محتمل است که از گروه خاصی آمده باشد، و هزینه‌های بدرده‌بندی‌کردن چقدر است.

تحلیل خوشه‌ای تحلیل خوشه‌ای سعی بر آن دارد که زمانی که تعداد گروه‌ها نامعلوم است مشاهده‌ها را در گروه‌هایی رده‌بندی کند. دو روش کلی برای تحلیل خوشه‌ای وجود دارد: خوشه‌بندی سلسله‌مراتبی که نیاز به تعریف یک فاصلهٔ بین‌نقطه‌ای و یک فاصلهٔ بین‌خوشه‌ای دارد، و روش‌های گرانیگاهی که در آن k نقطهٔ بذردانه انتخاب می‌شوند و داده‌ها بین k خوشه توزیع می‌شوند. سپس با استفاده از معیارهایی نظیر R^2، k خوشه با امکان ادغام به‌آرامی بهینه‌سازی می‌شوند. تحلیل خوشه‌ای یکی از روش‌های اصلی در داده‌کاوی است و به‌تفصیل در مطالعهٔ موردی ۴.۸ مورد مطالعه قرار می‌گیرد.

کمترین توان‌های دوم جزئی کمترین توان‌های دوم جزئی[1] (PLS) به‌طور گسترده‌ای در صنایع فرایندی برای مرتبط کردن یک ماتریس تبیینی یا متغیرهای فرایند با ماتریس دیگری از متغیرهای خروجی به‌کار می‌رود. به‌عنوان مثال، برای بررسی نرخ‌های جریان دمای رآکتور، می‌توان روش PLS را برای مرتبط‌کردن متغیرهای فرایند، X = زمان‌های تغذیه، آهنگ‌های اغتشاش، و مقدار اندازه‌گیری‌شده، به انواعی از Y =اندازه‌های کیفیت خروجی به‌کار برد.

PLS عموماً در شیمی‌سنجی به‌کار می‌رود که موضوع آن، در اساس، مطالعهٔ کاربردهای آماری در تحلیل داده‌های شیمیایی است. برای مروری بر ابزارهای رگرسیونی شیمی‌سنجی (مشتمل بر PLS و رگرسیون مؤلفه‌های اصلی)، نگاه کنید به مقاله‌ای از فرانک و فریدمن (۱۹۹۳). مثالی از یک کاربرد با استفاده از رگرسیون مؤلفه‌های اصلی به‌وسیلهٔ منچستر و همکاران (۱۹۹۹) ارائه می‌شود.

تحلیل سری زمانی

وقتی داده‌ها با زمان اندیس‌گذاری شده باشند، مشاهده‌های دنباله‌ای در یک سری زمانی محتملاً همبسته خواهند بود. برای درچنگ‌آوردن این سیمای زمینه‌ای، دو رویکرد مورد استفاده قرار می‌گیرند: مدل‌های زمان‌دامنه و تحلیل طیفی. مدلبندی زمان‌دامنه به‌دست باکس و جنکینز (۱۹۷۰) عمومیت

1. partial least squares

پیدا کرد و رایج‌تر است، به‌ویژه وقتی که پیش‌بینی مورد توجه باشد. تحلیل طیفی زمانی مناسب است که دوره‌ای‌بودن مورد توجه باشد. هر دو روش را می‌توان برای فرایندهای برداری به‌کار برد. برخی مراجع مفید دربارهٔ تحلیل سری زمانی عبارت‌اند از براکول و دیویس (۱۹۹۱)، باکس و همکاران (۱۹۹۴)، و پریستلی (۱۹۸۱). کاربردی از تحلیل سری زمانی در مطالعهٔ موردی ۴.۷ ارائه می‌شود.

مدل‌های ARIMA در رویکرد باکس ـ جنکینز، برآوردهای همبستگی‌نگار تابع خودهمبستگی[1] (ACF) و تابع خودهمبستگی جزئی[2] (PACF) رسم می‌شوند و برای انتخاب کاندیداهای مناسب از ردهٔ مدل‌های میانگین متحرک جمع‌بستهٔ اتورگرسیو[3] (ARIMA) مورد استفاده قرار می‌گیرند. مدل‌های ARIMAی ضربی را می‌توان برای مدلبندی سری‌های زمانی فصلی به‌کار برد.

تحلیل طیفی روش‌های تحلیل طیفی مبتنی بر تبدیل فوریهٔ گسسته[4] (DFT)اند که می‌توان آن‌ها را برای نمایش هر سری برحسب سینوسوارها به‌کار برد. در عمل، لازم است که این سری‌ها را روندزدایی کرد زیرا در DFT فرض می‌شود که یک سری متناهی خود را تکرار می‌کند. دوره‌نگار را می‌توان از DFT محاسبه کرد و برای آشکارسازی فراوانی‌های مهم در سری، از روش‌های تشخیصی عمده محسوب می‌شود. هموارسازی دوره‌نگار، برآوردی از طیف را در اختیار می‌گذارد که به‌طور یکتا یک فرایند مانای پیوسته را تعریف می‌کند.

سری‌های زمانی رسته‌ای مثالی از یک سری زمانی رسته‌ای، یک دنبالهٔ DNA است که می‌توان آن را برحسب جفت‌های پایه‌ای (A,T,C,G) که مارپیچ دوگانه را می‌سازند اندیس‌گذاری کرد. یک رویکرد جالب به تحلیل سری‌های زمانی رسته‌ای، روش‌شناسی پوشش طیفی ارائه‌شده به‌وسیلهٔ استوفر و همکاران (۱۹۹۳) است. رویکردی جایگزین، که کاربردهایی در تشخیص گفتار دارد، استفاده از زنجیرهای مارکوف پنهان است.

روش‌های تخصصی

مشاور آماری اغلب با پروژه‌های متعددی سروکار پیدا خواهد کرد که در آن‌ها به چیزی چندان فراتر از یک تحلیل استاندارد نیاز نیست. بنابراین، نسبتاً آسان است که به جای همراه‌شدن با پژوهش‌های روز و ابداع روش‌شناسی‌های جدید تا حدی تن‌آسایی پیشه کند. در این دام نیفتید. از مشاور

1. autocorrelation function 2. partial autocorrelation function 3. autoregressive integrated moving average 4. discrete Fourier transform

آماری انتظار می‌رود که خبره باشد و این جایگاه والا را نمی‌توان حفظ کرد مگر اینکه مشاور تلاش اکیدی در چاپ مقاله، شرکت در کنفرانس‌ها، و شاید مهم‌تر از همه، به‌روزماندن در فنون آماری با خواندن مقاله‌های مرتبط در مجلات علمی داشته باشد. پیوست الف فهرستی از برخی مجله‌های پژوهشی را در بر دارد که برای مشاور آماری جالب توجه خواهد بود. با فرض اینکه خواننده به قدر کافی متنبه شده باشد، به‌اختصار برخی روش‌شناسی‌های آماری را که ممکن است در بیشتر پروژه‌های پژوهش‌محور مورد نیاز باشد مورد بحث قرار می‌دهیم.

مدل‌های خطی تعمیم‌یافته در مدل خطی عمومی $y_i = \mu_i + \epsilon_i$، فرض می‌شود که μ_i تابعی خطی از متغیرهای تبیینی است. به عبارتی، $\mu_i = x_i'\beta$ که در آن x_i بردار وابسته به تمامین مشاهدهٔ k مشخصه است. توزیع خطای ϵ_i معمولاً $N(0\sigma^2)$ فرض می‌شود.
در یک مدل خطی تعمیم‌یافته، مجازیم که توزیع خطا را عام‌تر (اما درون ردهٔ توزیع‌های نمایی) در نظر بگیریم و یک تابع ربط وجود دارد به‌طوری که $g(\mu_i) = x_i'\beta$. دو مثال از مدل‌های خطی تعمیم‌یافته که در عمل عموماً با آن‌ها مواجه می‌شویم در زیر ذکر می‌شوند. یک مرجع کلاسیک دربارهٔ این موضوع مکولا و نلدر (۱۹۸۹) است. برای بحثی ساده‌تر نگاه کنید به دابسون (۱۹۹۰). اگرستی (۱۹۹۰) بحثی کامل از مدل‌های لگ‌خطی ارائه می‌کند. کولت (۱۹۹۱) و مورگان (۱۹۹۲) کتاب‌هایی هستند که به تحلیل داده‌های پاسخ دوتایی و چندایی می‌پردازند.

رگرسیون لوژستیک وقتی که متغیر پاسخ Y رسته‌ای است، بسیاری از مدل‌های پیشگویی، مسئله را به مدلی رگرسیونی برای پیشگویی $\pi_j = P[Y_i = j]$ مبدل می‌کنند که در آن زها، $j = 1, 2, \ldots, k$، سطح‌ها یا وضعیت‌های ممکن Y را نمایش می‌دهند. در رگرسیون لوژستیک فرض می‌شود که احتمال‌های Y را می‌توان به‌وسیلهٔ توزیع لوژستیک $f(w = x'\beta) = e^w/(1 + e^w)$ مدلبندی کرد.

مدل‌های لگ‌خطی وقتی که متغیر پاسخ، داده‌های شمارشی است، در این صورت توزیع پواسون در ترکیب با تبدیل لگاریتمی به مطالعهٔ مدل‌های لگ‌خطی برای جدول‌های پیشایندی منجر می‌شود. برای مشتریانی که چیزی بیش از یک آزمون استقلال ساده از ما می‌خواهند، مدلبندی لگ‌خطی مشابه با ANOVA برای جدول‌های پیشایندی است.

مدل‌های غیر خطی در برخی پدیده‌ها مانند رشد زیستی و واکنش‌های شیمیایی، دانش پیشین از یک رابطهٔ غیر خطی ممکن است به‌خوبی برقرار باشد. در این وضعیت‌ها، می‌توان

استفاده از یک مدل غیر خطی مناسب را در نظر گرفت. برای بحثی مفصل دربارهٔ نظریه و کاربردهای روش‌های رگرسیون غیر خطی، نگاه کنید به سبر و وایلد (۱۹۸۹).

روش‌های نوین رگرسیونی یک مدل جمعی تعمیم‌یافته به شکل $Y=\beta_0+\sum_{j=1}^{k} f_j(X_j)+\epsilon$ است که در آن $f_j(X_j)$ ها نامعلوم‌اند، اما فرض می‌شود تابع‌هایی هموار از متغیرهای تبیینی X_j باشند. این مطلب، مفهوم رگرسیون خطی را گسترش می‌دهد که به ابداع چندین روش نوین رگرسیونی منجر شده است: اسپلاین‌های هموارسازی، رگرسیون جستجوی تصویر، ACE، $AVAS$، و الگوریتم‌های $MARS$، و شبکه‌های عصبی. مطالعهٔ موردی ۳.۷ کاربردی از اسپلاین‌های هموارسازی را نشان می‌دهد و شبکه‌های عصبی در مطالعهٔ موردی ۲.۸ به‌کار رفته‌اند. برای جزئیات بیشتر (و مراجع مربوط) دربارهٔ همهٔ روش‌هایی که در بالا ذکر شده‌اند خواننده را به ونابلز و ریپلی (۱۹۹۴، فصل ۱۰) ارجاع می‌دهیم. جنبه‌های آماری برازش‌دادن مدل‌های جمعی تعمیم‌یافته در چمبرز و هستی (۱۹۹۳) مورد توجه قرار گرفته‌اند. مدل‌های درخت‌مبنا نیز در آخرین مرجع مینتی برکاربرایمن و همکاران (۱۹۸۴) مورد بحث قرار گرفته‌اند.

روش‌های استوار روش‌های استوار این امکان را به شیوه‌های آماری می‌دهند که در حضور داده‌های آلوده از قبیل داده‌های دورافتاده، اثرهای گردکردن، یا دوری از فرض‌های توزیعی قابل استفاده باشند. به این دلیل، مثلاً به‌کارگرفتن معادل استوار رگرسیون کمترین توان‌های دوم برای وارسی اینکه آیا نتایج سازگارند یا خیر، ارزشمند است. در غیر این صورت، لازم است که علت اختلاف را بررسی کنیم.

S-PLUS تابع‌های متنوعی مرتبط با روش‌های استوار و مقاوم در اختیار می‌گذارد. یک روش «مقاوم» ما را در برابر خطاهای فاحش (مثلاً میانه در مقابل میانگین نمونه‌ای) حفظ می‌کند. زمانی که داده‌ها آلوده نیستند یک روش «استوار» کارایی معقولی را القا می‌کند (مثلاً میانگین پیراسته). سری کتاب‌هایی که به‌وسیلهٔ هوگلین و همکاران (۱۹۸۳، ۱۹۸۵، ۱۹۹۱) نوشته شده، اطلاعاتی دربارهٔ حوزهٔ وسیعی از روش‌های مقاوم و استوار در اختیار می‌گذارد. روسی‌یو و لیروی (۱۹۸۷) یک متن عملی دربارهٔ روش‌های رگرسیون استوار است.

باید توجه کنیم که روش‌های استوار به‌آسانی به وضعیت‌های چندمتغیره گسترش نمی‌یابند (کدام «بردار» بزرگتر است؟)، و برخی فنون استوار چندمتغیره (مثلاً S-برآوردگرها) ممکن است از لحاظ محاسبه وقت‌گیر باشند. تابع `cov.mve()` در S-PLUS برآوردی استوار از ماتریس کوواریانس بر مبنای هذلولیوار با کمترین حجم در اختیار می‌گذارد.

فنون بازنمونه‌گیری ابزار اصلی زمینه‌ای بسیاری از شیوه‌های استنباط آماری، قضیهٔ حدی

مرکزی است. اما وقتی که اندازهٔ نمونه اجازهٔ برقراری نظریهٔ نرمال بودن مجانبی را نمی‌دهد، باید چه کنیم؟ می‌توان روش‌های ناپارامتری را به‌کار برد، اما برای نمونه‌های کوچک، این روش‌ها توان کمی دارند و نتایجی عاید می‌کنند که بیش از اندازه محافظه‌کارند. در این وضعیت، ممکن است به‌کاربردن روش‌های بازنمونه‌گیری از قبیل جک‌نایف یا شیوهٔ بوت‌استرپ سودمند باشد. فنون بازنمونه‌گیری پژوهشگر را قادر می‌سازد که خطای استاندارد، بازه‌های اطمینان، و توزیع‌ها را برای هر آماره برآورد کند. برای توصیف مفصل‌تر بوت‌استرپ‌کردن، نگاه کنید به اِفرون و تیبشیرانی (۱۹۹۳) و شاتو و تو (۱۹۹۵).

مدل‌های معادلهٔ ساختاری تحلیل مسیر با مسئلهٔ تخصیص اثرهای علّی سروکار دارد که به پدیدهٔ مورد نظر پژوهشگر منجر می‌شود. این روش‌شناسی اغلب در حوزهٔ روانسنجی به‌کار گرفته می‌شود، اما برای تعیین رابطهٔ بین متغیرها (که در بسیاری حالت‌ها آشکارا ذهنی است) به محقق تکیه دارد. فولر (۱۹۸۷) نظریهٔ آماری مرتبط با مدل‌های خطای اندازه‌گیری را مورد بحث قرار می‌دهد. مراجع اضافی را می‌توان در بخش PROC CALIS کتاب راهنمای SAS/STAT (SAS 1990) یافت.

۷.۳ طرح آزمایش‌ها

اصول طرح آزمایش‌های آماری در چارچوب روش‌های گردآوری داده‌ها (بخش ۱.۳) معرفی شد. در اینجا به‌اختصار طرح‌های خاصی را مورد بحث قرار می‌دهیم که مشاور آماری احتمالاً با آن‌ها مواجه خواهد شد. فهرست طرح‌های ما مطمئناً کامل نیست و مسئلهٔ داده‌های نامتعادل ممکن است در عمل پیش آید. ما به این موضوع در پایان بخش بازمی‌گردیم.

هدف طرح‌های آزمایشی آن است که تحلیل کارایی از واریانس را، زمانی که منابع تغییرپذیری معلومی را می‌توان به‌عنوان دلیل ذکر کرد، در اختیار گذارند. این طرح‌ها بر اندازه‌های نمونهٔ معقولی نیز متکی هستند. دو کتاب کلاسیک که ارزش ذکر دارند عبارت‌اند از کوکران و کاکس (۱۹۵۷) و کاکس (۱۹۵۸). مونتگومری (۱۹۹۷) کتاب استانداردی است که پوشش کلی مناسبی از طرح‌های آزمایشی فراهم می‌کند. باکس و همکاران (۱۹۷۸) به کاربرد عملی طرح‌های آزمایشی تأکید دارند.

طرح‌های بلوکی تصادفی

دسته‌ها، مردم، و زمان منابع رایج تغییرپذیری هستند که می‌توان از طریق طرح‌های بلوکی آن‌ها را کنترل کرد. این طرح‌ها امکان مقایسهٔ بهتری از تفاوت‌های تیماری را با حذف تغییرپذیری بین «بلوک‌ها» به وجود می‌آورند. اگر تغییرپذیری ناشی از بلوک‌ها، آن‌گونه که در آزمایش‌های کامل

تصادفی معمول است، کنترل نشوند، تغییرپذیری در خطای آزمایشی کل جذب خواهد شد که ممکن است بر تفاوت‌های تیماری معنی‌دار پرده بیفکند.

طرح‌های بلوکی کامل تصادفی این طرح، مفهوم آزمون t ای جفتی را تعمیم می‌دهد و همهٔ سطوح یک تیمار را به هر بلوک، که به‌عنوان منبع بالقوهٔ تغییرپذیری در بررسی شناخته می‌شود، تخصیص می‌دهد. در هر سطح تیمار در هر بلوک، یک مشاهده وجود دارد و ترتیبی که تیمارها انجام می‌شوند، درون هر بلوک تصادفی است.

مربع‌های لاتین اگر دو عامل بلوک‌بندی تعداد سطوح یکسانی به‌عنوان عامل تیمار، مثلاً I، داشته باشند، آنگاه می‌توان یک طرح مربع لاتین را بکار گرفت که نیاز به I^2 مشاهده دارد. نام این طرح به استفاده از حروف لاتین A، B، ... برای نشان‌دادن سطح‌های عامل تیمار اشاره دارد. نقشهٔ این طرح را می‌توان به‌صورت جدول مربع شکلی نمایش داد که سطرها و ستون‌ها سطوح دو عامل بلوک‌بندی را نشان می‌دهند. برای اینکه عامل تیمار را از دو منبع فرعی تغییرپذیری تفکیک کنیم، طرح باید متعامد باشد. این کار با اجبار به اینکه هر حرف لاتین (سطح تیمار) در هر سطر و ستون جدول مربع شکل تنها یک‌بار ظاهر شود، عملی می‌شود.

عامل سطر	عامل ستون C_1	C_2	C_3	C_4
R_1	A_α	B_β	C_γ	D_δ
R_2	B_δ	A_γ	D_β	C_α
R_3	C_β	A_α	A_δ	B_γ
R_4	D_γ	C_δ	B_α	A_β

مثال دو مربع لاتین برهم‌نهادهٔ 4×4 ارائه‌شده در بالا نشان می‌دهد که این امکان وجود دارد که عامل سومی را (با سطوحی که با حروف یونانی نشان داده می‌شود) وارد طرح کرد. این را طرح یونانی ـ لاتین می‌نامند که می‌توان از آن برای تحلیل اثر دو عامل تیمار، یا برای کنترل‌کردن یک عامل سوم فرعی تغییرپذیری استفاده کرد. باز هم به یک طرح متعامد نیاز داریم تا اثر هر عامل را تفکیک کنیم و این کار با برهم‌نهی دو مربع لاتین (متعامد)، به‌طوری که هر حرف یونانی دقیقاً یک‌بار با هر حرف لاتین ظاهر شود، عملی می‌شود مربع‌های یونانی‌ـلاتین به‌ازای همهٔ Iها، $I \geq 3$، موجودند بجز $I = 6$.

بلوک ناقص طرح RCB نیاز به این دارد که همهٔ سطوح عامل تیمار را به هر بلوک تخصیص دهیم. در برخی وضعیت‌ها، ممکن است این کار مقدور یا حتی مطلوب نباشد. مثال کلاسیک یک طرح بلوکی ناقص آزمایش موسوم به چشیدن شربت است: چندین نوع شربت قرار است به‌وسیلهٔ گروهی از داوران چشیده شوند. چون از توانایی تشخیص بین طعم‌ها پس از امتحان هر شربت جدید کاسته می‌شود، استفاده از یک طرح RCB نامطلوب است. به جای آن می‌توان یک طرح بلوکی ناقص را به‌کار گرفت که در آن تنها برخی سطوح تیمارها (شربتها) به هر بلوک (داور) تخصیص داده می‌شود. طرح متعادل است هرگاه هر دو سطح تیمار همراه با هم به تعداد دفعات مساوی در هر بلوک ظاهر شوند.

طرح کرت‌های خردشده طرح‌های RCB و مربع لاتین فنون کارایی در پرداختن به محدودیت‌های تصادفی‌سازی نسبت به یک عامل تیمار است. این طرح‌ها را می‌توان برای تحلیل بیش از یک عامل تیمار نیز به‌کار برد (نگاه کنید به مثال طرح مربع یونانی‌ـلاتین بالا) مشروط بر اینکه همهٔ ترکیب‌های سطوح تیمار در هر بلوک حاضر باشند. با این حال، در اینجا فرض می‌شود که ترتیب انجام‌دادن آزمایش در درون هر بلوک را می‌توان کاملاً تصادفی کرد. به‌عنوان مثال، اگر A دو سطح ($I = ۲$) و B سه سطح ($J = ۳$) داشته باشد، آنگاه مختار باشیم که تاسی را پرتاب کنیم تا ببینیم که کدام ترکیب خاص سطوح را اول اجرا کنیم.

در برخی آزمایش‌های چندعاملی، ممکن است از لحاظ اقتصادی موجه یا عملی نباشد که داده‌ها را با استفاده از تصادفی‌سازی کامل درون بلوک‌ها گردآوری کنیم. با استفاده از مثال بالا، فرض کنید که A نشان‌دهندهٔ دو دستور پخت کیک باشد که در سه ارتفاع مختلف (B)، در چهار اجاق (بلوک) پخته خواهد شد. برای تعیین اینکه آیا کیک خوب پخته شده است یا خیر، وزن کیک اندازه‌گیری خواهد شد. اگر هر اجاق برای سه کیک جا داشته باشد، عملی‌تر است که هر بار از یک دستور پخت استفاده کنیم و به قدر کافی مخلوط کیک درست کنیم که سه کیک بپزیم. در تصادفی‌کردن کامل، مجاز نیستیم که نحوهٔ پخت و خوردن کیک خود را دیکته کنیم!

طرح‌های کرت‌های خردشده و کرت‌های خردِ خردشده، تعمیم‌هایی از طرح بلوکی تصادفی‌اند که می‌توان آن‌ها را در آزمایش‌های چندعاملی، که در آن نیاز به تصادفی‌سازی مقید است، به خدمت گرفت. این اصطلاح‌ها از کاربرد آن‌ها در آزمایش‌های میدانی کشاورزی عاریت گرفته شده‌اند که در آن‌ها یک «کرت» نشان‌دهندهٔ یک قطعه زمین واقعی بود که به کرت‌های کوچکتری «خرد» می‌شد (در یک طرح «خردِ خرد»، کرت‌های کوچکتر به زیرزیرکرت‌ها خرد می‌شوند). چون تصادفی‌سازی تنها در درون هر سطح تیمار اصلی (A در مثال کیک) رخ می‌دهد، گوییم

اثرهای اصلی در یک طرح کرت‌های خردشده با بلوک‌ها اختلاط دارند؛ به این معنا که اثر روی پاسخ ناشی از هر عامل کنترل‌نشدهٔ دیگر را که با سطوح تیمار اصلی تغییر می‌کند نمی‌توان از اثر تیمار اصلی تشخیص داد. چون تیمار زیرکرت با بلوک‌ها اختلاط ندارد، عامل تیماری که بیشتر از همه در مطالعه مورد توجه است، معمولاً به زیرکرت‌ها تخصیص داده می‌شود.

تبصره‌ها

۱. اگر تفاوت‌های بلوکی مسئول بخش بزرگی از تغییرپذیری باشند که نمی‌توان آن‌ها را به حساب (سطوح) تیمار گذاشت، آنگاه طرح بلوکی تصادفی کامل در مقایسه با ANOVAی یک‌طرفه کارایی نسبی بالایی خواهد داشت. کارایی نسبی، عمدتاً اندازه‌ای از تعداد تکرارهای مورد نیاز در یک ANOVAی یک‌طرفه برای حصول حساسیتی به اندازهٔ طرح RCB است. اگر کارایی نسبی مثلاً ۳ باشد، یک ANOVAی یک‌طرفه را باید با مشاهداتی ۳ برابر اجرا کرد تا همان مقدار حساسیت آزمایش طرح RCB حاصل شود.

۲. اغلب نرم‌افزارهای آماری جدول ANOVAی دوطرفهٔ معمولی، دربرگیرندهٔ آزمون F، $F_B = \text{MSB/MSE}$، را مرتبط با اثر بلوکی در اختیار می‌گذارند. این آزمون را نباید معتبر دانست زیرا تصادفی‌سازی سطوح تیمار به داخل بلوک‌ها محدود شده است.

۳. جمعی‌بودن اثرهای تیمار و بلوکی را باید با بازبینی نمودار مانده در برابر برازش‌یافته ، یا با ساختن یک نمودار اثرهای متقابل امتحان کرد (نگاه کنید به ANOVAی دوطرفه). در برخی حالت‌ها، ناجمعی‌بودن ممکن است بر اثر دورافتاده‌ها، یا یک بستگی ضربی زمینه‌ای باشد. یک تبدیل توانی شاید برای بازیابی جمعی‌بودن، در صورتی که بستگی ضربی باشد، سودمند باشد.

۴. یکی از معایب طرح‌های بلوک‌بندی آن است که از درجه‌های آزادی خطا کاسته می‌شود. به‌عنوان مثال، در یک طرح مربع لاتین با $I = 4$ سطح تیمار، تنها ۶ درجه‌آزادی برای مؤلفهٔ خطا وجود دارد. این بدان معناست که برآورد تغییرپذیری وابسته به خطای آزمایشی (یعنی σ^2ی توزیع خطای مفروض $N(0, \sigma^2)$) به‌طور مؤثری مبتنی بر نمونه‌ای به اندازهٔ ۶ است که نسبتاً کوچک است. برای افزایش دقت برآورد σ^2، می‌توان تکرارهای کامل طرح را به‌کار گرفت[1]. با این حال، همان‌گونه که در تذکر اول ما مطرح شد، عموماً منظورکردن یک اثر بلوک‌بندی بالقوه توصیه می‌شود.

۱. حالتی خاص از طرح مربع لاتین ۲ × ۲ی تکراری، طرح متقاطع نامیده می‌شود. برای جزئیات بیشتر نگاه کنید به مونتگومری (۱۹۹۷).

۵. چندین طرح بلوک‌بندی دیگر نیز وجود دارد که ما آن‌ها را ذکر نکرده‌ایم، از قبیل مربع‌های یودن[1]، طرح‌های بلوکی ناقص جزئاً متعادل، و طرح‌های شبکه‌ای. هدف این طرح‌ها فراهم‌کردن جایگزین‌های «خوب»ی برای پژوهشگر است هنگامی که برخی قیدهای عملی مانع از به‌کارگیری یک طرح متعادل‌اند.

آزمایش‌های عاملی

در یک آزمایش عاملی، هر تکرار همهٔ ترکیب‌های سطوح عامل‌های تیمار را دربردارد. بنابراین، گویند عامل‌ها در یک طرح عاملی متقاطع هستند و دستکم دو تکرار برای در نظرگرفتن همهٔ اثرهای متقابل ممکن لازم است. در ساده‌ترین حالت که در آن تنها دو عامل طرح دخیل‌اند، تحلیل از شیوهٔ ANOVAی دوطرفه تبعیت می‌کند که قبلاً مورد بحث قرار گرفت. با این حال، وقتی چندین عامل دخیل‌اند، تعداد اثرهای متقابل به میزان چشمگیری افزایش می‌یابد ($۲^k - k - ۱$ برای k عامل) و اجرای یک طرح عاملی کامل با سطوح چندگانه به‌ازای هر تیمار اغلب غیر عملی است. با وجود این، در مرحلهٔ آغازین پروژه، ممکن است چندین عامل مورد نظر پژوهشگر باشند و به‌ویژه چند حالت خاص طرح عاملی کلی سودمند باشند.

۲^k عاملی طرح عاملی ۲^k کمترین تعداد ترکیب‌های تیمار را برای تحلیل k عامل در اختیار می‌گذارد. هر عامل تنها دو سطح دارد که به آن‌ها «زیاد» و «کم» اطلاق می‌شود، و می‌توان آن‌ها را با حضورداشتن یا نداشتن یک تیمار، یا دو چینش نمایندهٔ یک عامل کمّی نظیر دما، یا صرفاً دو حالت مختلف کیفی (مثلاً دو اپراتور) مشخص کرد. چون تنها دو سطح برای هر عامل در دست است، لازم است فرض کنیم که پاسخ، روی دامنهٔ سطوح انتخاب‌شده، خطی است. مقادیر کرانگین را نباید برای سطوح یک عامل به‌کار برد (نگاه کنید به مطالعهٔ موردی ۲.۷). توجه کنید که هیچ برآوردی از خطا از تنها یک تکرار طرح ۲^k به‌دست نخواهد آمد مگر اینکه اثرهای متقابل مرتبه‌های بالاتر را با هم ادغام کنیم تا برآوردی از تغییرپذیری «خطا» بسازیم.

اختلاط در طرح‌های ۲^k یک تکرار کامل طرح عاملی ۲^k همهٔ اثرهای متقابل مرتبه‌های بالاتر را شامل خواهد شد، اما ممکن است اجرای به‌یکبارهٔ آزمایش عملی نباشد. یک شیوهٔ رایج اجرای آزمایش ۲^k در بلوک‌های ۲^p است به‌طوری که هر بلوک شامل ۲^{k-p} اجرا باشد. این بدان معناست که برخی اثرها با بلوک‌ها اختلاط پیدا خواهند کرد و بنابراین، انتخاب اینکه کدام p اثر مستقل باید اختلاط داده شوند مهم است. توجه کنید که کل $۲^p - ۱$ اثر با بلوک‌ها اختلاط خواهند یافت که این‌ها اثرهای (تعمیم‌یافته) وابسته به p اثر انتخاب‌شدهٔ اولیه‌اند. در حالت کلی، اثرهای

1. Youden

منقابل مرتبه‌های بالاتر را برای اختلاط با بلوک‌ها انتخاب می‌کنیم چون محقق توجه کمی به این‌ها دارد (یک اثر متقابل چهار عاملی را چگونه تعبیر می‌کنیم؟). با این حال، لازم است دقت کنیم تا تضمین شود که اثرهای مهم (یعنی اثرهای اصلی و اثرهای متقابل دوعاملی) در اثرهای متقابل اضافی که با بلوک‌ها اختلاط خواهند یافت منظور نمی‌شوند. مونتگومری (۱۹۹۷) جدولی از طرح‌های بلوک‌بندی پیشنهادی و اثرهای انتخاب‌شده برای تولید بلوک‌ها فراهم کرده است.

عاملی کسری برای آزمایش‌های غربالگری، طرح عاملی کسری 2^k به‌طور گسترده‌ای به‌منظور شناسایی k عامل با بالاترین اثرها بکار می‌رود. این طرح‌ها تنها 2^{k-p} اجرای یک طرح 2^kی کامل را شامل می‌شوند که منتج به هم‌اثرسازی اثرهای تیمار با یکدیگر می‌شوند. این بدان معناست که برآورد یک اثر مورد نظر در واقع نمایش‌دهندهٔ ترکیب اثرها و اثر(های) هم‌اثرشدهٔ آن است. اصطلاح وضوح[1] برای توصیف نوع هم‌اثرسازی حاضر در یک طرح 2^{k-p} بکار می‌رود:

وضوح III اثرهای اصلی با اثرهای متقابل دوعاملی هم‌اثر می‌شوند.

وضوح IV اثرهای متقابل دوعاملی با یکدیگر هم‌اثر می‌شوند.

وضوح V اثرهای متقابل دوعاملی با اثرهای متقابل سه‌عاملی هم‌اثر می‌شوند.

3^k عاملی طرح عاملی 3^k تعمیم طرح 2^k است و این امکان را می‌دهد که k تیمار را مطالعه کنیم که در آن در هر عامل سه سطح مورد نظر است. مفاهیم اختلاط و تکرارهای عاملی را نیز می‌توان به طرح 3^k تعمیم داد.

طرح‌های مرکب پس از آنکه طرح‌های غربالگری اجرا و نتایج ارزیابی شدند، می‌توان از طرح‌های مرکب برای بررسی دقیق‌تر عامل‌های مهم استفاده کرد. این طرح‌ها امکان برآوردکردن تغییرپذیری را با افزودن نقاط تکرار در برخی موقعیت‌ها در فضای طرح فراهم می‌آورند. مطالعهٔ موردی ۲.۷ کاربردی از طرح مرکب مرکزی را در اختبار می‌گذارد.

طرح‌های آشیانی

در یک طرح آشیانی دو مرحله‌ای یا سلسله‌مراتبی، سطوح تیمار B در امتداد سطوح تیمار A مشابه‌اند اما یکسان نیستند. به‌عنوان مثال، دو تأمین‌کنندهٔ مختلف (عامل A) هر یک سه بسته (عامل B) مواد خام به‌تصادف انتخاب‌شده را که از نظر خلوص آزمون می‌شوند در اختیار می‌گذارند. در این حالت، بسته‌ها مختص یک تأمین‌کنندهٔ معین‌اند و سطوح B درون هر سطح A «آشیانی»اند. وضعیت در شکل ۸.۳ به تصویر کشیده شده است که در آن نماد پریم بر این تأکید دارد که هیچ

1. resolution



شکل ۸.۳ طرح‌بندی یک طرح آشیانی دومرحله‌ای.

ارتباطی بین بسته‌های تأمین‌کنندگان مختلف وجود ندارد ــ $B1$ و $B1'$ بستهٔ واحدی نیستند.

اثر آشیانی «B درون A» به‌صورت $B(A)$ نشان داده می‌شود که عامل والد درون پرانتزها ظاهر می‌شود.[۱] چون لزوماً هر سطح B با هر سطح A ظاهر نمی‌شود، اثر متقابلی بین این عامل‌ها وجود ندارد و تجزیهٔ مجموع توان‌های دوم به شکل SST=SSA+SSB(A)+SSE است. توسیع به طرح‌های آشیانی چندمرحله‌ای سرراست است و این امکان نیز وجود دارد که هم عامل‌های آشیانی و هم متقاطع را در یک آزمایش داشته باشیم.

تبصره‌ها

۱. در برخی حالت‌ها ممکن است، بر مبنای توصیف اولیهٔ مشتری از پروژه، نسبت به اینکه آیا عاملی متقاطع یا آشیانی است مطمئن نباشیم. باید از این مطلب اطمینان حاصل کرد زیرا تعبیر یک «اثر متقابل» معنی‌دار، در صورتی که عامل آشیانی را به غلط چنین تحلیل کنیم که گویی متقاطع است، ممکن است گمراه‌کننده باشد. مانند حالت طرح کرت خردشده، عامل‌های آشیانی را اغلب می‌توان با بحث و فحص با مشتری دربارهٔ روش دقیق گردآوری داده‌های به‌کار رفته در آزمایش تعیین کرد.

۲. طرح‌های آشیانی را می‌توان برای تحلیل مؤلفه‌های واریانس مرتبط با یک فرایند تولید به‌کار گرفت. شناسایی منابعی که بیشترین سهم را در تغییرپذیری در خروجی دارند، گامی مهم به سمت بهبود کیفیت است.

۳. در شکل ۸.۳، عامل آشیانی $B(A)$ اغلب اثری تصادفی خواهد بود، بنابراین $F_A = \text{MSA/MSB(A)}$ آمارهٔ مناسب برای آزمون‌کردن اثر عامل A (تأمین‌کننده) خواهد بود. در حالت کلی، آماره‌های F مناسب بستگی به این خواهند داشت که آیا

۱. در S-PLUS مدل‌های آشیانی به‌وسیلهٔ عملگر خط مورب / تعریف می‌شود که عامل آشیانی پس از خط مورب ظاهر می‌شود: A/B .

عامل ثابت است یا تصادفی و وقتی طرح دربرگیرندهٔ هم عامل‌های آشیانی و هم متقاطع باشد، دقت خاصی لازم است. خواننده را برای مروری کلی بر تحلیل طرح‌های آشیانی به مونتگومری (۱۹۹۷) ارجاع می‌دهیم.

طرح‌های نامتعادل

طرح نامتعادل زمانی مطرح می‌شود که تعدادی نابرابر از مشاهدات در هر اثر تیمار داشته باشیم. این وضعیت به دلیل پیش‌آمدن مشکلاتی در گردآوری داده‌ها (مثلاً، مردن بعضی نمونه‌ها ضمن آزمایش) رخ می‌دهد و منجر به از دست‌رفتن مشاهدات در آنچه در بدو امر می‌توانست یک طرح آزمایشی متعادل باشد می‌شود. از سوی دیگر، یک طرح نامتعادل ممکن است از روی عمد انتخاب شده باشد. اجرای برخی ترکیب‌های تیمار ممکن است گران‌تر باشد و بنابراین ممکن است مشاهداتی کمتر یا تعداد تکراری بیشتر برای عامل‌هایی که در مطالعه مورد توجه بیشتری هستند ترتیب داده شود.

مشکل اصلی با طرح‌های نامتعادل آن است که فنون تحلیل واریانس رایج مرتبط با طرح‌های متعادل قابل اعمال نیستند. این امر نتیجهٔ آن است که برخی ترکیب‌های تیمارها را می‌توان در یک آزمایش نامتعادل دقیق‌تر از بقیه برآورد کرد. در نتیجه، خاصیت تعامد اثرهای تیمارها که در طرح‌های متعادل حضور دارد به حالت نامتعادل تسری پیدا نمی‌کند و تحلیل را پیچیده می‌کند. به‌علاوه، فرض‌هایی که برای اثر تیمار آزمون می‌شوند، با فرض‌های مربوط به یک طرح متعادل تفاوت دارند و ممکن است به‌سادگی قابل تفسیر نباشند. در برخی حالت‌ها، فرضی که آزمون می‌شود ممکن است یکتا نباشد یا ارزش آن محل تردید باشد. این مطلب در رابطه با انواع مختلف SS و فرض‌های مربوط به مدل‌های ANOVA مورد بحث قرار می‌گیرد (نگاه کنید به ANOVAی دوطرفه).

در حالت کلی، باید یک بستهٔ نرم‌افزاری آماری جامع، یا بسته‌ای را که ویژهٔ طرح‌های آزمایشی است، برای تحلیل طرح‌های نامتعادل به‌کار گرفت (نرم‌افزارهای آماری در بخش بعدی مورد بحث قرارگرفته‌اند). با این حال، اگر داده‌ها به متعادل‌بودن «نزدیک» باشند، چندین روش تقریبی موجودند که می‌توان از آن‌ها استفاده کرد. این روش‌ها یک مسئلهٔ نامتعادل را به مسئلهٔ متعادل تبدیل می‌کنند که می‌توان آن را با اغلب بسته‌های نرم‌افزاری آماری تحلیل کرد.

تبصره‌ها

۱. صرف‌نظر از سهولت محاسباتی تحلیل، طرح‌های متعادل دو مزیت عمدهٔ دیگر را بر طرح‌های نامتعادل دارند:

- وقتی تعداد مشاهدات در هر اثر تیمار برابر است، توان آزمون ماکسیمم می‌شود.
- وقتی اندازه‌های نمونه‌ها برابرند، آزمون نسبت به انحراف از فرض برابری واریانس‌ها در سطوح تیمار کمتر حساس است.

۲. وقتی که خانه‌های گمشده (ترکیب‌های تیمار) وجود داشته باشند، روش‌های تقریبی به درد نمی‌خورند. در این حالت، یک تحلیل دقیق مورد نیاز است و نتایج را باید به‌دقت تفسیر کرد. توجه کنید که بسته‌های نرم‌افزاری مختلف، پارامترگذاری‌های مختلفی را به‌کار می‌برند. نقش ما به‌عنوان یک مشاور آماری در این نوع تحلیل‌ها مهم خواهد بود. کتاب سیل (۱۹۸۷) به تحلیل داده‌های نامتعادل می‌پردازد.

۸.۳ نرم‌افزارهای آماری

طیف گسترده‌ای از بسته‌ها و نرم‌افزارهای کاربردی وجود دارند که می‌توانند تحلیل‌های آماری انجام دهند. لازم است که مشاور آماری با قابلیت‌های چندین بستهٔ نرم‌افزاری آشنا و در استفاده از حداقل یک بستهٔ جامع زبردست شود. در این کتاب ما از SAS و S-PLUS (Mathsoft 2000) در مطالعه‌های موردی استفاده می‌کنیم و در بخش‌های آتی بحث مختصری از این بسته‌ها ارائه شده است. در اینجا مروری بر برخی نرم‌افزارهای آماری که در حال حاضر در دسترس‌اند و سطوح تحلیل آماری آن‌ها ارائه می‌کنیم.

سطح ۱

منویی. ممکن است برخی شیوه‌های آماری را دربرنداشته باشد. کیفیت تحلیل ممکن است متفاوت باشد. باید گرافیک‌های کاربردی خوبی را در اختیار گذارد. برای استفاده‌کنندهٔ تازه‌کار/دانشجو مناسب است. یادگیری آن نسبتاً آسان است. قیمت کم یا متوسط به کیفیت و شیوه‌های موجود در آن بستگی دارد.
مثال‌ها: JMP، Statgraphics، Statview، اکسل.

سطح ۲

نسخه‌های منویی ممکن است یک پنجرهٔ خط فرمان داشته باشند. رشته‌ای به‌قدر کافی جامع از شیوه‌های آماری را در اختیار می‌گذارند. کیفیت تحلیل عموماً خوب است. تعریف گرافیک‌ها به‌وسیلهٔ کاربر امکان‌پذیر است. اغلب غیرآمارداتان از آن استفاده می‌کنند. یادگرفتن نحو خط فرمان ممکن است زمان‌بر باشد. نسخه‌های کامل آن‌ها قیمت متوسطی دارد.
مثال‌ها: JMP (ویرایش ۴)، SPSS، MINITAB، Systat.

سطح ۳

برای کاربر خبره طراحی شده است. مبتنی بر خط فرمان است. برخی از آن‌ها ممکن است فقط مخصوص برخی انواع تحلیل‌ها باشند. کیفیت تحلیل می‌تواند عالی باشد. خروجی همیشه به شکل استاندارد ارائه نمی‌شود. تبدیل‌شدن به یک کاربر چیره‌دست زمان و تلاش می‌برد. برخی از بسته‌های جامع ممکن است (خیلی) گران باشند.

مثال‌ها: SAS، S-PLUS، R، GLIM، GENSTAT، BMDP، Enterprise Miner، Clementine، BUGS.

تبصره‌ها

۱. باید تأکید کنیم که رسته‌های «سطح» ما عمدتاً مبتنی بر این هستند که استفاده از نرم‌افزارهای آماری فهرست‌شده تا چه حد برای مشتری، دانشجو، یا مشاور آسان است. به‌عنوان مثال، تفاوت قابل ملاحظه‌ای بین JMP و اِکسل از لحاظ قابلیت آماری آن‌ها وجود دارد. هیچ مشاور آماری از اِکسل برای اجرای تحلیل آماری استفاده نمی‌کند، اما اِکسل به‌طور گسترده‌ای برای واردکردن داده‌ها مورد استفاده قرار می‌گیرد و می‌تواند علاوه بر تهیهٔ آمارهای خلاصه و نمودارها برخی شیوه‌های آماری را اجرا کند. بنابراین، مشاور نیاز به آگاهی از این دارد که اِکسل چه چیزی می‌تواند برای مشتری فراهم کند.[۱]

۲. تمایز بین سطح‌های ۱ و ۲ نیز گاهی شفاف نیست. JMP 3.x کاملاً منویی است و در اساس تنها یک مدول برای تحلیل سری زمانی کم دارد (این مدول در JMP 4.0 موجود است). با این حال، بسته‌ای است که آن را به مشتری‌یی که می‌خواهد در یک نرم‌افزار آماری «واقعی» سرمایه‌گذاری کند، توصیه می‌کنیم. این برنامه هم در مکینتاش و هم درPC پلت‌فرم[۲] اجرا می‌شود و به‌کاربردن آن آسان است. به‌کاربردن ویرایش‌های PCی SPSS و MINITAB نیز بسیار آسان است و یکی از این دو ممکن است انتخاب مناسبی برای مشتریان باشد (در رشته‌های بازرگانی، گرایش به استفاده از MINITAB و در رشته‌های علوم انسانی گرایش به استفاده از SPSS است).

۳. نرم‌افزار سطح ۳ آشکارا همان است که مشاور آماری برای اجرای تحلیل‌های آماری به آن‌ها نیاز دارد. SAS و S-PLUS مطمئناً بسته‌های نرم‌افزاری برتر محسوب می‌شوند، اما

---

۱. مشتریان گاهی کاملاً از اینکه اِکسل می‌تواند چقدر کار آماری برای آن‌ها انجام دهد، شگفت‌زده می‌شوند.

۲. ساختار اساسی یا حداقل یک کامپیوتر، بدون دستگاه‌های جنبی اضافی. معمولاً به تختهٔ مدار اصلی ورودی-خروجی و حافظهٔ اصلی به‌علاوهٔ سیستم عامل آن اطلاق می‌شود.

ویرایش‌های کامل آن‌ها بسیار گران‌اند. ویرایش‌های PCی SAS و S-PLUS در دسترس‌اند و R یک نرم‌افزار همگانی است که بسیار شبیه به S-PLUS است. دشواری عمده‌ای که اغلب پیش می‌آید آن است که مشتریان ممکن است از مشاهدهٔ خروجی‌های این بسته‌ها چندان راضی نباشند. آن‌ها ممکن است از این اطلاعات احساس سردرگمی و حتی یأس پیدا کنند. «آه، من از این پرینت هیچ سر در نمی‌آورم!» شاید لازم باشد که مشاور از شلوغی خروجی پیش از ارائهٔ نتایج به مشتری بکاهد. این موضوع خاص را در فصل بعدی مورد بحث قرار می‌دهیم.

SAS و *S-PLUS*

نرم‌افزارهای آماری که در همهٔ مطالعه‌های موردی از آن‌ها استفاده می‌کنیم عبارت‌اند از SAS و S-PLUS. هر دو نرم‌افزارهای کاربردی در سطح خبره‌ها هستند که برای زبردست‌شدن در آن‌ها نیاز به مقدار زیادی تلاش و زمان هست. پیوست ب آشنایی‌یی با SAS و S-PLUS در اختیار می‌گذارد، اما نامحتمل است که کاربر تازه‌کاری را به مهارت برساند. پس از خواننده انتظار داریم چه مقدار از کدهای SAS و S-PLUS را بداند؟ این سؤال را مختصراً پاسخ می‌دهیم. نخست به توصیف کوتاهی از این برنامه‌های کاربردی می‌پردازیم.

SAS می‌توان آن را تعاملی[1] اجرا کرد، اما عمدتاً یک برنامهٔ کاربردی پردازش دسته‌ای است که می‌تواند از عهدهٔ مقادیر بزرگی از داده‌ها به‌طرزی کارا برآید. در واقع، دستور مربوط به گام موسوم به DATA قسمت دشوار SAS است. ممکن است پیش از آنکه بتوان یک شیوهٔ آماری (گام "PROC") را به‌کار گرفت نیاز به دستکاری‌های گسترده‌ای روی داده‌ها باشد. خوشبختانه اغلب دستورهای PROC در SAS نام‌هایی دارند که هدف از آن‌ها را به‌روشنی مشخص می‌کنند. خوانندگانی که با نرم‌افزارهای کاربردی برنامه‌نویسی آماری نظیر SPSS، GENSTAT، یا GLIM آشنا باشند، نباید دشواری زیادی در فهم مقصود کلی کد در یک شیوهٔ SAS داشته باشند. برای خوانندگان این رسته، پیوست ب مرجع سودمندی در اختیار می‌گذارد. SAS اخیراً نرم‌افزار داده‌کاوی جدیدی به نام SAS Enterprise Miner معرفی کرده است که بیشتر برای تحلیل‌کردن و «کاویدن» مجموعه داده‌های بزرگ مناسب است. در مطالعهٔ موردی ۴.۸ روش‌شناسی داده‌کاوی آماری را معرفی می‌کنیم.

S-PLUS همیشه باید به‌صورت تعاملی اجرا شود. S-PLUS واقعاً یک محیط زبان برنامه‌نویسی مبتنی بر تعداد زیادی «تابع» از پیش تعریف شده است. برخی کاملاً مقدماتی‌اند: ()q

۱. صفت سیستمی است که می‌تواند همزمان با کار خود به فرمان‌های کاربر نیز پاسخ دهد.

برای خروج از S-PLUS؛ برخی دیگر می‌توانند یک تحلیل آماری کامل را تولید کنند: `lsfit(x,y)` برای انجام‌دادن رگرسیون کمترین توان‌های دوم. ایجاد تابع‌های کاربرخواه یکی از نقاط قوت اصلی S-PLUS است. یک تحلیل آماری در S-PLUS نوعاً مشتمل است بر تعامل مستقیم کاربر با داده‌ها در هر گام تحلیل، اعمال تغییرات کوچک، بهبوددادن مدل یا شیوهٔ آماری، و ایجاد تابع‌های کاربرخواه در صورت لزوم. S-PLUS به کاربر اجازهٔ اعمال کنترل دقیق روی یک شیء گرافیکی را نیز می‌دهد که آن را برای تولید گرافیک با کیفیت ارائه‌کرد، ایده‌آل می‌کند. R یک نرم‌افزار آماری همگانی است که از دیدگاه کاربر تقریباً با S-PLUS یکی است. تفاوت‌های عمده در این است که S-PLUS یک واسط گرافیکی کاربری[1] فراهم می‌کند، و برخی بسته‌های نرم‌افزاری دیگر را شامل می‌شود. به‌عنوان یک محصول تجاری، S-PLUS از کاربران خود پشتیبانی برخط می‌کند. همهٔ کدهای S-PLUS موجود در مثال‌های این کتاب، با R هم سازگارند.

مقایسهٔ SAS و S-PLUS

S-PLUS با SAS بسیار متفاوت است. در واقع، باید خیلی تلاش کنیم که یک عبارت زائدتر از این پیدا کنیم. تفاوت بنیادی در رویکرد آن‌ها به تحلیل آماری است:

SAS شیوه‌های آماری را فراهم می‌کند ــ برای کاربران

S-PLUS نیاز کاربر را رفع می‌کند ــ با شیوه‌های آماری.

این واقعاً موضوع مهمی است. ارائهٔ فهرستی از تفاوت‌های خاص تا حدی بی‌مورد است زیرا این تفاوت‌ها ذاتاً ناشی از رویکردهای متفاوت است. لازم است که مشاور آماری این تفاوت‌ها را بپذیرد و یاد بگیرد که چگونه به بهترین وجه از مزیت‌های هر یک از برنامه‌های کاربردی استفاده کند.

دیگر نرم‌افزارها در سطح خبره‌ها

SAS و S-PLUS تنها نرم‌افزارهای آماری موجود در سطح خبره‌ها نیستند، اما به‌طور گسترده‌ای شناخته شده و در سراسر دنیا مورد استفاده‌اند.

SPSS (بستهٔ آماری برای علوم اجتماعی) به‌طرز شگفت‌آوری عمدتاً مورد استفادهٔ متخصصان علوم اجتماعی است. آماردانان عموماً از SPSS دوری می‌کنند و خروجی SPSS تولیدشده به‌وسیلهٔ کاربران ناخبره را با شک و تردید می‌نگرند. در حالی که شک در خروجی تولیدشده به دست یک کاربر غیر خبره شاید هنوز دستخوش تغییر نشده باشد، اما کیفیت SPSS مطمئناً بهبودیافته

1. graphical user interface (GUI)

و امروز این انتقاد تا حدی غیر منصفانه است. از آنجا که SPSS منویی است، جایگاه خود را در دوره‌های مقدماتی آماری پیدا کرده است. Clementine نرم‌افزار داده‌کاوی بسیار موفقی است که به‌وسیلهٔ SPSS تولید شده است.

GENSTAT و GLIM بیرون از ایالات متحده بسیار معروف‌اند و هر دو نرم‌افزارهای آماری مورد توجهی هستند. GENSTAT یک برنامهٔ کاربردی عمومی است در حالی که GLIM ویژهٔ مدلبندی خطی تعمیم‌یافته (تعاملی) است.

XlispStat (Tierney 1991) و BUGS دو نرم‌افزار همگانی با هوادارانی جدی در محیط‌های دانشگاهی هستند. BUGS مخصوص تحلیل آماری بیزی است و XlipStat گرافیک‌های تعاملی فوق‌العاده خوبی دارد.

ارائهٔ کد SAS و S-PLUS

منظور از ارائهٔ کد SAS و S-PLUS آن است که فرض می‌کنیم خواننده با دستورهای برنامه‌نویسی SAS و S-PLUS و نیز قراردادها و نحو آن‌ها آشناست. البته، «فرض» اینکه خواننده با کد SAS و S-PLUS آشناست لزوماً به این معنی نیست که «انتظار» داریم خواننده با این کد آشنا باشد. برای آنکه طیف وسیعی از خوانندگان را در نظر بگیریم، برخی مصالحه‌ها به‌صورت زیر انجام شده‌اند.

غیر کاربر حیطهٔ مسئلهٔ مطرح‌شده در مطالعهٔ موردی و توصیف تحلیل، راهنمای اصلی شماست. قسمت اعظم کد برنامه احتمالاً برای شما بی‌معنی خواهد بود. آن را نادیده بگیرید! اگر کتاب را تا به اینجا خوانده باشید آن را اشتباهاً برنگزیده‌اید؛ علاقهٔ شما آشکارا در دیگر جنبه‌های فرایند مشاوره نهفته است.

کاربر تازه‌کار مقدمه و جزئیات موجود در پیوست ب تا حدی به شما کمک خواهد کرد. با این حال، احتمالاً برای بازسازی یک برنامهٔ کامل برای تحلیل مطالعهٔ موردی کافی نخواهد بود. لازم است کمک بگیرید، اما این کار ارزش تلاش را خواهد داشت.

کاربر نیمه‌مجرب نرم‌افزار خوب مستمراً در تحول و رشد است. ممکن است برخی قسمت‌های کد را تشخیص ندهید، اما امید می‌رود که مدت زیادی از زمان استفادهٔ شما از نرم‌افزار نگذشته باشد. توصیف تحلیل ممکن است در برخی حالت‌ها راهنمای عمدهٔ مورد نیاز شما باشد.

کاربر نیمه‌خبره شما نباید در درک مقصود کلی کد در شیوهٔ SAS مشکل زیادی داشته باشید. توصیف تحلیل و متن مسئله در یک مطالعهٔ موردی برای شما کفایت می‌کند تا تحلیل را

در نرم‌افزار دلخواه انتخابی خود تکرار کنید. کد S-PLUS هنوز هم ممکن است برای شما مرموز به نظر برسد.

کاربر خبره متوجه خواهید شد که کد برنامهٔ ما (به صورتی که نوشته شده) اغلب کاری بیش از آنچه ذکر می‌کنیم و کمتر از آنچه مورد نیاز است انجام می‌دهد. در بهترین حالت کد برنامه شاید یک یادآوری سودمند یا راهنمایی برای شما فراهم کند، اما احتمالاً دربارهٔ این برنامه‌های کاربردی بیش از ما می‌دانید!

مروری بر ابزارهای محاسباتی

مشاوران آماری به تعداد زیادی ابزارهای محاسباتی دسترسی دارند که در پلت‌فرم‌های مختلف و اینترنت در دسترس‌اند. ترکیب این ابزارها امکان‌های جدیدی به روی ما می‌گشاید اما در همان حال مشکلات جدیدی ایجاد می‌کند. به‌عنوان مثال، ممکن است مجموعه گراف‌هایی با کیفیت در فرمت PostScript تولید کرده باشیم، اما حالا نیاز داریم که آن‌ها را به اسلایدهایی برای ارائه با پاورپوینت تبدیل کنیم. به همین نحو، فرض کنید که با یک فایل داده‌های اِکسل کار را شروع کرده باشیم. دوست داریم که تحلیل اکتشافی داده‌ها را در S-PLUS انجام دهیم، اما لازم است که تحلیلی با استفاده از یک شیوهٔ SAS تولید کنیم. وضعیت‌هایی از این نوع اغلب در بسیاری از پروژه‌های مشاورهٔ آماری مطرح می‌شوند و ممکن است نیاز به انتقال داده‌ها بین کامپیوترها داشته باشند. گاهی لازم است که داده‌ها را در یک فایل «متن راسته» (ASCII) ذخیره و سپس، در صورت لزوم، برای تبدیل داده‌ها به فرمتی خاص، مقداری ویرایش دستی کنیم. با این حال، این کار ممکن است خیلی زمان‌بر باشد و اغلب بهتر است که از شیوه‌های خودکاری که به‌وسیلهٔ نرم‌افزارهای دیگر تهیه شده استفاده کنیم.

انتقال داده‌ها بین پلت‌فرم‌ها و برنامه‌های کاربردی

اغلب برنامه‌های کاربردی می‌توانند فایل‌های داده‌های متن راسته تولید کنند. معمولاً نیاز به حائل‌ها برای جداکردن فیلدهای داده‌ها داریم، اما باید این‌ها را به‌دقت انتخاب کرد. به‌عنوان مثال، حائل پیش‌فرض ممکن است یک «tab stop»، فاصلهٔ خالی، یا کاما باشد، اما اگر مجموعه‌داده‌ها فیلدی داشته باشد که فاصله‌های خالی و کاماها را به عنوان کاراکتر در نظر می‌گیرد (مثلاً، «Smith, James B.»)، آنگاه استفاده از کاما یا فاصلهٔ خالی به‌عنوان حائل آشکارا باعث ایجاد مشکلاتی خواهد شد. در این حالت‌ها، لازم است کاراکتر خاصی مانند «:» یا «;» را به‌عنوان حائل به‌کار ببریم. هم SAS و هم S-PLUS به‌کاربر اجازه می‌دهند که از هر کاراکتری به‌عنوان حائل استفاده کند.

فرمت‌های غیر متنی مانند اِکسل یا DBaseIV اغلب به این دلیل که بسیاری از برنامه‌های کاربردی این فرمت‌ها را درک می‌کنند به‌کار می‌روند. اکسل بسیار رایج است و برای اغلب پروژه‌ها واردکردن و ذخیره‌کردن داده‌ها به‌صورت یک فایل اکسل کافی است. اگر مجموعه‌داده‌ها حاوی نام متغیرها، برچسب‌ها، و اطلاعات فرمت‌بندی خاصی باشد، در این صورت، شاید بهتر این باشد که داده‌ها را در فرمتی مانند DBaseIV وارد و ذخیره کنیم. بعداً می‌توان فایل DBaseIV را به سیستم جدید انتقال داد و در برنامهٔ کاربردی نرم‌افزار با استفاده از فرمت جدید خواند و به این ترتیب همهٔ اطلاعات متغیر و مورد را حفظ کرد. یک مجموعه‌داده‌های داتمی SAS مثالی خوب از «پایگاه داده‌ها»یی است که شامل این نوع اطلاعات است و می‌توان آن را از روی داده‌های ذخیره‌شده در دیگر نرم‌افزارهای نظام مدیریت پایگاه داده‌ها[1] (DBMS) ذخیره کرد. حمل مستقیم یک مجموعه‌داده‌های بزرگ SAS بین کامپیوترها معمولاً عملی‌تر از تلاش برای بازسازی پایگاه داده‌ها از روی فایل فرمت‌شدهٔ متن راسته است. دراية مربوط به SAS و S-PLUS در پیوست ب مثال‌هایی را از خواندن و نوشتن انواع مختلف فایل‌های داده‌ها در SAS و S-PLUS در اختیار می‌گذارد.

صدور فایل‌های گرافیکی

گزارش‌های مشاورهٔ آماری اغلب نیاز به ضمیمه‌کردن نمایش‌های گرافیکی دارند. یک گزینه صرفاً این است که همهٔ نسخه‌های نمودارها را چاپ و آن‌ها را در گزارش درج کنیم. با این حال، این کار ممکن است تا حدی ناکارآمد و زمان‌بر باشد چون بی‌تردید به این معناست که تحلیل را باید، شاید چندبار، از نو تولید کرد. گزینه‌ای بهتر ذخیره‌کردن یک نمایش گرافیکی در یک فایل گرافیکی است. چندین روش برای انجام‌دادن این کار موجود است و موضوع اصلی آن است که از چه فرمتی استفاده کنیم.

در محیط‌های ویندوز و مکینتاش، کپی‌کردن و چسباندن بین برنامه‌های کاربردی متفاوت آسان است. بنابراین، اگر تصمیم بگیریم که نمودارهای خود را در یک فایل پاورپوینت نگاه داریم، هر زمان که نموداری در S-PLUS یا SAS تولید می‌کنیم، می‌توانیم آن را در یک فایل پاورپوینت کپی کنیم و بچسبانیم. از پاورپوینت می‌توانیم نمودارها را به بسیاری از برنامه‌های کاربردی دیگر و از روی پلت‌فرم‌ها منتقل کنیم.

SAS و S-PLUS هر دو تسهیلاتی برای صدور نمودارها به فایل‌ها در انواع گوناگونی از فرمت‌ها در اختیار می‌گذارند. توصیه می‌کنیم که فایل‌های گرافیکی را در فرمت GIF یا JPEG صادر کنید زیرا این فرمت‌ها در اغلب پلت‌فرم‌ها وجود دارند. فرمت (EPS)Encapsulated PostScript را می‌توان زمانی که نمودار به‌طور عمده برای چاپ مورد استفاده قرار می‌گیرد، به‌کار برد. یک مزیت

1. database management system

فرمت EPS آن است که نمودار در کد ASCII ذخیره می‌شود؛ یعنی فایل متن راسته‌ای که می‌توان آن را ویرایش کرد. نمودارهای PostScript را می‌توان در LaTeX (مانند کتاب حاضر) جا داد و این امکان وجود دارد که هر خروجی قابل چاپ را، به جای «چاپ»، با ارسال خروجی به پرینتر و «ذخیره‌کردن» (در فرمت EPS) تبدیل کرد. با این حال، هشیار باشید که تبدیل برخی انواع خروجی قابل چاپ، نظیر فایل‌های GIF یا JPEG ممکن است فایل‌های PostScript بسیار بزرگی ایجاد کند.

واسط‌ها و ترکیب‌های نرم‌افزاری

ترکیب دو (یا چند) برنامهٔ کاربردی نرم‌افزاری گاهی ممکن است نتایج بهتری تولید کند: بگذارید هر کدام بهترین کاری را که قادرند انجام دهند. از Xgobi می‌توان به عنوان واسط گرافیکی بین S-PLUS و R استفاده کرد و به این طریق S-PLUS را با توانایی‌های گرافیکی پویای پیشرفته مجهز کرد. واسط S-PLUS با Mathematica، SAS، و اکسل نیز به S-PLUS قدرت عمل بیشتری می‌دهد.

در فصل بعد، سهولت استفاده از نرم‌افزار آماری JMP را برای ارائهٔ نتایج مقدماتی به مشتریان خود، SAS را برای تحلیل آماری نهایی، و S-PLUS را برای تولید گرافیک با کیفیت ارائه‌کردی که مورد نظر مشتری است، مورد استفاده قرار می‌دهیم.

StatLib

StatLib یک منبع مهم اطلاعات نرم‌افزارهای آماری است: می‌توان به این منبع با استفاده از شبکه یا ftpای «ناشناس»[1] در سایت

```
Web: http//lib.stat.cmu.edu/
FTP: ftp .lib.stat.cmu.edu
```

دسترسی یافت. در بین نرم‌افزارهای موجود در StatLib، برنامه‌های S-PLUS و SAS و بسته‌های تهیه‌شده به‌وسیلهٔ کاربران، نرم‌افزار همگانی R، Xgobi و بسیاری دیگر را می‌توان یافت. ftpای ناشناس برای دانلودکردن برخی فایل‌ها و برنامه‌های کاربردی که (به‌صورت غیر فشرده) ممکن است بسیار بزرگ باشند، ضروری است.

ftpای ناشناس صرفاً به این معناست که از کلمهٔ anonymous به‌عنوان نام کاربری «guest» موقع دسترسی به سایت ftp استفاده کنیم. فیلد «گذرواژه» با واردکردن آدرس ای‌میل کامل می‌شود. پس از آنکه موفق به ورود شدیم (گاهی کاربرهای زیادی همزمان به سایت وارد شده‌اند) و فایل‌هایی را که می‌خواهیم دانلود کنیم، پیدا کردیم، باید اطمینان حاصل کنیم که مد انتقال برای

۱. ftpای ناشناس نوعی پایگاه ftp در اینترنت است که بر خلاف پایگاه‌های ftp دیگر، درخواست گذرواژه و نام کاربری نمی‌کند.

دانلودکردن فایل‌ها و برنامه‌های کاربردی فشرده در وضعیت «binary» است. مد «ASCII» را تنها باید هنگام دانلودکردن فایل‌های متن راسته مانند فایل‌های "README" به‌کار برد. آرشیو یا فرمت فشردگی معمولاً با پسوندی در انتهای نام فایل مشخص می‌شود. برخی پسوندهای رایج در زیر فهرست شده‌اند.

آرشیوها

sea. آرشیو خوداستخراج (ویندوز و مکینتاش). فایل‌های گروهی باید زمانی که فعال می‌شوند خود به خود استخراج شوند.

tar. آرشیو tape (UNIX). از `tar xf filename.tar` برای استخراج استفاده کنید.(ارجاع به «tape» جنبهٔ تاریخی دارد).

فشرده‌سازی

Z. (UNIX) از `uncompress filename.z` استفاده کنید.

gz. فایل gzip (UNIX). از `gunzip filename.gz` استفاده کنید. می‌توان آن را روی پلت‌فرم‌های ویندوز و مکینتاش از حالت فشرده خارج کرد.

zip. فایل zip (ویندوز و UNIX). در UNIX از `unzip filename.zip` استفاده کنید. در پلت‌فرم مکینتاش می‌توان آن را از حالت فشرده خارج کرد.

sit. فایل StuffIt (مکینتاش). روی ویندوز می‌توان آن را از حالت انباشته خارج کرد.

۴

یک پروژهٔ مشاوره از الف تا ی

در این فصل، هدف ما تلاش برای این است که فرایند مشاوره را به‌طور کامل برای پروژه‌ای خاص از مرحلهٔ اولین تماس با مشتری تا مرحلهٔ گزارش کتبی نهایی و پیگیری‌های تکمیلی بازسازی کنیم. مشاورهٔ واقعی در یک محیط دانشگاهی صورت گرفته که ما آن را SCP می‌نامیم (نگاه کنید به بخش ۵.۴.۱). برخی جنبه‌های این پروژه را تغییر داده‌ایم که هویت مشتری محفوظ، و حجم این ارائه‌کرد در حدی معقول بماند.

۱.۴ اطلاعات پیشین

ترتیب زمان و مکان نخستین جلسهٔ مشاوره از طریق تماس تلفنی با مشتری داده شد. پروژه به‌اختصار مورد بحث قرارگرفت و به مشتری گفتیم که اطلاعات مربوط، از قبیل نسخهٔ چاپی داده‌ها، را در دیدار پیش‌رو با خود به‌همراه بیاورد. این تماس فرصت حصول اطلاعاتی دربارهٔ پروژه را نصیب ما کرد:

- پروژه پساآزمایشی بود.

- مشتری از ما می‌خواست که تحلیل را انجام دهیم و نمودارها را آماده کنیم.
- پروژه موضوع مورد مطالعه در پایان‌نامهٔ مشتری بود.
- پروژه به روش‌های تدریس و شیوه‌های یادگیری مربوط می‌شد.
- آزمایش $n = ۸۷$ آموزگار را شامل می‌شد.

تبصره‌ها

پرسشی بدیهی این است که «چرا اطلاعات بیشتری گردآوری نکردیم؟» یک پاسخ صرفاً این است که در واقع از مشتری می‌خواهیم که از اول شروع کند و این مطلب اساساً به‌معنای تکرار هر چیزی است که پیش از جلسهٔ مشاورهٔ نخستین مورد بحث قرار گرفته بود. بنابه تجربه‌ای که داشتیم، تلاش برای کسب اطلاعات خیلی زیاد پیش از نخستین جلسه امکان سوءتفاهم‌ها را بیشتر می‌کند. این امر نه‌تنها از عینی‌گرایی که از ضرورت‌های مشاورهٔ آماری است می‌کاهد، بلکه کاملاً دشوار نیز هست که تصورات از پیش شکل‌گرفته در ذهن دربارهٔ داده‌ها و سمت و سوی تحلیل را از نو تعدیل کرد.

البته برقراری تماس با مشتری پیش از نخستین جلسهٔ مشاوره مزیت‌هایی دارد. با این کار فرایند ایجاد ارتباط آغاز می‌شود و این فرصت را در اختیار ما قرار می‌دهد که با ماهیت کلی پروژه آشنا شویم. در برخی حالت‌ها، ممکن است دریابیم که پروژه ارزش تعیین یک جلسهٔ مشاورهٔ نخستین را ندارد. بدیهی است که می‌خواهیم بحث خود را با مشتری طولانی‌تر کنیم تا مطمئن شویم که اطلاعاتی را که به ما داده است سوءتعبیر نکرده‌ایم. بنابراین، لازم است برای آگاهی از میزان و سطح پیچیدگی (آماری) موجود در پروژه اطلاعات کافی کسب کنیم. برخی دلایل برای اینکه چرا ممکن است ترتیب‌دادن یک جلسهٔ مشاورهٔ اولیه را به تعویق بیندازیم یا آن را منتفی کنیم عبارت‌اند از:

- در موارد حقوقی، ممکن است تضاد منافعی در ارتباط با پروژهٔ مشتری صرفاً به این دلیل که برای مشتری دیگری کاری در دست انجام داریم یا قبلاً انجام شده به وجود آید.
- اندازهٔ نمونه برای موجه‌بودن تحلیل آماری بسیار کوچک باشد.
- شاید بهتر باشد که مشتری برخی کارها را پیش از ترتیب‌دادن یک جلسهٔ ملاقات انجام دهد. گردآوری، واردکردن یا فرمت‌بندی داده‌ها.
- ما منابع کافی یا تخصص لازم را در ارتباط با تحلیل آماری مورد نیاز در پروژهٔ مشتری در اختیار نداریم. آیا می‌توانیم یک آزمایش بالینی طراحی و اجرا کنیم؟ آیا می‌دانیم یک مدل بازار را چگونه تعبیر کنیم؟

برای این پروژهٔ خاص، اطلاعات پیشینی که از مشتری به‌دست آوردیم اجازهٔ آن را می‌داد که جلسهٔ مشاوره را ترتیب دهیم و این امکان را فراهم آورد که نظرات زیر را در مورد پروژه شکل دهیم:

۱. ضروری است که نحوهٔ طراحی و اجرای آزمایش به‌دقت مشخص شود تا اطمینان حاصل شود که اصول طرح آزمایشی: کنترل، تصادفی‌سازی، و تکرار را برآورده می‌کند.

۲. تحلیل احتمالاً مستلزم شیوه‌های AVOVAی استاندارد و آزمون t خواهد بود. اگر داده‌ها مشکلی داشته باشند، ممکن است شیوه‌های ناپارامتری ضرورت پیدا کنند. در صورتی که عامل‌های زیادی دخیل باشند، شاید اندازهٔ نمونه موجب نگرانی باشد.

۳. ممکن است به نمودارهای باکیفیت نیاز پیدا کنیم زیرا این پروژه موضوع پایان‌نامهٔ مشتری بود. محتمل نیز هست که پروژه را بتوان یک بررسی مقدماتی تلقی کرد ــ ضروری است که نتایج معنی‌دار با احتیاط تعبیر و تفسیر شوند.

۴. نیاز خواهیم داشت بفهمیم که «سبک‌های یادگیری» واقعاً به چه معنایی هستند.

۲.۴ مسائل مالی

می‌توانیم به این پرسش از یک دیدگاه صرفاً اقتصادی نزدیک شویم که «چرا اطلاعات بیشتری گردآوری نکردیم؟» پاسخی به همان اندازه بدیهی این است که مشاور آماری انتظار دارد که اجرت صرف وقت و توصیه‌های حرفه‌ای‌اش را بپردازند! با وجود آنکه این لزوماً پاسخی نیست که خواننده انتظار آن را دارد، مطمئناً از نظر مشتری به صواب نزدیک‌تر خواهد بود که پیش از آنکه خود را متعهد به حضور در یک جلسهٔ مشاورهٔ اولیه بکند، حساب هزینه‌های مربوط را بکند. بنابراین، لازم است که به این پرسش مشتری پاسخ دهیم:

«چقدر می‌گیرید؟»

معمولاً ارائهٔ نرخ ساعتی متعارفی به مشتری کافی است اما برای پروژه‌های کوتاه‌مدت، مشتری اغلب هزینهٔ کل را می‌خواهد. آنها ممکن است علاقه‌مند نیز باشند که بدانند به چه «سرعتی» می‌توانیم تحلیل را کامل کنیم. برای اینکه به‌صورت واقع‌بینانه‌ای به این مباحث بپردازیم، ضروری است که جزئیات پروژه را به‌طور کامل بدانیم، جزئیاتی را که آشکارا می‌خواهیم تا جلسهٔ مشاورهٔ واقعی به تعویق بیفتند. پس چه پاسخی می‌دهیم؟ راه‌حل آن است که توافق مشتری را برای ترتیب‌دادن یک جلسهٔ مشاورهٔ اولیه، بدون اینکه دست و پای خود را با برآوردهای غیر واقعی وقت و هزینه ببندیم، جلب کنیم. برخی راهبردهای ممکن عبارت‌اند از:

- نخستین جلسهٔ مشاوره را به‌رایگان، اما تحت یک شرط غیر الزام‌آور، برگزار کنیم. این بدان معنی‌است که این گزینه را از پیش برای خود حفظ کنیم که طی این مدت از قبول پروژهٔ مشتری سرباز زنیم، اما از مشتری خواه به این گزینه عمل کنیم و خواه نکنیم هزینه‌ای دریافت نمی‌کنیم.

- به مشتری عقد قراردادی را پیشنهاد کنید. پس از جلسهٔ مشاورهٔ اولیه، به مشتری مبلغ مشخصی را برای پروژه ارائه می‌کنیم: پردازش داده‌ها، تحلیل، گزارش‌نویسی، و رسم نمودارهای لازم.

در برخی حالتهای نادر، بهترین گزینه ممکن است صرفاً این باشد که رویکردِ همین است و بس را در پیش گیریم. ممکن است مشتری را از دست بدهیم، اما ممکن است غریزه حکم کند که این بهترین کار است.

درپیش‌گرفتن این رویکرد، ممکن است هر زمان که دو رویکرد قبلی به شکست انجامیده باشد، ضروری باشد. اگر مشتری مایل نباشد که پروژه را رودررو مورد بحث قرار دهد، آنگاه واردشدن به پروژه احتمالاً ارزش نخواهد داشت.

تبصره‌ها

گرچه ممکن است بتوانیم مشکل مشتری را طی یک ساعت حل کنیم، گزینهٔ نخستین ساعت رایگان، مناسبِ وضعیتی است که اطلاعات پیشینی دربارهٔ پروژهٔ مشتری در دست داریم. این رویکرد، اغلب روش خوبی برای جلب سریع موافقت مشتری برای ترتیب‌دادن یک دیدار مشاوره‌ای اولیه و پرهیز از بحثی طولانی دربارهٔ نرخ‌ها و هزینه‌های خاص است. این مطالب را می‌توان به‌تفصیل در نخستین جلسهٔ مشاوره مورد بحث قرار داد.

برآوردکردن زمان کل (و بنابراین هزینهٔ کل) برای کامل‌کردن همهٔ جنبه‌های پروژهٔ مشتری کار آسانی نیست. پردازش داده‌ها و گزارش‌نویسی اغلب بیش از آنچه انتظار داریم زمان می‌برد. به‌کارگماردن دانشجویان تحصیلات تکمیلی قطعاً به قابل تحمل‌شدن هزینه‌های کل کمک می‌کند و به دانشجویان فرصت کسب تجربه را می‌دهد، اما نهایتاً خود ما مسئولیت تحلیل را به عهده داریم. بستن قرارداد در مورد پروژه‌های کوچک‌مقیاس، که در آن همهٔ الزامات پروژه را می‌توان به‌صراحت مشخص کرد، کار مناسبی است. لازم است بندها و شرایط قرارداد مکتوب شوند و هر دو طرف آن را امضا کنند.

گرچه برنامه‌های مشاورهٔ دانشگاهی این توانایی را دارند که دانشجویان دورهٔ تحصیلات تکمیلی را با انواع گسترده‌ای از پروژه‌ها درگیر کنند، مشتری‌ها گاهی بر این تصورند که از دانشجویان برای دادن «تجربه» به آن‌ها استفاده می‌شود. این مطلب قطعاً درست است و لازم است ما قدری انعطاف به‌خرج دهیم؛ دانشجویان اغلب آمادهٔ کارهای کوتاه‌مدت با دستمزدهای پایین‌اند مادامی که تجربه برای آن‌ها سودمند باشد. وظیفهٔ ما آشکارا این است که موارد افراط و تفریط را بپالاییم.

۳.۴ جلسهٔ I: نخستین دیدار

حال زمان آن است که برای نخستین بار چهره‌به‌چهره با مشتری خود ملاقات و به‌طور مناسب خود را برای این دیدار آماده کنیم. به‌ویژه باید:

- خونسرد باشید! مشتری برای گرفتن نظر ما به ما مراجعه می‌کند و احتمالاً بیشتر از ما احساس نگرانی دارد.
- اطمینان حاصل کنید که ملاقات در محیطی صورت پذیرد که موجب تعاملات شفاهی با مشتری با امکان تمرکز حواس است. سالن‌های گذران وقت دانشجویان، ترتیب‌دادن ملاقات طی ساعت‌هایی که برای رفع اشکال دانشجویان برنامه‌ریزی شده است، یا در حالی‌که جایی بر روی میز برای نگاه‌کردن به مطالب پرینت‌شده وجود ندارد، به‌روشنی محیط مناسبی برای مشاوره نیست.
- پوشش ظاهری شما به منوالی باشد که استانداردهای حرفه‌ای ارائهٔ خدمات را بازتاب دهد. این واقعیت که خدمات مشاوره‌ای خود را در محیطی دانشگاهی ارائه می‌کنیم نباید رسومات معمول کسب و کار را تغییر دهد.
- پروندهٔ مشتری را مقابل خود داشته باشید. این پرونده باید شامل یادداشت‌های ما از تماس‌های قبلی، فرم درج اطلاعات برای گردآوری خلاصهٔ پروژه (بعداً شکل ۲.۴ را ببینید)، و نیز کاغذ، قلم، و مداد[1] آماده برای یادداشت‌برداری ضمن جلسهٔ مشاوره باشد.
- وقت‌شناس باشید.
- همهٔ تمهیدات دیگر را از پیش در نظر بگیرید. مثلاً حضور کارکنان دیگر که شاید وجود آن‌ها در مشاوره لازم باشد: دانشجویان تحصیلات تکمیلی، مشاوران همکار، یا خبرگانی از رشته‌های علمی دیگر.

بر مبنای اطلاعات پیشینی که از مشتری گرفتیم، کاملاً محتمل به نظر آمد که یک دانشجوی تحصیلات تکمیلی می‌تواند بخش اعظم تحلیل موجود در پروژهٔ مشتری را انجام دهد. بنابراین، ترتیبی دادیم که دانشجو در این جلسهٔ مشاورهٔ آماری حاضر باشد. این کار فرصتی به دانشجو می‌دهد که در «فرایند مشاوره» شرکت کند و نیز به مشتری اجازه داد که دریابد دانشجو «که» بود. بعداً به اینکه چرا این کار مهم است می‌پردازیم.

علاوه بر آمادگی‌های معمولی فهرست‌شده در بالا، مورد مهم‌تر دیگری هست که باید خود را برای آن آماده کنیم. خود ما! به این معنا که همواره باید سعی کنیم که با نوعی دستورکار در ذهن با جلسهٔ مشاوره مواجه شویم. راهی مناسب برای فرمولبندی دستور جلسه پرسیدن سؤال زیر از خودمان است:

در این جلسه می‌خواهیم به چه دستاوردی برسیم؟

---

۱. حتماً از مشتری پیش از نوشتن چیزی بر روی مطالب پرینت‌شدهٔ آن‌ها اجازه بگیرید. ما یک‌بار این اشتباه را کردیم و مشتری نزدیک بود بلند شود و برود! خوشبختانه آن‌ها این کار را نکردند، اما این‌کار بقیهٔ مدت مشاوره را تنش‌آلود کرد.

با تمرین، آسانتر می‌شود که روند کلی جلسات مشاورهٔ خود را پیش‌بینی کنیم و دستورکار ما ممکن است صرفاً مشتمل بر یادداشت‌هایی باشد که به ذهن خود سپرده‌ایم. در غیر این‌صورت، باید دستورکار پیشنهادی خود را بنویسیم و طی جلسه آن را با خود همراه داشته باشیم.

پس، در این مثال چه دستاوردی می‌خواهیم داشته باشیم؟ از آنجا که مطالعهٔ مشتری نسبتاً کوچک بود (تنها $n = ۸۷$ مشاهده)، این امکان وجود داشت که در موقعیتی قرار گیریم که اجرای تحلیل آماری پروژه را پس از پایان جلسه شروع کنیم. برای نیل به این بهترین سناریوی ممکن، لازم است با موارد زیر در دستورکار خود به‌درستی و یکی پس از دیگری برخورد کنیم:

- تعریف روشنی از مسئله و متغیرهای مرتبط با پروژهٔ مشتری.
- اهداف مطالعه را می‌توان به کمک یک تحلیل آماری پشتیبانی کرد.
- سهم کارهای خاص لازم برای این پروژه را می‌توان به‌طور صریح بیان کرد.
- چارچوب زمانی و چگونگی پرداخت برای طرفین قابل قبول باشد.

تبصره‌ها

در این مثال خاص، توانستیم به همهٔ موارد بالا در یک جلسهٔ مشاوره بپردازیم. به این معنی که در پایان این جلسه به این هدف خود دست یافته بودیم که در وضعیتی قرار گیریم که تحلیل آماری پروژه را اجرا کنیم. این کار همیشه ممکن نیست، و مطمئناً نباید انتظار داشته باشیم که بتوانیم در یک جلسهٔ مشاوره، هر پروژهٔ مشتری را رفع و رجوع کنیم. در واقع، خوانندهٔ بصیر ممکن است از همان ابتدا تا حدودی نسبت به ادعای ما در این مثال تردید به‌خود راه دهد. آیا واقعاً همهٔ مواردی را که در زیر مطرح می‌کنیم در یک جلسه به پایان بردیم؟ ما به این سؤال در پایان ارائه‌کرد پاسخ خواهیم داد.

تماس اولیه

مشتری سر وقت در جلسهٔ مشاوره حاضر شد. اگر چنین نشود، چه اتفاقی می‌افتد؟ اگر مشتری زودتر بیاید (که معمولاً چنین است)، می‌توان معارفه‌ها را انجام داد، و در صورتی که لازم باشد، می‌توانیم از مشتری الف بخواهیم تا منتظر بماند که کار ما با مشتری ب به پایان برسد. هدف از این کار آن است که به مشتری الف اجازه دهد که نوع تعاملی را که به‌زودی با آن مواجه خواهد شد بشنود و ببیند. اگر مشتری دیر بیاید (بیش از ۲۰ تا ۳۰ دقیقه)، هنوز هم باید سعی کنیم که آن‌ها را در برنامهٔ همان روز خود بگنجانیم. معمولاً اینکه اطلاعاتی دربارهٔ پروژه بگیریم بهتر از این است که وقت ملاقات جدیدی تعیین کنیم. بنابراین، برنامهٔ ما باید مقداری انعطاف‌پذیری داشته باشد:

انباشتن تعداد زیادی مشتری در یک روز واحد به احتمال زیاد منجر به مشکلاتی در برنامه‌ریزی مجدد خواهد شد. در این مثال، گفتگوی ما با معارفهٔ اولیه آغاز می‌شود:

مشتری: سلام، من «آن مشتری دیگر» (الف) هستم.

مشاور: سلام. من مشاور ی هستم و ایشان «دانشجوی دستیار»ند که به ما در این پروژه کمک خواهند کرد. خوب، این‌طور که می‌بینم اطلاعاتی با خود آورده‌اید که نگاهی به آن‌ها بکنیم. اما قبل از هر چیز، شاید بهتر باشد که با توصیف خود شما از پروژه بحث را شروع کنیم...؟

مشتری: بسیار خوب...می‌خواستیم نشان دهیم که گنجاندن ترجیحات LSI در آموزش فناوری، به در یاد داشتن درازمدت کمک می‌کند. در این آزمایش، روش‌های سنتی را با ترجیحات شنیداری، دیداری، جنبشی، و لامسه‌ای شرکت‌کنندگان آمیختیم. ما در هر جلسه از فرمت‌های آموزشی یکسانی استفاده کردیم و ابزارهای ما پیش‌آزمون، پس‌آزمون، و نمرهٔ SDS بودند. اغلب شرکت‌کنندگان کارگاه آموزگاران زن مدارس ابتدایی بودند.... ما داده‌ها را گردآوری کرده‌ایم و من برخی پرینت‌های خروجی را برای شما آورده‌ام که ببینید....

مشاور: (حرف مشتری را قطع می‌کند)... خوب، اما قبل از این‌کار، می‌خواهم به جزئیاتی بازگردم دربارهٔ ...

مشتری: ... استاد راهنمای من چیزی دربارهٔ «تحلیل واریانس» گفت ــ و این دلیل حضور من در اینجاست!

مشاور: بسیار خوب، بگذارید ببینیم به ANOVA نیازی هست یا خیر.

تبصره‌ها

انتخاب زمانی که حرف مشتری را قطع می‌کنیم همواره آسان نیست. اگر زود وسط حرفش بپریم، ممکن است که مشتری احساس کند که فرصت منصفانه‌ای برای توضیح پروژه‌اش به او نمی‌دهیم. ارزش یادآوری دارد که از مشتری انتظار داریم که بعداً به توضیحات ما دربارهٔ برخی مباحث آماری به‌دقت گوش دهد. از سوی دیگر، مهم است که بحث را حول و حوش مطلب اصلی نگهداریم. ممکن است لازم باشد حرف مشتری را قطع کنیم تا مانع از دورشدن مشتری از موضوع شویم. این بدان معناست که نیاز داریم که پردازش «بسته‌ها»ی اطلاعاتی را آغاز کنیم. این «راهنمایی‌ها»ی شفاهی به کار این هدف دوگانه می‌آید که مانع از پیش‌افتادن بیش از حد مشتری از سؤال‌های ما و درک پروژه می‌شود، و به مشتری کمک می‌کند که کمی از آنچه واقعاً لازم است بدانیم آگاه شود.

در گفتگوی بالا توانستیم حرف مشتری را در نقطهٔ مناسبی قطع کنیم: توصیف مشتری از پروژه‌اش از همان اول چندین سؤال را پیش آورده بود که باید رفع و رجوع می‌شد؛ مشتری موضوع داده‌ها و پرینت‌های خروجی را پیش کشیده بود. در عمل، این فرصت اغلب به‌وسیلهٔ آن دسته از مشتریان فراهم می‌شود که به‌طور قابل‌درکی سعی می‌کنند به‌سرعت بحث را از موضوعات قبلی (توصیف پروژه) به وضعیت جاری پروژه (تحلیل داده‌ها) بکشانند.

پیش از مطرح‌کردن پرسش‌هایی که باید با مشتری حل و فصل کنیم، خوب است که اشارهٔ مشتری به ANOVA را بررسی کنیم. این مطلب نشانه‌ای از دانش آماری او را در اختیار ما گذاشت: او آن‌قدر می‌دانست که از ما کمک بگیرد، اما ما باید در پاسخ‌دادنمان دقت کنیم. استقبال از پیشنهاد روش‌شناختی مشتری در این مرحلهٔ آغازین عموماً عاقلانه نیست: ممکن است نهایتاً کارمان به آنجا بکشد که به او توضیح دهیم که پیشنهاد «عالی» او مناسب نیست و هر دوی ما تا حدی احساس حماقت کنیم. در چنین وضعیتی، باید سعی کنیم که روحیه‌بخش به‌نظر برسیم، اما از ایجاد تعهد برای خود برای عمل به روش‌های آماری پیشنهادی مشتری خودداری ورزیم. اگر پیشنهاد او درست از کار در آید، احساس خوبی به او دست می‌دهد؛ در غیر این‌صورت، دلیل مراجعهٔ او به ما در وهلهٔ اول همین است. حال به پرسش‌های خود بازگردیم.

۱. چند اصطلاح از طرف مشتری مطرح شد که به توضیح بیشتری نیاز دارند: LSI، روش‌های سنتی، ترجیحات جنبشی و لمسی، نمرهٔ SDS.

۲. دریادنگهداری درازمدت چگونه با این بررسی ارتباط پیدا می‌کند و چگونه اندازه‌گیری‌شده است؟ به جزئیات مربوط به ابزارهای پیش‌آزمون و پس‌آزمون نیاز خواهیم داشت.

۳. در این مطالعه تا چه حدی عامل‌های جنسیت و سطح مدرسه مورد توجه‌اند؟ لازم است به‌تفصیل دربارهٔ اندازه‌های نمونه‌ای دخیل در این رسته‌ها اطلاع داشته باشیم.

۴. نگرانی اصلی ما آن است که زمینه‌چینی طرح معمول «کنترل» در مقابل «تیمار»، از روی توصیف مشتری بدیهی به‌نظر نمی‌رسد. افزون بر آن، اشاره به فرمت‌های جلسه، عامل طرح بالقوهٔ پیچیده‌ای را به تحلیل اضافه می‌کند. لازم است از مشتری بخواهیم که فرمت طراحی روش‌های اجرایی را که در این مطالعه به‌کار گرفته شده‌اند دقیقاً توصیف کند.

تعریف مسئله

پس از مطرح‌کردن این پرسش‌ها با مشتری، اطلاعات زیر را به‌دست آوردیم. در طول مدت این بحث، پرینت‌های خروجی را نیز که مشتری با خود آورده بود، بازبینی کردیم.

طرح نمونه مرکب است از معلمان مدارس متوسطه و ابتدایی که به‌طور تصادفی به دو گروه تفکیک شده بودند: کنترل و آزمایشی. فرمت آموزشی سنتی (کتاب‌های درسی) در مورد گروه کنترل مورد استفاده قرار گرفته بود. برای گروه آزمایشی، آموزش سنتی با فعالیت‌هایی که با سبک‌های آموزشی ترجیحی شرکت‌کنندگان مناسبت داشت همراه شده بود. هر دو گروه به چهار دورهٔ تحصیلی (زیرگروه) تفکیک شده بودند و در هر دورهٔ تحصیلی فرمت آموزش یکسانی ارائه شده بود.

اجرا ابزار سیاههٔ سبک آموزشی (LSI)[1] در وهلهٔ اول برای رده‌بندی ترجیحات شرکت‌کنندگان به‌کار رفته بود. با در نظرگرفتن مقاصد این مطالعه، تنها یک ترجیح به هر شرکت‌کننده بر مبنای بالاترین نمرهٔ LSI او که حاصل از رسته‌های شنیداری (A)، دیداری (V)، جنبشی (K)، یا لمسی (T) بود تخصیص داده شد، مشروط بر اینکه نمرهٔ LSI بیشتر از ۵۰ بود؛ در غیر این‌صورت چنین تلقی شد که شرکت‌کننده هیچ ترجیحی (N) ندارد.

در گروه آزمایشی، فعالیت‌های خاص صوتی (A)، بینایی (V)، اجرای نقش (K)، و ساختن (T) برای همهٔ شرکت کنندگان، در رابطه با ترجیح تخصیص داده‌شده به آن‌ها، به‌کار گرفته شد. بنابراین، N شرکت‌کنندهٔ با ترجیح در فعالیت‌های خاص مربوط به بالاترین نمرهٔ LSI آن‌ها در رسته‌های A، V، K، T نیز درگیر بودند.

متغیرها برای مقاصد ارزیابی، سه اندازهٔ کمّی در این مطالعه به‌کار گرفته شدند که مرکب بودند از: یک پیش‌آزمون (PRE) که در آغاز دورهٔ آموزش ارائه شد، یک نمرهٔ مقیاس نگرش بر مبنای ابزار مقیاس افتراق نحوی[2] (SDS) که در پایان دورهٔ آموزش داده شد، و یک پس‌آزمون (POST) که یک ماه پس از آموزش ارائه شد. دیگر عامل‌های ثبت‌شده عبارت بودند از جنسیت [GENDER]، دورهٔ تحصیلی [SESSION]، و سطح مدرسه [SLEVEL]. جزئیات مربوط به این متغیرها در زیر خلاصه شده‌اند.

اندازه‌های کمّی

PRE	پیش‌آزمون:	رتبهٔ ماکسیمم= ۱۰۰
POST	پس‌آزمون:	رتبهٔ ماکسیمم= ۱۰۰
SDS	نمرهٔ نگرش:	رتبهٔ ماکسیمم= ۶۰

1. learning style inventory 2. semantic differential scale

عامل‌های رسته‌ای

GROUP	کنترل / آزمایشی
GENDER	زن / مرد
SLEVEL	ابتدایی / دبیرستانی
SESSION	S۱، S۲، S۳، S۴ (درون هر GROUP)
PREF	شنیداری، جنبشی، هیچ‌کدام، لمسی، دیداری

اندازه‌های نمونه‌ای ($N = ۸۷$)

کنترل	۴۳	آزمایشی	۴۴
زن	۷۰	مرد	۱۷
ابتدایی	۵۹	دبیرستانی	۲۸
ترجیحات:	$A = ۱۲$، $K = ۴$، $N = ۳۲$، $T = ۲۳$، $V = ۷$		
دورهٔ تحصیلی:	غیر قابل دسترسی ــ باید توسط مشتری فراهم شود		

مباحث و هدف‌های کلی

تاکنون، سمت و سو و هدف جلسهٔ مشاوره عمدتاً به نفع ما بوده است. ما اجزای مهم مطالعه را مشخص، و ویژگی‌های متغیرهای دخیل در آن را معین کرده‌ایم. اکنون باید به هدف مطالعهٔ مشتری بازگردیم. بیانی رسمی از فرض پژوهش[۱] که مشتری آن را به‌عنوان هدف این مطالعه اقامه کرده است، در زیر داده می‌شود.

$H_۱$ آموزگاران در جلسه‌های مهارت‌آموزی فناوری که از یک فعالیت پردازشی بهره می‌برند که با ترجیحات سبک یادگیری ادراکی‌شان مطابقت دارد، به‌طور معنی‌داری دریادنگهداری بلندمدت محتوایی بیشتری را در قیاس با آموزگارانی در آموزش سنتی که از فعالیت پردازشی استفاده نمی‌کنند از خود بروز می‌دهند.

این فرض در اساس به فرض $H_۰ : \mu_{\text{Cont}} = \mu_{\text{Expt}}$ در قالب اصطلاحات ما تبدیل می‌شود. آزمون‌های t و شیوه‌های ANOVA را می‌توان برای بررسی این فرض به‌کار برد. وظیفهٔ نخست ما توضیح این روش‌های آماری به مشتری است.

۱. برای اختصار، توجه خود را به بحث دریادنگهداری بلندمدت (پیش [PRE] و پس [POST]) محدود می‌کنیم. فرض پژوهشی مشابهی نیز برای تفاوت‌های نگرشی (SDS) اقامه شده است.

آزمون‌های t این آزمون‌ها برای ارزیابی این امر به‌کار می‌روند که آیا تفاوت معنی‌داری بین میانگین‌های مرتبط یا دو نمونهٔ مستقل وجود دارد یا خیر. به‌عنوان مثال، ما به این موضوع علاقه‌مندیم که آیا گروه آموزشی نمرهٔ متوسطِ POST به‌طور معنی‌دار بالاتری در قیاس با گروه کنترل داشته یا خیر.

ANOVA برای آزمودن تأثیرات چندین عامل به‌طور همزمان، می‌توان از یک شیوهٔ تحلیل واریانس استفاده کرد. به‌عنوان مثال، می‌توانیم این فرض را آزمون کنیم که آیا GENDER، SLEVEL و SESSION پس از آنکه اثر GROUP به‌حساب آورده شد نیز، بر POST تأثیر دارند یا خیر. شاید دبیران دبیرستانی در درون هریک از گروه‌ها بهتر عمل کرده باشند.

ذکر این نکته ارزشمند است که «توضیح» ما در حقیقت صرفاً بیانی از این است که چگونه می‌توانیم این روش را به‌کار ببریم که به کمک مثالی در زمینهٔ پروژهٔ مشتری شرح داده می‌شود. به یاد بیاورید که عملاً هنوز تحلیل را انجام نداده‌ایم. مطابق تجربیات ما، مطرح‌کردن مفاهیم مجرد استنباط آماری در غیاب نتایج ملموس، چیزی بیش از بحثی طولانی اما توخالی را به وجود نمی‌آورد.

مثال‌های پرآب و رنگی را که برای تشریح روش‌های آماری به‌کار می‌گیریم باید با ارزیابی واقع‌بینانهٔ مفروضات و مسائل بالقوه‌ای که ممکن است اعتبار آماری تحلیل را تحت تأثیر قرار دهند، به‌حالت اعتدال درآورد. به‌ویژه باید تأکید کنیم که معنی‌داری مستلزم علیت نیست. این مطلب همواره به آن سادگی که به‌نظر می‌رسد نیست. پرداختن به موضوع علیت اغلب نشان‌دهندهٔ مرز بین آن چیزی که مشتری می‌خواهد نتیجه بگیرد در مقابل آن چیزی است که تحلیل آماری عملاً در اختیار خواهد گذاشت. برخی از موضوعاتی که باید به‌تفصیل با مشتری مورد بحث قرار دهیم عبارت‌اند از:

۱. با فرض نتیجه‌ای معنی‌دار از آزمون t برای POST برحسب GROUP، تفسیر ما (از نتیجه) چیست؟

۲. درصورتی که PRE برحسب GROUP معنی‌دار باشد، چگونه باید عمل کنیم؟

۳. عدم توازن در سطوح GENDER و PREF (اندازهٔ رده‌ها) ممکن است تأثیر نامطلوبی بر شیوهٔ ANOVA داشته باشد.

۴. اگر SESSION عامل معنی‌داری از کار درآید، چه اتفاقی می‌افتد؟

۵. آیا تخصیص یک ترجیحِ تنها واقع‌بینانه است؟

تبصره‌ها

آن‌گونه که از فرض پژوهشی مشتری آشکار است، دریادنگهداری درازمدت به‌عنوان برآمد وابسته به یک اثر معنی‌دار GROUP بیان می‌شود. حال، قسمت مشکل مطلب مطرح می‌شود. آیا باید

اصرار کنیم که مشتری از این ارتباط دادنِ دریادنگهداری درازمدت با علیت دست بکشد؟ در این مثال، ما این‌کار را نکردیم.

به‌طور قطع می‌توانیم استدلال کنیم که یک نتیجهٔ معنی‌دار آماری تنها شواهدی حاکی از یک پیوند بین اثر «تیمار» (روش‌های آموزش با بهره‌گیری از ترجیح‌های سبک یادگیری) و اجرای پس‌آزمون در اختیار می‌گذارد؛ این نتیجه ثابت نمی‌کند که دریادنگهداری درازمدتی وجود داشته است. اما دریادنگهداری درازمدت، تفسیر مضمونی مشتری از آن چیزی است که عملکرد پس‌آزمون دلالت بر آن می‌کند که با توجه به طرح و هدف آزمایش معقول به‌نظر می‌رسد. خواه لزوماً با این نوع تفسیر ذهنی موافق باشیم یا خیر، باید دقت کنیم که دانش و تخصص خود را بر زمینهٔ علمی مشتری تحمیل نکنیم. مسئولیت ما حصول اطمینان از این امر است که مشتری به‌روشنی درک می‌کند که شواهد آماری فقط شواهد آماری است و دلیل و برهان نیست. مسئولیت ماهیت ذهنی یک تفسیر مرتبط با زمینهٔ موضوع آشکارا بر عهدهٔ مشتری است.

از آنجا که این مطالعه پساآزمایشی است، تمرکز عمدهٔ ارزیابی ما از نکات آماری مربوط به پروژهٔ مشتری، در شناسایی منابع بالقوهٔ خطای غیر تصادفی است. این بدان معنی است که آیا می‌توانیم هر نوع اثر اریبی را، که ممکن است به‌طور جدی اعتبار آماری تحلیل را به مخاطره افکند پیدا کنیم؟ در این وضعیت، که در آن به توصیف مشتری از طرح و اجرای آزمایش متکی هستیم، استفاده از رویکرد بازجویی سودمند خواهد بود.

مشتری تعبیر خود از یک نتیجهٔ فرضی را در اختیار ما می‌گذارد

که ما برای او مطرح کرده‌ایم، از قبیل یک آزمون *t*ی معنی‌دار برای POST در مقابل GROUP، و سپس سناریوهای «اگر ... چطور؟» را مطرح می‌کنیم.

هدف این تمرین آن نیست که به مشتری «پشت‌پا» بزنیم ــ هیچ آزمایشی بدون ایراد نیست ــ بلکه هدف ما انجام‌دادن یک ارزیابی مستقل و عینی از مطالعهٔ انجام‌شده به‌وسیلهٔ مشتری است. در بسیاری از حالت‌ها، هم مشتری و هم مشاور درک بسیار بهتری از نکات حاکم بر تحقیق به‌دست می‌آورند: ممکن است به نکته‌ای انگشت بگذاریم که مسئله‌ای مربوط به زمینهٔ موضوع را برای مشتری حل می‌کند، و مشتری ممکن است جزئیاتی دربارهٔ مطالعه را به‌خاطر بیاورد که از لحاظ آماری مهم از کار درمی‌آیند.

پس هر دوی ما از این مثال چه آموختیم؟ موارد زیر به نکاتی مربوط می‌شود که با مشتری به‌تفصیل مورد بحث قرار داده‌ایم.

۱. برخی نکات دخیل در تفسیر نتیجه‌ای معنی‌دار برای آزمون *t*ی POST برحسب GROUP عبارت‌اند از:

- همان‌گونه که در بررسی‌های کوچک‌مقیاس معمول است، نمایهٔ نمونه‌ای عموماً میزان تعمیم نتیجه‌ای معنی‌دار به جامعهٔ بزرگتر را محدود می‌کند. بنابراین، ضروری است که نتایج ما مبتنی بر مقدار پیشنهادی مرتبط با یک مطالعهٔ «مقدماتی» باشد.
- دلالت بر دریادنگهداری درازمدت بر مبنای یک نتیجهٔ معنی‌دار POST در مقابل GROUP به‌روشنی محل بحث و جدل است. این امر، یک تفسیر مربوط به زمینهٔ موضوع از اثر GROUP است؛ نتیجهٔ آماری ثابت نمی‌کند که این حکم لزوماً درست است. یک نکتهٔ ثانویه آن است که آیا یک ماه واقعاً شواهدی از دریادنگهداری «درازمدت» به دست می‌دهد؟ هر دو مورد به داوری مشتری واگذار شد.
- این آزمون t دانش شرکت‌کنندگان را تنها در نقطهٔ زمانی یک‌ماهه ارزیابی می‌کند. آیا مشتری می‌داند که شرکت‌کنندگان طی ماه پیش چه کرده‌اند؟ به‌عنوان مثال، فرض کنید که شرکت‌کنندگان در گروه آزمایشی چنان از آن فعالیت‌های ترجیحی مسخره سرخورده شده باشند که همه در یک دورهٔ فناوری فشرده که از روش‌های آموزشی سنتی بهره می‌برده ثبت‌نام کرده باشند[۱]
- نتیجهٔ POST در مقابل GROUP سطوح متفاوت دانش (خطوط مبنا) را که ممکن است پیش از آموزش بین شرکت‌کنندگان موجود بوده باشد، به‌حساب نمی‌آورد. پیش‌آزمون برای در نظرگرفتن تفاوت‌های خط مبنای ممکن به‌کار گرفته می‌شود.

۲. یک نتیجهٔ معنی‌دار PRE برحسب GROUP ممکن‌است برای مشتری‌مشکل‌آفرین باشد. از این مطلب‌چنین برداشت‌می‌شود که طرح‌مشتری تا حدی‌معیوب است و تفسیرعملکردپس‌آزمون را پیچیده‌می‌کند. البته هیچ‌یک از این‌ها لزوماً درست‌نیستند، اما نامعقول‌نیست که‌مشتریان پیش‌آزمون‌ها را با «امید یارشدن‌بخت» به‌خدمت بگیرند: آن‌ها می‌دانندکه پیش‌آزمون‌هامهم‌اند؛ تنها امیدشان به این است که درمطالعهٔ آن‌ها این‌آزمون‌هامعنی‌دارنباشند! در این وضعیت، بایدوقت بیشتری را صرف مشتری کنیم تا جنبه‌های آماری زیر را برای او روشن کنیم.

- از تصادفی‌سازی برای اجتناب از اریبیِ گزینش استفاده می‌شود. به‌عنوان مثال، ما هرگز در یک قرعه‌کشی بخت‌آزمایی، $\{۱, ۲, ۳, ۴, ۵\}$ را «تصادفی» تلقی نمی‌کنیم، اما احتمال وقوع این مجموعه به‌قدر هر مجموعهٔ ۵ عددیِ دیگری است! طرح مشتری،

۱. به ما این اطمینان داده شد که این امر اتفاق نیفتاده است. نکتهٔ مهم این است که ضرورت دارد که مشاور در ایجاد سناریوهای «اگر ...چطور؟» ابتکار به‌خرج دهد.

صرفاً به این دلیل که دو گروه با عملکردهای پیش‌آزمون متفاوت به‌دست آورده‌اند، معیوب نیست، بلکه این نتیجه به‌طور شانسی حاصل شده است.

- عملکرد پس‌آزمون را می‌توان با به‌کارگیری یک شیوهٔ آزمون tی جفت شده روی $\mathrm{DIFF} = \mathrm{POST} - \mathrm{PRE}$ ارزیابی کرد؛ به این معنی که تفاضل بین نمره‌های PRE و POST یک شرکت‌کننده را در نظر گیریم. در واقع این‌کار معقول‌تر است بدان لحاظ که برای هر شرکت‌کننده بستگی مثبتی بین این نمره‌ها موجود است (یعنی، PRE بالا $\Longleftarrow$ POST بالا). آزمون tی جفت‌شده این اثر خطِ مبنا را به‌حساب می‌آورد.
- آزمون tی DIFF برحسب GROUP مزیت دیگری نیز دارد. اگر طیف گسترده‌ای از نمره‌های POST درون هر گروه موجود باشد، این امر ممکن است اثر GROUP را بپوشاند. با دخالت‌دادن خط مبنای شرکت‌کننده، یک اثر معنادار GROUP را می‌توان آسان‌تر تشخیص داد.

۳. نامتعادل‌بودن در سطوح GENDER و PREF (اندازهٔ رده‌ها) ممکن است بر شیوهٔ ANOVA اثر نامطلوبی داشته باشد. امکان دارد که نتایجی غیر واقعی به‌دست آوریم که صرفاً معلول اندازهٔ کوچک نمونه باشند. توجه کنید که تعداد K و Vی شرکت‌کننده‌ها در این تحلیل توسط GROUP باز هم خُرد خواهد شد. یک نتیجهٔ محتمل‌تر آن است که این عامل‌ها معنی‌دار نباشند. این لزوماً بدان معنی نیست که این عامل‌ها فاقد اهمیت‌اند؛ بلکه صرفاً در این تحقیق اطلاعات کافی در دست نداریم که حکم معنی‌داری را اعلام کنیم. برای عامل PREF، ممکن است ترکیب‌کردن برخی سطوح، ارزشمند باشد. مشتری $\mathrm{A}+\mathrm{V}$ و $\mathrm{K}+\mathrm{T}$ را توصیه کرد.

۴. گرچه مقصود از این مطالعه فراهم‌کردن فرصت‌های آموزشی یکسان در چهار دورهٔ تحصیلی درون هر گروه بود، قطعاً امکان آن هست که سطوح عملکرد بین SESSIONها به‌طور معنی‌داری متغیر باشند. این مسئله از نگرانی‌های عمدهٔ مشتری به‌حساب نمی‌آمد.

۵. دو موضوع مرتبط با تخصیص ترجیح تنها یک سبک آموزشی عبارت بودند از:

- برخی شرکت‌کنندگان ممکن بود چندترجیحی باشند. بنابراین، تصمیم به تخصیص یک ترجیح خاص به این شرکت‌کنندگان، قدری ذهنی‌بودن را در مطالعه دخالت داد.
- شرکت‌کنندگان بدون ترجیح را باید جداگانه مورد تحلیل قرار داد[۱].

۱. این بخش مهمی از پروژه بود که ما آن را به‌عنوان تمرین واگذار کرده‌ایم.

کارهای مشخص

وقت آن رسیده است که چیزها را جمع و جور کنیم. ما بخش عمدهٔ وقتی را که برای این مشاوره برنامه‌ریزی شده بود صرف کرده‌ایم و حالا نیاز داریم که دستورکار بعد از جلسه را برای مشتری تنظیم کنیم. این کار بسیار مهم است: مشتری باید همیشه در جریان آنچه در پایان هر جلسهٔ مشاوره به‌آن رسیده‌ایم قرارگیرد. در این مثال، در موقعیتی هستیم که به‌کارهای مشخصی در پروژهٔ مشتری بپردازیم:

- آنچه برای مشتری انجام خواهیم داد.
- از مشتری انتظار چه کاری داریم؟
- تعیین چارچوب زمانی و هزینه‌ها.

مشاور مشتری از بحث ما دربارهٔ موضوعات و اهداف کلی می‌داند که رویکرد ما به تحلیل‌آماری پروژه چیست. بنابراین، مقصود عمدهٔ این دستورکار، فراهم‌آوردن شرحی از گام‌های تحلیل، و مشخص‌کردن نیازمندی‌های خاص مرتبط با پروژه برای مشتری است. بر مبنای این اطلاعات، یک چارچوب زمانی واقع‌بینانه برای تکمیل کار تعیین می‌کنیم.

۱. پردازش داده‌ها: انتقال پایگاه داده‌ها و تحلیل وارسی خطا. اصلاحات باید به‌وسیلهٔ مشتری صورت گیرد.

۲. تحلیل اکتشافی داده‌ها: آماره‌های خلاصه برای همهٔ متغیرها در پایگاه داده‌ها، به‌طور کلی و برحسب GROUP.

۳. تحلیل آماری:

آموزن‌های t: PRE، POST، و DIFF برحسب GROUP.

ANOVA: PRE، POST، و DIFF با همهٔ عامل‌های تبیینی: GROUP، SESSION، PREF، GENDER و SLEVEL.

۴. گزارش کتبی: مشتمل بر توضیحی از روش‌های آماری به‌کاررفته (درصورت تقاضای مشتری) خواهد بود.

۵. نمودارهای باکیفیت برای ارائه‌کرد: نمودارهای میله‌ای و بافت‌نگارها. نمودار پراکنش (یا نمودارهای معادل) برای ارائهٔ نتایج آزمون t. انتخاب نهایی با مشتری خواهد بود.

مشتری در این مثال، مشتری پایگاه داده‌هایی را که در یک فرمت الکترونیکی مناسب بود (فایل اکسل روی لوح فشرده) در اختیار ما قرار داد. متن چاپ‌شدهٔ آن‌ها که قبلاً بررسی کرده

بودیم به میزان کافی گویا بود، اما حاوی کدهای SESSION برای شرکت‌کنندگان نبود. چون نمونه نسبتاً کوچک بود، رویکردی ساده آن بود که از مشتری بخواهیم فهرست کدهای SESSION را (با کد شناسایی شرکت‌کننده) به ما ای‌میل کند.

درحالت کلی، مشتری‌ها مسئول گردآوری داده‌ها و مرحلهٔ ورود داده‌های مطالعه هستند. همچنین لازم است که آن‌ها قادر به تصحیح خطاهای ممکنی باشند که طی تحلیل وارسی خطای پایگاه داده‌ها به توسط مشاور شناسایی می‌شود. برای یاری در این مرحلهٔ آغازین تحلیل، پایگاه داده‌ها باید به موارد زیر پایبند باشد.

نرم‌افزار صفحهٔ گسترده اکسل (دیسک) یا سند متنی (ای‌میل).

فرمت مستطیلی. هر سطر یک مشاهده است و در هر ستون (متغیرها) یک داده داریم.

مقادیر گمشده یک کد خاص (مثلاً «۹-») را باید به جای مقادیر گمشده به‌کار برد. در پایگاه داده‌ها نباید هیچ دادهٔ خالی داشته باشد.

از کد درآوردن: کلیدی که همهٔ کدهای به‌کار رفته در پایگاه داده‌ها را فهرست می‌کند.

زمان و هزینه‌ها دخالت‌دادن دانشجویان تحصیلات تکمیلی در پروژه‌ها یکی از هدف‌ها(و مزیت‌ها)ی برنامه‌های مشاورهٔ دانشگاهی است. این پروژهٔ خاص به‌طور ایده‌آلی مناسب دانشجویان تحصیلات تکمیلی است و نقش اصلی ما نظارت بر پیشرفت و ارائهٔ کمک درصورت لزوم بود. البته، لازم است دربارهٔ قصد خود از به‌کارگیری دانشجویان تحصیلات تکمیلی با مشتری بحث کنیم. برخی از مباحث عمده عبارت‌اند از:

۱. مسئول نهایی پروژه خود ما هستیم، نه دانشجو.

۲. لازم است ارزش در مقابل هزینه، از نظر زمان‌بندیِ انعطاف‌پذیرتری برای مشتری سبک و سنگین شود. دانشجو ممکن است تعهدات دیگری مانند درس، امتحان، تکلیف، و کار موظف تدریس داشته باشد که به اندازهٔ پروژه برای او اهمیت داشته باشد.

۳. برای دانشجویان اغلب گزارش کتبی دشوارترین کار است و ممکن است چندبار بازنویسی ضرورت پیدا کند. از مشتری نباید برای این تجربهٔ آموزشی هزینه‌ای وصول کرد!

۴. ایجاد نمودارهای با کیفیت ارائه‌کرد ممکن است بسیار زمان‌بر باشد و برخی از دانشجویان ممکن است (هنوز) مهارت کافی برای تهیهٔ نمودارهای باکیفیت در نرم‌افزارهایی آماری مانند SAS و S-PLUS را کسب نکرده باشند.

۵. محرمانگی یا محدودیت‌های دیگر مرتبط با پروژه به دانشجو نیز تسری می‌یابد.

خلاصهٔ جلسه

در این مثال، در موقعیتی قرار داشتیم که چیزها را جمع و جورکنیم و به کارهای مشخص در پروژه بپردازیم. این بدان معنی است که اساساً آنچه راکه لازم بود برای پروژهٔ مشتری انجام شود و نیز ضرورت هر نوع گفتگوی بیشتر را، پیش از آنکه ملاقات بعدی انجام شود، به‌طور غیرمستقیم ـ از طریق تلفن، فکس، یا ای‌میل ـ حل و فصل کرده‌ایم. ملاقات بعدی ما زمانی انجام می‌شود که نتایج و گزارش اولیه تکمیل شده باشد. این‌ها را باید پیش از آن ملاقات، که بر تفسیر نتایج تمرکز خواهد یافت، به مشتری بفرستیم.

آیا واقعاً همهٔ این‌ها را در یک جلسه انجام دادیم؟

پاسخ این سؤال این است که بله ... و خیر. در واقع امر، برای ورود به موضوعات این پروژه با جزئیاتی که در اینجا ارائه شد بیش از یک جلسه مشاوره ضروری بود. پس چگونه می‌توانیم ادعا کنیم که یک جلسه لازم بود؟ ... طی جلسه، ما یادداشت‌های جامعی برداشتیم. به این معنی که برای بازپردازی برخی جزئیات دربارهٔ پروژه، پس از آنکه جلسهٔ مشاوره به اتمام رسیده بود، به یادداشت‌های خود تکیه کردیم. به این ترتیب، بیان ما از جلسهٔ مشاوره نتیجهٔ ادغام این دو منبع اطلاعات بود:

- اطلاعاتی که مستقیماً طی تعامل مستقیم با مشتری دریافت کردیم؛
- جزئیاتی که از روی یادداشت‌هایمان بازپردازی کردیم.

مزیت اصلی این رویکرد آن است که به ما اجازه می‌دهد که دید همه‌جانبهٔ خوبی از پروژهٔ مشتری به‌دست آوریم. فایدهٔ بدیهی آن برای مشتری این است که تنها یک‌بار ضرورت دارد که جزئیات پروژهٔ خود را توصیف کند.

جلسه‌های دیگر

حتی در پروژه‌های کوچک‌مقیاس، شاید ممکن نباشد که موضوعات مرتبط با پروژهٔ مشتری را در یک جلسهٔ مشاوره حل و فصل کنیم. ممکن است هنوز بحث دربارهٔ موضوعات و اهداف کلی پروژه ناقص باشد. بنابراین، حساب وقت اهمیت دارد زیرا همواره باید سعی کنیم که وقت کافی در انتهای جلسهٔ مشاوره برای تنظیم دستورکاری برای کارهای آتی مشتری بیندوزیم. این کار نباید منحصر به برنامه‌ریزی جلسه‌ای دیگر با مشتری باشد؛ آن‌ها استحقاق این را در وهلهٔ اول دارند که بدانند از این جلسه چه دستاوردی داشته‌اند.

خلاصه‌کردن نکات اصلی پروژهٔ مشتری، یا بیان اجمالی رویکرد آماری که درپیش خواهیم گرفت، راه‌هایی ساده و مؤثر برای انتقال این احساس به مشتری است که جلسه پرحاصل بوده است. همچنین

در این وضعیت، باید سعی کنیم که «وظیفه»ای نیز برای مشتری در نظر گیریم که پیش از جلسهٔ بعدی مشاوره انجام دهد. گرچه این کار به ماهیت خاص پروژه بستگی دارد، چند مثال متداول عبارت‌اند از:

- شروع به ایجاد پایگاه داده‌ها (یا قالبی برای چنین پایگاهی)
- تهیهٔ نسخه‌هایی از مراجع مربوط که مشتری به آن‌ها ارجاع داده است
- اصلاح پرسشنامه یا تهیهٔ پیش‌نویسی از آن برای مرور
- نوشتن پیش‌نویس دستورکار برای یک آزمایش
- گزارش‌کردن موضوعات مربوط به استاد مشاور یا راهنما.

دانشجویان

گاهی مشتریان از اینکه دانشجویان در پروژهٔ آن‌ها دخالت داده شوند بیمناک‌اند. اگر رفتار مشتری حاکی از این مطلب باشد، بی‌درنگ برای اطمینان‌بخشی به آن‌ها در خصوص این وضعیت قدم بردارید. نگرانی عمدهٔ آن‌ها ــ کیفیت تحلیل ــ ممکن است به‌طور بالقوه منجر به احساس منفی‌تر آن‌ها شود: می‌خواهیم آن‌ها را از سر وا کنیم؛ علاقهٔ جدی به پروژهٔ آن‌ها نداریم، و مسائلی از این قبیل. واداشتن دانشجویان تحصیلات تکمیلی به ارائهٔ مطلب طی جلسهٔ مشاوره قطعاً دراین‌باره سودمند خواهد بود. دانشجو آشکارا از داشتن شانس مشارکت در «فرایند مشاوره» سود خواهد برد، و مشتری می‌فهمد که این دانشجو «کیست». خواه دانشجو حضور داشته باشد یا نه، هنوز نیاز به تأکید بر مسئولیت خود در تحلیل به مشتری داریم.

نمودارها

نمودارهای تشخیصی معقول معمولاً به‌عنوان بخشی از خروجی شیوه‌های آماری تهیه می‌شوند. اضافه‌کردن عنوان‌ها، برچسب‌ها، و دیگر توضیحات، ممکن است نیازمند تلاش‌های بیشتری باشد، اما هرگونه پیرایش جدی آن‌ها برای مقاصد ارائه‌کرد وقت‌گیر خواهد بود. این امر موضوعی واقعاً کلیدی است و تصمیم ما برای تولید نمودارهایی با کیفیت ارائه‌کرد برای مشتری مستلزم توجه دقیق است.

در این مثال، دانشجو فرصت به‌دست‌آوردن مقداری تجربه با S-PLUS را داشت و بنابراین هزینه‌ای که از مشتری دریافت کرده‌ایم، بازتاب‌دهندهٔ هزینه‌های (وقت) صرف‌شده در ایجاد نمودارهای مورد تقاضا نبود. با این حال، توانایی دریافت بهای زیر قیمت از مشتری لطفی نیست که همیشه استطاعت انجام‌دادن آن را داشته باشیم، و چند نکته که باید در رابطه با نمودارهای با کیفیت ارائه‌کرد در نظر بگیریم عبارت‌اند از:

- درجهٔ کیفیت مورد نیاز چیست؟
- نمودارها چقدر پیچیده‌اند؟
- آیا لازم است نمودارها رنگی باشند؟
- جنبه‌های زیبایی؟
- چند نمودار مورد نیاز است؟

خروجی استاندارد نرم‌افزارهای آماری با عنوان‌ها، برچسب‌ها، و دیگر توضیحات مناسب ممکن است واقعاً کافی باشد. در غیر این صورت، باید از یک متال برای تعیین درجهٔ کیفیت مورد نظر استفاده کرد. در این وضعیت، باید تأکید کنیم که نمودارهای ساده می‌توانند کاملاً مؤثر باشند و پیرایش آن‌ها ساده‌تر است. نمودارهای چندمتغیره شامل متغیرهای گروه‌بندی، خم‌های تراز، یا نمایش‌های سه‌بعدی به تلاش‌های بیشتری نیاز دارند تا «درست ازکار درآیند». توجه کنید که باید از مقیاس خاکستری برای نمودارهایی که محتملاً از آن‌ها نسخهٔ فتوکپی تهیه خواهد شد استفاده کرد؛ موارد رنگی نیازمند امکانات بیشتر است و آنچه که ما در صفحهٔ مانیتور می‌بینیم لزوماً موقع چاپ به آن شکل بازتولید نمی‌شود.

در صورتی که مشتری طرح یا زیبایی نمودارهای حاصل ازکار ما را نپسندد، چه اتفاقی می‌افتد؟ لازم است قادر به انتقال این مطلب به مشتری باشیم که، پیش از ایجاد نمودار، این نمودارها به چه شکل خواهند بود. در نهایت، ایجاد یک نمودار خاص (منحصر به فرد) زمان می‌برد. تعداد «فرمت‌ها»ی نموداری خاص را باید در حداقل نگاه داشت و از هریک برای بیش از یک نمودار استفاده کرد.

۴.۴ تهیهٔ مستندات

جلسهٔ مشاوره به پایان رسیده است. مشتری با شرحی ازکارهایی که باید انجام شود ما را ترک کرده است و ما یک قرارداد تقریبی برای هزینهٔ خدماتمان به مشتری داده‌ایم. همان‌طورکه در هر نوع مبادلهٔ خدمات بر اساس قراردادها معمول است، این توافق باید به‌صورت کتبی رسمیت یابد و هر دو طرف آن را امضا کنند. خدمات مقررشده را صرفاً بر اساس توافق شفاهی انجام ندهید. در غیر این‌صورت، خود را برای مواجهه با این احساس آماده کنید که تلاش زیادی کرده‌اید و سر وقت کارها را انجام داده‌اید ــ و چیزی دریافت نکرده‌اید! شرحی از یک قرارداد ساده در شکل ۱.۴ ارائه شده است.

باید تأکید کنیم که این نوع قرارداد در واقع صرفاً یک «توافق شرافتی» است و برخی جزئیات را باید در آن به‌تفصیل آورد (مثلاً مبلغ پرداختی، موعد پرداخت، و غیره). نباید از آن به‌عنوان جایگزینی برای یک قرارداد تعهدآور حقوقی مناسب استفاده کرد. چنین قراردادهایی زمانی ضرورت پیدا می‌کند که لازم باشد موضوعات محرمانگی، حقوق معنوی، و مسئولیت در قالب اصطلاحات حقوقی به‌دقت

برنامهٔ مشاورهٔ آماری
اطلاعات سربرگ ب. م. آ.

قرارداد

تاریخ
نام و آدرس مشاور
نام و آدرس مشتری

عنوان پروژه
توصیف پروژه

مشتری با خدمات و شرایطی که ب. م. آ. در پیوست خلاصه پروژه فراهم کرده است موافق است. پس از تکمیل خدمات ب. م. آ. مشتری متعهد است که دستمزد را بپردازد....

امضا

الف. مشتری ی. مشاور

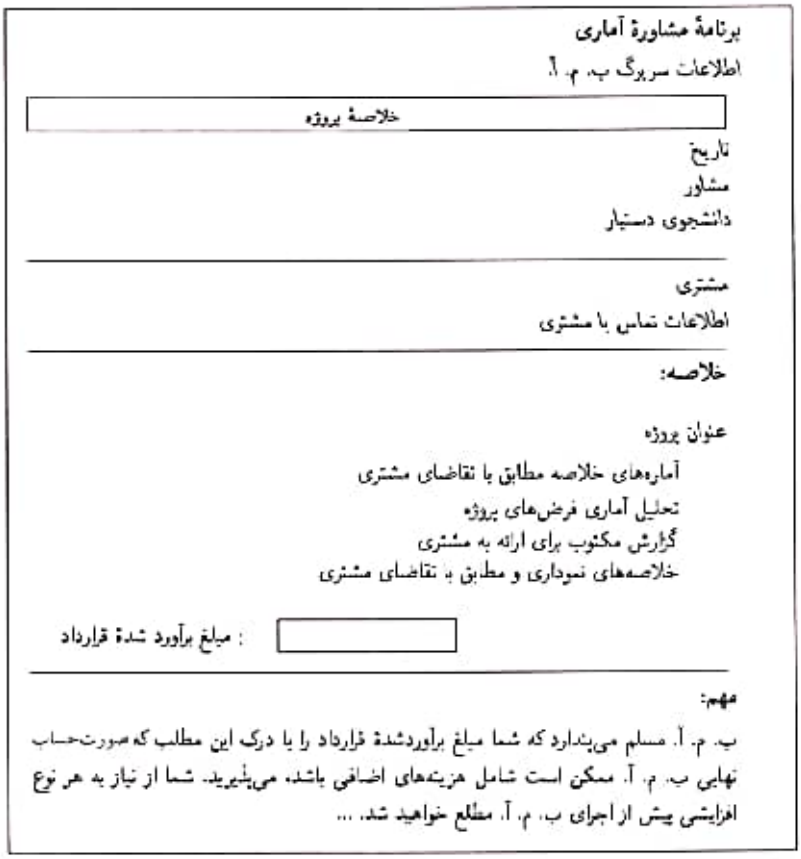
افزوده های ب. م. آ. م مشتری:

- تحلیل ب. م. آ. بر مبنای اطلاعات و پایگاه‌داده‌هایی است که مشتری در اختیار او می‌گذارد. تا آنجا که ب. م. آ. مطلع است، درستی پایگاه داده‌ها و اطلاعات داده‌شده توسط مشتری فاقد پیشداوری است.
- محاسبات آماری بر اساس قراردادی فرعی توسط یک دانشجوی تحصیلات تکمیلی آمار تحت نظارت مشاور ب. م. آ. انجام خواهد شد.
- ب. م. آ. خلاصه‌های نموداری با کیفیت ارائه‌کرد را در اختیار مشتری قرار خواهد داد.
- دیگر موارد ضروری.

شکل ۱.۴ طرح اجمالی نمونهٔ قرارداد.

مورد توجه قرارگیرند. شرکت‌های مشاوره‌ای کوچک ممکن است در عین حال بخواهند گروه مشتریان خود را، وقتی که قراردادهای فرعی پروژه‌های خود را به مشاوران مستقل واگذار می‌کنند، حفظ کنند.

شرح خلاصهٔ پروژه که در شکل ۲.۴ نشان داده شده از دو جهت به‌کار می‌آید: فهرستی از وظایف اصلی را که ما انجام خواهیم داد در اختیار مشتری می‌گذارد، و نوع اطلاعاتی را که می‌توانیم برای مقاصد ارجاع به‌کار بریم در اختیار ما قرار می‌دهد. این مستندات اطلاعات مربوط به مشتری و پروژه را که به آسانی از پایگاه داده‌های «پروندهٔ مشتری» ما قابل بازیابی است دربردارد. خلاصهٔ پروژه در


برنامهٔ مشاورهٔ آماری
اطلاعات سرپرگ ب. م. آ.

خلاصهٔ پروژه

تاریخ
مشاور
دانشجوی دستیار

مشتری
اطلاعات تماس با مشتری

خلاصه:

عنوان پروژه
آماره‌های خلاصه مطابق با تقاضای مشتری
تحلیل آماری فرض‌های پروژه
گزارش مکتوب برای ارائه به مشتری
خلاصه‌های نموداری و مطابق با تقاضای مشتری

: مبلغ برآورد شدهٔ قرارداد

مهم:
ب. م. آ. مسلم می‌پندارد که شما مبلغ برآوردشدهٔ قرارداد را با درک این مطلب که صورت‌حساب نهایی ب. م. آ. ممکن است شامل هزینه‌های اضافی باشد، می‌پذیرید. شما از نیاز به هر نوع افزایشی پیش از اجرای ب. م. آ. مطلع خواهید شد. ...

شکل ۲.۴ طرح اجمالی خلاصهٔ پروژه.

«پروندهٔ اصلی» مشتری نیز، که نهایتاً شامل همهٔ اسناد مرتبط با پروژه خواهد بود، قرار خواهد گرفت.

۵.۴ تحلیل پروژه

قرارداد امضا شد وکدهای ناموجود SESSION از مشتری دریافت شدند و یک نسخهٔ پشتیبان از پایگاه داده‌های اصلی تهیه شد. بخش اعظم تحلیل پروژه به دست دانشجوی تحصیلات تکمیلی انجام شد که مرکب از موارد زیر بود:

- پردازش داده‌ها

- تحلیل اکتشافی داده‌ها (EDA)
- تحلیل آماری
- نسخهٔ پیش‌نویس گزارش کتبی
- نمونه‌ای از نمودارها با کیفیت ارائه‌کرد.

پردازش داده‌ها

نخستین کار، دانلودکردن داده‌های مشتری از فرمت اکسل به چیزی سودمندتر برای مقاصد آماری بود. در این مثال، ما از بستهٔ نرم‌افزار آماری JMP استفاده کردیم که قادر به بارگذاری فایل‌های اکسل است.[۱] کدهای SESSION در پایگاه داده‌های JMP ادغام شد و یک نسخهٔ متن ساده (ASCII) خروجی‌ای برای استفادهٔ بعدی در SAS و S-PLUS بود.

آماره‌های خلاصهٔ عددی و نموداری که با JMP تهیه شد، حکایت از خطاهای آشکاری نمی‌کرد و با اطلاعات چاپ‌شده‌ای که طی جلسهٔ مشاوره وارسی کرده بودیم، تطابق داشت. بنابراین، می‌توانستیم مستقیماً به مرحلهٔ اکتشافی تحلیل پیش برویم.

تبصره‌ها

معمولاً این‌قدر خوش‌اقبال نیستیم. مشتری ما تصادفاً شخص تمام وکمالی بود و از اشتباهات رایجی که ممکن است در عمل پیش آید، پرهیز کرده بود (ن.ک. بخش ۲.۳، پردازش داده‌ها). ما همچنین پرینت داده‌های مشتری را دیده بودیم که ما را به انتخاب JMP ترغیب می‌کرد. Minitab و Statgraphics نمونه‌های دیگری از بسته‌های کاملاً جامع منویی هستند که می‌توانستیم به‌جای JMP از آن‌ها استفاده کنیم. پس چه چیزی ما را واداشت که JMP را انتخاب کنیم؟

۱. اندازهٔ نمونهٔ مطرح، نسبتاً کوچک بود.

۲. پایگاه داده‌ها در فرمت مستطیلی و بدون هیچ مقدار گمشده‌ای بود.

۳. واسط اشاره ــ و ــ کلیک‌کن و منویِ JMP اجرای تحلیل‌های اکتشافی و اولیه را بسیار آسان می‌سازد.

۴. نمودارهای تشخیصی استاندارد، مبنایی سودمند در اختیار مشتری می‌گذارد تا در مورد نمودارهای با کیفیت ارائه‌کرد تصمیم بگیرد.

---

۱. ویرایش ۴ JMP می‌تواند یک فایل اکسل را مستقیماً بارگذاری کند. JMP 3.x به نسخهٔ «ذخیره به‌صورت فایل متنی» فایل اکسل نیاز دارد.

آشکار است که ما در این مثال نسبتاً خوش‌اقبال بودیم. در یک وضعیت کلی چه باید کرد؟ جنبه‌های زیر را در نظر می‌گیریم.

متن ساده در اغلب وضعیت‌ها بهتر است پایگاه داده‌های مشتری را به فرمت متن ساده تبدیل کنیم و از SAS یا S-PLUS برای انجام داده‌پردازی استفاده کنیم. یکی از بزرگترین مزیت‌های فایل‌های متن ساده آن است که می‌توان آن‌ها را (تقریباً) در هر برنامهٔ کاربردی، صرف‌نظر از طرح زیربنایی پلت‌فرم به‌کار رفته، بارگذاری کرد. به‌ویژه، در صورت استفاده از پیوست‌های متن ساده در ای‌میل تضمین می‌شود که هر دو طرف قادر به خواندن اطلاعات خواهند بود. یک عیب آشکار آن این است که فرمت‌بندی خاص پایگاه داده‌های اصلی را از دست می‌دهیم. به این ترتیب، کارکردن مؤثر با فایل‌های متن ساده نیاز به ویرایشگر خوبی مانند Emacs ــ و البته دانش خوبی از طرز استفاده از آن ــ دارد.

پایگاه داده‌های مشتری در این وضعیت «نامحتمل»[1]، که مشتری فایل متن ساده در اختیار ما بگذارد، باید از مشتری بخواهیم که پایگاه داده‌های خود را با چه فرمتی به ما بدهد؟ اِکسل عموماً در دسترس همه است، قابل استفادهٔ مبتدیان است، و از دید ما، به آسانی قابل تبدیل به فرمت متن ساده است:

$$File \rightarrow Saveas... \rightarrow Text(Tab\ delimited)$$

اگر شما، یا مشتری شما، ترجیح می‌دهید از چیزی غیر از صفحه‌گستردهٔ اکسل استفاده کنید، اطمینان حاصل کنید که هر دو طرف برنامهٔ کاربردی مورد نظر را داشته باشند[2]. مترجم‌ها همیشه، به‌ویژه در کامپیوترهای مختلف، خوب کار نمی‌کنند.

کدبندی بسیاری از مشکلات داده‌پردازی از کدبندی نامناسب نشأت می‌گیرند. به‌عنوان مثال، پرسش‌های پیمایشی را، که به پاسخ‌دهنده این اجازه را می‌دهند که بیش از یک گزینه را انتخاب کند، باید با متغیرهای ظاهری جداگانه برای هر گزینه دوباره کدبندی کرد. به همین‌نحو، می‌توان از یک متغیر ظاهری برای کدبندی اظهارنظرهای تشریحی یا پاسخ‌های داده‌شده به «سایر موارد؛ توضیح دهید» به‌کار برد، زیرا این موارد معمولاً از لحاظ غیر آماری مورد توجه‌اند. در مطالعات موردی ارائه‌شده در قسمت II به این موضوع بازمی‌گردیم.

۱. مشتریان اغلب از صفحه گسترده‌هایی نظیر اِکسل استفاده می‌کنند، اما در واقع درک نمی‌کنند که مشاوران آماری عموماً اغلب ضروری می‌بینند که این فرمت‌بندی بسیار عالی را کنار بگذارند.

۲. یکی از بدترین تجربه‌های ما موردی بود که مشتری مقدار زیادی داده را در برنامه‌ای کاربردی وارد کرده بود که کاملاً از دور خارج شده بود. در نهایت توانستیم یک برنامه در C بنویسیم که پایگاه داده‌ها را مستقیماً از کد ASCII با پایهٔ ۸ استخراج کند!

تحلیل اکتشافی داده‌ها (EDA)

ابزارهای تشخیصی عددی و نموداری از JMP وجود دو مورد دادهٔ دورافتادهٔ بالقوه را در گروه کنترل برای POST (نمرات پایین) خاطرنشان کرد. با این‌حال، این شرکت‌کننده‌ها هر دو نمرات PRE پایینی داشتند و هیچ دادهٔ دورافتاده‌ای در متغیر حاصل، یعنی $\mathrm{DIFF} = \mathrm{POST} - \mathrm{PRE}$ حضور نداشت. برای متغیرهای رسته‌ای، PREF و GENDER بیشترین ناهمخوانی را در اندازه‌های رده‌ای خود نشان دادند. بنابراین، CPREF با ادغام رده‌های PREF، $\mathrm{A+V}$ و $\mathrm{K+T}$ ایجاد شد که $\mathrm{CPREF = N}$ ردهٔ سوم را تشکیل می‌داد.

ما از JMP برای انجام آزمون‌های t و برازش انواع مدل‌های ANOVA (ن.ک. تحلیل آماری در زیر) استفاده کردیم. غرض اصلی از این تمرین به‌دست‌آوردن مقداری بینش دربارهٔ نتایجی بود که با استفاده از SAS به‌تفصیل بررسی خواهد شد، و نیز برای ارزیابی مفروضات زمینه‌ای این شیوه‌های آماری. آن‌گونه که از شکل ۳.۲ در فصل ۲ دیده می‌شود، توزیع نمرات POST برحسب Group غیر نرمال‌بودن را آشکار می‌کند. با این‌حال، آزمون tی دونمونه‌ای، در برابر دورشدن از فرض‌های توزیعی، استوار است.

نتایج معنی‌دار نسبت به عامل GROUP تنها برای POST و DIFF به‌دست آمد. هیچ‌یک از عامل‌های دیگر در مدل‌های ANOVA (از جمله CPREF) معنی‌دار نبودند. نتایج JMP و نمودارهای تشخیصی برای بررسی در جلسهٔ آتی ما با مشتری چاپ و ذخیره شدند.

تبصره‌ها

در واقع می‌توانستیم تحلیل پروژه را تنها با استفاده از JMP کامل کنیم، چون نتایج بالا به‌طور اساسی پایهٔ گزارش کتبی را تشکیل می‌داد. با این حال، یک عیب سیستم‌های منویی مانند JMP آن است که بازتولید خروجی معمولاً مستلزم تکرار همهٔ عملیات انجام‌شده با واسطه‌های اشاره ــوــ کلیک‌کن است که قبلاً اجرا کرده بودیم. این کار برای مقاصد اکتشافی خوب است، اما وقتی بخواهیم تمامی تحلیل را تکرار کنیم، به‌سرعت نسبتاً کسل‌کننده و ناکارآمد می‌شود[۱].

برای گزارش‌نویسی، عیب جدی‌تر آن است که خروجی JMP را نمی‌توان به‌صورت فایل متن ساده ذخیره کرد. بنابراین، نمی‌توانیم بخش‌هایی از «فایل» خروجی را ویرایش، حذف، یا اصلاح کنیم تا برای منظور ما مناسب شود. همچنین نمی‌توانیم ای‌میلی (قابل خواندن) از خروجی JMP را به مشتری خود بفرستیم. با در نظرگرفتن همهٔ این‌ها چرا به فکر استفاده از JMP افتادیم؟ پاسخ صرفاً این است

۱. یک‌بار این اشتباه را کردیم که تنها نسخهٔ خود از خروجی JMP را به مشتری دادیم. متأسفانه کمی در توضیحات خود دربارهٔ نتایج اشتیاق به خرج دادیم و تذکرات سودمند متعددی برای آن نوشتیم. در پایان جلسهٔ مشاوره مشتری از ما درخواست نسخهٔ پاکیزه‌تری از خروجی را کرد!

که JMP ابزار مناسبی برای چیزی بود که می‌خواستیم در مرحلهٔ اکتشافی این تحلیل به‌دست آوریم.

۱. استفاده از آن آسان است (و دانشجویان جدید در مدت کمی آن را یاد می‌گیرند).

۲. آماره‌های خلاصهٔ تشخیصی و نمودارهای آن خوب‌اند.

۳. طیف جامعی از روش‌های آماری را در اختیار می‌گذارد.

نرم‌افزارهای کاربردی آماری دیگری نیز هستند که سه خواستهٔ بالا را برآورده می‌کنند و می‌توان از آن‌ها به‌جای JMP استفاده کرد: Minitab و Statgraphics دو موردی هستند که قبلاً از آن‌ها یاد کرده‌ایم. نکتهٔ اصلی آن است که مشاور آماری واقعاً نیاز به این دارد که به‌راحتی با بیش از یک برنامهٔ کاربردی کار کند. استفاده از نقاط قوت و مزایای نرم‌افزارهای مختلف به پرثمرترشدن فرایند مشاورهٔ آماری سود می‌رساند.

تحلیل آماری

آزمون‌های t و مدل‌های AVOVA که در بالا مطرح شدند با استفاده از SAS مجدداً اجرا شدند. البته نتایجی که قبلاً به‌دست آوردیم تغییری نکردند، اما حالا می‌شد خروجی SAS را در نسخهٔ اولیهٔ گزارش کتبی وارد کرد. کد SAS زیر مراحل اساسی تحلیل را نشان می‌دهد.

کد SAS برای تحلیل داده‌های مشتری

```
data a ;
  infile 'client.dat' ;
  input id grp $ session $ pref $ gender $ slevel $
        pre post sds ;
proc freq ;
  tables grp session pref gender slevel
         (session pref gender slevel)*grp ;
proc means ;
  var sds pre post ;
proc ttest ;
  class grp ;
  var sds pre post ;
proc glm ;
  class grp session pref gender slevel ;
  model sds pre post = grp session(grp) pref gender slevel;
```

گرچه در برنامهٔ SAS بالا دیده نمی‌شود، متغیرهای DIFF و CPREF ایجاد شدند و چندین صورت متفاوت از مدل‌های ANOVA مورد مطالعه قرار گرفتند. مانند تحلیل JMP، تنها عامل

معنی‌دار در مدل‌های POST و DIFF، عامل GROUP بود. مقایسهٔ میانگین‌ها و ابزارهای تشخیصی مانده‌ها پشتیبانی بیشتری از این نتیجه به‌عمل آوردند. با قانع شدن دربارهٔ اینکه تحلیل آماری دیگر کامل شده است، ما دو صورت از خروجی برنامهٔ SAS را به‌دست آوردیم:

ب.م.آ. [۱] صورتی برای خود ما که شامل طیف کاملی از آماره‌های خلاصه، امتحان‌های تشخیصی، و متغیرهای اضافی نظیر DIFF و CPREF در آزمون‌های t و مدل‌های ANOVA بود.

مشتری صورتی برای مشتری که فقط نتایج مربوط را شامل می‌شد. این نسخه ویرایش و در گزارش اولیه گنجانده شد.

تبصره‌ها

همان‌طورکه در فصل ۳ (بخش ۸.۳، نرم‌افزارهای آماری) خاطرنشان کردیم، فرض می‌کنیم که خواننده با دستورهای برنامه‌نویسی SAS و نیز قراردادهای استاندارد (نظیر نیاز به $ grp وقتی ستون grp شامل مقادیر کاراکتری است) آشنا باشد. گرچه با هر نوع فرض دلخواه خود دربارهٔ خواننده مشکلی برای ما پیش نمی‌آید، در مورد مشتریان وضعیت این‌گونه نیست؛ از آن‌ها نمی‌توان انتظار داشت که کد SAS را بفهمند! با این‌حال، از آن‌ها انتظار می‌رود که خروجی تولیدشده به‌وسیلهٔ کد SAS را بفهمند. برای استفادهٔ خوانندگان بی‌تکلف، به‌اختصار شرح می‌دهیم که این برنامهٔ SAS چه کاری انجام می‌دهد.

خلاصهٔ برنامهٔ SAS

`data a`؛ نُه ستون داده‌ها را از فایل «`client.dat`» می‌خواند.

`proc freq`؛ جدول‌های فراوانی را تولید می‌کند (۴ جدول دوطرفه، همه در مقابل grp).

`proc means`؛ میانگین‌ها، واریانس، و غیره را برای متغیرهای کمّی محاسبه می‌کند.

`proc ttest`؛ آزمون‌های tی SDS، PRE، و POST را برحسب GROUP انجام می‌دهد.

`proc glm`؛ مدل‌های ANOVA را برای SDS، PRE، و POST برازش می‌دهد[۲].

در این مثال، عامل SESSION درون GROUP «آشیانی» شده بود که در دستور مربوط به شیوهٔ مدل‌های خطی عمومی، یعنی `proc glm`، با `session(grp)` نشان داده می‌شود.

۱. برنامهٔ مشاورهٔ آماری (SCP).

۲. مطابق این نوشته، `proc glm`، یک تحلیل MANOVA را علاوه بر مدل‌های ANOVA (ی تک‌متغیره)، به‌صورتی که خواسته شده، اجرا می‌کند. گرچه ما تحلیل MANOVA را بررسی کردیم، آن را در نسخهٔ خروجی مشتری ویرایش کردیم.

این مطلب از این حقیقت برمی‌آید که دورهٔ تحصیلی S1 گروه کنترل با دورهٔ تحصیلی S1 در گروه آزمایشی ارتباطی ندارد. مقایسهٔ رده‌های مختلف دورهٔ تحصیلی در امتداد گروه‌ها آشکارا بی‌معنی است زیرا «برچسب»های دورهٔ تحصیلی به‌طور دلخواه انتخاب شده‌اند.

گزارش مقدماتی

قصد ما این است که پیش‌نویس گزارشی دربرگیرندهٔ نتیجه‌گیری‌های خود را بر مبنای نتایج گنجانده‌شده در نسخهٔ خروجی SAS مشتری به او ای‌میل کنیم. برای اجتناب از پیچیدگی‌ها و سردرگمی‌های بالقوهٔ مشتری، ما ویرایش‌هایی اضافی بر روی خروجی انجام دادیم. گرچه موارد زیر خاصِ SAS هستند، اما نشان می‌دهند که تلاش برای آسان‌کردن کارها به کمی جد و جهد نیاز دارد (ن.ک. بخش ۷.۴، گزارش نهایی، جدول‌های ۳ و ۴).

۱. آزمون F برای واریانس‌های همگن، که در زیر هر نتیجهٔ آزمون t داده شده بود، حذف شد.

۲. مجموع‌توان‌های دوم‌نوع I برای تأثیرات عامل‌های انفرادی در مدل‌های ANOVA برداشته‌شدند.

پس‌از آنکه نتایج ارائه‌شده در این خروجی را به سرانجام بردیم، وظیفهٔ کامل‌کردن گزارش و پرداختن به موضوعات زیر باقی مانده بود. این‌کارها به‌صورتی که در **گزارش نهایی** (بخش ۷.۴) نشان داده شدند، رفع و رجوع شدند.

- معنی‌داری. کدام نتایج معنی‌دار بودند و چرا؟ (Pـمقدارها)
- SESSION (GROUP) در خروجی ANOVA دیده می‌شود. دلیل استفاده از اثر آشیانی دورهٔ تحصیلی چه بوده است؟
- مجموع توان‌های دوم نوع III. چرا از این‌ها برای آزمون یک عامل انفرادی استفاده شد؟

پیش‌نویس گزارش کامل شد و همراه با تقاضایی برای برنامه‌ریزی ملاقات بعدی ای‌میل شد. مشتری به این تقاضا پاسخ داد و نیز اضافه کرد که آن‌ها سؤال‌هایی دربارهٔ Pـمقدارها و معنی‌داری دارند که مایل‌اند در جلسهٔ آتی مطرح کنند. پیش از این ملاقات، تنها وظیفهٔ باقیماندهٔ ما آماده‌کردن نمونه‌ای از نمودار با کیفیتِ ارائه‌کرد برای مشتری بود.

تبصره‌ها

چرا به خود زحمت ویرایش خروجی را دادیم؟ به دو دلیل:

۱. اگر مشتری چیزی را از خروجی نفهمد، باید آن را برای او توضیح دهیم.

۲. در هر صورت لازم بود که این کار را برای گزارش نهایی انجام دهیم.

برای لحظه‌ای به سراغ دلیل اول می‌رویم و فرض می‌کنیم که اطلاعات را ویرایش نکرده بودیم. در این صورت توضیح ما چه می‌بود؟ «نگران آن نباش. اهمیتی ندارد!» —[مشتری] «پس چه نیازی هست که آن را داشته باشیم؟»... شاید می‌توانستیم پاسخ اول خود را دوباره مطرح کنیم. نکته این است که مشتری‌ها نگران می‌شوند. آن‌ها راهی برای تشخیص نامربوط‌بودن یک مطلب ندارند و این همان دلیل مراجعهٔ آن‌ها به ماست.

نمودارهای با کیفیت ارائه‌کرد

ما از S-PLUS برای تولید نمودارهایی با کیفیتِ ارائه‌کرد برای مشتری استفاده کردیم. گرچه امیدوار بودیم که بتوانیم نسخهٔ PostScript نمودار را به مشتری بفرستیم، آن‌ها منابع لازم (چاپگر PostScript) برای چاپ فایل را نداشتند. کد S-PLUS زیر شرحی بسیار اجمالی از نحوهٔ تولید شکل ۳.۲ در فصل ۲ را نشان می‌دهد.

کد S-PLUS برای تولید شکل ۲.۳

```
yy <- read.table("client.dat")    # read in data to S-PLUS
y  <- yy$post                     # extract the POST scores
ys <- yy$grp                      # extract the GROUP codes
g.hist2.fun(y,ys,...)          # call our customized graphics
                               # function (options not shown)
```

طرح اجمالی تابع گرافیکی مشتری‌مدار

```
g.hist2.fun <- function(y,ys,signif=T,...)
{
  yb <- split(y,ys)                # split POST by GROUP
  xm <- c(mean(yb[[1]]),           # get POST means by GROUP
          mean(yb[[2]]))

  par(new=F,mar=c(6,4,3,4)+0.1)  # Figure Region Margins

  # ------------------------ This is the key step:
                                   # Create 2 sub-Figure
  fg <- list(c(0,1,0.4,1),         # Regions. One takes 60%
             c(0,1,0,0.6))         # of the top; the other
                                   # takes 60% of the bottom
```

```
  for(i in 1:2){                  # Loop through the two
       par(new=T,fig=fg[[i]])     # histgram plots, using
       hist(yb[[i]], ... )        # the sub-Figure Regions.
       legend( ... )              # Add Legend (indexed) and
       abline(v=xm[i])            # put vertical line at mean
  }
                                  # If significant, put this
  if(signif)                      # text on plot (position
      text( ... ,                 # arguments not shown)
           "** Means Differ Significantly ** ")

  polygon( ... )                  # Draw the bowtie polygon

  box()                           # Box around everything
  title( ... )                    # Title and subtitles

  invisible()                     # Make plot invisible ??
}
```

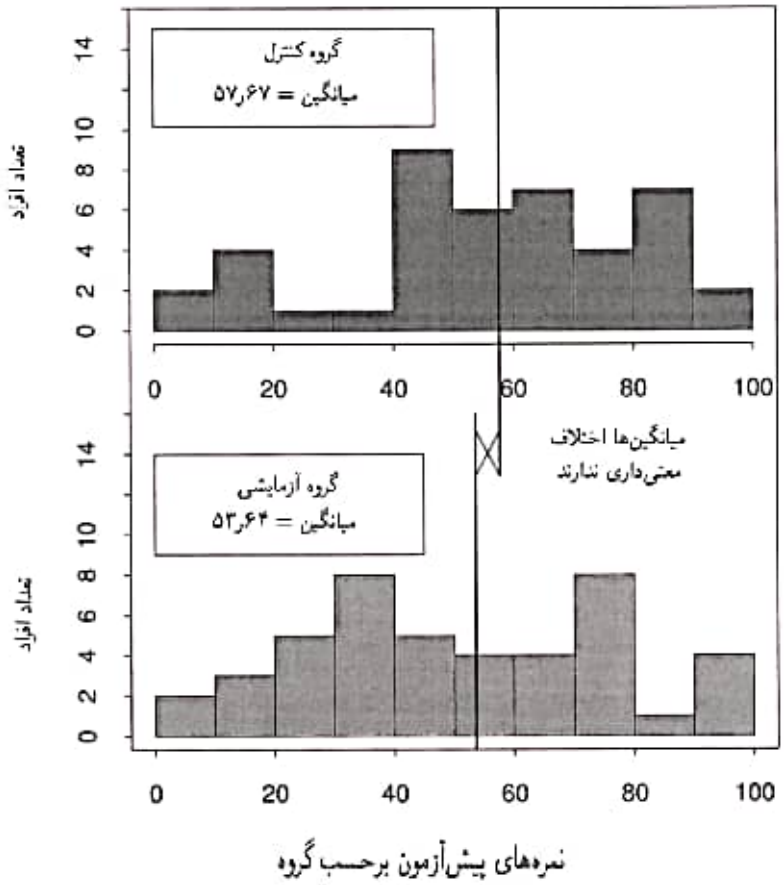
این تابع برای ایجاد نمودارهای بافت‌نگار برای نتایج آزمون t‌ی غیر معنی‌دار PRE و SDS برحسب GROUP نیز به‌کار رفت. با قراردادن signif=F در گزینه‌های سطر g.hist2.fun(yy pre,ys,signif=F,...) می‌توان از مرحلهٔ text() در موقع ایجاد نمودار PRE برحسب GROUP صرف‌نظر کرد. گزینهٔ پیش‌فرض این بود که متن به صورتی چاپ شود که در شکل ۳.۴ نشان داده شده است. البته گزینه‌ها و شناسه‌های دیگر بسیاری برای تولید این نمودار مورد نیاز بود، اما مراحل اصلی در بالا نشان داده شده‌اند. توجه کنید که دستور invisible() واقعاً نمودار را نامرئی نمی‌کند. S-PLUS خوش ندارد که تابعی مقداری را به‌عنوان مایه‌ازا برنگرداند. invisible() یک مقدار «هیچ‌چیز» خاص را به‌عنوان مایه‌ازا بازپس می‌دهد که از ناخشنودی آن جلوگیری می‌کند.

از آنجا که نمودارهای PRE و POST برحسب GROUP راه بسیار مؤثری برای نشان‌دادن تصویری نتایج آزمون t برای مشتری در اختیار ما می‌گذارند، ما شکل ۳.۴ را نیز برای جلسهٔ مشاوره آتی خود تولید کردیم.

۶.۴ نشست II: ارائهٔ نتایج

مشتری سر وقت برای حضور در جلسهٔ مشاوره رسید، در حالی که گزارش اولیه را در دست داشت و آمادهٔ طرح سؤال بود...(جایی در میانهٔ راه، ما خوش و بش خود را انجام دادیم!)

مشتری: من گزارش را به‌طور کامل خوانده‌ام و متأسفانه خیلی سؤال دارم که از شما بپرسم...



شکل ۳.۴ نمایش آزمون *t*ی غیرمعنی‌دار.

مشاور: (بدون مکث، فوراً حرف‌هایش را قطع می‌کند) واقعاً؟ شاید یک گزارش اشتباهی را برای شما فرستاده باشم! جداً این‌طور نیست! اگر سؤال‌های بسیاری نمی‌داشتید، در آن صورت برای من مایهٔ تعجب بود. بنابراین، کاری که به‌نظرم می‌توانیم انجام دهیم، آن است که ابتدا مقداری از خروجی را که **دانشجوی دستیار** تهیه کرده به شما نشان دهیم و ببینیم که آیا، همچنان که پیش می‌رویم، برخی از پرسش‌های شما را پاسخ می‌دهد یا خیر. بعداً به گزارشی که با خود آورده‌اید نگاه می‌کنیم. چطور است؟

مشتری: بسیار خوب... باشد. اما آیا می‌توانم قبل از شروع فقط یک سؤال بپرسم؟

مشاور: البته.

مشتری: بخش‌های «برابر/ نابرابر» در جدول ۳ حسابی گیجم کرده‌اند... جدول‌های ANOVA هم همین‌طور، نمی‌دانم باید به کدام Pـمقدارها نگاه کنم؟ و ...

مشاور: (حرف مشتری را قطع می‌کند) ...اُو، سؤالی طولانی است؛ من فکر می‌کردم که شاید آن بخش‌ها را کمی گیج‌کننده بیابید. اما اگر اول با بخش‌های ساده‌تر شروع کنیم، نتایج آزمون t و ANOVA به آن بدی که حالا به‌نظر می‌رسند، نخواهند بود. پیش از آنکه شروع کنیم، سؤال دیگری هم دارید؟

مشتری: خوب، هوم ... نه.

آه، دعوت مشتری به آرامش و به تعویق‌انداختن سؤال‌هایش در شروع جلسه، مشکل‌ترین بخش (تفسیر) جلسهٔ مشاوره است. باید به‌طور معقولی در تنظیم دستور جلسهٔ ملاقات محکم عمل کنیم، اما زور نگوییم. بنابراین، باید به مشتری اجازه دهیم که «فقط یک ...» سؤال اجتناب‌ناپذیر خود را بپرسد. پاسخ ما به این سؤال باید دوباره نوعی تکرار دستور جلسه باشد. مشتریان معمولاً تنها به یک دور برخورد از این نوع نیاز دارند تا تشخیص دهند ما دربارهٔ موارد مربوط به دستور جلسه مصمم هستیم. افزودن کمی بذله‌گویی (اگر شرایط اجازه دهند ــ این‌کار را از روی اجبار نکنید) می‌تواند به کاهش تلخی سماجتی که به‌خرج می‌دهیم کمک کند.

دستور جلسه

دستور جلسهٔ ما برای این جلسه نسبتاً سرراست است: نتایج را ارائه می‌کنیم و به مشتری معنی آن‌ها را می‌گوییم. بسیارخوب ...آری و نه. دستور جلسه مطمئناً بدیهی است، اما ارائهٔ نتایج به‌گونه‌ای که مشتری واقعاً آن‌ها را بفهمد، نیاز به آمادگی خوبی دارد. صرف اینکه به مشتری بگوییم نتیجه‌ای معنی‌دار است، به بالابردن درک مشتری از تحلیل، کمکی نمی‌کند. آن‌ها به احتمال فراوان یکی دو روز بعد با ما تماس می‌گیرند و از ما می‌پرسند که چرا فلان یا بهمان نتیجه معنی‌دار شد. نتیجه: یک جلسهٔ مشاورهٔ هدررفته هم برای ما و هم مشتری.

پس چگونه باید خود را آماده کنیم؟ همان‌گونه که می‌توان انتظار داشت، توضیحات بصری بهترین سازوکار را برای سهولت فهم در اختیار می‌گذارند. هرجا که ممکن باشد، در خروجی آماری از نمودار به‌جای جدول استفاده کنید. مزیت نمودار این است که حتی نمایش‌های پیچیده (مانند نمودارهای موزاییکی) را می‌توان بسیار آسان‌تر از جدول‌ها تعبیر و تفسیر کرد. آشکارسازی الگوهای

شاخص در یک جدول نیاز به تجربه دارد؛ چیزی که روشن است آن است که وقت کافی برای ایجاد آن در یک مشتری که تازه وارد بازی شده است نداریم.

نمودارها عالی‌اند، اما تنها بخشی از ماجرا محسوب می‌شوند. به‌عنوان مثال، تا حدی دشوار است که جدول‌های ANOVA را به‌صورت نموداری به‌نمایش درآورد. در مقاطعی لازم داریم به جدول‌ها اشاره کنیم و بنابراین بخشی مهم از آمادگی ما تهیهٔ ارائه‌کردی منطقی و منسجم است. در این مثال، هدف عمدهٔ دستور کار ما پیش‌روی آرام از نمایش‌های نموداری به خروجی عددی بود. بنابراین، دستور جلسه عبارت است از دنبالهٔ

۱. آماره‌های خلاصهٔ تک‌متغیره. نمودارهای میله‌ای و بافت‌نگارها.

۲. آماره‌های دومتغیره. جدول فراوانی دوطرفه.

۳. آزمون‌های t. نمودارهای با کیفیت ارائه‌کرد که برای PRE و POST برحسب نتایج GROUP تولید کرده‌ایم.

۴. آزمون‌های t. خروجی SAS که با ارجاع به این نمودارها توضیح داده می‌شوند.

۵. جدول‌های ANOVA. بدون نمودار.

از آنجا که دانشجو روی پروژه کار کرده بود، این کار فرصت خوبی برای دانشجو بود که در رابطه با ارائه‌کرد تجربه کسب کند و مستقیماً با مشتری وارد تعامل شود. در این‌حالت، لطف خاص کار این نیز هست که توجه مشتری را برای مدتی از ما دور خواهد کرد، با توجه به اینکه ما پوده‌ایم که آن‌ها را به پذیرش دستور جلسه برای ملاقات واداشته‌ایم.

ارائه‌کرد دانشجو

دانشجو کار را با نشان‌دادن آماره‌های خلاصهٔ تک‌متغیره به مشتری آغاز کرد: نمودارهای میله‌ای، بافت‌نگارها، و جدول‌های فراوانی. آن‌گونه که انتظار می‌رفت، مشتری هیچ مشکلی در دنبال‌کردن این بخش از تحلیل نداشت و هدف اصلی این بود که بحث را دربارهٔ اینکه کدام نمودارها از نظر مشتری باید به نمودارهایی با کیفیت ارائه‌کرد تبدیل شوند شروع کنیم.

دانشجو لازم می‌دید که به‌طور مختصر دربارهٔ سطر و ستون خانه‌ها و درصدهای کل در جدول‌های دوطرفه توضیح دهد، اما به‌نظر مشتری این فرمت بیش از حد نیاز آن‌ها بود. ما ادغام GROUP در قالب جدول‌های یک‌طرفه را به او پیشنهاد کردیم (ن.ک. جدول ۱، صفحهٔ ۲۰۹). عالی بود! نظر مشتری تفکیک نمودارهای میله‌ای برحسب GROUP برای متغیرهای PREF و GENDER بود (نمودار میله‌ای PREF برحسب GROUP در شکل ۱.۳، فصل ۳، نشان داده شده است).

تنها مشکل ما با خروجی میانگین‌ها تصمیم دراین‌باره بود که چگونه متوسط نمرات را برای SDS، PRE، و POST برحسب GROUP به‌صورت نموداری ارائه کنیم. ما سه نمودار میله‌ای جداگانه را برای هر متغیر، با دو میله روی هر نمودار، پیشنهاد کردیم که ارتفاع میله‌ها متناظر با میانگین نمرات از روی دو گروه بود. مشتری با این پیشنهاد برای SDS موافقت کرد (چون برای آن مقیاس دیگری در نظر گرفته شده بود)، اما به امکان ادغام‌کردن PRE و POST علاقه نشان می‌داد. این کار قطعاً مقدور بود (و همان «فرمت نمودار»های میله‌ای PREF و GENDER برحسب GROUP مورد اشاره در بالا را داشت)، اما از این واهمه داشتیم که میله‌های POST به‌لحاظ تصویری چنین نموداری را تحت‌الشعاع قرار دهند؛ بدین معنی که ممکن بود اثر اختلاف غیر معنی‌دار مهم بین میله‌های PRE از بین برود. مشتری این نکته را دریافت، اما کاملاً متقاعد نشد. بنابراین، هر دو صورت را تولید کردیم.

نتایج آزمون t

در این مرحله، مشتری کاملاً خونسرد بود و به نظر می‌رسید هیچ مشکلی در درک ارائه‌کرد دانشجو ندارد. پس از آن نمودارهای بافت‌نگار نتایج آزمون t را برای POST برحسب GROUP (شکل ۳.۲، فصل ۲) و PRE برحسب GROUP (شکل ۳.۴) را مطرح کردیم. مشتری از اوضاع خشنود بود و هنگامی که نتایج آزمون tی PRE و POST در جدول ۳، صفحهٔ ۲۱۱ را در رابطه با این دو نمودار به بحث گذاشتیم، به نظر می‌رسید دیگر نگران نبود.

مفهوم Pـمقدار در ابتدا برحسب میزان محتمل‌بودن این امر مورد بحث واقع شد که، به فرض آنکه واقعاً هیچ مزیتی در دخالت‌دادن ترجیح‌های سبک آموزشی وجود نداشته باشد، آزمایش یک پاپیون (اختلاف بین میانگین‌ها) تولید می‌کرد که در بافت‌نگارها مشاهده شده بود. نتیجهٔ POST به‌طور آشکاری در تأیید این فرض به‌نظر نمی‌آمد، در حالی که نتیجهٔ PRE چنین بود. مقدار متداول ۵٪ (Pـمقدار< ۰۵ر۰) به‌عنوان معیار استانداردی برای ارزیابی معنی‌داری معرفی شد. در حالی که وجود این «عدد جادویی» برای ارزیابی معنی‌داری مطمئناً برای مشتری تازگی نداشت، تعبیر و تفسیر نتایج PRE و POST در قالب اصطلاحات آماری آشکارا بخش مشکل کار بود. بنابراین، ما به‌آرامی و به‌دقت راه خود را از نتیجهٔ «آماری»، که در مورد هر یک از این نتایج به‌کار رفته بود، طی کردیم. برای امتحان اینکه در این کار توفیق داشته‌ایم یا نه، از مشتری خواستیم که نتیجهٔ SDS را تفسیر کند. مشتری در تفسیر مرتبط با زمینهٔ پروژه به‌طرز شگفت‌آوری خوب بود تنها به کمی تلنگر نیاز داشت تا بررسی کند که آیا انحراف استانداردها قابل قیاس‌اند یا خیر.

نتایج ANOVA

به نظر می‌آمد که گذار آرام ولی پیگیرانه از خروجی نموداری به خروجی عددی کار خود را کرده است. مشتری حالا برای نتایج ANOVA (جدول ۴، صفحه ۲۱۲) آماده بود. ما در ابتدا موضوع عامل آشیانی‌بودن SESSION را در مدل

```
POST = GROUP + SESSION(GROUP) + PREF + GENDER + SLEVEL
```

مورد بحث قرار دادیم. مشتری موافق بود که دورهٔ تحصیلی S1 گروه کنترل هیچ ارتباطی با S1 گروه آزمایش ندارد، بنابراین، آن اختلاف‌های دورهٔ تحصیلی تنها در درون گروه‌ها مدخلیت داشتند. این امر حضور SESSION(GROUP) را در خروجی ANOVA توضیح می‌داد. گام بعدی مطرح‌کردن ایدهٔ ANOVA به‌صورت یک فرایند دو مرحله‌ای بود.

۱. آیا هیچ‌یک از عامل‌ها اثری دارند؟ (آزمون F مدل کلی).

۲. اگر چنین است، کدام‌یک از آن‌ها؟ (آزمون‌های F نوع III).

گام ۱ به کمک نتایج PRE و SDS تحکیم شد؛ لزومی نداشت از این فراتر رویم. با این‌کار، نتیجهٔ POST باقی می‌ماند. مدل معنی‌دار بود (Pـمقدار= ۰٫۰۲۷۰)؛ بنابراین معقول بود که آزمون‌های F نوع III را بازبینی کنیم. ما صرفاً به مشتری گفتیم که این به‌اصطلاح آزمون‌های F «نوع III» به دو دلیل ضرورت دارند: هیچ‌یک از عامل‌ها تعداد برابری شرکت‌کننده در رده‌های مربوط به خود نداشتند (طرح نامتعادل بود)؛ و می‌خواستیم ببینیم که آیا یک عامل حتی پس از اقامهٔ دلیل دیگر عامل‌ها (مجموع‌های توان‌های دوم جزئی) همچنان مهم بود یا خیر. ما نسبت به اینکه مشتری واقعاً دلیل دوم را درک کرده باشد تردید داشتیم، اما آن‌ها هیچ مشکلی در تشخیص اینکه GROUP تنها عامل معنی‌دار در مدل POST بود نداشتند.

نتایج

مشتری تا حدی از اینکه عامل PREF معنی‌دار از کار در نیامد، اظهار یأس می‌کرد (که بر تردید ما در بالا افزود). ما خود خاطرنشان کردیم که نبودن یک اثر PREF معنی‌دار (یا CPREF) لزوماً به این معنی نبود که این عامل بی‌اهمیت است، بلکه به این معنی است که، از آزمایش مشتری، شواهد برای تأیید این نتیجه در رابطه با ترجیح‌های فردی ناکافی است. از آنجا که GROUP معنی‌دار بود، نتیجه آن شد که ما به‌طور معقولی می‌توانستیم چنین استنتاج کنیم که دخالت‌دادن سبک‌های آموزشی بر نمرات POST تأثیری داشت. اینکه کدام یک از ترجیح‌های خاص مهم‌تر بودند، با این بررسی مشخص نمی‌شد. این‌کار ظاهراً به مشتری کمک کرد و از یأس آن‌ها دربارهٔ PREF کاست.

ما سعی کردیم (با ملایمت) دربارهٔ موضوعات مطرح با پیوند دادن نتیجهٔ POST، با دریادنگهداری درازمدت، چیزهایی به مشتری یادآوری کنیم، اما این بحث زودگذری بود. تعبیر مشتری از عملکرد پس‌آزمون برحسب دریادنگهداری بلندمدت بود. در عوض، ما بحث را به محدودیت‌های استنباط و میزان قابل تعمیم بودن نتایج ما به نمایهٔ جامعه‌ای بزرگتر برگرداندیم. مشتری تصدیق کرد که نمونهٔ آن‌ها را نباید نمایندهٔ جامعهٔ عام معلمان ابتدایی و دبیرستان تلقی کرد، و این مطلب منطبق بر آن پیشنهاد ما بود که آزمایش مشتری را باید به‌عنوان یک بررسی مقدماتی در نظر گرفت.

خاتمه

مشتری از آنچه در این جلسه حاصل شد، بسیار خرسند بود و درخواست دیگری کرد: «آیا می‌توانیم روش‌ها و نتایج آن‌ها را در پیش‌نویس فصل (مربوط به پایان‌نامهٔ آن‌ها) مرور کنیم؟» لازم بود که کار نوشتن را خود مشتری انجام دهد. ما موافقت کردیم که جنبه‌های آماری[1] آن فصل را بازبینی کنیم. جلسه را با بحث مختصر و جمع‌بندی آنچه از پروژه باقی مانده بود به‌پایان بردیم.

- تکرار رسم نمودارهای خاصی که باید در فرمت کیفیت ارائه‌کرد تولید می‌شدند.
- بازنگری گزارش اولیه برحسب درخواست‌های مشتری.
- ارسال گزارش نهایی، صورت‌حساب، و نمودارها به مشتری.
- چگونگی پرداخت.
- بازبینی مستمر فصل روش‌ها و نتایج فصل مشتری.

تبصره‌ها

و به این ترتیب پروژهٔ دیگری به پایان خود نزدیک می‌شود. کار هنوز هم ضعف‌هایی دارد که باید برطرف شوند، اما دیگر تماس رودررویی با مشتری نخواهیم داشت. بنابر تجارب خود، این مرحلهٔ خاص از فرایند مشاوره جایی است که به احتمال زیاد «احساس رهاگشتگی»[2] شروع می‌شود (اصطلاح دقیق پزشکی احتمالاً کلمهٔ سندروم را هم به نوعی در بر دارد، اما برای آنچه در نظر داریم همین کفایت می‌کند). اشارهٔ ما در اینجا به چیست؟ نباید از این تلاش‌های خود احساس رضایت کنیم، و شاید هم به تلاش‌های خود ببالیم؟ البته. اما کار به پایان رسیده، و گاهی «بریدن»

۱. تأکید بر «آماری» این مطلب را روشن می‌کند که ما فقط این جنبه را بازبینی خواهیم کرد. ما فصل پایان‌نامهٔ مشتری را بازنویسی نخواهیم کرد.

۲. let-down effect، بیماری که پس‌از پایان دوره‌ای پر از تنش و کار سخت و بعد از فراغت از آن در شخص ظهور می‌کند.م

از این واقعیت شاید دشوار باشد. در این وضعیت، راهبردی خوب این است که سعی کنیم خود را از آن پروژهٔ خاص، مثلاً با کار روی چیزی کاملاً متفاوت، یا صرفاً مقداری دست‌کشیدن از کار، جدا کنیم. از سوی دیگر، «احساس رهاگشتگی» (در صورت وجود) می‌تواند تجربهٔ بسیار مثبتی هم باشد: سرانجام، کار را تقریباً به پایان برده‌ایم.

نکتهٔ اصلی که می‌خواهیم بر آن تأکید کنیم، آن است که تعامل با مشتری مقدار زیادی انرژی هیجانی ما را مصرف می‌کند. اینکه پس از رفتن مشتری چقدر تحلیل رفته‌ایم، آشکارا به چیزهای بسیاری بستگی دارد و ممکن است اصلاً هیچ آثاری باقی نمانده باشد. فقط اگر فشاری حس می‌کنید، تعجب نکنید.

اتمام پروژه

در گزارش تجدید نظر شد و همهٔ نمودارهای با کیفیت ارائه‌کرد آماده شدند. علاوه بر قولی که به مشتری برای بازبینی روش‌ها و فصل نتایج پس از نوشتن آن داده بودیم، آماده‌کردن مدارک زیر پروژه را تکمیل می‌کرد. مثال‌هایی در شکل‌های مشخص‌شده ارائه شده‌اند.

- نامهٔ روکش[1] (شکل ۴.۴)
- صورت‌حساب (شکل ۵.۴)
- صفحهٔ عنوان (شکل ۶.۴)

تبصره‌ها

اضافه‌کردن «اطلاعات عنوان» (درجهٔ و سمت علمی) به نامهٔ همراه، روشن می‌کند که ما واجد شرایط برای انجام‌دادن این نوع تحلیل آماری هستیم. ممکن است مشتری تنها کسی نباشد که گزارش ما را می‌خواند.

نمونهٔ صورت‌حساب تنها به‌منظور ارائهٔ مثال تهیه شده است و شرح کامل هزینه‌ها (ساعت‌های کاری) مرتبط با پروژهٔ واقعی را منعکس نمی‌کند و در پی آن نیز نیستیم که بگوییم این فرمتی «استاندارد» برای صورت‌حساب‌هاست. با این حال با توجه به «صورت ظاهری» آن نکات زیر را متذکر می‌شویم:

- نرخ ساعتی (نشان داده نشده) هر نوع هزینهٔ سربار را شامل می‌شود، مگر اینکه لازم باشد آن را جداگانه منظور کنیم.
- کلمهٔ رایگان را به نفع مشتری افزوده‌ایم. مشتری می‌تواند ملاحظه کند که ما نخستین جلسهٔ یک‌ساعته را به احترام او مجانی برگزار کرده‌ایم و برای ۵ دقیقه‌ای که برای تبدیل و پردازش فایل اکسل آن‌ها صرف کرده‌ایم هزینه‌ای دریافت نکرده‌ایم.

1. cover letter

برنامهٔ مشاورهٔ آماری
اطلاعات سربرگ بنگاه م. آ.

تاریخ نامه

نام و آدرس مشتری

RE: گزارش ب. م. آ.، صورت‌حساب، و نمودارها

پروژه: پروژهٔ تحلیل پایان‌نامه:
اثر دخالت دادن ترجیحات سبک یادگیری مفهومی
بر دریادنگهداری بلندمدت و نگرش‌های
دبیران ابتدایی و
دبیرستانی نسبت به فناوری در ایجاد مهارت‌های حرفه‌ای

مشتری گرامی،

گزارش ب.م. آ.، صورت‌حساب، و خلاصه‌های نموداری برای پروژهٔ بالا به پیوست ارسال می‌شود.

لطفاً چک قابل پرداخت خود را به ... بنویسید.

در صورت هرگونه سؤالی، لطفاً با من تماس بگیرید.

ارادتمند،

ی. مشاور

اطلاعات عنوان

شکل ۴.۴ مثالی از یک نامهٔ روکش.

- ما هزینه‌ها را برحسب واحد ساعت منظور کرده‌ایم که همواره عملی نیست. بنابراین، برای پرهیز از کسرها یا دو رقم اعشار (بی‌مزه)، برخی مشاوران از واحد یک‌دهم ساعت (شش دقیقه) ــ و لابد یک دستگاه ساعت خوب ــ استفاده می‌کنند.

۷.۴ گزارش نهایی

برنامهٔ مشاورهٔ آماری

برنامهٔ مشاورهٔ آماری
اطلاعات سرویس ب. م. آ.

صورت‌حساب

تاریخ
مشتری
مشاور
پروژه

خدمات		ساعت	مبلغ
قرارداد فرعی			
آماده‌سازی داده‌ها		۱	رایگان
محاسبات آماری		۲	
تهیه مستندات یافته‌ها		۳	
جمع کل قراردادهای فرعی ...	برای هر ساعت	۵	
مشاور برنامهٔ آماری :			
قرار ملاقات	تاریخ	۱	
تحلیل آماری		۱	رایگان
آماده‌سازی گزارش		۱	
قرار ملاقات	تاریخ	۱	
نمودارهای یا کیفیت ارائه‌کرد		۲	
جمع کل مشاور برنامه آماری...	برای هر ساعت	۵	
جمع کل قرارداد برنامهٔ مشاورهٔ آماری		۱۰	

شکل ۵.۴ نمونه‌ای از صورت‌حساب.

مشاوران: ی. مشاور و الف. دانشجو
تاریخ: ۱ ژانویهٔ ۲۰۰۰
مشتری: الف. مشتری
پروژه: پروژه‌پایان‌نامه

۱. مقدمه:

هدف این مطالعه، تحقیق دربارهٔ این بود که آیا آموزش متکی بر ترجیح‌های سبک یادگیری افراد، دریادنگهداری مطالب آموزشی داده‌شده را بهبود می‌بخشد یا خیر.



پروژهٔ تحلیل آماری پایان‌نامه

اثرهای دخالت دادن ترجیحات سبک یادگیری بر دریادنگهداری بلندمدت و نگرش‌های معلمان ابتدایی و دبیران دبیرستانی نسبت به فناوری در ایجاد مهارت‌های حرفه‌ای

گزارش آماده‌شده برای مشتری الف.

توسط

ی، مشاور و الف، دانشجو

برنامهٔ مشاورهٔ آماری (ب. م. آ.)

۱ ژانویه ۲۰۰۰

خلاصهٔ اجرایی

از نتایج تحلیل ب. م. آ. چنین برمی‌آید که ارتقای روش‌های تدریس بر مبنای ترجیحات سبک یادگیری، دریادنگهداری بلندمدت مطالب آموخته‌شده را بهبود می‌بخشد. مطالعه‌ای دیگر برای ارزیابی درست اثر طولیِ دریادنگهداری ضروری است.

شکل ۶.۴ صفحهٔ عنوان برای گزارش نهایی.

۱.۱ طرح تحقیق:

نمونه مرکب بود از معلمان مدارس ابتدایی و متوسطه که به‌طور تصادفی به دو گروه تخصیص داده شدند: کنترل و آزمایشی. فرمت آموزش سنتی در مورد گروه کنترل به‌کار گرفته شد. برای گروه آزمایشی، آموزش سنتی با فعالیت‌های خاص مناسب با سبک‌های آموزشی ترجیحی شرکت‌کنندگان ترکیب شد. هر دو گروه به چهار دورهٔ تحصیلی تفکیک شدند و هر دوره فرمت‌های آموزشی یکسانی را دریافت کرد.

۲.۱ متغیرها:

ابزار سیاههٔ سبک یادگیری (LSI) در ابتدا برای رده‌بندی ترجیحات شرکت‌کنندگان (PREF) برای هر دو گروه (GROUP) به‌کار گرفته شد. برای مقاصد ارزیابی، سه اندازهٔ کمّی در این بررسی به‌کار گرفته شدند که عبارت بودند از: یک پیش‌آزمون (PRE)، یک نمرهٔ مقیاس نگرش (SDS)، و یک پس‌آزمون (POST) که یک ماه پس از آموزش ارائه شد. عامل‌های ثبت‌شدهٔ دیگر عبارت

بودند از GENDER، SESSION، و سطح مدرسه (SLEVEL). جزئیات مربوط به این سه متغیر در زیر خلاصه شده‌اند:

اندازه‌های کمّی

PRE	پیش‌آزمون:	رتبهٔ ماکسیمم= ۱۰۰
POST	پس‌آزمون:	رتبهٔ ماکسیمم= ۱۰۰
SDS	نمرهٔ نگرش:	رتبهٔ ماکسیمم= ۶۰

عامل‌های رسته‌ای:

GROUP	کنترل / آزمایشی
GENDER	مرد / زن
SLEVEL	سطح ابتدایی / دبیرستانی
SESSION	S1، S2، S3، S4 چهار دورهٔ تحصیلی درون هر GROUP
PREF	شنیداری، هیچ ترجیحی، لمسی، جنبشی، دیداری کدگذاری‌شده به‌صورت **A**، **N**، **T**، **K**، **V**

توجه: برای افرادی با PREF=N در GROUP=آزمایشی، ترجیح روش آموزش بر مبنای بالاترین نمرهٔ LSI نسبت به رسته‌های A، T، K، و V تخصیص داده شد.

۲. روش‌شناسی:

تحلیل آماری این آزمایش به‌وسیلهٔ ب.م.آ. با استفاده از بستهٔ نرم‌افزاری آماری SAS[1] انجام شد. سه شیوهٔ آماری در این تحلیل به‌کار رفت: از تحلیل اکتشافی داده‌ها (EDA) برای تلخیص داده‌ها استفاده شد، آزمون‌های t برای آشکارسازی اختلاف بین متوسط نمرات آزمون بین گروه‌های کنترل و آزمایشی به‌کار رفتند، و از تحلیل واریانس (ANOVA) برای آشکارسازی اثرهای عامل‌های معنی‌دار استفاده شد. جزئیات مربوط به روش‌شناسی و تفسیر این شیوه‌های آماری در زیر به‌اختصار مورد بحث قرار گرفته است. اطلاعات بیشتر در [۲] داده شده است.

۱.۲ تحلیل اکتشافی داده‌ها:

فنون EDA برای تلخیص داده‌ها مورد استفاده قرار می‌گیرند. جدول‌های فراوانی، نمودارهای

میله‌ای، و بافت‌نگارها به‌طرز مؤثری توزیع متغیرهای تحت‌بررسی را به نمایش می‌گذارند. نمودارهای میله‌ای برای متغیرهای رسته‌ای به‌خدمت گرفته شدند: GROUP، SESSION، SLEVEL، PREF، و GENDER؛ بافت‌نگارها برای متغیرهای کمّی: PRE، POST، و SDS.

۲.۲ آزمون‌های t:

از آزمون‌های t برای ارزیابی این امر استفاده شد که آیا اختلاف معنی‌داری بین میانگین‌های مرتبط با دو نمونهٔ مستقل (یعنی متوسط نمره‌های POST گروه آزمایشی در مقابل گروه کنترل) وجود دارد یا خیر. معنی‌داری مبتنی بر Pـمقدار مرتبط با آزمون t است. بنابه قرارداد، یک Pـمقدار کوچکتر از ۰٫۰۵ (۵٪) را فراهم‌کنندهٔ شواهد کافی مبنی بر یک اختلاف معنی‌دار در نظر می‌گیرند. یک Pـمقدار کوچکتر از ۱٪ شواهدی قوی از نتیجه‌ای به‌لحاظ آماری معنی‌دار به‌دست می‌دهد.

۳.۲ ANOVA:

برای آزمون همزمان اثرهای چندین عامل می‌توان از یک مدل تحلیل واریانس استفاده کرد. به‌عنوان مثال، مدلی مناسب برای آزمودن اینکه هر یک از عامل‌های مربوط اثری بر POST خواهد داشت یا خیر به‌گونهٔ زیر خواهد بود:

```
POST = GROUP + SESSION(GROUP) + PREF + GENDER + SLEVEL
```

که در آن POST متغیر پاسخ است و GROUP، SESSION(GROUP)، PREF، GENDER، و SLEVEL همگی عامل‌هایی هستند که ممکن است بالقوه بر پاسخ اثر بگذارند. در این بررسی، SESSION عاملی «آشیانی» نامیده می‌شود زیرا سطوح آن درون هر GROUP تخصیص داده شده‌اند. بنابراین، اثر آماری SESSION به‌عنوان عامل آشیانی SESSION(GROUP) ارزیابی می‌شود.

معنی‌داری کلی این مدل در ابتدا برای تعیین آن به‌کار رفت که هر یک از عامل‌ها اثری معنی‌دار بر پاسخ دارند یا خیر. برای تعیین اثر هر عامل فردی، مجموع‌های توان‌های دوم (SS) نوع III به‌کار رفت. این مجموع‌ها اثر فردی یک عامل را، وقتی سهم همهٔ عامل‌های دیگر قبلاً به‌حساب آورده شده باشند، نشان می‌دهند.

۳. یافته‌ها:

جدول ۱ مرکب است از چندین جدول فراوانی که زیرگروه‌های شرکت‌کنندگان را نسبت به عامل‌های رسته‌ای به‌دست می‌دهد. جدول‌های ۲الف و ۲ب آماره‌های خلاصه را برای متغیرهای کلی PRE، POST و SDS و نسبت به GROUP ارائه می‌کنند.

آزمون‌های t در جدول ۳ ارائه می‌شوند. معلوم شد که وقتی SDS و PRE در نظر گرفته می‌شوند، هیچ اختلاف معنی‌داری بین گروه‌های آزمایشی و کنترل وجود ندارد (Pـمقدار< ۰٫۰۵). از این مطلب چنین برمی‌آید که اختلاف معنی‌داری بین گروه کنترل و آزمایشی، مادام که دانش و نگرش پیشین مطرح است، وجود ندارد. با این‌حال، POST قویاً معنی‌دار است که از آن چنین مستفاد می‌شود که میانگین‌های دو گروه، مادام که معلوماتِ دریادنگه‌داشته‌شده مطرح است، به‌طور معنی‌داری اختلاف دارند.

نتایج تحلیل واریانس برای متغیرهای PRE، SDS و POST در جدول ۴ ارائه می‌شوند. عامل‌های یکسانی در هر مدل به شرح زیر به‌کار گرفته شده‌اند.

```
(1) PRE = GROUP  SESSION(GROUP)  PREF  GENDER  SLEVEL
(2) SDS = GROUP  SESSION(GROUP)  PREF  GENDER  SLEVEL
(3) POST = GROUP  SESSION(GROUP)  PREF  GENDER  SLEVEL
```

از این مدل‌ها تنها مدل (۳) نتیجه‌ای معنی‌دار با Pـمقدار ۰٫۲۷ به‌دست داد. از Pـمقدارهای ($Pr > F$) مرتبط با SS نوع III، معلوم شد که تنها GROUP با Pـمقدار ۰٫۰۰۰۶ معنی‌دار است.

۴. نتیجه‌گیری‌ها

ب.م.آ (SCP) دریافت که نتایج معنی‌داری برای متغیرهای PRE و SDS وجود ندارد که از آن چنین مستفاد می‌شود که شرکت‌کنندگان، در مجموع، با هم اختلاف معنی‌داری نسبت به این متغیرها نداشتند. یک اثر معنی‌دار GROUP برای متغیر POST وجود داشت که از آن چنین برداشت می‌شود که ترکیب روش‌های آموزشی از طریق استفاده از ترجیحات سبک یادگیری، دریادنگهداری مطالب را برای گروه‌های مشابه بهبود خواهد بخشید. نتیجه‌گیری ب.م.آ. این است که تحقیق به‌عمل‌آمده توسط مشتری، نتایج معنی‌داری به وجود آورده است که ارزش پیگیری در یک بررسی با مقیاس وسیع‌تر را دارد.

مراجع

[1] SAS Institute Inc. (1990) *SAS/STAT User's Guide*. Version 6. Cary. NC.

[2] Moore, D. S. and McCabe, G. P. (1993)*Introduction to the Practice of Statistics*. 2nd ed. Freeman Press, NY.

پیوست ۱: جدول‌ها

فهرست جدول‌ها

جدول ۱: جدول‌های فراوانی برای متغیرهای رسته‌ای

جدول ۲الف: آماره‌های خلاصهٔ نمره‌های آزمون

جدول ۲ب: آماره‌های خلاصهٔ نمرات آزمون برحسب GROUP

جدول ۳: نتایج آزمون t

جدول ۴: نتایج ANOVA

جدول ۱: جدول‌های فراوانی برای متغیرهای رسته‌ای

تعداد کل: ۸۷

گروه	تعداد	درصد
کنترل	۴۳	۴۹٫۴
آزمایشی	۴۴	۵۰٫۶

دورهٔ تحصیلی	کنترل	آزمایشی	تعداد	درصد
S۱	۱۱	۱۱	۲۲	۲۵٫۳
S۲	۱۱	۱۲	۲۳	۲۶٫۴
S۳	۱۱	۱۱	۲۲	۲۵٫۳
S۴	۱۰	۱۰	۲۰	۲۳٫۰

ترجیح	کنترل	آزمایشی	تعداد	درصد
A	۱۲	۹	۲۱	۲۴٫۱
K	۲	۲	۴	۴٫۶
N	۱۷	۱۵	۳۲	۳۶٫۸
T	۱۱	۱۲	۲۳	۲۶٫۴
V	۱	۶	۷	۸٫۰

جنسیت	کنترل	آزمایشی	تعداد	درصد
F	۳۳	۳۷	۷۰	۸۰٫۵
M	۱۰	۶	۱۷	۱۹٫۵

دورهٔ تحصیلی	کنترل	آزمایشی	تعداد	درصد
E	۳۱	۲۸	۵۹	۶۷٫۸
H	۱۲	۱۶	۲۸	۳۲٫۲

جدول ۲الف:

آماره‌های خلاصهٔ نمره‌های آزمون

تعداد کل= ۸۷

	PRE	POST	SDS
Min.	5.00	36.00	21.00
1st Qu.	40.00	76.00	42.50
Median	55.00	80.00	50.00
Mean	55.63	80.87	48.05
3rd Qu.	75.00	90.00	55.00
Max.	100.00	100.00	60.00
Std. Dev.	25.31	12.11	8.56

جدول ۲ب:

آماره‌های خلاصهٔ نمره‌های آزمون برای GROUP

تعداد کل: کنترل = ۴۳، آزمایشی = ۴۴

	کنترل			آزمایشی		
	PRE	POST	SDS	PRE	POST	SDS
Min.	5.00	36.00	21.00	5.00	64.00	21.00
1st Qu.	45.00	72.00	42.00	35.00	80.00	44.25
Median	60.00	80.00	49.00	50.00	88.00	51.50
Mean	57.67	76.00	47.12	53.64	85.64	48.95
3rd Qu.	80.00	84.00	55.00	75.00	96.00	44.25
Max.	95.00	96.00	59.00	100.00	100.00	60.00
Std. Dev.	24.67	11.58	8.48	26.04	10.75	8.64

جدول ۳:

نتایج آزمون t

اختلاف در نمره‌های آزمون متوسط برحسب GROUP

تعداد کل ۸۷: کنترل=۴۳، آزمایشی=۴۴

Variable: SDS

GROUP	N	میانگین	انحراف استاندارد	خطای استاندارد	مینیمم	ماکسیمم
Control	43	47.116	8.477	1.293	21	59
Experiment	44	48.954	8.637	1.302	21	60

Variances	T	DF	Prob> |T|
Unequal	-1.0018	85.0	0.3193
Equal	-1.0016	85.0	0.3194

Variable: PRE

GROUP	N	Mean	Std Dev	Std Error	Min	Max
Control	43	57.674	24.673	3.763	5	95
Experiment	44	53.636	26.045	3.926	5	100

Variances	T	DF	Prob> |T|
Unequal	0.7425	84.9	0.4598
Equal	0.7421	85.0	0.4601

Variable: POST (*** Significant)

GROUP	N	Mean	Std Dev	Std Error	Min	Max
Control	43	76.000	11.580	1.766	36	96
Experiment	44	85.636	10.751	1.621	64	100

Variances	T	DF	Prob> |T|	
Unequal	-4.0202	84.2	0.0001	***
Equal	-4.0237	85.0	0.0001	***

جدول ۴:

نتایج ANOVA

شیوهٔ مدل‌های خطی کلی

تعداد مشاهده‌ها=۸۷

Model 1: SDS
Model 2: PRE
Model 3: POST
} = GROUP + SESSION(GROUP)+ PREF + GENDER + SLEVEL

مدل ۱: SDS (هیچ عاملی معنی دار نیست)

Source	DF	Sum of Squares	Mean Square	F-Value	$Pr > F$
Model	13	876.6456	67.4343	0.91	0.5492
Error	73	5423.1705	74.2900		
Corrected Total	86	6299.8161			

Source	DF	Type III SS	Mean Square	F-Value	$Pr > F$
GROUP	1	65.0331	65.0331	0.88	0.3526
SESSION(GROUP)	6	707.4197	117.9033	1.59	0.1631
PREF	4	80.4978	20.1245	0.27	0.8958
GENDER	1	5.4309	5.4309	0.07	0.7876
SLEVEL	1	10.6014	10.6014	0.14	0.7067

مدل ۲: PRE (هیچ عاملی معنی دار نیست)

Source	DF	Sum of Squares	Mean Square	F-Value	$Pr > F$
Model	13	12281.0661	944.6974	1.61	0.1017
Error	73	42809.1638	586.4269		
Corrected Total	86	55090.2299			

Source	DF	Type III SS	Mean Square	F-Value	$Pr > F$
GROUP	1	23.3249	23.3249	0.04	0.8425
SESSION(GROUP)	6	4339.8085	723.3014	1.23	0.2993
PREF	4	4465.2543	1116.3136	1.90	0.1189
GENDER	1	6.4008	6.4008	0.01	0.9171
SLEVEL	1	689.2741	689.2741	1.18	0.2819

Model 3: مدل ۳: POST (عامل GROUP ** معنی دار است)

Source	DF	Sum of Squares	Mean Square	F-Value	$Pr > F$	
Model	13	3389.7547	260.7504	2.06	0.0270	**
Error	73	9231.8545	126.4638			
Corrected Total	86	12621.6092				

Source	DF	Type III SS	Mean Square	F-Value	$Pr > F$	
GROUP	1	1644.6794	1644.6794	13.01	0.0006	**
SESSION(GROUP)	6	880.1432	146.6905	1.16	0.3372	
PREF	4	295.3252	73.8313	0.58	0.6753	
GENDER	1	50.6251	50.6251	0.40	0.5289	
SLEVEL	1	12.4934	12.4934	0.10	0.7542	

پیوست ۲: داده‌ها

GROUP = کنترل							GROUP = آزمایشی						
Order: SESSION, PREF, GENDER, SLEVEL, PRE, POST, SDS													
S1	V	F	H	25	64	51	S1	A	F	H	40	96	52
S1	K	F	E	65	64	21	S1	T	F	E	5	64	29
S1	A	F	E	5	60	41	S1	N	F	E	20	64	54
S1	A	M	E	60	76	56	S1	N	M	E	95	92	57
S1	T	M	E	65	76	51	S1	T	F	H	75	92	53
S1	T	M	E	95	88	55	S1	A	F	E	70	100	46
S1	T	F	H	85	84	50	S1	T	F	E	40	84	21
S1	N	F	E	85	72	42	S1	N	M	H	80	96	37
S1	A	M	E	65	76	34	S1	N	F	E	25	76	52
S1	N	F	H	50	68	42	S1	V	F	E	20	92	38
S1	N	M	E	5	36	36	S1	V	F	E	5	96	44
S2	T	F	H	65	80	56	S2	A	F	E	15	76	54
S2	T	F	E	15	76	47	S2	N	F	H	50	80	48
S2	A	M	E	20	80	44	S2	V	M	H	50	88	57
S2	A	F	E	85	76	51	S2	N	F	H	25	88	41
S2	A	F	E	20	80	56	S2	T	F	H	50	96	48
S2	T	F	E	45	80	58	S2	N	F	E	80	92	49
S2	N	F	E	55	76	54	S2	N	M	E	60	68	43
S2	A	F	H	50	80	42	S2	T	F	E	40	92	52
S2	A	M	H	15	44	39	S2	N	F	E	35	96	32
S2	N	M	E	90	88	52	S2	N	F	E	90	92	60
S2	A	F	E	55	72	44	S2	V	F	H	25	88	52
							S2	A	F	H	50	80	42
S3	N	F	H	50	64	41	S3	A	F	H	70	72	56
S3	N	F	H	45	80	50	S3	T	F	H	35	68	49
S3	N	F	E	90	88	40	S3	N	F	H	75	96	48
S3	N	F	E	70	80	49	S3	A	M	H	25	100	45
S3	T	F	E	85	84	59	S3	N	F	H	40	84	56
S3	T	F	H	70	76	46	S3	K	F	E	75	96	60
S3	N	F	E	80	68	59	S3	N	F	E	80	96	51
S3	T	F	H	45	88	52	S3	A	F	E	45	92	45
S3	N	F	H	80	80	42	S3	T	F	E	95	80	57
S3	N	F	E	75	84	56	S3	T	F	E	60	88	37
S3	N	F	H	70	84	32	S3	T	F	E	100	96	59
S4	N	F	E	50	80	50	S4	A	F	E	80	92	56
S4	N	M	E	45	64	43	S4	T	M	H	65	92	58
S4	N	F	E	85	96	58	S4	A	F	H	30	80	52
S4	A	F	E	95	80	56	S4	N	F	E	60	64	49
S4	A	F	E	75	84	44	S4	K	F	E	70	64	60
S4	T	M	E	45	60	35	S4	N	M	E	40	80	54
S4	T	F	E	55	88	58	S4	V	F	E	35	80	52
S4	K	F	E	60	76	45	S4	T	F	E	100	96	51
S4	N	F	E	55	88	49	S4	V	F	E	60	84	43
S4	A	F	E	35	80	40	S4	T	F	E	75	80	55

۸.۴ پس‌نوشت

ما فصل روش‌ها و یافته‌های مشتری را که با بازبینی آن موافقت کرده‌بودیم، همراه با یک چک پرداخت وجه دریافت کردیم (مشتریان می‌توانند دست به نقطهٔ حساس نیز بگذارند!). در این نوع پروژه‌ها، باز گذاشتن امکان دسترسی مشتری به ما برای برخی خدمات پس از کامل شدن پروژه، روابط عمومی خوبی است و معمولاً به تلاش و وقت زیادی نیاز ندارد (اگر غیر از این باشد، همواره می‌توانیم هزینه‌های آن‌ها را دریافت کنیم). در زیر، نسخهٔ اصلاح شده‌ای از بازبینی با ای‌میل، که به‌صورت خودمانی و با صدا کردن اسم کوچک تهیه شده است، ملاحظه می‌شود. ارجاع به جدول ۱ مربوط به سند مشتری است.

الف،

چند نکته یا پیشنهاد برای روش‌های فصل شما...

[شمارهٔ صفحه یا سطر بر اساس سندی که به من فرستاده‌ای]

امیدوارم سودمند باشد.

با احترام،

ی.

صفحهٔ ۱: متغیرهای وابسته، میانگین آزمودنی‌ها بودند...

متغیرهای وابسته در حقیقت نمرات «واقعی»اند زیرا تنها یک آزمون POST/SDS برگزار شد؛ یعنی، شما این نمرات را از چندین آزمون POST متوسط‌گیری نکردید.

صفحهٔ ۳: ... و (ج) تحلیل‌واریانس (*ANOVA*) ...

آنچه نوشته‌اید بسیار خوب به نظر می‌آید. به خاطر داشته باشید که ANOVA زمانی به‌کار گرفته می‌شود که:

- یک عامل بیشتر از دو سطح داشته باشد.
- اثر همزمان چندین عامل را باید در نظر بگیریم.

توجه کنید که شیوه‌ای که واقعاً به‌کار برده‌ایم عبارت بود از: مدل‌های خطی کلی (GLM). ــتفاوت بین GLM و ANOVA صرفاً این است که GLM «اندازه‌های نابرابر خانه‌ها» را به‌حساب می‌آورد (یعنی، GLM تعداد متفاوت مرد/زن، مدرسهٔ ابتدایی/متوسطه، و غیره را در نظر می‌گیرد)... درواقع، GLM چیزی بیش از یک تحلیل ANOVAی «درست» نبوده است.

صفحهٔ ۳: [سطر ۴ ↑] ...گروه‌های آزمایشی (Pـمقدار<۰٫۰۵).

پیشنهاد می‌کنم که مفهوم Pـمقدار راگسترش دهید ... شاید بهتر باشد نخستین جمله را به‌صورت زیر به پایان برسانید:

«... گروه آزمایشی.» سپس چیزی شبیه به جملهٔ زیر را اضافه کنید:

برای ارزیابی معنی‌داربودن (یا نبودن) آماری مرتبط با آزمونی خاص، نظیر آزمون tی دونمونه‌ای به‌کار رفته در جدول ۱، می‌توانیم از معیار «Pـمقدار» استفاده کنیم. در عرف آماری، استاندارد ۰٫۰۵ را به‌عنوان معیار نقطهٔ برش برای ارزیابی معنی‌داری پذیرفته‌اند (Pـمقدار< ۰٫۰۵). همان‌طور که از جدول ۱ می‌توان دید، هر دو[1] مقدار فهرست‌شده تحت «$|T|$ < احتمال» از ۰٫۰۵ تجاوز می‌کنند. این مطلب مستلزم آن است که گروه‌های کنترل و آزمایشی برحسب ... (به‌طور معنی‌داری) متفاوت نباشند.

آزمون tی دونمونه‌ای را می‌توان تحت این فرض انجام داد که تغییرپذیری نمرات پیش‌آزمون درون گروه‌های کنترل و آزمایشی یکسان هستند (که در جدول ۱ با «برابر» نشان داده شده). گرچه این فرض را می‌توان در این مطالعه معقول دانست (و این با انحراف استانداردهای نمونه‌ای قابل توجه از دو گروه تأیید می‌شود)، آزمون tی دونمونه‌ای را می‌توان بدون در نظرگرفتن این فرض نیز انجام داد که تغییرپذیری یکسان («نابرابر») است. در هر یک از دو حالت، معیار Pـمقدار را می‌توان برای ارزیابی معنی‌داری آماری به‌کار برد.

صفحهٔ ۴: [سطر ۱۰] ... متفاوت ($p < ۰٫۰۰۰۴$).

Pـمقدار برابر است با ۰٫۰۰۰۴ (و نه کمتر از ۰٫۰۰۰۴). جمله را به‌صورت زیر بازنویسی کنید:

« ... متفاوت (Pـمقدار ۰٫۰۰۰۴ (*))».

(*) [پانوشت احتمالی.]

این نتیجه «قویاً» معنی‌دار تلقی می‌شود ... (چون کوچکتر از ۰٫۰۱ است که اغلب برای توصیف معنی‌داری ورای معیار استاندارد ۰٫۰۵ به‌کار می‌رود).

صفحهٔ ۵: [سطر ۱۲] اشتباه تایپی: (Pـمقدار ۰٫۰۰۳۱) ← (Pـمقدار ۰٫۰۰۱۳).

صفحهٔ ۶: به نظر خوب می‌رسد.

در فصل بعدی شما می‌توانستید قید کنید «سمت و سوهای تحقیق بیشتر» تا متغیر SDS را بررسی کنید. شما آشکارا در این بررسی داده‌های کافی برای اینکه یک تحلیل پر

۱. [یک پانوشت احتمالی برای توضیح اینکه چرا دو آزمون t داریم.]

محتوی از SDS برحسب PREF انجام دهید، نداشته‌اید (پس از تفکیک SDS برحسب GROUP با حذف PREF = N، عدهٔ کمی در گروه‌های K و V که واقعاً کافی نبودند حضور داشته‌اند) ... شاید در پایان‌نامهٔ بعدی شما؟

پرسش‌ها

۱. یکی از موضوعات اصلی که در ارائه‌کرد خود از این مثال مطالعهٔ موردی کنار گذاشتیم، مربوط به PREF = N شرکت‌کننده بود.

مهم. ممکن است بخواهید پرسش بعدی را پیش از آغاز هر یک از بخش‌های این پرسش بخوانید.

الف) دربارهٔ مباحث آماری مرتبط با PREF = N رده با مشتری بحث کنید. به خاطر بیاورید که به این شرکت‌کنندگان در گروه آزمایشی یک ترجیح «تخصیص داده شد».

ب) رویکردی بدیهی آن است که متغیر پاسخ را نسبت به زیرمجموعه‌های مجزا تحلیل کنیم: PREF = N و PREF ≠ N. چه مشکلی پیش می‌آید؟ آیا رویکرد بهتری به تحلیل وجود دارد؟

ج) تحلیل مطرح‌شده در بالا را انجام دهید. آیا نتیجه‌گیری‌های شما برای دو زیرمجموعه متفاوت‌اند؟ باید باشند؟

۲. فرض کنید از شما خواسته‌اند که یک صورت‌حساب تفصیلی برای مشتری ایجاد کنید. تحلیل مسئله PREF = N ذکرشده در بالا در اینجا بسیار مفید خواهد بود: برای هر کاری که انجام می‌دهید، سند تهیه کنید، و زمانی را که صرف هر کار می‌کنید یادداشت کنید. به روالی دیگر، صرفاً از ارائه‌کرد ما از این مطالعهٔ موردی به‌طور کامل به‌عنوان مبنای ایجاد صورت‌حساب استفاده کنید. فرض کنید که هر دو جلسهٔ مشاوره یک ساعت کامل وقت برده است.

الف) نخستین گام، قلم به قلم ذکرکردن هر مؤلفهٔ مرتبط با هزینهٔ فرایند مشاوره است. کجا به‌طور فعال مشغول کار روی پروژه شدیم؟ چه کاری انجام دادیم؟ تذکر: هزینه‌های سرمایه‌ای (کلان) را حساب نکنید. مشتری قصد ندارد که هزینهٔ چاپگر جدید شما و هزینهٔ مشاورهٔ شما را بپردازد!

ب) به هر مؤلفه یک واحد زمان اضافه کنید (از یک‌دهم ساعت استفاده کنید). چیزهای کوچکی از قبیل ۱۲ دقیقه‌ای را که صرف ترتیب‌دادن اولین جلسهٔ ملاقات و ارائهٔ اطلاعات اولیه کردیم از یاد نبرید. لازم خواهد شد که زمان صرف‌شده برای بسیاری

از این مؤلفه‌ها را برآورد کنید. [بنابراین، ارزش تهیهٔ مستند هر مرحله (و زمان صرف‌شده) در تحلیل خود از مسئلهٔ PREF = N را درمی‌یابید].

ج) نرخ ساعتی خود را، با فرض اینکه همهٔ کارها را خود انجام داده‌اید، در نظر بگیرید. نرخ شما به‌طور معمول سرباری را شامل می‌شود که هزینه‌های اتفاقی را جبران می‌کند و می‌توان از آن برای بهبود موارد سرمایه‌ای استفاده کرد. هزینهٔ کل را برای مشتری محاسبه کنید. آیا این مبلغ برای $n = ۸۷$ مشاهده واقعی است؟

د) این هزینهٔ کل را با برگزاری رایگان نخستین ساعت مشاوره و تخصیص کارهای مناسب به یک «دستیار» (با نرخ ساعتی کمتر) تعدیل کنید. هزینهٔ کل تعدیل‌شده را بر $n = ۸۷$ تقسیم کنید. این هزینهٔ برحسب هر مشاهده چه زمانی به‌درد می‌خورد؟ چه زمانی به درد نمی‌خورد؟

۳. با توجه به تذکر ما در صفحهٔ ۱۹۲ شاید تعجب‌آور باشد که ما دراید‌های «واریانس نابرابر» را در گزارش‌های مقدماتی و نهایی حفظ کردیم. توجه کنید که مقادیر `std Dev` در جدول ۳ برای گروه‌های کنترل و آزمایشی درون هر متغیر پاسخ بسیار شبیه به یکدیگرند.

الف) آیا دلیلی در دست هست که چرا نباید دراید‌های واریانس نابرابر را حذف می‌کردیم؟

ب) به‌نظر می‌رسد که هیچ آزمونی برای بررسی نرمال‌بودن در رابطه با آزمون t یا تحلیل‌های ANOVA اجرا نشده است. چگونه نتایج این آزمون را به مشتری توضیح می‌دهید؟

(i) پیامدهای یک نتیجهٔ معنی‌دار چیست؟ (یعنی فرض نرمال‌بودن رد می‌شود.)

(ii) موضوع «استواری» را چگونه وارد توضیح خود می‌کنید؟

۴. برای کاربران S-PLUS.

الف) این داده‌ها را با استفاده از S-PLUS تحلیل کنید. توجه کنید که جدول‌های ۲الف و ۲ب واقعاً با استفاده از تابع `summary()` در نرم‌افزار S-PLUS تولید شده بود. تفاوت خروجی نتایج آزمون t و ANOVA از S-PLUS در مقایسه با SAS در چیست؟

ب) تابع `g.hist2.fun()` شرح داده‌شده در صفحهٔ ۱۹۴ را کامل کنید. هدف آن است که بتوانیم یا نمودار بافت‌نگار PRE برحسب GROUP یا POST برحسب GROUP (شکل ۳.۲) را تنها با استفاده از شناسه‌های داده‌شده در خط گزینه‌های تابع `g.hist۲.fun()` شما تولید کنیم.

ج) مشتری بافت‌نگارهایی با کیفیت ارائه‌کرد را برای چندین متغیر درخواست کرد. مشتری می‌خواست، در هر شکل، مقدار میانگین و انحراف استاندارد متغیر بر روی نمودار (درون ناحیهٔ شکل) با برچسب‌های «Standard Deviation =< value >» و «Mean =< value >» درج شود. گزینه‌های زیر را در نظر بگیرید.

(i) یک تابع عمومی تولید کنید که به‌طور خودکار جایی «مناسب» برای قراردادن این متن در ناحیهٔ شکلِ هر بافت‌نگار تعیین کند.

(ii) با آزمون و خطا، متن را در جای مناسبی درون ناحیهٔ شکل هر بافت‌نگار قرار دهید.

کدام رویکرد را باید اتخاذ کنیم؟

قسمت II

مطالعه‌های موردی

۵

آشنایی با مطالعه‌های موردی

بهترین راه برای یادگیری مشاورهٔ آماری انجام‌دادن آن است! در حقیقت، هیچ جانشینی برای این کار وجود ندارد و اشتباه‌کردن بخش مهمی از این فرایند است. قبلاً چندین اشتباه خود را در فصل قبل ذکر کردیم؛ نگران نباشید، هنوز بسیاری از آن‌ها برای شما مانده که مرتکب شوید. این بحث ما را به مطالعه‌های موردی می‌کشاند. منظور از آن‌ها فراهم‌آوردن این فرصت برای خواننده است که کمی تجربه کسب کند ــ حداقل در قالب یادگیری تحلیل‌های آماری لازم در انواع مختلف مسائل مشاوره‌ای.

مقصود از این مقدمه مروری کلی برای خواننده از شیوه‌ای است که برای ارائهٔ مطالعه‌های موردی در قسمت II به‌کار می‌بریم. مطالعه‌های موردی به سه گروه تفکیک شده‌اند که مشخصاً به‌عنوان گروه I، گروه II، وگروه III برچسب خورده‌اند و سه فصل آتی را تشکیل می‌دهند. فرمت ارائه و جزئیات مربوط به رویکرد کلی به‌کارگرفته‌شده در این مطالعه‌های موردی در بخش بعدی توضیح داده می‌شوند. سطح پیچیدگی و نوع تحلیل آماری مورد نیاز در هر گروه نیز مورد بحث قرارگرفته است. فصل آخرِ قسمت II مرکب از گردایه‌ای از مطالعه‌های موردی گروه‌بندی‌نشده است که به‌صورت تمرین ارائه می‌شود.

۱.۵ فرمت ارائهٔ مطالعه‌های موردی

عنوان مطالعهٔ موردی در عنوان بخش می‌آید. پس از آن بلافاصله کادری ظاهر می‌شود که حاوی اطلاعات مربوط به داده‌ها، متغیرها، و روش‌های آماری به‌کاررفته در تحلیل است. این کادر ممکن است گاهی اطلاعات اضافی هم داشته باشد. در این صورت، جزئیات خاص مرتبط با آن مطالعهٔ موردی در زیربخشی که بعد از آن می‌آید ارائه می‌شود.

طرحی اجمالی از فرمت ارائه در زیر آمده است. تنها به‌منظور ارائهٔ مثال، از داده‌های گردآوری‌شده در زمینهٔ مطالعهٔ ازکارافتادن باتری که توسط NASA انجام‌شده استفاده می‌کنیم. داده‌ها را می‌توان در جانسون و ویچرن (۱۹۹۸، جدول ۴.۷) یافت. ترتیب و محتوای اصلی زیربخش‌های مورد اشاره در این طرح اجمالی در مورد همهٔ مطالعه‌های موردی عمومیت دارند.

فرمت مطالعهٔ موردی

عنوان آزمایش از کار افتادن باتری ناسا

کادر

منبع	جانسون و ویچرن (۱۹۹۸)
روش‌ها	تحلیل اکتشافی داده‌ها رویه‌های پاسخ تحلیل کوواریانس
داده‌ها	آزمایش طراحی‌شده با ۵ متغیر کنترل و پاسخ: چرخه‌های ازکارافتادن

زیربخش‌ها:

---

۱. زمینه «یک آزمایش برنامه‌ریزی‌شده»
زمینهٔ مسئله. پرسش‌هایی که باید مورد نظر قرار گیرند و مباحث آماری اصلی.

۲. داده‌ها توصیف و فرمت پایگاه داده‌ها. ویژگی‌های متغیرهای مرتبط.

۳. روش‌ها جزئیات شیوه‌های آماری که می‌توان در تحلیل به‌کار برد. در این مثال:
- تحلیل اکتشافی داده‌ها
- رویه‌های پاسخ
- تحلیل کوواریانس

۴. تحلیل پرسش‌ها چند نتیجهٔ مقدماتی. خروجی‌های SAS و S-PLUS.
کامل‌کردن تحلیل. گزارش.

۲.۵ جزئیات مطالعهٔ موردی

روشن است که مقصود از این فرمت ارائه، بازتولید یک مطالعهٔ موردی به سبک فصل ۴ نیست. در حالی که شکلی «مختصرشده» از یک مطالعهٔ موردی عیب‌هایی دارد، مزیت‌های مهم آن در این است که:

۱. تعداد بسیار بیشتری مطالعهٔ موردی را می‌توان ارائه کرد.

۲. زمینهٔ مسئله را می‌توان به‌طُرز مؤثری توصیف کرد.

۳. پردازش داده‌ها را می‌توان به‌صورت کلی‌تری انجام داد.

۴. روش‌شناسی آماری مناسب را می‌توان به‌تفصیل مورد بحث قرار داد.

۵. تحلیل مقدماتی را می‌توان به نتایج مربوط معطوف کرد.

۶. پرسش‌ها پایه‌ای آموزشی برای مطالعهٔ موردی فراهم می‌کنند.

عیب عمده این است که خواننده در جریان گفتگویی که در شکل مختصرشدهٔ مطالعهٔ موردی حاصل شده قرار ندارد. با این‌حال، دستورالعمل‌ها و اطلاعاتی که در قسمت I در اختیار گذاشته‌ایم، جبران این مصالحه را می‌کند و به خوانندگان اجازه می‌دهد که همان نوع «مختصر» را در پروژه‌های مشاوره‌ای خود انجام دهند.

زمینه

کاربرد آماری مربوط به این مطالعهٔ موردی با عنوان این زیربخش مشخص می‌شود. عنوانِ «یک آزمایش برنامه‌ریزی‌شده» چندان تخیل‌آمیز نیست (و در این مثال تا حدی با زیاده‌گویی همراه است)، اما، ممکن است در مواقعی که عنوان[1] مطالعهٔ موردی کمی رمزآلود! به‌نظر می‌رسد سودمند باشد.

هدف اصلی این زیربخش توصیف زمینهٔ مسئله‌ای است که این مطالعهٔ موردی به آن می‌پردازد. از آنجا که همهٔ مطالعه‌های موردی مبتنی بر پروژه‌های واقعی‌اند، به‌دلایل آشکار، برخی تغییرات در مسائل واقعی به‌عمل آمده است. با این حال، سعی کرده‌ایم که روح تحقیق اولیه را در توصیف خود از مطالعهٔ موردی حفظ کنیم و اطلاعات زمینه‌ای لازم را برای آنکه مضمون مسئله به‌قدر کافی کامل و معنی‌دار باشد فراهم کرده‌ایم.

---

۱. این عنوان در فهرست مندرجات جزو عناوین بخش‌ها آمده است، بنابراین دلیل انتخاب عنوان «کاربردی» مناسب برای این زیربخش محسوب می‌شود.

در این مثال، کاربردهای اولیهٔ ماهواره‌ها انگیزهٔ پیدایش پیل‌های باتری نقره‌روی شد. آزمون‌های زمان از کار افتادن برای ارزیابی قابلیت اعتماد این پیل‌ها تحت شرایط مختلف انجام شد.

داده‌ها

مجموعه‌داده‌های همهٔ مطالعه‌های موردی را می‌توان از وب‌سایت‌هایی که در پیوست ب (مجموعه‌داده‌ها) فهرست شده‌اند دانلود کرد.

نحوهٔ ارائهٔ داده‌ها، برای هر مطالعهٔ موردی، متفاوت است. مجموعه‌داده‌های کوچک در این زیربخش بازتولید می‌شوند، اما برای مجموعه‌داده‌های بزرگتر، تنها توصیفی از متغیرهای مربوط را فراهم می‌کنیم.

توجه به این نکته اهمیت دارد که اغلب مجموعه‌داده‌ها در وضعیت «پاکیزه»‌ای که ارائه می‌کنیم نیستند. ممکن است قبلاً ساعت‌ها برای تصحیح و تنظیم داده‌ها صرف شده باشد تا داده‌ها به این مرحله برسند. بار دیگر، این کار بخشی از مصالحه‌ای است که انجام داده‌ایم. با این حال، همواره باید یک تحلیل اکتشافی داده‌ها صورت داد، زیرا داده‌های دورافتادهٔ بالقوه را کنار نگذاشته‌ایم یا از سایر جهات داده‌ها را «بهداشتی» نکرده‌ایم.

پرسش‌های اصلی مورد توجه در هر مطالعهٔ موردی معمولاً در زیربخش زمینه در بالا مورد بحث قرار می‌گیرند. گاهی شاید مناسب‌تر باشد که این پرسش‌ها را پس از آنکه مجموعه‌داده‌ها و متغیرها ارائه شدند مورد توجه قرار دهیم. به‌عنوان مثال، اغلب آسان‌تر است که مباحث آماری را نسبت به متغیری خاص مورد مطالعه قرار دهیم تا در چارچوبی کلی. در این مثال، پیشنهاد می‌کنیم که تبدیلی از Y، که Y قبلاً به کمک جدول‌های زیر تعریف شده است، انجام دهیم.

جدول ۱.I

X1	نرخ شارژ (آمپر)
X2	نرخ خالی‌شدن شارژ (آمپر)
X3	عمق خالی‌شدن شارژ (٪ آمپر بر ساعت)
X4	دما (سانتی‌گراد)
X5	پایان ولتاژ شارژ (ولت)
Y	چرخه‌های از کار افتادگی (متغیر پاسخ)

جدول ۲.I داده‌های ازکارافتادگی باتری

Obs	X_1	X_2	X_3	X_4	X_5	Y
1	0.375	3.13	60.0	40	2.00	101
2	1.000	3.13	76.8	30	1.99	141
3	1.000	3.13	60.0	20	2.00	96
4	1.000	3.13	60.0	20	1.98	125
5	1.625	3.13	43.2	10	2.01	43
6	1.625	3.13	60.0	20	2.00	16
7	1.625	3.13	60.0	20	2.02	188
8	0.375	5.00	76.8	10	2.01	10
9	1.000	5.00	43.2	10	1.99	3
10	1.000	5.00	43.2	30	2.01	386
11	1.000	5.00	100.0	20	2.00	45
12	1.625	5.00	76.8	10	1.99	2
13	0.375	1.25	76.8	10	2.01	76
14	1.000	1.25	43.2	10	1.99	78
15	1.000	1.25	76.8	30	2.00	160
16	1.000	1.25	60.0	0	2.00	3
17	1.625	1.25	43.2	30	1.99	216
18	1.625	1.25	60.0	20	2.00	73
19	0.375	3.13	76.8	30	1.99	314
20	0.375	3.13	60.0	20	2.00	170

مسئلهٔ آماری مورد نظر حالا دیگر روشن است:
آن مقادیر X_1، ... ، X_5 را پیدا کنید که Y را بهینه کنند.

روش‌شناسی

روش‌های آماری به‌کار گرفته‌شده در تحلیل، سطح گروهی مطالعهٔ موردی را مشخص خواهد کرد. از چشم‌انداز یک مشاور آماری، این موارد چنین تعریف می‌شوند:

گروه I مطالعه‌های موردی ساده‌ای که برای تحلیل، تنها به روش‌های آماری استاندارد نیاز دارند (ن. ک. فصل ۳). پیداکردن جواب آن‌ها سرراست است. با این حال، ممکن است برخی نکات آماری و علمی جالبی در آن‌ها نیاز به بحث داشته باشند.

گروه II مطالعه‌های موردی پیچیده‌تری که مستلزم روش‌هایی از قبیل رگرسیون لوژستیک، مدلبندی سری‌های زمانی، و طرح‌های عاملی‌اند. این پیچیدگی ممکن است تا حدی ناشی از اندازهٔ مجموعه‌داده‌ها باشد. مسئلهٔ آماری عموماً خوش‌تعریف است، اما قلمرو وسیع‌تری از مطالعه‌های موردی گروه I دارند. ممکن است نیاز به چندین راه‌حل داشته باشیم.

گروه III مطالعه‌های موردی پژوهش‌محور که برای تحلیل، اغلب نیاز به روش‌های چندمتغیره یا

روش‌های آماری تخصصی دارند. چندین مرحلهٔ تحلیل نیز ممکن است برای به‌دست‌آوردن نتایج مناسب ضروری باشد. ممکن است برای مسئلهٔ آماری مورد بحث، «پاسخی» وجود نداشته باشد.

البته تفاوت قابل ملاحظه‌ای بین آنچه مشاور باید برای حل‌وفصل مسائل آماری ارائه‌شده در مطالعهٔ موردی بداند و آنچه از منظر آموزشی مورد نیاز است وجود دارد. بنابر تجارب ما، سطح پیچیدگی آماری ناشی از این گروه‌بندی‌ها، مانعی برای دانشجویان نیست. در واقع، یادگیری استفاده از یک روش آماری ناآشنا برای نخستین‌بار را می‌توان تا حدی یک وضعیت خاص محسوب کرد. شاید مهم‌تر از آن، تحلیل اکتشافی یک مطالعهٔ موردی پیچیده اغلب می‌تواند تجربهٔ یادگیری واقعی‌تری باشد.

بنابراین، هدف از این زیربخش پرداختن به آن جنبه‌های آماری مربوط به روش‌شناسی آماری است که می‌توان از آن‌ها در تحلیل مطالعهٔ موردی استفاده کرد. علاوه بر فراهم‌آوردن بحثی مفصل از فنون آماری مناسب مورد نیاز برای انجام یک تحلیل جامع، بر مؤلفه‌های خاص مرتبط با تحلیل اکتشافی داده‌ها نیز تأکید می‌کنیم. در برخی حالت‌ها، ممکن است صورت‌ها یا گزینه‌های دیگری بر رویکرد جامع مورد بحث پیشنهاد شود.

در این مثال، روش‌شناسی رویهٔ پاسخ به‌تفصیل مورد بحث قرار خواهد گرفت. ما تنها شرحی مختصر از مباحث اصلی این بحث را در زیر می‌آوریم. تذکر این نکته ضروری است که تمرکز اولیه بر این بحث می‌تواند آن را به مسئلهٔ جالب توجه انجام‌دادن یک تحلیل نموداری اکتشافی داده‌های چندمتغیره بکشاند.

- تحلیل اکتشافی داده‌ها

  تبدیل پایدارکنندهٔ واریانس (مثلاً $\log Y$)

  نمودار خم تراز Y در برابر X_1، X_2 (متغیرهای طرح)

  نمودارهای دار بستی: Y در برابر X_1، X_2 (خم تراز) برحسب X_3، X_4، X_5 (رسته‌ای‌شده)

- تحلیل رویهٔ پاسخ

  رگرسیون چندجمله‌ای

  تحلیل کانونی و طرح‌های مرکب

  تفسیر بهینه‌سازی

تحلیل مقدماتی

در این زیربخش، نتایج مربوط را، که از به‌کاربردن روش‌های آماری بحث‌شده در بالا حاصل شده است، ارائه می‌کنیم. شرحی از کد SAS و S-PLUS مورد استفاده برای تولید این نتایج به‌طور طبیعی در اینجا آورده می‌شود. در حالت کلی، تحلیل کاملی از مطالعهٔ موردی را ارائه نمی‌کنیم. برای این کار چند دلیل داریم، اما مهمترین آن‌ها این است:

بهترین راه یادگیری مشاورهٔ آماری انجام‌دادن آن است.

بنابراین، پس از همهٔ این مطالب، یک رویکردِ یا غرق شو ــ یا شنا کن داریم. البته همه‌اش هم این نیست. هدف ما فراهم‌کردن سری‌یی از مطالعه‌های موردی پیش‌رونده برای خواننده است. بدین معنی که تحلیل‌های مقدماتی مرتبط با مطالعه‌های موردی قبلی روش‌های خاصی را که در نتایج ارائه‌شده در مطالعهٔ موردی جاری بر آن‌ها تأکید شده تکمیل می‌کنند.

در این مثال، فرض می‌کنیم که خواننده با مفاهیم کلی و نتایجی که از به‌کار بستن یک تحلیل رگرسیونی استاندارد (روش‌های رگرسیونی در یک مطالعهٔ موردی پیشین ارائه شده‌اند) حاصل می‌شوند آشنا باشد. بنابراین، تمرکز عمدهٔ تحلیل مقدماتی ارائه‌شده در اینجا به نتایج حاصل از برازش یک مدل رویهٔ پاسخ معطوف خواهدبود. بحثی خلاصه از این نتایج نیز ارائه خواهد شد.

پرسش‌ها

یک هدف از بحث خلاصه، اشاره به جهتی کلی برای تحلیل بیشتر است. ممکن است جنبه‌های خاصی از مطالعهٔ موردی نیز وجود داشته باشند که مورد توجه باشند و امکان تحقیق بیشتر را بدهند. برای پرداختن به این مسائل، سری پرسش‌هایی را تدارک دیده‌ایم. این‌کار مبنایی آموزشی برای مطالعهٔ موردی فراهم می‌آورد و خواننده را قادر می‌سازد که تحلیل آماری مطالعهٔ موردی را تکمیل کند. به‌طور ایده‌آل، تحلیل کامل‌شده باید به شکل یک گزارش کتبی با تأکید خاص بر تعبیر زمینه‌ای یافته‌ها، یعنی «نتیجه‌گیری‌ها»ی گزارش، تدوین شود. ما تصمیم را به صلاحدید خود خواننده وامی‌گذاریم.

پس بدون های و هوی بیشتر، برویم سرِ مشاوره!

۶

مطالعه‌های موردی گروه I

در این فصل، چهار مطالعهٔ موردی را ارائه می‌کنیم که در آن‌ها روش و رویکرد به تحلیل ساده است، اما زمینهٔ مسئله‌ها برخی مباحث جالب آماری و علمی را پیش می‌آورد. در برخی از مطالعه‌های موردی، ارائهٔ تعبیری پرمعنی از نتایج آماری ممکن است واقعاً دشوار باشد. مطالعه‌های موردی بخش به بخش در زیر فهرست شده‌اند.

۱.۶ تبعیض در ارتقای شغلی

جدول‌های پیشایندی: آزمون‌های پیوند

۲.۶ موردِ مرسولهٔ گمشده

بررسی نمونه‌ای: طرح و تحلیل

۳.۶ دستگاهی برای کاهش آلاینده‌های موتور

آزمون‌های t و تحلیل واریانس

۴.۶ روان‌شناسی معکوس

آماره‌های خلاصه: همبستگی‌ها

۱.۶ تبعیض در ارتقای شغلی

روش‌ها	تحلیل جدول پیشایندی آزمون خی‌دو آزمون دقیق فیشر
داده‌ها	دو جدول پیشایندی ارتقاها در برابر پشتیبانی از کاندیداها

۱.۱.۶ ادعایی مبنی بر شواهد آماری

ادارهٔ حوادث اجتماعی[1] کشور آز[2،3] با یک شکایت تبعیض روبه‌روست. خواهان کارمند عالی‌رتبهٔ این اداره است که پس از انتخابات گذشته برای انتخاب نخست‌وزیر ارتقای شغلی نیافته است. او معتقد است که دلیل ارتقانگرفتن او آن است که تنها داوطلبانی ارتقا یافته‌اند که به نفع مبارزات انتخاباتی کاندیدای برنده برای نخست‌وزیری فعالیتی داشته‌اند. داوطلبان دیگری همچون او، که به‌نفع کاندیدای برنده در مبارزات انتخاباتی وارد عمل نشده‌اند، ارتقا نیافته‌اند. خلاصه‌ای از حقایق مربوط به شرح زیر است.

یافته:

۱۰ کارمند اداره، از جمله خواهان، در انتظار ارتقا بودند. هفت نفر ارتقای شغلی یافتند و، از این هفت نفر، شش نفر به مبارزات انتخاباتی کاندیدای برنده برای نخست‌وزیری کمک مالی کرده بودند. چهار نفر باقیمانده به مبارزات انتخاباتی نخست‌وزیر کمک نکرده بودند.

---

1. Department of Social Events

۲. در موارد تبعیض، مشخصات واقعی مطرح‌شده محرمانه‌اند و باید ناشناس باقی بمانند؛ بنابراین، از عنوان ساختگی برای اداره استفاده شده است.

۳. Oz، ظاهراً مؤلفان نام کشور آز را از عنوان کتاب The Wonderful Wizard of Oz نوشتهٔ L. Frank Baum عاریت گرفته‌اند.

شیوهٔ ارتقا:

شیوهٔ ارتقا مبتنی بر امتیازات حاصل از یک آزمون خدمات اجتماعی استاندارد است که همهٔ ۱۰ داوطلب در آن شرکت کرده بودند. امتیازها رتبه‌بندی شده و شیوهٔ ارتقای شغلی شرط می‌کند که، برای هر ارتقا، داوطلب موفق باید از بین آن‌هایی انتخاب شود که در حال حاضر در بین سه داوطلب با بالاترین رتبه‌ها قرار دارند (با منظورکردن همهٔ افرادی که امتیاز برابر دارند). این کار متوالیاً انجام می‌شود تا اینکه همهٔ موقعیت‌های شغلی موجود پر شوند.

ادعای خواهان:

رتبهٔ خواهان ۴ بود و او اعتقاد دارد که این رتبه بیش از حد کفایت برای ارتقای او از بین ۷ موقعیت ارتقای موجود بوده است. از آنجا که فعالیت‌های انتخاباتی جزو الزامات ارتقای شغل نبوده است، خواهان ادعا می‌کند که تبعیض رخ داده است.

اطلاعات بیشتر:

از کارمندانی که داوطلب ارتقای شغلی نبودند، سؤالی در این مورد به‌عمل آمد که آیا احساس تغییری مثبت یا منفی در وضعیت شغلی آن‌ها پس از انتخابات روی داده است یا خیر. همچنین از آن‌ها پرسیده شد که آیا به مبارزات انتخاباتی کاندیدای برنده کمک مالی کرده‌اند یا خیر.

داده‌ها

داده‌ها در سه جدول نشان داده شده‌اند. جدول ۱.۶ خلاصهٔ نتایج ارتقا بر اساس مشارکت در مبارزات انتخاباتی برای ۱۰ داوطلب است. آن‌گونه که از جدول ۲.۶ می‌توان دید، خواهان (D) تنها پس از آنکه نخستین موقعیت ارتقا پر می‌شود واجد شرایط برای ارتقا می‌شود.

جدول ۱.۶ مشارکت برحسب ارتقا

	ترفیع‌یافته	ترفیع‌نیافته
به برنده کمک کرده است	۶	۰
به برنده کمک نکرده است	۱	۳

جدول ۲.۶ رتبهٔ داوطلب (۱= بالاترین)

کاندیدا	A	B	C	D	E	F	G	H	I	J
رتبه	۱	۲	۳	۴	۵	۶	۷	۸	۸	۸

جدول ۳.۶ مشارکت برحسب تغییر در وضعیت شغلی

	مثبت	منفی	نامعلوم
به برنده کمک کرده است	۴	۰	۲
به برنده کمک نکرده است	۰	۷	۰
نامعلوم	۱	۹	۲

در جدول ۳.۶ پاسخ‌های دریافت‌شده از کارمندانی که داوطلب ارتقا نبوده‌اند خلاصه شده است. رستهٔ «نامعلوم» بدین معنی است که هیچ پاسخی به آن سؤال دریافت نشده است.

روش‌شناسی

۲.۱.۶ تحلیل جدول پیشایندی

به جدول‌های ۱.۶ و ۳.۶ معمولاً جدول پیشایندی اطلاق می‌شود و روش استاندارد تحلیل آن‌ها این است که یک آزمون استقلال انجام دهیم. این مطلب در فصل ۳ (نگاه کنید به جدول‌های پیشایندی) مورد بحث قرار گرفت و با ارزیابی این موضوع سروکار دارد که آیا شواهدی از یک پیوند آماری بین سطوح دو متغیر رسته‌ای وجود دارد یا خیر. در این مطالعهٔ موردی از رویکرد رایج استفاده می‌کنیم که محاسبهٔ آمارهٔ خی‌دوی پی‌یرسون یا، درمورد نمونه‌های کوچک، آزمون دقیق فیشر است. هر دوی این شیوه‌ها در فصل ۳ بحث شده‌اند. نکات اصلی در زیر خلاصه شده‌اند.

آزمون خی‌دو این آزمون را می‌توان در اغلب کتاب‌های آمار مقدماتی یافت (مثلاً نگاه کنید به اوت (۱۹۹۳)) و آزمونی برای فرض صفرِ نبودِ پیوند (استقلال) بین دو متغیر در اختیار می‌گذارد. از آزمون خی‌دو تنها زمانی باید استفاده کرد که اندازهٔ نمونه به میزان

قابل ملاحظه‌ای بزرگ باشد. در عمل، معیار رایج آن است که نباید بیش از ۲۰٪ خانه‌های جدول تعداد **مورد انتظاری** کمتر از ۵ داشته باشند.

آزمون دقیق فیشر این آزمون در بدو امر در حالت جدول پیشایندی ۲ × ۲ به‌کار گرفته شد: دو عامل رسته‌ای، هریک در دو سطح. با این حال، برای نمونه‌های کوچک، Pـمقدارها به‌شدت گسسته‌سازی خواهند شد و همواره این سؤال مطرح است که آیا تحلیل آماری نمونه‌های کوچک لزوماً معنی‌دار است یا خیر. آزمون دقیق فیشر برای استقلال را می‌توان به هر جدول پیشایندی $_{\text{ستون}}n \times {}_{\text{سطر}}n$ تعمیم داد، هرچند که این کار در برخی حالت‌ها با هزینهٔ محاسباتی نسبتاً قابل توجهی همراه است. این آزمون آزمونی نیست که بتوان آن را صرفاً برای مقاصد اکتشافی روی چندین جدول پیشایندی با اندازه‌های دلخواه اجرا کرد. آزمون دقیق فیشر در اگرستی (۱۹۹۰) بحث شده است.

دیگر آزمون‌ها آزمون‌ها و اندازه‌های پیوند متعدد دیگری وجود دارند که می‌توان از آن‌ها برای جدول‌های پیشایندی استفاده کرد. برای خلاصه‌ای مفصل‌تر، خواننده به بحث ارائه‌شده در بخش PROC FREQ راهنمای SAS (SAS ۱۹۹۰) ارجاع داده می‌شود.

۳.۱.۶ تعبیر معنی‌داری

مشکلی که در موارد تبعیض پیش می‌آید آن است که حتی اگر یافته‌ای را به لحاظ آماری معنی‌دار[۱] تلقی کنیم، «ثابت نمی‌شود» که تبعیض رخ داده است. ما بر این نکتهٔ مهم با سرمشق معروف زیر از استنباط آماری تأکید می‌ورزیم.

> پیوند آماری بین دو عامل دلالت بر این **نمی‌کند** که لزوماً رابطه‌ای علّی بین این عامل‌ها وجود دارد.

در تحلیل جدول پیشایندی، مضمون و تعبیر مسئله ممکن است به‌راحتی تحلیل را تحت تأثیر قرار دهد و در برخی موارد ممکن است پیوندهایی را مطرح کنند که گمراه‌کننده یا حتی بی‌معنی‌اند. بنابراین، یک نتیجهٔ معنی‌دار بر پایهٔ Pـمقداری از هر یک از آزمون‌های بالا را باید با احتیاط بسیاری تعبیر و تفسیر کرد. این مطلب به‌ویژه در وضعیت‌هایی از قبیل موارد تبعیض مهم است که در آن

۱. داعیهٔ مدرسهٔ هیزل‌وود (Hazelwood) علیه ایالات متحده [(1977) 433 U.S. 299, 311 n.17]، در دیوان عالی کشور: اگر اختلاف بین مقدار مورد انتظار و مقدار مشاهده‌شده بزرگتر از دو یا سه انحراف استاندارد باشد، آنگاه این فرض که کارمندان بدون توجه به نژاد استخدام شده‌اند در معرض تردید قرار خواهد گرفت.

یک داوری قضایی ممکن است پیامدهایی جدی داشته باشد. از یک منظر آماری، برخی نکات اصلی که ضروری است در انجام‌دادن تحلیل جدول‌های پیشایندی در نظر گرفته شوند عبارت‌اند از:

۱. آیا دو عامل به‌طور معنی‌داری به‌هم مرتبط‌اند؟

۲. عامل‌های اختلاط بالقوهٔ دخیل کدام‌ها هستند؟

۳. آیا زمانی که عامل‌هایی دیگر در تحلیل دخالت داده می‌شوند، هنوز هم پیوندی بین دو عامل وجود دارد؟

۴. تعبیری از یک پیوند معنی‌دار، در صورتی که عاملی دربردارندهٔ سطحی مانند «سایر» باشد، ممکن است مناسب نباشد.

۵. چه نوع مطالعه‌ای مطرح است؟ مطالعه‌های آینده‌نگر، گذشته‌نگر، یا مقطعی به بیش از صرفاً یک آزمون برای استقلال نیاز دارند.

۶. آیا شمارها وزنی برابر دارند؟ دو عامل ممکن است توضیح‌دهندهٔ اختلاف‌های مهم بین شرکت‌کنندگان در بررسی نباشند.

۷. تحلیل تا چه اندازه نسبت به تغییرات کوچک در شمارش خانه‌ها حساس است؟

۸. هنگام تبدیل متغیری پیوسته به یک عامل رسته‌ای، چه معیارهایی در تعیین نقاط انفصال به‌کار رفته است؟

مثال ۱.۶ پارادوکس سیمپسون[1]

مثال زیر پارادوکس سیمپسون را تشریح می‌کند که در آن سوی پیوند، با واردکردن عامل سومی در تحلیل، برعکس می‌شود. در مقاله‌ای به قلم رادلت (۱۹۸۱) (نگاه کنید به اگرستی (۱۹۹۰)، بخش ۲.۲.۵)، اثر خصیصه‌های نژادی بر این مطلب مطالعه شده است که آیا افرادی که به جرم آدمکشی محکوم شده‌اند مجازات اعدام دریافت می‌کنند یا خیر. یافته‌ها در جدول ۴.۶ داده شده است.

در کل، می‌توان ملاحظه کرد که $\frac{۱۹}{۱۶۰} = ۱۲٪$ متهمان سفید در برابر $\frac{۱۷}{۱۶۶} = ۱۰٪$ سیاهان مجازات اعدام دریافت کرده‌اند. با این حال، اگر پیوند بین نژاد متهم و رأی مجازات اعدام را اکنون نسبت به نژاد قربانی در نظر بگیریم، به نتیجه‌ای معکوس می‌رسیم. بدین معنی که وقتی قربانی سفیدپوست است، $\frac{۱۱}{۶۳} = ۱۷٫۵٪$ متهمان سیاه در قیاس با $\frac{۱۹}{۱۵۱} = ۱۲٫۶٪$ متهمان

1. Simpson's paradox

جدول ۴.۶ پارادوکس سیمپسون

نژاد متهم	نژاد قربانی	مجازات اعدام دریافت کرد	
		بلی	خیر
سفید	سفید	۱۹	۱۳۲
	سیاه	۰	۹
سیاه	سفید	۱۱	۵۲
	سیاه	۶	۹۷

سفیدپوست محکوم به اعدام شده‌اند. همچنین وقتی که نژاد قربانی سیاه است، پیوندی در همان جهت برقرار است (۰٪ برای متهمان سفیدپوست در مقابل $\frac{۶}{۱۰۳} = ۵٫۸٪$ برای متهمان سیاه‌پوست).

۴.۱.۶ تحلیل مقدماتی

برنامهٔ SAS زیر را می‌توان برای تحلیل جدول ۱.۶ به‌کار برد. خروجی این کد در زیر داده شده است. در اینجا، promote=1 نشان‌دهندهٔ داوطلبانی است که ارتقای شغلی یافته‌اند و promote=2 نشان‌دهندهٔ داوطلبانی است که ارتقا نیافته‌اند. به همین نحو، contrib=1 نشان‌دهندهٔ داوطلبانی است که به کاندیدای برنده کمک کرده‌اند و contrib=2 نشان‌دهندهٔ کسانی است که کمک نکرده‌اند. متغیر kount فراوانی‌ها را برای ترکیب‌های سطوح promote و contrib به‌دست می‌دهد.

کد *SAS* برای تحلیل جدول ۱.۶

```
data a;
   input promote contrib kount ;
   cards;
   1 1 6
   1 2 1
   2 1 0
   2 2 3
   ;
proc freq ;
   tables promote*contrib / chisq exact ;
   weight kount ;
endsas ;
```

خروجی از برنامه *SAS*

TABLE OF PROMOTE BY CONTRIB

Frequency / Percent / Row Pct / Col Pct	1	2	Total
1	6	1	7
	60.00	10.00	70.00
	85.71	14.29	
	100.00	25.00	
2	0	3	3
	0.00	30.00	30.00
	0.00	100.00	
	0.00	75.00	
Total	6	4	10
	60.00	40.00	100.00

STATISTICS FOR TABLE OF PROMOTE BY CONTRIB

Statistic	DF	Value	Prob
Chi-Square	1	6.429	0.011
Likelihood Ratio Chi-Square	1	7.719	0.005
Continuity Adj. Chi-Square	1	3.353	0.067
Mantel-Haenszel Chi-Square	1	5.786	0.016
Fisher's Exact Test (Left)			1.000
(Right)			0.033
(2-Tail)			0.033
Phi Coefficient		0.802	
Contingency Coefficient		0.626	
Cramer's V		0.802	

۵.۱.۶ خلاصه

این مطالعهٔ موردی دشواری‌های پیداکردن معیارهایی عینی راکه بتوان به کمک آن‌ها موارد تبعیض را ارزیابی کرد تشریح می‌کند. ممکن است چندین نتیجهٔ متضاد به‌دست آید که نتوان آن‌ها را به‌آسانی صرفاً بر اساس روش‌های آماری رفع و رجوع کرد. پرسش‌های زیر برخی از مباحثی را نشان می‌دهد که در یک گزارش تفصیلی باید به آن‌ها پرداخته شود.

پرسش‌ها

۱. از Pـمقدار حاصل از خروجی بالا چنین برمی‌آید که شواهدی برای رد فرض صفرِ استقلال وجود دارد. تعبیر شما از این نتیجهٔ معنی‌دار آماری چیست؟

۲. گرچه جدول ۳.۶ را می‌توان به سبکی مشابه تحلیل کرد، نتیجه‌ای معنی‌دار شواهدی دال بر تبعیض در اختیار نمی‌گذارد. چرا؟

۳. با رجوع به جدول ۲.۶ و با فرض اینکه هریک از موارد ارتقای شغلی به‌تصادف از بین سه داوطلب با بالاترین رتبه‌های موجود انتخاب شده‌اند، احتمال اینکه داوطلب D یکی از هفت نفر ارتقایافته نباشد، چقدر است؟

۴. از آنجا که این احتمال از کمترین مقدار استاندارد، یعنی ۵٪، بزرگتر است، خواهان باید استدلال کند که فرض «تصادفی‌بودن» ناموجه است. آیا استدلال خواهان معتبر است، و از آن دربارهٔ شیوهٔ انتخاب اعمال‌شده در ارتقای شغلی چه برمی‌آید؟

۵. برخی عامل‌های اختلاط بالقوه (متغیرهای دیگر) که در تحلیل‌های جدول‌های ۱.۶ و ۲.۶ به‌حساب آورده نشده‌اند، کدام‌اند؟ آیا توصیف شیوهٔ ارتقا کامل بوده است؟

۶. صرف‌نظر از نتایج آزمون آماری، جالب است که به فراوانی‌های بالا در دو خانهٔ اصلی در جدول ۳.۶ توجه کنیم. آیا از این امر چنین برنمی‌آید که ممکن است اختلافی در تعبیر زمینه‌ای دو رستهٔ نامعلوم وجود داشته باشد؟

۲.۶ مورد مرسولهٔ گمشده

روش‌ها	بررسی نمونه‌ای نمایه‌های جامعه
داده‌ها	فهرست کدهای پستی داده‌های مطالعه داده‌های ادارهٔ پست

۱.۲.۶ تحلیل بررسی نمونه‌ای

دربارهٔ یک توزیع‌کنندهٔ خاص مجله، که او را خانم ن مجله خواهیم نامید، کسی متوجه افزوده‌شدن صفری اضافه به عدد چهاررقمی مسیریابی که در کد پستی ایالات متحده (موسوم به عدد «Zip+4») به‌کار می‌رود نشد تا اینکه شکایتی از فردی دربارهٔ مقاله‌ای مندرج در آن مجله

دریافت شد. این شخص به وصول ۲۰ نسخه از یک شمارهٔ مجله معترض بود، به‌ویژه اینکه او هرگز حتی مشترک مجله هم نبوده!... به‌وضوح، چیزی سرجای خود نبود.

معلوم شد که تغییر در عدد مسیریابی منجر به آن شده بود که به‌ناچار چندین شمارهٔ مجله (به‌اشتباه) به نشانی واحدی ارسال شود. از آنجا که مجله در سراسر کشور توزیع می‌شد، توزیع‌کنندهٔ مجله علاقه‌مند بود که برآورد واقع‌بینانه‌ای از تعداد مجله‌هایی که تحویل نشده بود به‌دست آورد. برای انجام‌دادن این کار، یک بررسی تلفنی انجام شد تا مشخص شود که وقتی شمارهٔ مسیریابی اشتباه باشد چه اتفاقی می‌افتد. این اطلاعات برای به‌دست‌آوردن برآوردی از نسبت مرسوله‌های تحویل‌نشده مناسب بود.

روش‌شناسی

طراحی مطالعه در واقع مرحلهٔ حیاتی در این نوع تحلیل است. کیفیت داده‌هایی که گردآوری خواهند شد آشکارا به طرح مطالعه و بنابراین به هرگونه تحلیل آماری این داده‌ها بستگی دارد. از این رو مهم است که همان بار اول آن را درست به‌دست آورد. گرچه پرسش‌نامهٔ واقعی در این مطالعهٔ موردی بسیار ساده است، بسیاری از موضوعات مورد توجه در فصل ۳ (بررسی‌های نمونه‌ای را ببینید) اینجا مصداق پیدا می‌کنند. بحث زیر به شکل گزارشی در بردارندهٔ دستورکار و طراحی پرسشنامه برای بررسی پیشنهادی ارائه شده است. هدف بررسی تلفنی به‌دست‌آوردن اطلاعات کافی برای این منظور است که بتوان نسبت مرسوله‌های تحویل‌نشده از مجلهٔ مورد بحث را ارزیابی کرد.

۲.۲.۶ طرحوارهٔ بررسی

۱. برنامهٔ بررسی

این بررسی یک بررسی تلفنی خطاب به مدیران نامه‌بران آن اداره‌های پست است که از آن‌ها مرسوله‌هایی به نشانی نادرست تحویل شده است. از آنجا که کدهای پستی با اداره‌های پست متناظرند، واحدهای بررسی، اداره‌های پست‌اند. دستورکار زیر پیشنهاد می‌شود.

۱.۱ طبقه‌بندی

یک طبقه‌بندی از فهرست کدهای پستی تولید خواهد شد. هدف از این‌کار، پیشگیری از بیش نمونه‌گیری از آن اداره‌های پست است که تعداد کمی مجله توزیع می‌کنند که در اکثریت‌اند. افزون بر آن، این گروه‌بندی زیرجامعه‌ها را همگن‌تر خواهد ساخت زیرا ادارات پستِ با تعداد مشابهی ازاقلام، بیشتر شبیه هم خواهند بود.

متغیر گروه‌بندی: متغیر گروه‌بندی، تعداد اقلام با نشانی‌های اشتباه است که به هر ادارهٔ پست تحویل شده است.

روش گروه‌بندی: گروه‌ها چنان انتخاب خواهند شد که تعداد کل اقلام در هر گروه تقریباً یکسان باشد.

تعداد گروه‌ها: تعداد گروه‌ها به داده‌های مربوط به اقلام هر ادارهٔ پست بستگی خواهد داشت اما در هر حالت باید بین ۵ و ۱۰ باشند.

۲.۱ نمونه‌گیری تصادفی

با استفاده از نرم‌افزار S-PLUS، نمونه‌هایی تصادفی از کدهای پستی برای هر گروه تولید خواهد شد. برای هر گروه، نمونه مرکب از فهرستی از کدهای پستی با ترتیبی تصادفی خواهد بود.

اندازهٔ نمونه: اندازهٔ نمونه برای هر گروه تقریباً یکسان خواهد بود. با این کار، تعداد کدهای پستی لازم برای نمونه‌گیری به حداقل خواهد رسید. برای هر طبقه، نمونه‌ای از نسبت تحویل‌شده را مشاهده می‌کنیم که میانگین و واریانسی دارد. اگر واریانس‌های بین طبقه‌ها بسیار متفاوت باشند، آنگاه بهتر است اندازهٔ نمونه متناسب با ریشهٔ دوم واریانس باشد، اما اگر هیچ دانش پیشین یا شواهدی از این امر نداشته باشیم، آنگاه اندازه‌های نمونه باید برای همهٔ طبقه‌ها یکسان باشند.

۳.۱ خطاهای استاندارد

خاطر نشان می‌کنیم که هدف بررسی آن است که نسبت مرسوله‌های تحویل‌نشده را با انحراف استانداردی از ۲٫۵ تا ۵٪ اندازه‌گیری کنیم. این بدان معنی است که اندازهٔ نمونه باید بین ۱۰۰ تا ۴۰۰، بسته به مقدار نسبتی باشد که می‌خواهیم برآورد کنیم. ۱۰۰ متناظر با ۱۰٪ مرسوله‌های تحویل‌نشده است در حالی که ۴۰۰ متناظر با ۵۰٪ است.

۴.۱ نمونه‌گیری دو مرحله‌ای

به محض اینکه ۱۰۰ پرسشنامهٔ اول را به‌دست می‌آوریم، پاسخ‌ها را ارزیابی می‌کنیم و تصمیم می‌گیریم که آیا نمونه‌گیری بیشتری مورد نیاز است یا خیر.

۲. پرسشنامه

۱.۲ چه کسی باید به پرسشنامه پاسخ دهد

پرسشنامه باید به‌وسیلهٔ کسی که مسئول اداره‌کردن نامه‌بران است یا یک نامه‌بر ارشد پاسخ داده شود.

۲.۲ معرفی پرسشنامه

مصاحبه‌گر هدف از بررسی را به شکل زیر مطرح خواهد کرد.

> شرکت StatCon[1] بررسی‌یی را از طرف خانم ن مجله دربارهٔ مشکلی مربوط به مرسوله‌های تحویل‌نشده که طی ... رخ داده است انجام می‌دهد.

۱. خلاصه‌نامی ساختگی برای مشاورهٔ آماری-م.

۳.۲ پرسشنامه

سؤال ۱.
آیا با این مشکل خاص آشنا هستید؟ بلی/خیر

اگر پاسخ به سؤال ۱ خیر است به سؤال ۴ بروید.
اگر پاسخ به سؤال ۱ بلی است به سؤال ۲ بروید.

سؤال ۲.
آیا به خاطر می‌آورید که مرسوله را چه‌کار کرده‌اید؟ بلی/خیر

اگر پاسخ به سؤال ۲ خیر است به سؤال ۴ بروید.
اگر پاسخ به سؤال ۲ بلی است به سؤال ۳ بروید.

سؤال ۳.
اعداد مربوط به نسبت‌های زیر را تکمیل کنید.

(الف) تحویل به نشانی درست ————
(ب) مرجوع به فرستنده ————
(ج) دورریخته‌شده ————
پایان.

سؤال ۴.
اگر به چنان مرسوله‌ای برخورد کرده بودید، چه می‌کردید؟

(الف) تحویل به نشانی درست.
(ب) مرجوع به فرستنده.
(ج) دورریختن.
(د) نمی‌دانم.
پایان.

۳.۲.۶ تحلیل مقدماتی

داده‌های حاصل از این بررسی در دو فایل درج شده‌اند: `c62.survey.dat` مشتمل بر پاسخ‌های داده‌شده به بررسی است و `c62.surinfo.dat` مشتمل بر اطلاعاتی دربارهٔ نمونهٔ مورد استفاده است. متغیرهای موجود در `c62.surinfo.dat` عبارت‌اند از:

`group` = متغیر گروه‌بندی برای کد پستی ادارهٔ پست

`id` = کد پستی ادارهٔ پست

x = تعداد مجله‌های فرستاده‌شده به id

state = کد ایالتی ادارهٔ پست مربوط

نخستین کار این است که داده‌های حاصل از بررسی پردازش شوند. پاسخ‌های گمشده در فایل c62.survey.dat با نماد نقطه «.» کدبندی شده‌اند و باید بین آن‌ها و مقادیر گمشدهٔ واقعی تمایز قائل شد. در کد SAS زیر، هر گزینهٔ سؤال ۳ به‌عنوان متغیر جداگانه‌ای وارد شده است. این کار امکان آن را به ما می‌دهد که گزینه‌های با پاسخ‌های گمشده را به‌درستی با کد ۰ از نو کدبندی کنیم یا در حالتی که یک مقدار گمشدهٔ واقعی است آن را حذف کنیم. پس از پردازش داده‌ها، یک جدول‌بندی از فراوانی پاسخ‌ها صورت گرفت. خلاصه‌ای از نتایج در زیر ارائه شده است.

کد SAS برای پردازش داده‌های بررسی

```
data survey ;
  infile 'c62.survey.dat';
  input id q1 q2 q3a q3b q3c q4 $;
  if q3a eq . and ( q3b ne . or q3c ne . ) then q3a = 0;
  if q3b eq . and ( q3a ne . or q3c ne . ) then q3b = 0;
  if q3c eq . and ( q3a ne . or q3b ne . ) then q3c = 0;
  if q3b eq . and q3a eq . and q3c eq . and q4 eq " "
      then delete;
data infosur;
  infile 'c62.surinfo.dat';
  input group id x state $;
```

خلاصهٔ پاسخ‌ها به پرسشنامه

سؤال ۱: آیا از وقوع این مشکل خاص باخبرید؟

پاسخ	فراوانی	درصد
خیر	۳۶۱	۸۶٫۶
بله	۵۶	۱۳٫۴
کل	۴۱۷	۱۰۰

سؤال ۲: آیا به خاطر می‌آورید که مرسوله را چه‌کار کردید؟

پاسخ	فراوانی	درصد
خیر	۲۳	۴۱٫۱
بله	۳۳	۵۸٫۹
کل	۵۶	۱۰۰

سؤال ۳: چه درصدی...

(الف) به نشانی درست تحویل شد؟

پاسخ	فراوانی	درصد
۰	۱۰	۳۰٫۳
۵۰	۴	۱۲٫۱
۷۵	۲	۶٫۱
۸۰	۱	۳٫۰
۹۵	۲	۶٫۱
۹۹	۱	۳٫۰
۱۰۰	۱۳	۳۹٫۴
کل	۳۳	۱۰۰

(ب) به فرستنده مرجوع شد؟

پاسخ	فراوانی	درصد
۰	۲۹	۸۷٫۹
۵	۱	۳٫۰
۳۳	۱	۳٫۰
۱۰۰	۲	۶٫۱
کل	۳۳	۱۰۰

(ج) دور ریخته شد؟

پاسخ	فراوانی	درصد
۰	۱۶	۴۸٫۵
۱	۱	۳٫۰
۵	۱	۳٫۰
۲۰	۱	۳٫۰
۲۵	۲	۶٫۱
۵۰	۴	۱۲٫۱
۶۶	۱	۳٫۰
۱۰۰	۷	۲۱٫۲
کل	۳۳	۱۰۰

سؤال ۴: اگر به چنان مرسوله‌ای برخورد کرده بودید چه می‌کردید؟

(الف) تحویل به نشانی‌های درست.

(ب) مرجوع به فرستنده.
(ج) دوررختن.
(د) نمی‌دانم.

پاسخ	فراوانی	درصد
(الف)	۳۰۹	۸۰٫۵
(ب)	۱۱	۲٫۹
(ج)	۵۳	۱۳٫۸
(د)	۱۱	۲٫۹
کل	۳۸۴	۱۰۰

جدول‌های بالا خلاصه‌ای از پاسخ‌ها را در بررسی نشان می‌دهند. پاسخ‌ها به سؤال‌های ۱ و ۲ حاکی از آن است که تنها ۵۶ نفر از سرپرستان نامه‌بران از مشکلات باخبر بودند و از بین آن‌ها فقط ۳۳ نفر به یاد داشتند که با مرسوله‌ها چه کرده‌اند. در نتیجه، اطلاعات حاصل از سؤال ۴ بسیار اهمیت پیدا می‌کند. از سوی دیگر، اطلاعات حاصل از سؤال ۴ محتملاً کم‌دقت‌تر از اطلاعات حاصل از سؤال ۳ است. در تحلیل نهایی باید این مطلب را در نظر گرفت.

۴.۲.۶ خلاصه

در این مطالعهٔ موردی، فرایند اجرای یک بررسی را با شروع از طراحی بررسی و پرسشنامه توصیف کردیم و آن را با گردآوری و تحلیل داده‌ها دنبال کردیم. پرسش‌های زیر برخی از مراحلی را که می‌توان برای تولید گزارش نهایی دنبال کرد مطرح می‌کنند.

پرسش‌ها

۱. قدم بعدی محاسبهٔ نسبت مرسوله‌های تحویل‌شده است. برای انجام‌دادن این کار، متغیری مانند qtot را می‌توان به‌صورت زیر ایجاد کرد:

```
if q3a ne . then  qtot = q3a ;
else if q4 eq 'a' then qtot = 100;
else if q4 eq 'b' or q4 eq 'c' then  qtot = 0;
else if q4 eq 'd' then  qtot = .;
```

۲. پس از ادغام دو مجموعه‌داده، نتیجه می‌شود که درصد برآوردشدهٔ مجله‌های تحویل‌شده (برحسب کدپستی) عبارت است از x * qtot/100.

۳. فرض کنید که ntot نشان‌دهندهٔ نتیجهٔ بالا باشد. می‌توان از آن برآوردی برای تعداد کل مجله‌ها محاسبه کرد.

۴. با انجام‌دادن این کار برحسب group می‌توان تغییرپذیری بین درصدهای تحویل را ملاحظه کرد.

۵. یک گزارش شامل جدول خلاصه‌ای خواهد بود که به‌صورت بالا ساخته می‌شود و اختلاف‌های بین ntot و qtot درصدهای تحویل کلی در آن مورد بحث قرار می‌گیرد.

۶. راهی دیگر برای رویکرد به این مسئله، محاسبهٔ برآوردهای جداگانهٔ درصد تحویل‌شده به‌ترتیب با استفاده از پاسخ‌های سؤال‌های ۳ و ۴ است. چگونه دو برآورد را با هم ترکیب می‌کنید تا برآوردی نهایی تولید کنید؟

۳.۶ دستگاهی برای کاهش آلاینده‌های موتور

روش‌ها	آزمون‌های t تحلیل واریانس
داده‌ها	۱۳ روز آزمایش روی زن

۱.۳.۶ آزمون‌کردن ادعای یک تولیدگر

تولیدگران اغلب «ادعا می‌کنند» که محصول آن‌ها بر مبنای آزمون‌های آزمایشگاهی (در صورتی که مطابق دستورات از آن‌ها استفاده شود)، از عهدهٔ کار معینی بر خواهد آمد. با این حال، ممکن است سؤال‌هایی دربارهٔ روش‌های آزمون مورد استفادهٔ تولیدگران مطرح باشد. این مطلب ممکن است ناشی از نتایج متناقض حاصل از آزمون‌های مستقل باشد، یا ممکن است شواهد آماری کافی برای تأیید ادعای تولیدگر در دست نباشد.

در این مطالعهٔ موردی، تحقیق در مورد ادعایی مطرح‌شده از سوی تولیدگری را ارائه می‌کنیم مبنی بر اینکه وقتی دستگاه وی روی موتور یک وسیلهٔ نقلیهٔ موتوری نصب شود، موجب کاهش (تا حد صفر) انتشار مستمر هیدروکربن و مونوکسیدکربن (CO) و افزایش مستمر دیوکسیدکربن (CO_2) خواهد شد. برای بررسی صحت این ادعا، آزمون‌هایی توسط StatCon Enterprizes با استفاده از دستگاه شرکت روی موارد زیر انجام شد:

۱. یک لیموزین یوگو متعلق به تولیدگر؛

۲. یک بنتلی[1] متعلق به تیم بازبینی تحقیق مستقل.

1. Bentley

جدول ۵.۶ متغیرها در c63.dat

متغیر	تعریف
HC	هیدروکربن (جزء در هر میلیون)
CO	مونوکسیدکربن (٪ حجم)
CO2	دیوکسیدکربن (٪ حجم)
DAY	روز آزمایش: ۱، ۲، ...، ۱۳
REP	تکرار: ۱، ۲، ۳، ۴
DEV	ابزار: ۱ر۰ (نصب شده/نشده)

در اینجا با استفاده از خودروی تولیدگر توجه خود را به تحلیل معطوف می‌کنیم. آزمایش دوم به‌عنوان یک تمرین مطالعهٔ موردی در فصل ۹ به عهدهٔ خواننده واگذار شده است.

توصیف داده‌های خودرو

در این مطالعه، آلاینده‌های هیدروکربن (جزء در میلیون[1])، دیوکسیدکربن (٪ حجم)، و مونوکسیدکربن (٪ حجم) برای خودروی متعلق به تولیدگر اندازه‌گیری شد. خودرو تحت شرایط یکسان در ۱۳ روز مختلف راه برده شد و در هر یک از روزهای آزمون، آزمایش ۴ بار تکرار شد. دستگاه تا پایان روز دوم اندازه‌گیری‌ها روی خودرو نصب نشد و پس از آن تا ۹ روز دیگر روی خودرو باقی‌ماند و در این مدت اندازه‌گیری‌ها انجام شد. سپس دستگاه از خودرو برداشته شد و ۲ مجموعه‌اندازه‌گیری دیگر انجام شد. کامل‌شدن این مطالعه حدود شش‌ماه طول کشید. جدول ۵.۶ خلاصه‌ای از متغیرهای مورد نظر در مجموعه‌داده‌های c63.dat را به‌دست می‌دهد.

روش‌شناسی

برای تحلیل این داده‌ها می‌توان چندین رویکرد را در پیش گرفت. هدف اصلی این مطالعه آن است که ببینیم آیا اثری ناشی از DEV روی همه یا یکی از سه آلاینده (متغیرهای پاسخ) وجود دارد یا خیر. رویکردی ساده این است که از یک آزمون t ی دونمونه‌ای برای تحلیل اثر DEV بر انتشار هر آلاینده استفاده کنیم. گرچه، در این رویکرد فرض می‌شود که عامل‌های DAY و

1. ppm

REP به لحاظ آماری معنی‌دار نیستند. برای منظورکردن اثر این عوامل می‌توان از مدل تحلیل واریانس استفاده کرد. جزئیات مربوط به آزمون‌های t و مدل‌های ANOVA در فصل ۳ آمده‌اند.

۲.۳.۶ آزمون‌های t

اثر DEV را می‌توان با استفاده از آزمون tی دونمونه‌ای برای هر آلاینده تحلیل کرد. با این کار فرض می‌شود که می‌توان اثرهای DAYو REP را نادیده گرفت که نیاز به وارسی خواهد داشت.

SAS به‌طور خودکار صورت‌های برابر و نابرابر آزمون tی دونمونه‌ای را در اختیار می‌گذارد. نتیجهٔ نابرابر را زمانی باید در نظر گرفت که انحراف‌های استاندارد به‌میزان فاحشی از هم متفاوت باشند (مثلاً یکی چهار برابر دیگری باشد). از آنجا که آزمون t در حالت واریانس نابرابر تنها به‌صورت تقریبی است، نتایج را باید با احتیاط تعبیر کرد و با آزمون‌های ناپارامتری از قبیل آزمون رتبهٔ علامت‌دار ویلکاکسون مقایسه کرد. پیوست ج جزئیاتی از این آزمون‌ها را در اختیار می‌گذارد.

۳.۳.۶ تحلیل واریانس

اثرهای عامل‌های DAY و REP را می‌توان با واردکردن آن‌ها در یک مدل ANOVA برای هر آلاینده بررسی کرد. می‌توان اثرهای متقابل را نیز افزود، گرچه ممکن است در صورت کاسته‌شدن بیش از اندازهٔ درجه‌های آزادی خطا نتایج نامربوطی به‌دست آید.

از آنجا که تکمیل‌کردن این مطالعه سه ماه به طول انجامید، از یک اثر معنی‌دار DAY چنین بر خواهد آمد که نوعی عامل «زمانی» دخالت داشته است. یک اثر معنی‌دار REPممکن است به‌طور بالقوه ناشی از ترتیب تکرار آزمایش‌هایی باشد که در هر روز از آزمون انجام شده‌اند. لازم است، برای تعیین ماهیت اثر معنی‌داری، تحلیل‌های پساآزمایشی انجام شوند.

۴.۳.۶ تحلیل مقدماتی

ما علاقه‌مند به برآورد اثر دستگاه بر میانگین آلاینده‌ها هستیم، به این معنی که آیا تفاوت‌هایی بین آلاینده‌ها در روزهایی که دستگاه نصب شده بود و روزهایی که دستگاه نصب نشده بود وجود دارد یا خیر. در جدول ۶.۶ داده‌های میانگین آلاینده‌ها برای این خودرو خلاصه شده و جدول ۷.۶ دربردارندهٔ نتایج آزمون tی دونمونه‌ای برای هر آلاینده است. خروجی لازم برای این جدول‌ها از کد SAS زیر تولید شده‌اند.

جدول ۶.۶ میانگین آلاینده‌ها برای خودروی تولیدگر

ابزار	روز	هیدروکربن (جزء در میلیون)	CO۲(٪)	CO(٪)
وجود دارد	۱	۱۵٫۳۰	۱۵٫۸۳	۰٫۲۵
	۲	۱۰٫۷۵	۱۴٫۴۱	۰٫۱۱
وجود دارد	۳	۵٫۴۶	۱۴٫۸۵	۰٫۱۱
	۴	۶٫۰۴	۱۳٫۴۸	۰٫۱۶
	۵	۸٫۶۶	۱۴٫۶۲	۰٫۰۸
	۶	۱۳٫۱۸	۱۳٫۳۸	۰٫۱۳
	۷	۲۵٫۸۹	۱۳٫۷۵	۰٫۲۱
	۸	۴٫۴۵	۱۵٫۹۵	۰٫۲۰
	۹	۱۳٫۱۸	۱۵٫۵۲	۰٫۱۰
	۱۰	۱٫۱۱	۱۵٫۴۲	۰٫۱۱
	۱۱	۱۲٫۷۱	۱۴٫۸۵	۰٫۱۴
وجود ندارد	۱۲	۱۱٫۳۱	۱۵٫۶۱	۰٫۱۳
	۱۳	۱۳٫۵۴	۱۴٫۲۴	۰٫۱۵
وجود ندارد	متوسط	۱۲٫۷۲	۱۵٫۰۲	۰٫۱۶
وجود دارد	متوسط	۱۰٫۰۷	۱۴٫۶۵	۰٫۱۴

کد SAS برای تحلیل آزمون t

```
data a ;
      infile 'c63.dat' ;
      input day hc co2 co ;
      dev=0 ;
      if days < 3 or day > 11 then dev=1 ;
proc means ;
      var hc co2 co ;
      by day ;
proc ttest ;
      class dev ;
      var hc co2 co ;
```

از نتایج حاصل از این آزمون tی دونمونه‌ای چنین برمی‌آید که شواهد کافی در دست نیست که نتیجه بگیریم یک تفاوت معنی‌دار آماری در میانگین آلاینده‌ها نسبت به وجود یا نبود دستگاه وجود دارد.

جدول ۷.۶ نتایج آزمون t برای خودروی تولینگر

متغیر: HC

DEV	N	میانگین	انحراف استاندارد	خطای استاندارد
وجود ندارد	۱۶	۱۲٫۷۲۴	۴٫۱۷۲	۱٫۰۴۳
وجود دارد	۳۶	۱۰٫۰۷۶	۷٫۷۶۷	۱٫۲۹۴

واریانس‌ها	T	DF	$prob > \|T\|$
نابرابر	۱٫۵۹۳	۴۸٫۰	۰٫۱۱۷۸
برابر	۱٫۲۷۹	۵۰٫۰	۰٫۲۰۶۸

متغیر: CO_2

DEV	N	میانگین	انحراف استاندارد	خطای استاندارد
وجود ندارد	۱۶	۱۵٫۰۲۲	۱٫۲۱۱	۰٫۳۰۳
وجود دارد	۳۶	۱۴٫۶۴۷	۱٫۲۶۲	۰٫۲۱۰

واریانس‌ها	T	DF	$prob > \|T\|$
نابرابر	۱٫۰۱۶	۳۰٫۰	۰٫۳۱۷۸
برابر	۱٫۰۰۰	۵۰٫۰	۰٫۳۲۲۲

متغیر: CO

DEV	N	میانگین	انحراف استاندارد	خطای استاندارد
وجود ندارد	۱۶	۰٫۱۵۹	۰٫۰۹۹	۰٫۰۲۵
وجود دارد	۳۶	۰٫۱۳۸	۰٫۰۷۴	۰٫۰۱۲

واریانس‌ها	T	DF	$prob > \|T\|$
نابرابر	۰٫۷۷۴	۲۲٫۸	۰٫۴۴۶۸
برابر	۰٫۸۶۵	۵۰٫۰	۰٫۳۹۱۰

۵.۳.۶ خلاصه

تولیدگر مدعی شد که انتشار آلاینده‌های هیدروکربن و CO (CO_2) پس از نصب دستگاه روی موتور، با روشن‌شدن موتور (یا افزوده‌شدن بر مسافت طی‌شده) و پس از یک دورهٔ اولیهٔ پایدارسازی (۱۰٪ مقدار مسافت طی‌شده توسط خودرو از آغاز)، باید به‌طور یکنواخت کاهش (افزایش) یابد. با این‌حال، با نگاهی به جدول ۶.۶ روشن می‌شود که این داده‌ها دال بر صحت این ادعا نیستند. تولیدگر همچنین مدعی شد که به محض اینکه این دستگاه بر موتور نصب شود، موتور به‌طور غیر قابل برگشتی تغییر می‌کند! اگر دستگاه برداشته شود، اثرهای آن باقی می‌ماند.

باید توجه کنیم که تغییرپذیری در اندازه‌گیری‌های آلاینده‌ها، که در اینجا مورد تحلیل قرار گرفت، ممکن است ناشی از عامل‌های دیگری از قبیل هوا و طول مدتی باشد که موتور بین آزمون‌ها بدون کار مانده است، عامل‌هایی که در آزمایش کنترل نشده بودند. به علاوه، کمیت آلاینده‌های هیدروکربن اندازه‌گیری‌شده، در اغلب موارد، کمتر از دقت وسیلهٔ اندازه‌گیری هستند. چند موضوع دیگر در زیر ارائه می‌شوند.

پرسش‌ها

۱. از یک مدل ANOVA برای وارسی اینکه آیا عامل‌های DAY یا REP مهم‌اند یا خیر استفاده کنید.

۲. دقت وسیلهٔ اندازه‌گیری ۱۲± جزء در میلیون برای هیدروکربن‌ها، ۰٫۵±٪ برای CO_2 و ۰٫۶±٪ برای CO است. این امر چه تأثیری بر تحلیل آماری دارد؟

۳. تأثیر احتمالی موارد زیر را بر نتایج آزمایشی مورد بحث قرار دهید.

- روزهای آزمون لزوماً پشت‌سرهم نبودند.
- خودرو تقریباً ۷۰۰ تا ۲۰۰۰ مایل بین آزمون‌ها راه برده شد.
- آزمون‌کردن در نوامبر شروع شد و در ماه می همان سال خاتمه یافت.
- پس از ۷ روز شمع جدیدی روی خودرو نصب شد.

۴. پیامد بالقوهٔ ناشی از این حقیقت که وسیلهٔ اندازه‌گیری در هر روز اندازه‌گیری از تو تنظیم شود چیست؟

۵. برای بررسی این ادعای تولیدگر که، به محض اینکه دستگاه روی موتور نصب شود موتور به‌طور غیر قابل بازگشتی تغییر می‌یابد، می‌توانیم یک آزمون tی دونمونه‌ای برای برابری میانگین آلاینده‌ها با استفاده از داده‌های حاصل از روزهای ۱ و ۲ در مقابل داده‌های روزهای ۱۲ و ۱۳ انجام دهیم. آیا شواهدی برای تأیید این ادعا وجود دارد؟

۶. رویکردی دیگر، در نظرگرفتن هر سه آلاینده به‌طور همزمان است. این کار منجر به آمارهٔ T^2ی هتلینگ یا تحلیل داده‌ها با استفاده از یک مدل MANOVA می‌شود. مزیت بالقوهٔ این رویکرد آن است که همبستگی بین پاسخ‌ها به‌حساب آورده می‌شود که همان چیزی است که مطابق ادعای تولیدگر انتظار آن می‌رود. آیا رویکرد چندمتغیره تفاوتی ایجاد می‌کند؟

۴.۶ روان‌شناسی معکوس

روش‌ها	تحلیل پیمایشی ارزیاب همبستگی
داده‌ها	ابزار PANSS ۳۰ رستهٔ ارزیابی ۷۲ پاسخگو

۱.۴.۶ یک آزمایش مشاهده‌ای

نتایج یک مطالعهٔ بزرگ دربارهٔ جلوگیری از خودکشی حاکی از شواهدی است مبنی بر اینکه یک داروی ضد روان‌پریشی درکمک به جلوگیری از خودکشی در برخی بیماران مبتلا به شیزوفرنی مؤثر بوده است. با این حال، برای مؤثر بودن این درمان، لازم است که پزشک قادر باشد وضعیت روانی و قصد خودکشی بیمار را به‌دقت ارزیابی کند. برای این منظور، می‌توان چندین نوع مقیاس ارزیابی را به‌عنوان ابزارهای اندازه‌گیری مورد استفاده قرار داد. در این مطالعهٔ موردی، نتایج حاصل ازکارگاه یک روزه‌ای را که برای ابزار PANSS (مقیاس سندروم مثبت و منفی [1]) ترتیب داده شده بود، در نظر می‌گیریم.

شرح مختصر کارگاه:
درکل ۷۲ پزشک از جامعهٔ بین‌المللی در این کارگاه شرکت کردند. مؤلفه‌های اصلی جلسهٔ آموزشی عبارت بودند از:

۱. پزشکان ابتدا با ابزار PANSS آشنایی پیدا کردند که ۳۰ نشانهٔ روان‌شناختی را در فهرست داشت (نگاه کنید به جدول ۸.۶): ۷ مورد به‌عنوان منفی، ۷ مورد به‌عنوان مثبت، و باقی ۱۶ مورد به‌عنوان عام تخصیص یافتند. یک مقیاس ۱ (کم) تا ۷ (زیاد) برای ارزیابی هر نشانه مطابق با درجهٔ روان‌پریشی ظهوریافته از طرف هر بیمار برای آن نشانهٔ خاص به‌کارگرفته شد.

1. positive and negative syndrome scale

۲. مرحلهٔ بعدی کارگاه در برگیرندهٔ همهٔ پزشکان بود که یک ویدئوی ۳۰ دقیقه‌ای از یک بیمار را مشاهده می‌کردند و از ابزار PANSS برای ارزیابی نشانه‌های روان‌پریشی بیمار استفاده می‌کردند.

۳. ابزار PANSS به سه زبان در دسترس بود و در خاتمهٔ ویدئو، ارزیابی پزشک با ارزیابی یک «خبره»، که به آن ارزیابی کلیدی اطلاق می‌شد، مقایسه می‌شد. پس از استراحتی کوتاه، کارگاه با ارائهٔ نتایج پایان یافت.

داده‌ها

جدول ۸.۶ فهرستی از ۳۰ نشانهٔ روان‌شناختی را ارائه می‌کند که طی بخش ویدئویی کارگاه ارزیابی شده بودند. هر نشانه بر مبنای مقیاسی از ۱ (کم) تا ۷ (زیاد) به‌وسیلهٔ پزشکانی که تنها با یک کد RATER [ارزیاب] و انتخاب زبان آن‌ها (LANG) مشخص شده بودند ارزیابی شد. داده‌های کارگاه PANSS و ارزیابی‌های کلیدی انجام‌شده توسط فرد خبره (که با RATER=0 شناخته می‌شود) در مجموعه‌داده‌های c64.dat گنجانده‌شده‌اند. فرمت این مجموعه‌داده‌ها در زیر نشان داده شده است.

```
RATER  LANG  P1 P2 ... P7  N1  N2 ... N7  G1  G2 ... G15 g16
  0      E   4  2 ... 2   5   4 ... 2   3   4 ... 3    1
 101     F   4  1 ... 1   5   3 ... 3   3   5 ... 2    5
  :      :
172      E   3  2 ... 2   5   3 ...  2  2   4 ...  4   3
```

همان‌گونه که از این خلاصه ملاحظه می‌شود، نخستین سطر c64.dat در واقع «عنوان ستون‌ها» است که درایه‌های کاراکتری آن‌ها متناظر با نام‌های متغیرهای داده‌شده در جدول ۸.۶ هستند. این سطر خاص برای واردکردن داده‌ها به S-PLUS سودمند خواهد بود زیرا قراردادن گزینهٔ header=T در تابع read.table() در S-PLUS این نام‌ها را قادر می‌سازد که به ستون‌های data frame حاصل تخصیص داده شوند. در دیگر نرم‌افزارهای آماری، لازم است که پیش از واردکردن داده‌ها از این سطر صرف‌نظر یا آن را حذف کنیم.

روش‌شناسی

نوع تحلیل آماری لازم در مطالعه‌های مشاهده‌ای اغلب نیازمند چیزی بیش از فراهم‌کردن خلاصهٔ توصیفی خوبی از یافته‌ها است. جدول‌های خلاصه، نمایش‌های نموداری، و تحلیل‌های آماری ساده، روش‌های مؤثری هستند که آماردان مشاور می‌تواند برای این منظور به‌کار بگیرد.

جدول ۸.۶ تعریف‌های متغیر PANSS

متغیر	تعریف	
RATER	کد شناسایی ارزیاب RATER = 0 ارزیاب خبره است	
LANG	زبان مورد استفاده E=انگلیسی، F=فرانسه، I=ایتالیایی	
P1	هذیان	نشانه‌های مثبت
P2	درهم‌ریختگی فکری	
P3	رفتار توهمی	
P4	برانگیختگی	
P5	خودبزرگ‌بینی	
P6	سوءظن/احساس آزار و ایذا	
P7	عناد	
N1	افت هیجانی	نشانه‌های منفی
N2	گوشه‌گیری هیجانی	
N3	ضعف ارتباطی	
N4	گوشه‌گیری اجتماعی با بی‌تفاوتی/انفعالی	
N5	دشواری در تفکر انتزاعی	
N6	ضعف خودانگیختگی	
N7	تفکر قالبی	
G1	جسمانی	نشانه‌های عمومی
G2	اضطراب	
G3	احساس گناه	
G4	تنش	
G5	اطوارگرایی و ژست‌گیری	
G6	افسردگی	
G7	کندی حرکتی	
G8	فقدان روحیهٔ همکاری	
G9	افکار غیرعادی	
G10	گم‌گشتگی	
G11	بی‌توجهی	
G12	ضعف داوری و بینش	
G13	اختلال اراده	
G14	ضعف در کنترل قوای محرک آنی	
G15	شیفتگی	

البته این بدان معنی است که فرض می‌کنیم وقت کافی برای پردازش داده‌ها، انجام‌دادن تحلیل، و نیز گردآوردن نتایج در قالب فرمت ارائه‌کرد مناسبی برای مشتری خود داشته باشیم... مثلاً در حدود ۲۰ دقیقه؟ به‌عبارت دیگر، طی زمان تنفس «کوتاه»ی در کارگاه PANSS!

۲.۴.۶ آمار: بازاری با هر نوع کالا

قراردادن «فی‌الفور» بازخورد در اختیار پزشکان دربارهٔ نحوهٔ عملکرد آن‌ها با ابزار PANSS مستلزم برنامه‌ریزی خوب و آمادگی قبلی است. از منظر مشاور آماری، تفاوت اصلی در این تحلیل آن است که همه‌چیز ــ از پردازش داده‌ها تا ارائهٔ نتایج ــ نیاز به برنامه‌ریزی قبلی دارد. اما، بدون داده‌ها! آشکار است که این از آن قبیل رویکردها نیست که بدون درک درستی از مسئله سعی در انجام‌دادن آن داشته باشیم. به برخی از موضوعات اصلی دخیل در پیش‌برنامه‌ریزی یک تحلیل آماری می‌پردازیم.

پردازش داده‌ها ضروری است که ساختار و فرمت داده‌هایی که باید گردآوری شوند نسبتاً ساده باشند تا مرحلهٔ پردازش داده‌ها به‌طرز کارآمدی کامل شود. از بحث ما در فصل ۳، بخش ۲.۳، نتیجه می‌شود که

- ایجاد فرمتی خاص برای ورود داده‌ها اساسی است. مقادیر گمشده را باید به‌طور مناسب شناسایی کرد به‌طوری که بتوان در مرحلهٔ تحلیل به نحو مناسب به آن‌ها پرداخت. به همین نحو، مدخل‌های تاریخ [1] و ساعت (در صورت وجود) باید به‌درستی تعبیر شوند.
- اشکال عمدهٔ یک تحلیل از پیش برنامه‌ریزی‌شده آن است که کیفیت داده‌ها نامعلوم است. اینجاست که یک تحلیلِ از پیش‌برنامه‌ریزی‌شده احتمالاً به شکست می‌انجامد؛ به‌ویژه اگر بازبینی‌های معمول خطا و کیفیت داده‌ها را در برنامهٔ تحلیل از قلم بیندازیم.

در این مثال، مجموعه‌داده‌ها مرکب از یک متغیر گروه‌بندی رسته‌ای (LANG)، و ۳۰ متغیر ترتیبی متناظر با ارزیابی‌های PANSS است. از آنجا که این نوع از داده‌ها را می‌توان به‌راحتی از طریق یک دستگاه اسکن‌کننده وارد کرد، کیفیت داده‌ها نباید مشکل‌ساز شود.

آزمون‌کردن برنامه ضرورت دارد که برنامهٔ تحلیل به‌دقت با استفاده از داده‌های صوری آزمون

۱. با اینکه مشکل Y2K (منظور مشکل تغییر سال ۲۰۰۰ یا کامپیوترها بودـم.) چندان جدی از کار در نیامد، امروزه بسته‌های نرم‌افزار آماری از قبیل SAS یک تاریخ مقطع برای فرمت‌های سالی دو رقمی بکار می‌برند. بنابراین، کسی که در ۱۹۲۱ متولد شده ممکن است نسبت به سن خود جوان به نظر برسد (به دلیل مشتبه‌شدن ۱۹۲۱ با ۲۰۲۱ـم.)، به‌ویژه آن‌هایی که در واقع هنوز به دنیا نیامده‌اند!

شود. به‌ویژه، باید اثر مقادیر گمشده را بازبینی کرد زیرا ممکن است منجر به بروز مشکلات و خطاهای غیرمترقبه در کاری شود که در غیر این صورت یک برنامهٔ تحلیلِ به‌درستی کدبندی‌شده خواهد بود. در اینجا توجه خود را به S-PLUS معطوف و فرض می‌کنیم که دستگاه اسکن‌کننده مجموعه‌داده‌هایی تولید خواهد کرد که می‌توان آن‌ها را با S-PLUS خواند. اطلاعات بیشتر دربارهٔ S-PLUS را می‌توان در پیوست ب.۲ یافت.

- گزینهٔ `na.rm=T` در تابع‌هایی نظیر `mean(x)` و `sum(x)` برای حذف مقادیر گمشده پیش از محاسبهٔ بردار `x` ضروری است، در غیر این صورت این تابع‌ها در صورتی که هر مؤلفهٔ `x` گمشده باشد، `NA` (گمشده) را بازمی‌گردانند.
- مقادیر گمشده همچنین ممکن است «عبارت‌های منطقی» را نامعتبر کنند. عبارت‌های منطقی در گزاره‌هایی از قبیل `if (a && b) { ... }` ظاهر می‌شوند که وقتی عبارت منطقی `"a && b"` (`a` و `b`) به‌عنوان `TRUE` ارزیابی شود اجرا می‌شوند. اگر `"a"` گمشده باشد، عبارت منطقی را نمی‌توان به‌صورت `TRUE` یا `FALSE` ارزیابی کرد و خطایی رخ می‌دهد.
- یک جلسهٔ S-PLUS نوعاً شامل ایجاد تابع‌هایی به‌وسیلهٔ کاربر است که برخی وظایف مرتبط با تحلیل را انجام می‌دهد. هر جا که عملی باشد، این تابع‌ها باید به‌صورت عام نوشته شوند و باید گزینه‌های پیش‌فرض به‌جای کمیت‌های کد اختصاصی[1] به‌کار گرفته شوند.

S-PLUS رشته‌ای جامع از تابع‌های آماری و نموداری در اختیار می‌گذارد، بنابراین اغلب تابع‌هایی که ایجاد می‌کنیم معمولاً با دستکاری داده‌ها و تنظیم گزینه‌ها برای یک شیوهٔ آماری یا نمایش نموداری خاص سروکار پیدا می‌کند. در بخش تحلیل مقدماتی این مطالعهٔ موردی، برخی از مؤلفه‌هایی را که در یک برنامهٔ تحلیل S-PLUS برای مجموعه داده‌های PANSS گنجانده می‌شود، به‌طور خلاصه شرح داده‌ایم.

تولید خروجی در خروجی تولیدشده به‌وسیلهٔ برنامهٔ تحلیل، نیاز به آن است که اهداف مطالعه برآورده شوند. معمولاً اهداف مطالعه، با استفاده از رویکردی تکرارشونده به تحلیل، مورد تحقیق قرار می‌گیرد. این بدان معنی است که یک تحلیل مقدماتی برای شناسایی نتایج مهم یا جالب انجام می‌شود که به نوبهٔ خود منجر به تحلیل مجدد می‌شود. در یک برنامهٔ

1. *hard-coded*

از پیش‌برنامه‌ریزی‌شده، فرصت تحلیل مجدد محدود است و اهداف مطالعه به‌طور اساسی لزوماً باید به کمک خروجی اولیهٔ تولیدشده به‌وسیلهٔ برنامهٔ تحلیل برآورده شوند. اتخاذ دو تصمیم مهم نسبت به خروجی تولید شده به‌وسیلهٔ برنامهٔ تحلیل، ضروری است.

- نتایج خاصی که باید تولیدشوند کدام‌اند؟
- چه اجزایی از برنامه نیاز به انعطاف‌پذیری دارند؟

برای مطالعهٔ PANSS، تصمیم‌گیری دربارهٔ نتایج خاصی که باید تولید شوند حائز اهمیت است، زیرا ۳۰ ارزیابی و یک متغیر گروه‌بندی موسوم به LANG وجود دارد. نتیجه می‌شود که تعداد نسبتاً قابل توجهی از جدول‌های فراوانی چندطرفه، همبستگی‌ها، آزمون‌های t، و آزمون‌های ناپارامتری را می‌توان از این متغیرها تولید کرد. به همان کثرت می‌توان به نمایش‌های نموداری نیز دست یافت. با اینکه همهٔ این نتایج را می‌توان بعداً به‌تفصیل بازبینی کرد، برای دستیابی به هدف فراهم‌آوردن بازخورد برای پزشکان تنها باید مناسب‌ترین آماره‌های خلاصه و نمایش‌ها را تولید کرد.

پرسش دوم این موضوع را مطرح می‌کند که برنامهٔ تحلیل S-PLUS چگونه باید تدوین شود. این کار ممکن است به شخصی که واقعاً تحلیل را در کارگاه انجام می‌دهد بستگی داشته باشد. به‌عنوان مثال، کل برنامه را می‌توان در یک تابع مانند `run.anx()` قرار داد که همهٔ خروجی‌ها را تولید و آن را فرمت می‌کند (بعداً مورد بحث قرار می‌گیرد). این رویکرد هیچ انعطاف‌پذیری‌ای ندارد، اما می‌توان از آن برای تولید همهٔ خروجی‌های اولیه‌ای استفاده کرد که شخص می‌تواند آماده‌کردن ارائه‌کرد را با آن شروع کند. در همان حال، «تحلیلگر» کارگاه می‌تواند برخی اجزای برنامه را با گزینه‌های مختلفی از قبیل مقایسهٔ قبول/رد ارزیابی پزشکان با ارزیابی‌های فرد خبره مجدداً اجرا کند. ما به این نکتهٔ آخر در بخش تحلیل مقدماتی بازمی‌گردیم.

ارائه‌کرد خروجی تولیدشده به‌وسیلهٔ برنامهٔ تحلیل قطعاً نیازمند به شکلی از «بازبسته‌بندی» پیش از آن است که بتوان نتایج را به مشتری، یا در حالت این مطالعهٔ موردی، به پزشکان حاضر در کارگاه ارائه کرد. نتایج عددی‌ای که S-PLUS تولید می‌کند نیاز به کمترین فرمت را دارد و خروجی برخی از شیوه‌های آماری S-PLUS برای غیرآمارداناتان کاملاً رمزی خواهد بود. برای مقاصد ارائه‌کرد، می‌توان از تابع‌هایی مانند `table()` و `cat()` برای فرمت مجدد برخی از خروجی‌های عددی استفاده کرد.

گرچه فرمت‌کردن مجدد و ویرایش خروجی عددی ممکن است ضروری باشد، نمایش‌های نموداری پیش‌فرض که توسط S-PLUS تولید می‌شوند، معمولاً بسیارخوب‌اند و باید از آن‌ها در ارائه‌کرد بهره برد. یک رویکرد استفاده از پاورپوینت است (نگاه کنید به فصل ۲،

استفاده از پاورپوینت برای ارائه‌کردها) که می‌تواند نمودارهای S-PLUS و نیز گزیده‌هایی از خروجی عددی S-PLUS را در خود جا دهد. این اسلایدهای پاورپوینت را می‌توان مجدداً پیش از کارگاه به کمک داده‌های صوری تنظیم و امتحان کرد.

۳.۴.۶ داده‌های ترتیبی

ابزارهای اندازه‌گیری مبتنی بر مقیاس‌ها به‌صورت گسترده‌ای در بسیاری از حوزه‌های دانش به‌کار می‌روند و اصطلاح روان‌سنجی با تحلیل آماری چنان ابزارهایی مرتبط است. نظیر ارزیابی‌های PANSS، داده‌های تولیدشده به‌وسیلهٔ ابزارهای اندازه‌گیری مبتنی بر مقیاس، ترتیبی یا «رتبه‌بندی»های با کد عدد صحیح هستند (۱، ۲، ...، ۷ در مثال PANSS). این کدهای عددی هم‌معنی رتبه‌ها هستند؛ این کار همچنین اجازهٔ محاسبهٔ آماره‌های خلاصه از قبیل میانگین را می‌دهد که می‌توان آن را به‌عنوان رتبهٔ «متوسط» تعبیر کرد.

در عمل، اغلب از تعبیر رتبه صرف‌نظر می‌شود (یا حتی در نظر گرفته نمی‌شود) و با داده‌های ترتیبی مانند به‌اصطلاح «نمره‌های خام» رفتار می‌شود. گذر سریع به داده‌های پیوسته زمانی پیش می‌آید که این نمره‌های خام به «نمره‌های استاندارد» تبدیل می‌شوند (شیوه‌های استانداردسازی به‌نحوی (؟) نمره‌ها را بهتر می‌کنند). در هر حالت، پس از آنکه همهٔ آزمون‌های t و همبستگی (حاصل‌ضربی) به اتمام رسیدند، هیچ تحلیلی بدون کاربرد نوعی فن چندمتغیره مانند تحلیل عاملی کامل نخواهد بود.... روشن است که مؤلفان [این کتاب] با این رویکرد موافق نیستند.

مشکلی که در رفتار با داده‌های ترتیبی به‌عنوان نمره داریم آن است که متریک ثابتی بر آنچه اساساً انتخاب دلخواهی برای کدبندی با اعداد صحیح بوده اعمال شده است؛ بدین معنی که می‌توانستیم سیستم رتبه‌بندی A، B، ...، G را نیز برای نمایش ارزیابی‌های PANSS به‌کار ببریم. در حالی که این کدبندی به‌لحاظ عددی مناسب نیست، بر این نکته تأکید دارد که «فاصله‌ها»ی بین کدهای سطح لزوماً خوش‌تعریف نیستند: $A - B$ همان صورتِ با عدد کدبندی‌شدهٔ ۱ − ۲ است، اما آیا حاصل آن "1" است؟ رتبه‌بندی با حروف نیز دوگان‌بودن نوع داده‌ها را، که داده‌های ترتیبی ــ رسته‌ای و رتبه‌دار ــ واجد آن هستند، تشریح می‌کند. از روش‌هایی که در فصل ۳ مورد بحث قرارگرفت چنین برمی‌آید که

- در شکل رسته‌ای، ارزیابی‌های PANSS به داده‌های شمارشی تبدیل می‌شوند: جدول‌های فراوانی و خلاصه، آزمون‌های بزرگ نمونه‌ای برای نسبت‌ها، و آزمون‌های پیوند را می‌توان اجرا کرد.
- در حالت رتبه‌دار، می‌توان آزمون‌های رتبهٔ ناپارامتری، همبستگی‌های اسپیرمن، و مقایسه‌های مساحت زیر خم (آزمون‌های t) را انجام‌داد.

۴.۴.۶ تحلیل مقدماتی

هدف از این مطالعه ارزیابی برنامهٔ آموزشی برای ابزار PANSS بود. برگزارکنندگان کارگاه به این مطلب نیز علاقه‌مند بودند که آیا «زبان» تأثیری بر نحوهٔ عملکرد پزشکان دارد یا خیر.

کد زیر برخی از مراحلی را که در تولید یک تحلیل خلاصه از داده‌های PANSS با استفاده از S-PLUS مطرح است تشریح می‌کند. هدف از کد ارائه‌شده در هر مرحله باید برای اغلب خوانندگان آشکار باشد. پیوست ب.۲ جزئیات بیشتری دربارهٔ S-PLUS در اختیار می‌گذارد.

تحلیل *S-PLUS*

```
#------------------------------------------------------------
# Read in the data from "c64.dat" using 1st row as column
# labels. Treat periods as missing values.

  panss <- read.table("c64.dat",header=T,na=".")
  print(panss)         # Output the data
  panss$P1             # Output the "P1" ratings
  stem(panss$P1)       # Stemplot of P1 (Check the data ?)
```

تابع `read.table()` مجموعه‌داده‌های PANSS "`c64.dat`" را می‌خواند و شیء S-PLUS به نام "`panss`" را تولید می‌کند که نوع خاصی از list به نام "`data frame`" است. این کار امکان آن را فراهم می‌کند که به ستون P1 از ارزیابی‌ها به‌عنوان مؤلفهٔ list ارجاع داده شود: `panss$P1`. یک نمودار ساقه و برگ، روش ساده‌ای برای وارسی داده‌های یک متغیر در اختیار می‌گذارد؛ وارسی همهٔ داده‌ها باید به کمک تابع کاربرساختهٔ "`chk(panss)`" انجام شود.

```
#------------------------------------------------------------
# Recalibrate the physicians' ratings to those of the Expert
  a <- panss[-1,]    # Copy of panss, excluding the 1st row
# Simple loop. "0" entries now indicate a correct rating.
  for(i in 1:72)
     a[i,-c(1,2)] <- panss[i+1,-c(1,2)] - panss[1,-c(1,2)]
```

از آنجا که لازم است که مقایسه‌های ما نسبت به‌نمره‌های ارزیاب خبره (RATER=0) انجام شود،واسنجی ارزیابی‌های پزشکان نسبت به‌ارزیابی‌های خبره سودمند خواهد بود. عمل واسنجی توسط دستور "`for`" که در بالا نشان داده شده انجام می‌شود. یک ارزیابی «درست» برای هر نشانه اکنون به‌وسیلهٔ یک درایهٔ "`a`" در data frame موسوم به "`a`" نشان داده می‌شود. باقیماندهٔ تحلیل نسبت به ارزیابی‌های واسنجیده انجام خواهد شد.

```
#-------------------------------------------------------------
# Graphical summaries.
  X11()          # Specify a graphics device (UNIX)
  hist(a$P1)     # Histogram of the recalibrated P1 ratings.
# Parallel Boxplots of the PANSS ratings.
  boxplot(a[,-c(1,2)])
# Boxplots of mean across all symptoms by Language
  plot.factor(apply(a[,3:32],1,mean,na.rm=T) ~ a$LANG)
```

لازم است که یک ابزار نموداری برای تولید نمایش‌های نموداری بالا مشخص کنیم. X11() یک ابزار گرافیکی استاندارد برای استفادهٔ تعاملی است. برای گرفتن خروجی این نمودارها در یک فایل PostScript در فرمت ESP (encapsulated PostScript)، ابزار گرافیکی postscript() را مشخص کنید. برای کنترل خروجی گرافیکی، چندین گزینه نظیر onefile=T وجود دارند که فایل PostScript جداگانه‌ای برای هر نمودار تولید می‌کند.

از آنجا که "a" همهٔ صفات panss (از جمله برچسب‌های ستون‌ها) را به ارث می‌برد، boxplot() یک نمودار جعبه‌ای موازی با ارزیابی‌های واسنجیده را که به‌وسیلهٔ نشانه برچسب‌گذاری‌شده تولید می‌کند. نتیجه در شکل ۱.۶ نشان داده شده است. ستون‌های یک data frame را نیز می‌توان به‌عنوان متغیرها در مشخص‌سازی «مدل» به‌کار گرفت (S-PLUS از کاراکتر "~" در فرمول مدل استفاده می‌کند). بنابراین، تابع plot.factor() نمودارهای جعبه‌ای موازی با ارزیابی متوسط (میانگین) را برای هر ۳۰ نشانه برحسب گروه‌بندی LANG تولید می‌کند. تابع apply(*, 1, mean, *) برای به‌دست‌آوردن ارزیابی میانگین برای هر پزشک ضروری است. به عبارت دیگر به وسیلهٔ ۱ = «سطرها». نمودار حاصل در شکل ۲.۶ نشان داده شده است.

```
#-------------------------------------------------------------
# Numerical summaries.
# Overall summary ("RATER" column excluded). Includes mean,
#   median, min, max, std dev, and quartiles.
# Same summary, but "split" by Language (hence exclude the
#   "LANG" column as well as the "RATER" column)
  summary(a[,-1])
  lapply(split(a[,-c(1,2)],a$LANG),summary)
#-------------------------------------------------------------
# Spearman Correlation between P1 and P2 ratings.
# Overall and within the "E" group only
  cor.test(a$P1, a$P2, method="s")
  cor.test(a$P1[a$LANG == "E"], a$P2[a$LANG == "E"],
           method="s")
```

نمایش‌های نموداری مروری کلی بر نتایج را در اختیار می‌گذارد. برای جزئیات خاص، به

خلاصه‌های عددی نیاز داریم.تابع summary() آماره‌هایی‌ساده (که‌در بالا نشان داده شده) برای هر ستون data frame در اختیار می‌گذارد. تابع lapply() را می‌توان برای به‌دست آوردن خلاصه‌هایی برای listهای پیچیده‌تر از قبیل split(a[,-c(1,2)], a$LANG) بکار برد. آماره‌های توصیفی ساده‌ای از قبیل همبستگی‌های (مبتنی بر رتبهٔ) اسپیرمن را نیز می‌توان به‌صورتی که خاطرنشان شده به‌دست آورد. هر دوی این روش‌ها در تحلیلی مفصل‌تر از داده‌های PANSS مناسب خواهد بود.

```
#------------------------------------------------------------
# Overall summary table.
  table(rep(names(a)[-c(1,2)],72),
        as.vector(t(as.matrix(a[,-c(1,2)]))))
```

```
     -4 -3 -2 -1  0  1  2  3 4 5 6
  G1  0  0 36 26  5  5  0  0 0 0 0
 G10  0  0  0  0 64  7  1  0 0 0 0
 G11  0  0  0 26 24 17  3  2 0 0 0
 G12  0  0  0 19 20 27  6  0 0 0 0
 G13  0  0  9 12 28 18  5  0 0 0 0
 G14  0  0  0  0 44 19  8  1 0 0 0
 G15  0  0 11 21 13 17  9  1 0 0 0
 G16  0  0  0  0  2  3 40 11 9 6 1
  G2  0  1  4 43  8 16  0  0 0 0 0
  G3  1  1  6 16 32 14  2  0 0 0 0
  G4  0  0 24 31 14  2  1  0 0 0 0
  G5  0  0 44  8 20  0  0  0 0 0 0
  G6  0  0  0  5 36 11 16  4 0 0 0
  G7  0  1  2 20 35  7  6  1 0 0 0
  G8  0  0 31 28 13  0  0  0 0 0 0
  G9  0  0  7  7 28 10 18  0 1 0 0
  N1  1  1  4  9 37 12  8  0 0 0 0
  N2  0  0  3 14 39 14  2  0 0 0 0
  N3  0  3 12 37 13  6  1  0 0 0 0
  N4  0  0  0  2 12 17 32  9 0 0 0
  N5  0  0  1  9 28 25  8  1 0 0 0
  N6  0  2  2 17 20 27  4  0 0 0 0
  N7  0  0  0  9 24 27  9  3 0 0 0
  P1  0  4  3 13 28 14 10  0 0 0 0
  P2  0  0  0 16 25 18 10  2 1 0 0
  P3  0  1  4 13 43 11  0  0 0 0 0
  P4  0  0  0  0 59  9  3  1 0 0 0
  P5  0  0  0 27 27  7  4  5 1 1 0
  P6  0  1 11 20 20 13  7  0 0 0 0
  P7  0  0  1 51 15  4  0  1 0 0 0
```

برای مقاصد ارائه‌کرد، نوع فشرده‌تری از فرمت «خلاصه» ارجحیت دارد. دستور table() بالا جدول‌بندی خلاصهٔ اولیه‌ای از ارزیابی‌های PANSS واسنجیدهٔ کلی در اختیار می‌گذارد. آن‌گونه که از خروجی ویرایش‌نشده پیداست، نیاز به این است که ترتیب‌بندی حرفی‌ـعددی پیش‌فرض نشانه‌های PANSS در فرمت «خوانندهٔپسند»تری مرتب شود. می‌توان برای هریک از گروه‌های LANG نیز مانند حالت همبستگی جدول‌هایی تولید کرد. به‌عنوان مثال، با قراردادن a[a$LANG=="E",] به جای "a"ی مذکور، جدول‌های خلاصهٔ "E" تولید خواهد شد.

```
#-----------------------------------------------------------
# Did they pass 7
pass.fun<-function(a, np=5, nn=5, ng=10) {
# Default options pass/fail criteria:
# Pass = at least 5 of 7 within abs(1) of key ("0") for
#        both the N and P symptoms and at least 10 of 16
#        within abs(1) of key for the G symptoms
#        (Missing values count against the physician)
  b <- matrix(0,72,4)
  r1 <- 3:9      #   "Positive" symptoms (columns)
  r2 <- 10:16    #   "Negative"
  r3 <- 17:32    #   "Generic"

  for(i in 1:72) {
    b[i,1] <- sum(abs(a[i,r1])<= 1, na.rm=T)
                - sum(is.na(a[i,r1]))
    b[i,2] <- sum(abs(a[i,r2])<= 1, na.rm=T)
                - sum(is.na(a[i,r2]))
    b[i,3] <- sum(abs(a[i,r3])<= 1, na.rm=T)
                - sum(is.na(a[i,r3]))
    b[i,4] <- 1
    if( b[i,1] < np ) b[i,4] <- 0
    if( b[i,2] < nn ) b[i,4] <- 0
    if( b[i,3] < ng ) b[i,4] <- 0
  }
  return(b)
}
```

مانند هر «آزمون» دیگر، مردم می‌خواهند بدانند که آیا پذیرفته شده‌اند یا خیر و برای تعیین آن معیاری لازم است. در اینجا، انعطاف‌پذیری در برنامهٔ تحلیل ضرورت دارد. هر استاندارد قبولی که پیش از کارگاه وضع شود، ممکن است خیلی سخت‌گیرانه[۱] از کار درآید و برگزارکنندگان کارگاه ممکن است بخواهند که معیار قبولی را در محل تعیین کنند.

۱. پرداختن به موضوع شکست کار ساده‌ای نیست. اگر کسی رد نشود، آیا واقعاً لزومی به کارگاه بود؟ اگر تعداد کمی مردود شوند، بار مسئولیت را می‌توان به دوش افراد انداخت که خود را بهبود بخشند؟ اگر افراد زیادی مردود شوند، این تصور پیش می‌آید که خود کارگاه موفقیت‌آمیز نبوده است.

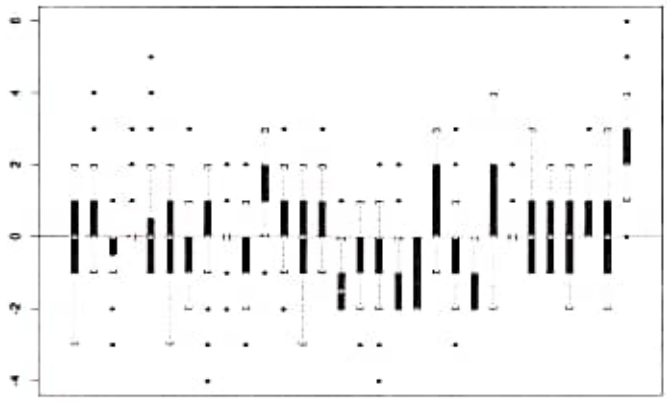
صورت پیش‌فرض این تابع "pass.fun(a)" از گزینه‌های مشخص‌شده برای محاسبهٔ تعداد پزشکانی که پذیرفته شده‌اند استفاده خواهد کرد. برای بالابردن استاندارد نشانه‌های «عام» به ۱۲ تا از ۱۶ نشانه، این تابع را به‌صورت "pass.fun(a,ng=12)" اجرامی‌کنیم. در هر حالت، نتیجهٔ قبول/رد در ستون آخر ماتریس "b" ظاهر می‌شود که در آن رد= ۰ و قبول= ۱.

```
#-----------------------------------
# Fisher's exact test.
table(a$LANG,b[,4])
   0  1
E 13 35
F  3 11
I  5  5
fisher.test(a$LANG,b[,4])
        Fisher's exact test
data:  a$LANG and b[, 4]
p-value = 0.3147
alternative hypothesis: two.sided
```

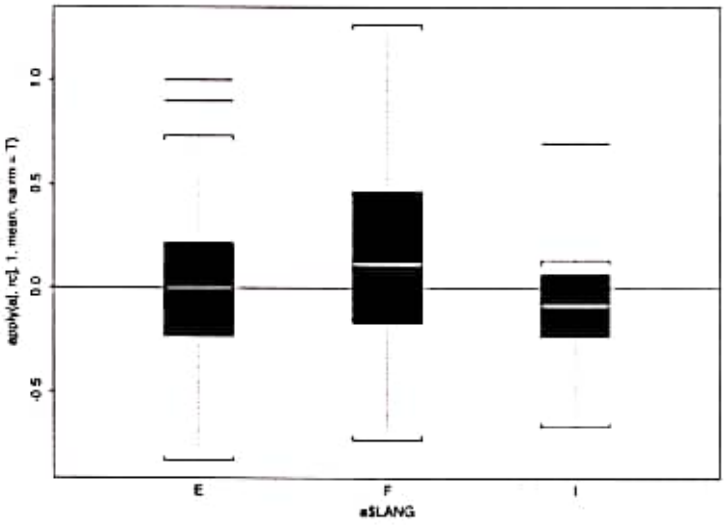
قسمت نهایی برنامهٔ تحلیل ما به این سؤال می‌پردازد که آیا زبان تأثیری در عملکرد پزشکان دارد یا خیر. آزمون دقیق فیشر در حالتی که تنها تعداد کمی از پزشکان از یکی از فرمت‌های زبانی برای ابزار PANSS استفاده می‌کنند به‌کار برده می‌شود.

نتایج *S-PLUS*

از نمودارهای جعبه‌ای موازی شکل ۱.۶ می‌توان دریافت که اغلب پزشکان در محدودهٔ ۲± از



شکل ۱.۶ نمودارهای جعبه‌ای موازی برای ارزیابی‌های PNASS.



شکل ۲.۶ میانگین ارزیابی PNASS برحسب زبان.

ارزیابی خبره قرار داشتند. با این حال، برخی عدم توافق‌های بسیار زیاد در تعدادی از نشانه‌ها وجود داشت که قابل توجه‌ترین استثنا نشانهٔ G16 («دوری‌گزینی اجتماعی آگاهانه») بود که توسط اغلب پزشکان بالا ارزیابی شد. از آنجا که برخی نشانه‌ها ممکن است در تعیین قصد خودکشی مهم‌تر باشند، این اطلاع به‌طور آشکاری سودمند است.

از شکل ۲.۶ ظاهراً چنین برمی‌آید که شرکت‌کنندگان فرانسوی و ایتالیایی عملکردی متفاوت با همتایان انگلیسی خود داشته‌اند. به‌ویژه، ارزیابی‌های فرانسوی‌ها، به‌طور متوسط، بالاتر از ارزیابی خبره و ارزیابی‌های ایتالیایی‌ها پایین‌تر از آن بود. یک عامل اختلاط آن است که ارزیاب خبره، انگلیسی بود. به‌علاوه، تنها ۱۴ شرکت‌کنندهٔ فرانسوی و ۱۰ شرکت‌کنندهٔ ایتالیایی در قیاس با ۴۸ شرکت‌کنندهٔ انگلیسی حضور داشتند.

از نتیجهٔ آزمون دقیق فیشر چنین برمی‌آید که زبان تأثیری بر عملکرد پزشکان نداشته است. با این حال، این مطلب ممکن است تا حدی ناشی از اندازه‌های نمونهٔ کوچک گروه‌های فرانسوی و ایتالیایی باشد. علی‌رغم فقدان معنی‌داری، از تفاوت‌های توزیعی چنین برمی‌آید که عامل زبان، دستور تحقیق بیشتر را صادر می‌کند.

تحلیل SAS

انجام‌دادن این تحلیل خلاصه با SAS کمی پیچیده‌تر است زیرا نمی‌توانیم مستقیماً به عناصر سطری متغیرها دسترسی داشته باشیم. به مقداری کار زبردستانه بر روی داده‌ها برای مقایسهٔ نمره‌های ارزیاب خبره با نمره‌های دیگر پزشکان نیاز داریم. با فرض اینکه «عنوان ستون» (سطر ۱) از مجموعه‌داده‌های c64.dat حذف شده است[1]، کد SAS زیر رویکردی برای این مسئله در اختیار می‌گذارد.

```
data a ;
  infile 'c64.dat' ;
  if _N_ = 1 then
     do ;
        input key $ klang $ kp1-kp7 kn1-kn7 kg1-kg16 ;
        retain  ;
     end ;
  else
     input rater $ lang $ p1-p7 n1-n7 g1-g16 ;
```

مجموعه‌داده‌های SAS به نام "a" حالا ۶۴ ستون دارد: کد شناسایی ارزیاب خبره، زبان، و ۳۰ نمرهٔ ارزیابی، و پس از آن‌ها همان متغیرها (با نام‌های متفاوت) برای ۷۲ ارزیاب دیگر. مقصود از دستور "retain" آن است که نتایج ارزیاب خبره را نگاه داریم و آن‌ها را، پیش از واردکردن ارزیابی‌های هر پزشک جدید، در ۳۲ مکان نخست تکرار کنیم.

حال می‌توانیم عمل‌های معمولی ستونی را در SAS اجرا کنیم و تفاوت‌های بین ارزیاب خبره و نمره‌های ارزیاب‌های دیگر را به‌دست آوریم. این کار طی مرحلهٔ زیر بر روی داده‌ها انجام می‌شود. استفاده از نام خاص "_LAST_" صرفاً به معنی آن است که به جای مندرجات جاری مجموعه‌داده‌های SAS به نام "a"، داده‌های حاصل از این مرحله جایگزین خواهد شد. نخستین سطر پیش از ایجاد متغیرهای معرف تفاوت‌ها کنار گذاشته شده و تنها کد ارزیاب، زبان، و تفاوت‌ها نگاه داشته می‌شوند.

```
data _LAST_ ;
  set a ;
  if _N_ > 1 ;

  dp1 = p1 - kp1 ;  ...  ; dp7 = p7 - kp7 ;
  dn1 = n1 - kn1 ;  ...  ; dn7 = n7 - kn7 ;
  dg1 = g1 - kg1 ;  ...  ; dg16 = g16 - kg16 ;

  keep rater lang dp1-dp7 dn1-dn7 dg1-dg16 ;
```

۱. به روش دیگر، به سطر infile عبارت add firstobs=2 را اضافه کنید.

مجموعه‌داده‌ها حالا برای تحلیل آماده‌اند. جدول‌های فراوانی و آزمون‌های همبستگی را می‌توان به روش معمول از طریق `proc freq` و `proc corr` انجام داد. دستور "`by lang ;`" را می‌توان برای به‌دست‌آوردن خلاصه‌ها برحسب زبان به‌کار برد.

۵.۴.۶ خلاصه

محدودیت‌های زمانی، به‌وضوح، میزانی را که مشاور آماری می‌تواند تحقیق را دنبال کند محدود می‌کند و اغلب باید مصالحه‌هایی صورت گیرد. گرچه شرایط این مطالعهٔ موردی تا حدی خاص بودند، آمادگی خوب و درک کامل مسئله جزو مؤلفه‌های اساسی این تحلیل از پیش‌برنامه‌ریزی شده بود که تحلیل «در محل»، نتایج کارگاه PANSS را تسهیل کرد.

ما بر اهمیت دانش زمینه‌ای تأکید کرده‌ایم زیرا برخی مشتری‌ها تصور می کنند که مشاور می‌تواند «تحلیل کاملی» از پروژهٔ آن‌ها را فی‌الفور در اختیارشان بگذارد. «این صرفاً بررسی ساده‌ای با ۲۰ سؤال است ... نمی‌توانیم همین حالا آن را انجام دهیم؟» پاسخ به این سؤال ممکن است بدیهی به نظر برسد (خیر!)، زیرا ابتدا لازم است که، پیش از «انجام‌دادن» تحلیل، هدف بررسی را بفهمیم و برای خود مسجل کنیم که داده‌ها چگونه گردآوری شده‌اند. وضعیتی مشکل‌تر زمانی پیش می‌آید که مشتری فقط می‌خواهد بداند که «چگونه» برنامه‌نویسی تحلیل را انجام دهد ــ آن‌ها خود می‌توانند تحلیل واقعی را انجام دهند؛ تنها به برنامهٔ کامپیوتری نیاز دارند. مجدداً خاطرنشان می‌کنیم که این مطالعهٔ موردی تشریح می‌کند که ایجاد یک برنامهٔ تحلیل، تمرین بدیهی‌یی نیست.

پرسش‌ها

۱. یک جنبهٔ ذهنی مطالعهٔ PANSS معیار قبول/رد است. تحقیق کنید که آیا یک اثر معنی‌دار زبان را می‌توان با جرح و تعدیل این معیار به‌دست آورد یا خیر. اگر چنین است، چگونه نتایج متناقض را توضیح می‌دهید؟

۲. راهی دیگر برای پردازش مجموعه‌داده‌های PANSS آن است که ستون‌های ارزیابی را «پشته» کنیم؛ به این معنی که یک ستون تنهای RESP مرکب از همهٔ ارزیابی‌ها ایجاد کنیم به‌طوری که SYMP مشخص‌کنندهٔ نشانهٔ واقعی باشد. کدهای RATER و LANG حالا برای هر پزشک ۳۰ بار تکرار می‌شود. مزیت‌های انجام‌دادن این کار چیست؟

۳. برنامهٔ تحلیل S-PLUS یا / و SAS را کامل کنید. این کار باید شامل جدول‌های خلاصه به تفکیک گروه‌بندی زبان باشد و تعداد و درصد پزشکان پایین‌تر از، بالاتر از، و در محدودهٔ ±۱ ارزیاب خبره را فراهم کند.

۷

مطالعه‌های موردی از گروه II

در این فصل، چهار مطالعهٔ موردی را که نیازمند روش‌های آماری عام‌تری مانند رگرسیون لوژستیک، مدلبندی سری‌های زمانی، و طرح‌های عاملی‌اند ارائه می‌کنیم. مسئلهٔ آماری همچنان خوش‌تعریف است اما میدان عمل گسترده‌تری نسبت به مطالعه‌های موردی گروه I دارد. شاید لازم باشد چندین راه‌حل را مورد ارزیابی قراردهیم. این مطالعه‌های موردی در زیر به‌ترتیب بخش‌ها فهرست شده‌اند.

- **۱.۷ بررسی دم‌تکاندن**

  تحلیل پروبیت

- **۲.۷ آیا طعم خوبی دارد؟**

  طرح‌های عاملی

- **۳.۷ هزینه‌کرد در شهرداری‌های نیویورک**

  روش‌های رگرسیونی: پیشگویی

- ۴.۷ اندازه‌گیری زمان کیفیت

تحلیل سری‌های زمانی

۱.۷ بررسی دم‌تکاندن

روش‌ها	تحلیل پروبیت: ED50
داده‌ها	مطالعهٔ تیمار دارویی بر موش‌ها پاسخ: شمار حرکت‌های دم: «Flicks»

۱.۱.۷ آمار پیش‌بالینی

این بررسی دربارهٔ شدت اثرهای متقابل دارویی بین مورفین و ماری‌جوانااست. در فصل ۱ دربارهٔ فرایند فراوری داروها بحث کردیم و یکی از مراحل مهم آن تعیین حداقل میزان دز داروست که منجر به اثربخشی می‌شود. گاهی این کار ممکن است متضمن برآوردکردن ترکیب‌های دارویی باشد که با حداقل میزان دارو به حداکثر اثربخشی نائل می‌شود.

آزمایش مطرح در این مطالعهٔ موردی عبارت از دادن ترکیبی از دو دارو به موش‌ها در دزهای متفاوت است. این داروها مورفین و ماری‌جوانا هستند و از آن‌ها برای تسکین درد استفاده می‌شود. برای سنجش اثر آن‌ها آزمون‌های «دم تکاندن[1]» برای دو دارو و ترکیب‌های آن‌ها ترتیب داده می‌شود. هدف تشخیص این است که آیا اثر متقابل، یا سینرژی[2]، بین این دو دارو موجود است یا خیر. به‌علاوه، به محاسبهٔ بازه‌های اطمینان برای آن ترکیب دُزی علاقه‌مندیم که به ۵۰٪ پاسخ منجر شود؛ یعنی، دز موثری که انتظار می‌رود ۵۰٪ موردها به آن پاسخ دهند. این میزان دز به مقدار $ED50$ موسوم است.

داده‌ها

داده‌ها مرکب از یک متغیر پاسخ P است که نسبت تعداد موش‌هایی (از بین ۱۰ موش) را می‌دهد که دم‌های خود را پس از قرارگرفتن در معرض یک محرک گرمایی به‌ازای ترکیب دارویی مفروض تکان داده‌اند. متغیرهای پیشگو عبارت‌اند از دزهای مورفین [MORPHINE] و ماری‌جوانا [DEL9] برحسب میلی‌گرم (mg).

۱. آزمون دم‌تکاندن (flick tail test) آزمونی برای بررسی عکس‌العمل حیوانات به درد است. این آزمون در تحقیقات اساسی درد و برای سنجش میزان تأثیر داروهای ضد درد از طریق مشاهدهٔ واکنش آن‌ها به گرمای شدید مورد استفاده قرار می‌گیرد. برای این کار یک پرتو گرمازا روی دم حیوان متمرکز می‌شود و تایمری شروع به کار می‌کند. وقتی حیوان دم خود را (شدیداً) تکان می‌دهد، تایمر متوقف و زمان ثبت‌شده به‌عنوان معیاری برای آستانهٔ درد حیوان بکار می‌رود.م.

2. synergy

جدول ۱.۷ تعریف‌های متغیر

متغیر	تعریف
OBS	عدد اندیس
SET	یکی از چهار ترکیب ممکن
MORPHINE	دز سولفات مورفین (mg/kg) تزریق‌شده به موش‌های مورد مطالعه. دامنهٔ آن از ۰ تا ۰٫۸ است.
DEL9	دز Delta9-THC (mg/kg) تزریق‌شده به موش‌های مورد مطالعه. دامنهٔ آن از ۰ تا ۱۶ است.
FLICK	تعداد موش‌هایی که دم خود را بعد از اعمال یک محرک گرمایی از زیر تکان داده‌اند.
REPS	تعداد تکرارها (با موش‌های متفاوت)
P	FLICK/REPS

آزمایش برای ۲۰ ترکیب مختلف دارو تکرارشده و یافته‌ها در جدول ۲.۷ نشان داده شده‌اند. در شش آزمایش، به موش‌ها ترکیبی از هر دارو تزریق شد، در حالی که در ۱۳ تا از آزمایش‌ها به موش‌ها تنها یک دارو داده شد. یکی از آزمایش‌ها جنبهٔ کنترل داشت که در آن به ۱۰ موش هیچ دارویی داده نشد.

روش‌شناسی

۲.۱.۷ رگرسیون لوژستیک

آزمایش‌های لازم برای مطالعهٔ اثربخشی یا سمّی‌بودن یک دارو پاسخی تولید می‌کند که اغلب یک متغیر دودویی (۱: پاسخ، ۰: بی‌پاسخی) یا دوجمله‌ای (k تا از n تا پاسخ‌اند) است. برای این نوع داده‌ها مرسوم است که یک مدل رگرسیون لوژستیک

$$\log it(p) = \log\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 Dose$$

برازش داده شود که در آن p احتمال پاسخ در دز $[Dose]$ مفروض است. متغیر $Dose$ را می‌توان در مقیاس لگاریتمی به‌صورت $\log(Dose)$ نشان داد زیرا دزهای دارویی به‌طور نوعی با تصاعد نمایی انتخاب می‌شوند. این امر در آزمایش‌هایی متداول است که دامنهٔ معقولی برای دزها

جدول ۲.۷ داده‌های دم تکاندن

OBS	SET	MORPHINE	DEL9	FLICK	REPS	P
1	O	0.0	0.0	0	10	0.0
2	A	1.0	0.0	0	10	0.0
3	A	2.0	0.0	2	10	0.2
4	A	3.0	0.0	4	10	0.4
5	A	4.0	0.0	4	10	0.4
6	A	5.0	0.0	6	10	0.6
7	A	6.0	0.0	9	10	0.9
8	A	7.0	0.0	9	10	0.9
9	A	8.0	0.0	8	10	0.8
10	B	0.0	1.0	0	10	0.0
11	B	0.0	2.0	0	10	0.0
12	B	0.0	4.0	9	10	0.9
13	B	0.0	8.0	9	10	0.9
14	B	0.0	16.0	10	10	1.0
15	AB	0.5	0.5	0	10	0.0
16	AB	1.0	1.0	2	10	0.2
17	AB	1.5	1.5	7	10	0.7
18	AB	2.0	2.0	9	10	0.9
19	AB	2.5	2.5	9	10	0.9
20	AB	3.0	3.0	9	10	0.9

مشخص نیست و آزمایشگران سعی می‌کنند که پاسخ‌هایی برای دزها در چندین مقیاس به‌دست آورند.

برای هر مقدار دز، آزمایش n_i بار تکرار می‌شود و تعداد پاسخ‌های مشاهده‌شده r_i است. پارامترهای مدل (β_0, β_1) را می‌توان به کمک برآورد ماکسیمم درستنمایی (MLE) برآورد کرد که آن‌ها را با $(\hat{\beta}_0, \hat{\beta}_1)$ نشان می‌دهیم. می‌توان نشان داد که $(\hat{\beta}_0, \hat{\beta}_1)$ تقریباً دارای توزیع نرمال با ماتریس کوواریانس V است که در آن $V_{00} = Var(\hat{\beta}_0)$، $V_{11} = Var(\hat{\beta}_1)$، $V_{01} = V_{10} = Cov(\hat{\beta}_1, \hat{\beta}_0)$.

یکی از کمیت‌های مورد توجه، میانهٔ دز مؤثر $ED50$ یا دز لازم برای به‌دست‌آوردن نسبت اثربخشی معادل با $p = \%50$ است. از معادلهٔ مدل به‌دست می‌آوریم

$$\log\left(\frac{p}{1-p}\right) = \log\left(\frac{1/2}{1-1/2}\right) = 0 = \beta_0 + \beta_1 ED50$$

و بنابراین $ED50 = -\beta_0/\beta_1$. برآورد ماکسیمم درستنمایی $ED50$ برابر است با $-\hat{\beta}_0/\hat{\beta}_1$.

به‌ازای سایر مقادیر p، $ED۱۰۰p$ را به‌عنوان دز p٪ اثربخشی تعریف می‌کنیم. MLE برابر است با

$$\widehat{ED۱۰۰p} = [\log(p/(۱-p)) - \hat{\beta}_۰]/\hat{\beta}_۱$$

به محض اینکه برآوردهای $ED۵۰$ها را به‌دست آوردیم، بازه‌های اطمینانی برای آن‌ها محاسبه می‌کنیم. برای این منظور، قضیهٔ فیلر[1] را به‌کار می‌بریم که ما را قادر می‌سازد تا بازه‌های اطمینانی برای نسبت‌های متغیرهای تصادفی گاوسی به‌دست آوریم. فرض کنید که $\gamma = z_{\alpha/۲}V_{۱۱}/\hat{\beta}^۲_۱$؛ در این صورت بازهٔ اطمینان $۱۰۰(۱-\alpha)$٪ برای $ED۵۰$ عبارت است از

$$\widehat{ED۵۰} + \frac{\gamma}{۱-\gamma}(\widehat{ED۵۰} + \frac{V_{۱۰}}{V_{۱۱}}) \pm \frac{z_{\alpha/۲}}{(۱-\gamma)\hat{\beta}_۱}K$$

که در آن

$$K^۲ = V_{۰۰} + ۲\widehat{ED۵۰}V_{۱۰} + \widehat{ED۵۰}^۲V_{۱۱} - \gamma(V_{۰۰} - V^۲_{۱۰}/V_{۱۱})$$

ماتریس $V_{(ij)}$ها از خروجی رگرسیون لوژستیک به‌دست می‌آید. این معادله را می‌توان به‌حالت spی کلی با قراردادن $ED۱۰۰p$ به جای $ED۵۰$ در معادلهٔ بالا نیز گسترش داد.

۳.۱.۷ رگرسیون لوژستیک چندمتغیره

در این مطالعهٔ موردی، مدل لوژستیک با دو پیشگو و یک جملهٔ اثر متقابل را نیز در نظر می‌گیریم.

$$logit(p) = \log\left(\frac{p}{۱-p}\right) = \beta_۰ + \beta_۱Dose۱ + \beta_۲Dose۲ + \beta_۳Dose۱ * Dose۲.$$

پاسخ در حالا رویه‌ای از دومتغیر است؛ یعنی $Dose۱$، $Dose۲$، و $ED۵۰$ یک نقطه نیست، بلکه خم ترازی به ارتفاع ۰٫۵ روی رویه است. خم $ED۵۰$ را می‌توان با دادن مقداری از $Dose۱$ به‌ازای هر مقدار ثابت $Dose۲$ محاسبه کرد که عبارت است از

$$Dose۱ = \frac{-\beta_۰ - \beta_۲Dose۲}{\beta_۱ + \beta_۳Dose۲}.$$

اگر این خم محدب باشد، نشانهٔ آن است که سینرژی وجود دارد.

1. Fieller

۴.۱.۷ تحلیل مقدماتی

برای تحلیل این داده‌ها استفاده از PROC PROBIT در نرم‌افزار SAS را توصیه می‌کنیم. PROC PROBIT برآوردگرهای ماکسیمم درستنمایی را برای همهٔ پارامترها و بازه‌های اطمینان را برای $ED۱۰۰p$ با استفاده از قضیهٔ فیلر به‌ازای هر p محاسبه می‌کند. تحلیل پیچیده‌تری را می‌توان با استفاده از شیوهٔ glm() در S-PLUS به‌صورت زیر به‌دست آورد.

یادداشت‌های مربوط به $S-PLUS$

برخی از تابع‌های S-PLUS مورد استفاده به‌صورت زیرند.

glm	تابعی مقدماتی برای مدل‌های خطی تعمیم‌یافته (این تابع شامل رگرسیون لوژستیک است).
contour	تابعی برای رسم نمودار خم تراز.
predict.glm	تابعی برای محاسبهٔ مقدارهای پیشگویی.

تحلیل برای مجموعه داده‌های مثال

```
# Create the dataset flick from the file "c71.dat"
flick <- read.table("c71.dat",head=T)

# Run the logistic regression for the model with
# interaction term
flick.glm <- glm(p ~ morphine*del9,w=rep(10,20),
                 data=flick, family=binomial)

# Create the dataset with the grid points for the
# contour plot
morphine <- sort(unique(flick$morphine))
del9 <- sort(unique(flick$del9))
pred <- data.frame(morphine=rep(morphine,10),
                   del9=rep(del9,rep(12,10)))

# Evaluate the estimated model at the grid points
phat <- predict.glm(flick.glm, pred, type="response")
dim(phat) <- c(12,10)

# Graph the ED50 and ED75 curves.
contour(morphine,del9, phat, levels=c(0.5,0.75))
```

اگر بین مورفین و Del9 سینرژی وجود داشته باشد، نمودار باید خم ترازهایی محدب را نشان دهد.

۵.۱.۷ خلاصه

این مطالعهٔ موردی، روش‌شناسی آزمایش‌های پاسخ به دز را تشریح می‌کند. قضیهٔ فیلر در اینجا برای محاسبهٔ بازه‌های اطمینان برای نسبت‌ها بسیار سودمند است، اما کاربردهای دیگر بسیاری هم دارد که در اینجا نیامده است.

پرسش‌ها

چند پرسش جالب که در این مطالعه باید به‌آن‌ها پاسخ دهیم عبارت‌انداز:

۱. $ED50$ها را برای هر یک از دو دارو و برای ترکیبی از این دو محاسبه کنید. یک گزینهٔ دیگر این است که پاسخ را با استفاده از $\log(Dose)$ یا $\log(1 + Dose)$ مدلبندی کنید. کدام مدل بهترین برازش را در اختیار می‌گذارد؟

۲. با استفاده از قضیهٔ فیلر، بازه‌های اطمینانی برای $ED50$ ها محاسبه کنید.

۳. به ترکیب‌های داروها نگاه کنید. آیا نشانی از سینرژی دارند؟
همان‌طور که در بخش بالا بیان شد، راهی برای نشان‌دادن سینرژی، تولید یک نمودار خم تراز از رویهٔ مقادیر پیشگویی‌شدهٔ P به کمک مدل، در ناحیهٔ دز $0 < MorphineDose < 8$ و $0 < Del9 < 16$ است. به‌خصوص می‌توان خم تراز متناظر با $ED50$ را رسم کرد. خم‌های تراز محدب نشان‌دهندهٔ سینرژی‌اند.

۴. ترکیبی خطی از داروها ارائه دهید که بالاترین میزان اثربخشی را تولید می‌کنند.

۲.۷ آیا طعم خوبی دارد؟

روش‌ها	طرح عاملی طرح نقطه‌مرکزی
داده‌ها	۵ عامل طرح ۴ متغیر پاسخ

۱.۲.۷ طرح‌های عاملی در علم تغذیه

این مطالعهٔ موردی خاص، مثال خوبی از طرح پژوهشی در حال اجرایی را نشان می‌دهد که در آن نقش آموزشی مشاور آماری اثر مهمی طی مراحل آغازین تحقیق داشت. مشاورهٔ آغازین ما با مشتری عملاً چندین سال پیش از آزمایشی که در اینجا ارائه شده انجام گرفت. با در نظرگرفتن مقاصد تشریحی، بگذارید مختصراً مراحل را تا مسئلهٔ آغازین بازپی‌گیریم.

آغاز مسئله

طرح پژوهشی مشتری به بررسی ویژگی‌های قابل اندازه‌گیری یک محصول غذایی از قبیل رنگ، غلظت و میزان رطوبت موجود در آن مربوط می‌شود. این متغیرهای پاسخ با عبوردادن مادهٔ‌غذایی (ذرت) از درون یک «اکسترودر[۱]» به‌دست آمده‌اند که می‌شد این دستگاه را تحت شرایط مختلف تنظیم کرد. مثلاً اپراتور دستگاه می‌توانست سطح دمای هر یک از اتاقک‌های فر، میزان آب و رنگ اضافه‌شده به ترکیب ذرت، و سرعت عبور را تغییر دهد. برای هر متغیر کنترل یا عامل، حداقل سه سطح بالقوهٔ مورد توجه مشتری وجود داشت. مشتری به‌خوبی آگاه بود که این موضوع ممکن بود به آزمایشی با چندین هزار (!) عامل ترکیبی منجر شود. همین امر دلیل او برای خواستن مشاورهٔ آماری بود.

حتی در این مرحلهٔ اولیهٔ جلسهٔ مشاوره، روشن بود که باید مشتری را با رویکرد آماری در خصوص انجام‌دادن آزمایش آشنا کنیم. با این حال، به جای معطوف‌کردن تمرکز بحث به مباحث آماری، نخستین گام ما این بود که با مشتری کار کنیم و عامل‌ها و سطوحی را که در اولویت قرار نداشتند حذف کنیم. این کار به ما کمک می‌کرد که درک بهتری از جزئیات مرتبط با پروژهٔ مشتری به‌دست آوریم. پس از تمام‌شدن کار حذف‌کردن‌های اولیه، عامل‌های باقیمانده‌ای که همچنان با آن‌ها مواجه بودیم ۱۰۰۰ ترکیب در اختیار ما قرار می‌داد و دیگر آمادگی آن را داشتیم که شروع به متقاعدکردن مشتری خود دربارهٔ نیاز به در نظرگرفتن طرح‌های آزمایشی دیگری کنیم. این کار چندان طول نکشید.

در این مرحله بود که مشتری خاطرنشان کرد که به نفعشان خواهد بود که در هر هفته دو آزمایه انجام دهند! نیازی به گفتن نیست که طرحی دیگر بی‌درنگ در نظر گرفته شد که مطالعه را به ۱۶ آزمایه کاهش می‌داد. آن‌طور که معلوم شد ۲۰٪

۱. extruder: دستگاهی که مواد خمیری را با فشار از قالب‌هایی که شکل‌های مختلفی دارند عبور می‌دهد تا آن‌ها را به شکل لایه‌های نازک، نوار، یا لوله‌ای‌شکل در آورد.ـم.

از این ترکیب‌های آزمایه‌های عاملیِ کسری اولیه، ناپایدار بودند و به نتیجه‌ای منجر نمی‌شدند.

وضع حال

صنعت غذای میان‌وعده‌ای به‌شدت رقابتی است و ایجاد بازار موفقیت‌آمیز برای یک محصول مستلزم چیزی بیش از صرفاً یک ویژگی «خوش‌طعمی» است. از آنجا که ممکن است ارزش غذایی برخی از محصولات تا حدی محل تردید باشد، بهترکردن ظاهر، بو، و دیگر «کیفیت»های محصول غذایی حائز اهمیت است. بسته‌بندی پرجلا و تبلیغات پرزرق‌وبرق قطعاً کمک‌کننده است، اما هیچ‌کس حاضر نیست چیزی بخورد که بوی بدی می‌دهد!

هدف پروژهٔ مشتری ما آن بود که تلاش کند شرایطی را بهینه کند که تحت آن شرایط برخی مشخصه‌های کیفی غذا بهینه می‌شوند. مانند آنچه در مورد اغلب مشتریان اتفاق می‌افتد، پرهیز از رویکرد «هر بار، یک عامل» در انجام‌دادن آزمایش ممکن است دشوار باشد و در این وضعیت است که نقش آموزشی مشاور آماری اهمیت می‌یابد. منظور آن است که مشتری راهنمایی شود تا درک بهتری از ماهیت **تکراری** آزمایش‌های آماری به‌دست آورد.

طرح «غربالگری» عاملی کسری اولیه‌ای که در بالا به‌کار بردیم، به‌روشنی آشکار کرد که میزان بالای رطوبت موجود و سطوح بالای دما منجر به شکست می‌شود. بر مبنای این نتایج، مشتری توانست تیم آزمایشگران را در مورد مزیت‌های رویکرد آماری متقاعد کند. در «تکرار» بعدی پروژه از یک طرح عاملی کسری با سه سطح در هر عامل استفاده شد. این کار بعداً با استفاده از طرح‌های مرکب مرکزی ظرافت بیشتری پیدا کرد. نتایج حاصل از آزمایش مرکب مرکزی در این مطالعهٔ موردی ارائه شده‌اند.

داده‌ها

متغیرهای کنترل (عامل‌ها) و چهار متغیر پاسخ (خروجی‌ها) که در آزمایش عبوردادن از دستگاه اکسترودر اندازه‌گیری شده‌اند در جدول زیر فهرست شده‌اند. ترکیب‌های مختلف سطوح عامل برای هر بار اجرای آزمایش به‌کار برده شدند که در آن نرخ پایه[1] به مقدار محصول غذایی (ذرت) اشاره دارد که در دستگاه اکسترودر قرار داده شد. آب و رنگ (که به‌صورت درصد محاسبه شده‌اند) به ذرت اضافه شد و ترکیب حاصل به‌آرامی در حالی که به درون دستگاه اکسترودر با فشار وارد می‌شد، «پخته شد» (دما). ساروکار فشار اساساً مرکب از یک محفظهٔ «پیچی» چرخنده بود و بنابراین سرعت پیچ معیاری برای آن است که ترکیب با چه سرعتی از دستگاه عبور داده می‌شود.

1. Base Rate

متغیرها در طرح مرکب مرکزی

نمره	رطوبت	سفتی	غلظت	خروجی‌ها		
سطوح عامل			عامل‌ها			
A	سرعت پیچ	(rpm)	۵۵۰	۵۹۰	۶۲۵	۶۶۵	۷۰۰
B	نرخ پایه	(lbs)	۱٫۵	۱٫۷	۱٫۹	۲٫۱	۲٫۳
E	آب اضافه‌شده	(٪)	۱۴	۱۷	۲۰	۲۳	۲۶
G	رنگ اضافه‌شده	(٪)	۰٫۰۰	۰٫۷۵	۱٫۵	۲٫۲۵	۳٫۰۰
H	دما	(فارنهایت)	۲۴۰	۲۵۵	۲۷۰	۲۸۵	۳۰۰

نتایج حاصل از آزمایش در مجموعه‌دادهٔ c72.dat گنجانده شده‌اند. همان‌طور که از جدول ۳.۷ آشکار است، طرح مرکب مرکزی به‌کاررفته متشکل از نیم تکرار 2^5 عاملی (۱۶ بار اجرا) است که ۱۰ به‌اصطلاح نقطهٔ محوری و ۴ نقطهٔ مرکزی به آن افزوده شده‌اند که در مجموع ۳۰ بار اجرا را حاصل می‌کند. حال می‌خواهیم ببینیم همهٔ این‌ها واقعاً به چه معناست!

روش‌شناسی

هدف این مطالعه تلاش برای پیداکردن ترکیبی از شرایط عمل مرتبط با عامل‌های کنتی A، B، E، G، و H بود که خروجی چهار متغیر پاسخ، یعنی غلظت، سفتی، رطوبت، و مزه را بهینه می‌کرد. برای تحلیل این نوع مسئله‌ها، می‌توان روش‌شناسی رویهٔ پاسخ را به‌کار برد.

۲.۲.۷ روش‌شناسی رویهٔ پاسخ

مفهوم یک رویهٔ پاسخ به نمایش مدل به‌صورت

$$y = g(x_1, x_2, \ldots, x_k) + \epsilon \qquad (۱.۷)$$

بازمی‌گردد که در آن y پاسخ مشاهده شده است، ϵ مؤلفهٔ خطای تصادفی است، و $g(x_1, x_2, \ldots, x_k)$ تابعی از سطوح متغیرهای کنترل کنتی x_1، x_2، ...، x_k است. بنابراین، مقدار مورد انتظار پاسخ به کمک «رویه»ی $g(x_1, x_2, \ldots, x_k)$ تعریف می‌شود. توجه کنید که وقتی $k = 2$، رویهٔ پاسخ را می‌توان به‌صورت نموداری به کمک یک نمودار خم تراز نمایش داد.

در حالت کلی، شکل واقعی رابطه تابعی بین y و متغیرهای کنترل نامعلوم خواهد بود. در نتیجه، مرحلهٔ اول از یک تحلیل رویهٔ پاسخ یافتن تقریب مناسبی برای تابع $g(x_1, x_2, \ldots, x_k)$

جدول ۳.۷ آزمایش طرح مرکب مرکزی

دکتوریل (۲^۵ تیم‌تکرار)					خروجی‌ها			
A	B	E	G	H	صلفت	سفتی	رطوبت	طعم
590	1.7	17	0.75	285	197.725	0.18350	11.300	337
665	1.7	17	0.75	255	187.940	0.19565	11.175	340
590	2.1	17	0.75	255	198.290	0.21460	11.125	337
665	2.1	17	0.75	285	184.620	0.22895	11.050	339
590	1.7	23	0.75	255	303.340	0.16855	14.000	326
665	1.7	23	0.75	285	280.615	0.16540	13.600	329
590	2.1	23	0.75	285	280.265	0.19550	13.525	326
665	2.1	23	0.75	255	269.440	0.19385	13.375	329
590	1.7	17	2.25	255	218.215	0.19575	11.575	333
665	1.7	17	2.25	285	202.620	0.22005	11.125	338
590	2.1	17	2.25	285	220.495	0.22495	11.150	337
665	2.1	17	2.25	255	214.700	0.22175	11.425	339
590	1.7	23	2.25	285	312.415	0.15585	13.975	321
665	1.7	23	2.25	255	306.810	0.16190	13.775	322
590	2.1	23	2.25	255	315.135	0.18230	13.900	320
665	2.1	23	2.25	285	290.955	0.19585	13.300	325
نقاط محوری								
550	1.9	20	1.50	270	265.405	0.18465	12.625	329
700	1.9	20	1.50	270	238.715	0.19395	12.200	334
625	1.5	20	1.50	270	245.055	0.15315	12.525	331
625	2.3	20	1.50	270	243.790	0.20715	12.225	331
625	1.9	14	1.50	270	167.655	0.22575	9.750	341
625	1.9	26	1.50	270	336.835	0.16625	14.575	321
625	1.9	20	0.00	270	224.930	0.18630	12.025	336
625	1.9	20	3.00	270	261.510	0.18990	13.125	331
625	1.9	20	1.50	240	255.015	0.19100	12.625	330
625	1.9	20	1.50	300	246.800	0.21305	12.275	331
نقاط مرکزی								
625	1.9	20	1.50	270	253.840	0.18490	12.625	331
625	1.9	20	1.50	270	248.495	0.18800	12.675	333
625	1.9	20	1.50	270	237.240	0.19405	12.325	339
625	1.9	20	1.50	270	247.310	0.20325	12.600	332

است که در (۱.۷) ظاهر می‌شود. از آنجا که هدف عمده تعیین شرایط عمل بهینه برای فرایند تحت مطالعه است، یک تقریب چندجمله‌ای خطی یا درجهٔ دوم اغلب برای این منظور کافی است. به این تقریب‌ها به‌ترتیب مدل‌های مرتبهٔ اول و مرتبهٔ دوم اطلاق می‌شود:

$$\text{مرتبهٔ اول} \quad y = \beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k + \epsilon$$

$$\text{مرتبهٔ دوم} \quad y = \beta_0 + \Sigma_{i=1}^{k} \beta_i x_i + \Sigma_{i=1}^{k} \Sigma_{j=1}^{k} \beta_{ij} x_i x_j + \epsilon$$

آشکار است که می‌توان انتظار داشت که این تقریب‌ها تنها در درون یک ناحیهٔ «موضعی» از

فضایی که به‌وسیلهٔ متغیرهای کنترل تنیده‌شده به‌خوبی کار می‌کنند و هدف این نیست که محقق درکی از $(x_1, x_2, \ldots, x_k)g$ی واقعی زمینه‌ای فرایند به‌دست آورد. با این حال، می‌توان آن‌ها را به‌طور مؤثری برای مقاصد بهینه‌کردن یک فرایند به سبکی «تکراری» به‌کار گرفت. این بدان معناست که سطح روش‌شناسی رویهٔ پاسخ یک شیوهٔ دنباله‌ای است: نتایج حاصل از یک مرحله از آزمایش، محقق را به سمت ناحیه‌ای که آزمایش «بعدی» باید در آنجا انجام شود هدایت می‌کند.

روش تندترین صعود

سریع‌ترین مسیر به بالای یک تپه دنبال‌کردن مسیر تندترین صعود است. این مسیر در عین حال از لحاظ بدنی طاقت‌فرساترین مسیر نیز هست! خوشبختانه، کامپیوتر ماست که عمل بالارفتن را انجام می‌دهد و در چارچوب (زمینهٔ) تحلیل رویهٔ پاسخ اتخاذ این راهبرد، شیوه‌ای به‌صرفه را برای یافتن بهینه به‌دست می‌دهد.[۱]

طی مراحل آغازین تحقیق، موقعیت ناحیه‌ای که شامل بهینه است نامعلوم است. برای ساده‌کردن تعبیر نتایج، اغلب یک مدل مرتبهٔ اول به‌کار برده می‌شود. در این حالت، جهت تندترین صعود متناظر با جهتی است که در آن $\hat{y}$ سریع‌ترین افزایش را دارد. می‌توان نشان داد که گام‌های برداشته‌شده در امتداد این مسیر با بردار متشکل از ضرایب رگرسیونی برازش‌یافتهٔ $x\hat{\beta}$ متناسب‌اند. اندازهٔ واقعی قدمی که باید برداشته شود باید توسط محقق و بر مبنای تجربهٔ او از فرایند تحت مطالعه تعیین شود. در عمل، آزمایش‌ها در امتداد این مسیر اجرا می‌شوند تا وقتی که کاهشی حاصل شود یا آنکه به حد شرایط عملِ معقولی برسیم.

به محض آنکه ناحیهٔ شامل بهینه شناسایی شود، می‌توان از مدل‌های ماهرانه‌تری برای تعیین بهینهٔ واقعی استفاده کرد. از تقریب درجهٔ دوم معمولاً پیش از این مرحله استفاده نمی‌شود زیرا طرح‌های آزمایشی به‌کار رفته برای این مدل‌ها نیاز به آن دارند که آزمایش‌های بیشتری به مرحلهٔ اجرا گذاشته شوند.

۳.۲.۷ طرح‌های مرتبهٔ اول و دوم

تقریب‌های خطی و درجهٔ دومی را که در یک تحلیل رویهٔ پاسخ به‌کار می‌روند می‌توان به‌صورت یک مسئلهٔ «رگرسیون» استاندارد برازش داد: محقق صرفاً مقادیری به قدر کافی به‌عنوان متغیرهای کنترل انتخاب می‌کند تا پارامترهای رگرسیونی را به‌طور یکتا برآورد کند. رویکردی کاراتر آن است که مقادیر متغیرهای کنترل را مطابق با یک طرح آزمایشی برگزینیم.

---

۱. مینیمم‌کردن مقدار مورد انتظار پاسخ معادل است با رو به پایین‌رفتن به سمت کف دره. در این صورت روش تندترین نزول را می‌توان به‌کار برد.

عامل‌های کدشده

فرض بر این است که متغیرها یا عامل‌های کنترلِ وابسته به یک تحلیل رویهٔ پاسخ، کمّی و مستقیماً به‌وسیلهٔ محقق قابل کنترل‌اند. این بدان معناست که هر سطحی در دامنهٔ ممکن یک عامل، قابل حصول است. با اینکه ممکن است اثرهای درجه‌بندی در عمل محدودیت‌هایی ایجاد کنند، نتیجهٔ امر این است که سطوح اصلی یک عامل را می‌توان «کدبندی» کرد. شکل رایج تبدیل کدکنندهٔ اعمال‌شده بر یک عامل، مثلاً ξ_j، عبارت است از

$$\text{مقدار کدشده} = x_j = \frac{\xi_j - \frac{1}{2}(\min \xi_j + \max \xi_j)}{\frac{1}{2}(\max \xi_j - \min \xi_j)} \qquad (۲.۷)$$

که سطوح عامل ξ_j را به بازهٔ $[-۱, ۱]$ می‌نگارد. اصطلاح متغیر طبیعی گاه برای تمیزدادن بین واحدهای اندازه‌گیری اصلی یک عامل و صورت کدشدهٔ (بدون واحد) آن به‌کار می‌رود. در بحث زیر، $x_۱, \ldots, x_k$ عامل‌های کدشده را نشان می‌دهند.

مدل مرتبهٔ اول

ردهٔ طرح‌های متعامد که واریانس ضرایب رگرسیونی مدل مرتبهٔ اول را مینیمم می‌کند مشتمل بر ۲^k طرح عاملی و ۲^{k-p} طرح عاملی کسری است که در آن‌ها اثرهای اصلی با یکدیگر هم‌اثر نشده‌اند. این طرح‌ها این امکان را می‌دهند که بتوان k عامل را با کمترین تعداد اجراهای ممکن مورد بررسی قرار داد زیرا تنها دو سطح از هر عامل مطرح‌اند. برحسب عامل‌های کدشدهٔ ما، این دو سطح با ± ۱ نشان داده می‌شوند. این طرح‌ها در فصل ۳ به‌اختصار مورد بحث قرار گرفته‌اند (نگاه کنید به آزمایش‌های برنامه‌ریزی‌شده)، و در بسیاری از کتاب‌های درسی آماری به‌طور گسترده به آن‌ها پرداخته می‌شود. به‌طور خاص، خوانندگان را برای جزئیات بیشتر دربارهٔ طرح‌های ۲^k عاملی و طرح‌های مرتبط به مونتگومری (۱۹۹۷) و باکس و همکاران (۱۹۷۸) ارجاع می‌دهیم.

با این حال یک تکرار یک طرح ۲^k توان برآورد خطای آزمایشی را ندارد مگر اینکه برخی اجراها تکرار شوند. به جای تکرار اجراها، روشی استاندارد عبارت از گسترده‌کردن ۲^k عاملی با افزودن چندین مشاهده در «مرکز» طرح است. برحسب عامل‌های کدشده، این مرکزها عبارت‌اند از نقاط $x_j = ۰$، $j = ۱, ۲, \ldots, k$. افزودن نقاط مرکزی نه خاصیت متعامدبودن طرح را تغییر می‌دهد و نه بر ضرایب رگرسیونی β_j وابسته به عامل‌های x_j تأثیر می‌گذارد.

طرح متعامد دیگری که نیازمند تنها $k+1$ اجرا برای k متغیر است سادک[1] است. این طرح عبارت از یک طرح چندوجهی منتظم با $k+1$ رأس در k بعد است. به‌عنوان مثال، با $k=2$ عامل، طرح سادکی مثلثی متساوی‌الاضلاع است؛ به‌ازای $k=3$ این طرح یک چهاروجهی منتظم است.

مدل مرتبهٔ دوم

برای برازاندن یک مدل مرتبهٔ دوم، هر عامل باید حداقل سه سطح داشته باشد. این شرط آشکارا با یک طرح 3^k عاملی برآورده و نیز با افزودن نقاط مرکزی بر طرح 2^k حاصل می‌شود. با این حال، تعداد اجراهای لازم در یک طرح 3^k با افزایش k به‌سرعت افزایش می‌یابد و هیچ‌یک از این طرح‌های عاملی طرح‌هایی دَوران‌پذیر نیستند. ما منظور از «دَوران‌پذیر» را در بحث خود دربارهٔ طرح مرکب مرکزی در زیر تعریف می‌کنیم. قبل از این کار، باید منظور از «تحلیل کانونی» یک رویهٔ پاسخ را مشخص کنیم.

تحلیل کانونی

تحلیل مدل مرتبهٔ دوم برازش‌یافته برای یک رویهٔ پاسخ را تحلیل کانونی می‌نامند که دلیل اصلی استفاده از عامل‌های کدشده است. این بدان دلیل است که تحلیل کانونی یک رویهٔ پاسخ نسبت به تغییرات مکانی و مقیاسی ناوردا نیست. در نتیجه، کدبندیِ تعریف‌شده به‌وسیلهٔ (۲.۷) به هر عامل وزنی برابر می‌دهد و سطح‌های عامل‌های متفاوت را قابل مقایسه می‌سازد.

نتایج یک تحلیل کانونی ماهیت رویهٔ پاسخ را در جواب مانای وابسته به مدل درجهٔ دوم مشخص‌سازی می‌کنند. برای دریافتن این مطلب، خاطرنشان می‌کنیم که مدل مرتبهٔ دوم برازش‌داده‌شده را می‌توان به شکل

$$\hat{y} = \hat{\beta}_0 + \boldsymbol{x}'\boldsymbol{b} + \boldsymbol{x}'\boldsymbol{B}\boldsymbol{x}$$

نوشت که در آن $\boldsymbol{x}$ بردار $k \times 1$ عامل‌های کدشده، $\boldsymbol{b}$ بردار $k \times 1$ ضرایب رگرسیونی مرتبهٔ اول برازش‌داده‌شده، و $\boldsymbol{B}$ ماتریس ضرایب درجهٔ دوم است. بهینه‌کردن پاسخ پیشگویی‌شده نسبت به $\boldsymbol{x}$ جواب مانای

$$\boldsymbol{x}_s = -\frac{1}{2}\boldsymbol{B}^{-1}\boldsymbol{b} \;\Rightarrow\; \hat{y}_s = \hat{\beta}_0 + \frac{1}{2}\boldsymbol{x}_s'\boldsymbol{b}$$

را می‌دهد که در آن $\hat{y}_s$ پاسخ پیشگویی‌شده در نقطهٔ مانااست.

1. simplex

نقطهٔ مانا به کمک علامت و اندازهٔ مقدارهای ویژهٔ وابسته به ماتریس B مشخص‌سازی می‌شود. با نشان‌دادن مقدارهای ویژهٔ B با λ_1، λ_2، ... ، λ_k، ماهیت رویهٔ پاسخ در نقطهٔ مانا را می‌توان به‌صورت زیر معین کرد:

ماکسیمم اگر همهٔ λ_jها منفی باشند، رویهٔ پاسخ یک ماکسیمم موضعی در نقطهٔ مانا دارد.

مینیمم اگر همهٔ λ_jها مثبت باشند، رویه یک مینیمم موضعی دارد.

نقطهٔ زینی اگر λ_jها علامت آمیخته داشته باشند، نقطهٔ مانا یک «نقطهٔ زینی» نامیده می‌شود. این بدان معناست که رویهٔ پاسخ (با دورشدن از نقطهٔ مانا) در یک جهت افزایش می‌یابد، اما در جهت دیگر کاهش می‌یابد.

ستیغ مانا اگر یک یا تعداد بیشتری از λ_jها به صفر نزدیک باشند، آنگاه رویهٔ پاسخ، در یک ناحیه، «تخت» است یا یک ستیغ مانا دارد. این بدان معناست که متغیر پاسخ نسبت به تغییرات در ترکیب معینی از سطوح عامل‌ها روی آن ناحیه یا در امتداد ستیغ حساس نیست.

طرح مرکب مرکزی

پراستفاده‌ترین طرح برای برازش یک مدل رویهٔ پاسخ مرتبهٔ دوم طرح مرکب مرکزی است. پنج سطح از هر عامل در این طرح‌ها که متشکل از موارد زیرند مورد نیاز است.

عاملی یک طرح 2^k یا 2^{k-p} عاملی کسری برای دو سطح از عامل. این سطوح با نمادگذاری معمول ± 1 کدبندی می‌شوند. اثرهای اصلی نباید در طرح 2^{k-p} با یکدیگر هم‌اثر شوند.

نقاط محوری مجموعاً $2k$ مشاهدهٔ دیگر با اختصاص دو اجرا به دو سطح کرانگین هر عامل انجام می‌شود. در هر حالت، دو سطح به‌صورت $\pm\alpha$ کدبندی می‌شوند و بقیهٔ عامل‌ها در سطح نقطه‌مرکزی (کدشده) قرار داده می‌شوند: $x_j = 0$. انتخاب α در زیر مورد بحث قرار می‌گیرد.

نقاط مرکزی مجموعاً n_0 اجرای نقطه‌مرکزی انجام می‌شود و همهٔ عامل‌ها در سطح نقاط میانی خود قرار داده می‌شوند. این امر با کد نقطه‌مرکزی $(0, 0, \ldots, 0)$ متناظر است.

طرح مرکب مرکزی را می‌توان از یک طرح عاملی مرتبهٔ اول 2^k یا 2^{k-p} بازسازی کرد. با انتخاب مناسب α و n_0، خاصیت‌های زیر برقرارند.

دَوران‌پذیر یک طرح را دوران‌پذیر نامند هرگاه واریانس $\hat{y}$ به‌ازای همهٔ xهای، تنها به فاصلهٔ آن نقطه از مرکز طرح بستگی داشته باشد و به جهت آن بستگی نداشته باشد.

فرض کنید $n_f = 2^{k-p}$ نشان‌دهندهٔ تعداد نقاط به‌کار رفته در بخش عاملی (کسری) طرح باشد. در این صورت، قراردادن $\alpha = \sqrt[4]{n_f}$ (ریشهٔ چهارم n_f) موجب دوران‌پذیرشدن طرح مرکب مرکزی می‌شود[1].

متعامد در یک طرح مرکب مرکزی متعامد، تعداد نقاط مرکزی به‌نحوی انتخاب می‌شود که برآوردهای پارامتری مرتبهٔ دوم، کمترین همبستگی را با دیگر برآوردهای پارامترها داشته باشند. انتخاب n_0 به k یعنی تعداد عامل‌ها و این امر بستگی دارد که آیا بلوک‌بندی به‌کار رفته است یا خیر.

دقت یکنواخت در یک طرح با دقت یکنواخت، واریانس $\hat{y}$ در مبدأ با واریانس در یک فاصلهٔ واحد از مبدأ یکسان است. یک طرح مرکب مرکزی را می‌توان با انتخاب مناسب n_0 به طرحی با دقت یکنواخت تبدیل کرد.

جدولی از چند طرح مرکب مرکزی در پیوست پ، جدول ج.۱۰، ارائه شده است. درایه‌های این جدول به کمک نرم‌افزار آماری JMP به‌دست آمده است. علاوه بر انتخاب n_0 و طرح‌های بلوک‌بندی، مقادیر αیی را که اثرها را ناهمبسته (متعامد) می‌کنند درج کرده‌ایم. این مقدارها با α_o نشان داده شده‌اند و در مورد سه نوع طرح مرکب مرکزی متعامد به‌کار می‌روند: متعامد، بلوکی، و با دقت یکنواخت.

طرح باکس-بنکن[2]

طرحی جایگزین برای طرح مرکب مرکزی طرح باکس-بنکن است که در آن، به‌ازای هر عامل، تنها به سه سطح نیاز داریم. این طرح اساساً یک طرح عاملی 3^k است، اما عملاً تنها نقاط میانی به اجرا درمی‌آیند. بنابراین، در مقایسه با طرح مرکب مرکزی، عدم حتمیت بیشتری در مجاورت رأس‌های ابرمکعب در یک طرح باکس-بنکن وجود دارد.

۴.۲.۷ ملاحظات عملی

بحث خود دربارهٔ روش‌شناسی رویهٔ پاسخ را با چند تبصره و تذکر سودمند به پایان می‌بریم.

- مفهوم ریاضی اصطلاح «بهینه» اشاره به ماکسیمم (یا مینیمم)کردن یک تابع پاسخ نسبت به متغیرهای کنترل دارد. در عمل، شکل واقعی رابطهٔ تابعی نامعلوم است و تقریب‌ها تنها شاید در یک ناحیهٔ نسبتاً کوچک سودمند باشند.

۱. از لحاظ هندسی، مقدار α متناظر است با شعاع آن‌کره‌ای که ابرمکعب تولیدشده بوسیلهٔ 2^{k-p} عاملی را محاط می‌کند. در ساده‌ترین حالت ($k = 2$)، نقاط محوری در نقاط قطب‌نما (N، E، S، و W) بر روی دایره‌ای به شعاع $\alpha = \sqrt{2} = \sqrt[4]{4}$، که مربع با رأس‌های $(\pm 1, \pm 1)$ را محاط می‌کند، قرار دارند. از آنجا که رأس‌های مربع با 2^2 نقطهٔ عاملی متناظرند، آشکار است که طرح دوران‌پذیر است.

2. Behnken

- بهینهٔ حاصل از یک تقریب «موضعی» ممکن است خارج از ناحیهٔ امکان‌پذیر بیفتد. این امر قیودی را بر شیوهٔ بهینه‌سازی اعمال می‌کند و ممکن است لازم باشد برای تعیین ترکیب بهینهٔ سطوح عامل از روش‌های برنامه‌ریزی غیرخطی استفاده کنیم.
- برای تحلیل این نوع از مسائل بهینه‌سازی به رویکردی «دنباله‌ای» نیاز داریم. به عبارت دیگر، ضرورت دارد که محقق آمادگی اجرای چندین مجموعه از آزمایش‌ها را داشته باشد تا ناحیهٔ بهینگی را تعیین کند.
- ترکیب بهینهٔ عاملی حاصل از آزمایش‌های کوچک‌مقیاس ممکن است در فرایند «مقیاس‌بالا» به دلیل تغییرپذیری ذاتی در یک عمل تام‌مقیاس[1] بهینه نباشد.[2]
- ترکیب بهینهٔ عاملی برای یک متغیر پاسخ ممکن است منتج به پاسخی ضعیف برای متغیر خروجی دیگری شود. در این صورت محقق باید در جستجوی شرایط نابهینه‌ای باشد که پاسخ‌های قابل پذیرشی برای همهٔ متغیرهای خروجی در اختیار می‌گذارد.

۵.۲.۷ تحلیل مقدماتی

شیوهٔ RSREGدر SAS مشخصاً برای تحلیل آزمایش‌های رویهٔ پاسخ طراحی شده است. در کد زیر، رطوبت را به‌عنوان متغیر پاسخ Y اختیار کرده‌ایم.

برنامه *SAS*

```
data a ;
      infile 'c72.dat' ;
      input a b e g h density thick moist flavor ;
      y = moist ;

proc rsreg data=a  ;
      model y = a b e g h  ;
```

شیوهٔ RSREGبر این مجموعه‌داده‌ها اعمال می‌شود و خروجی شامل نتایج برازش یک مدل مرتبهٔ دوم به Y است. به عبارت دیگر، جمله‌های درجهٔ دوم و حاصل‌ضرب به‌طور خودکار به کمک SAS از فهرست عامل‌های مشخص‌شده در دستور مدل تولید شده‌اند. پیش از برازش‌دادن مدل

1. full-scale

۲. این تغییرات ممکن است به این امر نیز منجر شوند که فرایند از وضعیت بهینه «رانده» شود. فن عمل تکاملی (*evolutionary operation*=EVOP)، مطرح‌شده توسط باکس (۱۹۵۷) و باکس و درپر (۱۹۶۹)، روشی برای حفظ شرایط عمل بهینه در اختیار می‌گذارد.

درجهٔ دوم، SAS از (۲.۷) برای کدبندی هر عامل استفاده می‌کند. خروجی کامل این برنامهٔ SAS در زیر ارائه شده است.

خروجی *SAS*

Coding Coefficients for the Independent Variables

Factor	Subtracted off	Divided by
a	625.000000	75.000000
b	1.900000	0.400000
e	20.000000	6.000000
g	1.500000	1.500000
h	270.000000	30.000000

Response Surface for Variable y

Response Mean	12.451667
Root MSE	0.201209
R-Square	0.9902
Coefficient of Variation	1.6159

Regression	DF	Type I Sum of Squares	R-Square	F Value	Pr > F
Linear	5	36.573050	0.9795	180.67	<.0001
Quadratic	5	0.205313	0.0055	1.01	0.4626
Crossproduct	10	0.197187	0.0053	0.49	0.8611
Total Model	20	36.975551	0.9902	45.67	<.0001

Residual	DF	Sum of Squares	Mean Square
Total Error	9	0.364366	0.040485

Parameter	DF	Estimate	Standard Error	t Value	Pr > |t|	Parameter Estimate from Coded Data
Intercept	1	-11.525846	28.875689	-0.40	0.6991	12.526067
a	1	0.011770	0.045124	0.26	0.8001	-0.209533
b	1	1.536970	7.412558	0.21	0.8404	-0.192992
e	1	1.240421	0.475572	2.61	0.0284	2.436631
g	1	1.679162	1.834098	0.92	0.3838	0.274575
h	1	0.023274	0.115973	0.20	0.8454	-0.168621
a*a	1	-0.000009458	0.000027669	-0.34	0.7403	-0.053200
b*a	1	0.005113	0.006702	0.76	0.4651	0.153384
b*b	1	-0.566433	0.974196	-0.58	0.5752	-0.090629
e*a	1	-0.000538	0.000447	-1.20	0.2592	-0.242141
e*b	1	-0.085937	0.083837	-1.03	0.3321	-0.206250
e*e	1	-0.008420	0.004330	-1.94	0.0837	-0.303129

g*a	1	-0.000663	0.001787	-0.37	0.7192	-0.074612
g*b	1	0.135417	0.335349	0.40	0.6958	0.081250
g*e	1	-0.004861	0.022357	-0.22	0.8327	-0.043750
g*g	1	0.048609	0.069276	0.70	0.5006	0.109371
h*a	1	-0.000002589	0.000089360	-0.03	0.9775	-0.005825
h*b	1	-0.005729	0.016767	-0.34	0.7404	-0.058750
h*e	1	0.000034722	0.001118	0.03	0.9759	0.006250
h*g	1	-0.005139	0.004471	-1.15	0.2801	-0.231250
h*h	1	-0.000017366	0.000173	-0.10	0.9223	-0.015629

Factor	DF	Sum of Squares	Mean Square	F Value	Pr > F
a	6	0.363290	0.060548	1.50	0.2817
b	6	0.306766	0.051128	1.26	0.3612
e	6	35.722186	5.953698	147.06	<.0001
g	6	0.534401	0.089067	2.20	0.1385
h	6	0.229543	0.038257	0.94	0.5091

Canonical Analysis of Response Surface Based on Coded Data

Factor	Critical Value Coded	Uncoded
a	7.022171	1151.662799
b	4.738991	3.795596
e	-0.257964	18.452219
g	-2.134848	-1.702273
h	-1.383836	228.484934

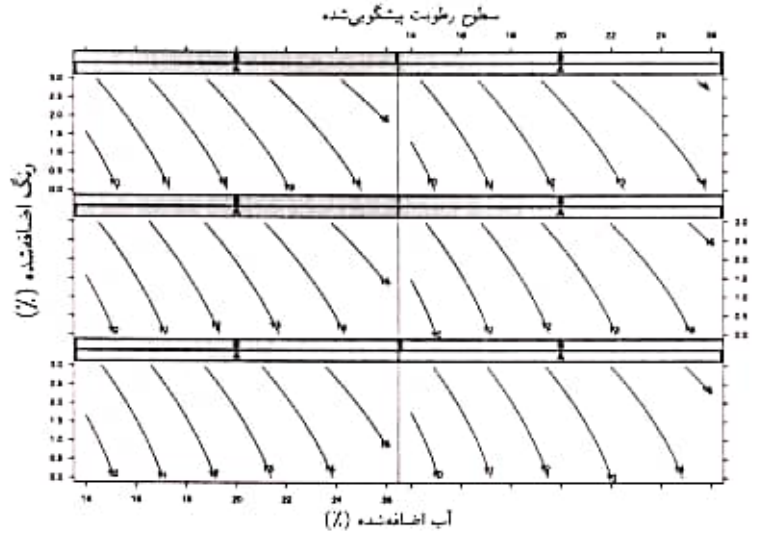
Predicted value at stationary point: 10.842389

Eigenvalues	Eigenvectors a	b	e	g	h
0.191566	-0.018442	0.205180	-0.078416	0.836458	-0.501747
0.073780	0.733600	0.541105	-0.374166	-0.168158	-0.027548
-0.084988	0.011803	0.111807	-0.095366	0.481977	0.863692
-0.162026	0.611921	-0.766940	0.005267	0.192547	-0.015948
-0.371550	0.294796	0.253789	0.919091	0.051815	0.035686

Stationary point is a saddle point.

از تحلیل نتایج واریانس ظاهراً چنین فهمیده می‌شود که رویهٔ پاسخ خمیدگی معنی‌داری ندارد و اینکه رطوبت موجود در مادهٔ غذایی اساساً به کمک عامل E، یعنی درصد آب اضافه‌شده، معین می‌شود.[۱] از مباحث تشخیصی رگرسیونی وابسته به هر جمله از مدل، شواهد ضعیفی

۱. با اینکه این مطلب از لحاظ شهودی کاملاً موجه است، رویکرد آماری نه تنها مؤید این نتیجهٔ «شهودی» است، بلکه از آن چنین برمی‌آید که رطوبت موجود را می‌توان تماماً بوسیلهٔ عامل E کنترل کرد. تفاوت اساسی آن است که نتیجهٔ «آماری» بر روش‌های عینی استوار است.



شکل ۱.۷ خم تراز سطح رطوبت با استفاده از نمودارهای داربستی.

(Pr>|t|=0.0837) از اثر درجهٔ دوم ناشی از عامل E پدیدار می‌شود. بنابراین، رابطهٔ بین عامل E و رطوبت موجود لزوماً خطی نیست.

تحلیل کانونی آشکار می‌کند که جواب «بهینه» یک نقطهٔ زینی است. همچنین این امر را نشان می‌دهد که حصول بهینه غیرممکن است زیرا عامل G نیاز به افزودن یک مقدار منفی رنگ است (که در عمل چندان ساده نیست). البته در این حالت نتایج تحلیل واریانس قبلاً نشان داده‌اند که رطوبت به هیچ‌یک از عامل‌ها بجز E حساس نیست و لذا شرایط بهینه‌ای را که از تحلیل کانونی برمی‌آید باید به‌عنوان مصنوعات ریاضی بهینه‌سازی تلقی کرد.

یک رویهٔ پاسخ را می‌توان به‌راحتی با استفاده از نمودارهای خم تراز نشان داد. در این مثال، ابتدا مدل برازش‌داده‌شده را برای تولید مقدارهای رطوبت پیشگویی‌شده روی مشبکه‌ای از نقاط تعریف‌شده به‌وسیلهٔ عامل‌ها به‌کار بردیم و سپس از دستگاه نمودارهای «داربستی[1]» در S-PLUS برای تولید دنباله‌ای از نمودارهای خم تراز شرطی استفاده کردیم. نتایج در شکل ۱.۷ نشان داده شده‌اند. همان‌گونه که می‌توان انتظار داشت، به نظر می‌رسد که رویه از نواری به نوار دیگر تغییر کمی می‌کند که با نتیجه‌گیری‌های قبلی ما از خروجی عددی تطابق دارد. جزئیات مربوط به این نمودار بعداً مورد بحث قرار می‌گیرد.

1. trellis

جزئیات نمودار خم تراز

عامل‌های E و G برای تولیدکردن مشبکهٔ اصلی نقاط به‌ازای هر نمودار خم تراز «انفرادی» انتخاب شدند. از دو سطح A (۵۹۰، ۶۶۰) و سه سطح B (۱٫۱، ۱٫۹، ۱٫۷) برای تدارک دنبالهٔ نمودارهای «نوار» شرطی استفاده شد. عامل H در مقدار بهینهٔ ۲۲۸ ثابت نگه داشته شد که دمای عمل قابل حصولی است و در عین حال این امر صدور آسان پاسخ‌های پیشگویی‌شده را روی مشبکه میسر کرد. در کد SAS زیر فرض می‌شود که مجموعه‌داده‌های "a" که قبلاً آن را تولید کردیم در دسترس‌اند:

```
data b ;
     set a end=eof ;  * copy the dataset "a" into "b"  ;
     output ;         * when eof = end-of-file reached ;
                      * start the grid of points ;

     if eof then
        do ;
           y = . ;    * set the response Y as missing  ;
           h = 228 ;  * fix factor H at the value 228  ;

                      * begin the A, B Panel loop      ;

           do a = 590 to 660 by 70;    * 2 levels of A ;
           do b = 1.7 to 2.1 by 0.2 ;  * 3 levels of B ;

                      * begin the E, G main grid loop  ;

             do e = 14 to 26 by 1 ;
             do g = 0  to 3  by 0.5 ;

                output ;  * output the current value of ;
                          * Y,H,A,B,E for every G loop  ;

     end ; end ; end ; end ; end ;
*-------------------------------------------------------;
      * the nonmissing observations are needed to fit   ;
      * the RS model which is then used to predict the  ;
      * missing Y's over the grid. The predicted values ;
      * are saved to the dataset "b2"                   ;

proc rsreg data=b out=b2 noprint ;
      model y = a b e g h / predict ;
```

```
*-------------------------------------------------------;
data _LAST_ ;        * set up dataset "b2" for exporting ;
  set b2 ;           * to the file "f0.dat" (extension   ;
  if h = 228 ;       * added by SAS). Only output the    ;
  file f0 ;          * predicted values (when H=228)     ;
  put a b e g y ;    * hence no need to output factor H  ;
```

فایل "f0.dat" را می‌توان حالا مستقیماً در S-PLUS وارد کرد و تابع زیر برای رسم نمودارهای خم تراز داربستی ایجاد شده است. توجه کنید که باید یک ابزار گرافیکی را پیش از آنکه g.cont() اجرا شود مشخص کرد. پیوست ب.۲ جزئیات بیشتری دربارهٔ S-PLUS در اختیار می‌گذارد (شاید بحث ارائه‌شده پس از مطالعهٔ موردی ۴.۶، روان‌شناسی معکوس، نیز برای خواننده سودمند باشد).

```
g.cont <- function(file="f0.dat") {
   xx <- read.table(file,col.names=c("A","B","E","G","Y"),
                    row.names=NULL)
   attach(xx)
   ans <- contourplot(Y ~ E * G | A * B , data = xx,
               xlab = "Water Added (%)",
               ylab = "Dye Added (%)",
               main="Predicted Moisture Surface")
   detach()
   ans
}
```

برای آنکه از نمودارهای داربستی استفاده کنیم، لازم است یک data frame ایجاد و سپس آن را "attach" کنیم به‌طوری که بتوان نام‌های ستونی را به‌عنوان متغیرهای یک دستور مدل به‌کار برد. در این مثال، ما از تابع نمودارهای گرافیکی داربستی contourplot() با یک دستور مدل استفاده کردیم که E و G را به‌عنوان متغیرهای مشبکه مشخص می‌کند تا در نمودار خم تراز اصلی برای Y به‌کار برده شود. این نمودار خم تراز روی تعداد ترکیب‌های وابسته به متغیرهای (A و B) ظاهرشده پس از علامت "|" شرطی است. نمایش تکمیل‌شده سپس به‌عنوان شئ گرافیکی S-PLUS با نام "ans" برای رسم نمودار بازگردانده می‌شود.

در شکل ۱.۷، سطوح عامل‌های A و Bبه کمک یک خط کوچک در دو نوار سایه خورده نشان داده شده‌اند. گرچه دیدن خط نشانگر تا حدی مشکل است، سه نمودار سمت چپ همگی با A در سطح پایین (۵۹۰) متناظرند؛ سه نمودار سمت راست نیز همه با $A = ۶۶۰$ متناظرند. با خواندن نمودارها از بالا به پایین در هر طرف، خط نشانگر نوار B متناظر با سطوح B است و از بالا (۱٫۲)، به وسط (۱٫۹) و به سمت پایین (۱٫۷) امتداد دارد.

۶.۲.۷ خلاصه

گام بعدی، تحلیل متغیرهای پاسخ دیگر به شیوه‌ای مشابه است تا دریابیم که آیا بهینهٔ کلی مناسبی موجود است یا خیر. در صورت موجود نبودن، مشاور آماری می‌تواند از این نتایج برای توصیه به مشتری در این مورد که آزمایش بعدی کجا انجام شود استفاده کند. گرچه این مطلب به‌عنوان تمرینی به عهدهٔ خوانندهٔ علاقه‌مند گذاشته می‌شود، باید تأکید کنیم که تجربهٔ مشتری نقش مهمی در تعیین اینکه آزمایش بعدی باید کجا انجام شود بازی می‌کند.

در پروژه‌های تحقیقاتی در جریان، از قبیل آنچه در این مطالعهٔ موردی ارائه شده است، نقش آموزشی مشاور می‌تواند به قدر تدارک یک طرح کارا اهمیت داشته باشد. در این مثال، توانستیم مشتری را نسبت به مزیت‌های اتخاذ یک رویکرد تکراری در تحقیق متقاعد کنیم. در نتیجه، مشتری قادر بود دنباله‌ای تکاملی از آزمایش‌های کوچک‌مقیاس را با استفاده از طرح‌های تخصصی کارا برای نیل به اهداف تحقیق خود و اصلاح آن‌ها اجرا کند.

روش‌شناسی رویهٔ پاسخ فن آماری بسیار سودمندی است، اما مستلزم توجه قابل ملاحظه‌ای به جزئیات طراحی و اجرای آزمایش‌هاست. با مروری قهقرایی، خطایی که در به وجودآوردن طرح غربالگری عاملی کسری آغازین برای مشتری مرتکب شدیم نامیدن سطوح هر عامل به‌صورت «بالا» و «پایین» بود. ۲۰٪ شکست‌ها آشکارا ناشی از گمراه‌کردن مشتری از سوی ما در ایجاد این تصور بود که این سطوح باید نزدیک کرانگین‌های دامنهٔ هر یک از عامل‌ها باشند. آموزش دو لبه دارد... دفعهٔ بعد از اصطلاحاتی نظیر «بالا/پایین» یا «±۱»، تا زمانی که دامنهٔ عامل به‌خوبی به‌وسیلهٔ مشتری معین نشده باشد، استفاده نخواهیم کرد.

پرسش‌ها

۱. کدبندی پیش‌فرض (۲.۷) مورد استفاده در SAS، کل طرح را دوباره به‌گونه‌ای مقیاس‌بندی می‌کند که نقطه‌های محوری در نقاط ۱± قرارگیرند. این کارکدبندی منقش[1] نامیده می‌شود که کاربر می‌تواند آن را با گزینهٔ

```
model y = ... / nocode ;
```

منتفی کند. در این آزمایش، سطوح واقعی عامل A تفاوتی جزئی با کدبندی ۲±، ۱±، ۰ مورد انتظار در یک طرح مرکب مرکزی پنج‌عاملی (نیم تکرار) دارد. فایل جدیدی شامل کدبندی مورد انتظار برای عامل‌ها ایجاد کنید و تحلیل را با استفاده از گزینهٔ «/nocode» مجدداً اجرا کنید. آیا این امر تفاوتی در یافته‌ها یا نتیجه‌گیری‌ها ایجاد می‌کند؟

1. *inscribed* coding

۲. متغیرهای پاسخ دیگری را مورد بررسی قرار دهید. آیا بهینهٔ کلی مناسبی در درون ناحیهٔ تنیده‌شده به‌وسیلهٔ عامل‌های این آزمایش موجود است؟ اگر نیست، آزمایش بعدی کجا باید انجام شود؟ (برای این منظور می‌توان پس از مشخص‌ساختن مدل، دستور: `ridge max` را اضافه کرد.)

۳.۷ هزینه‌کرد در شهرداری‌های نیویورک

روش‌ها	رگرسیون خطی چندگانه
داده‌ها	مالیات بر دارایی

۱.۳.۷ مدلبندی رگرسیونی

هدف این مطالعهٔ موردی فراهم‌کردن برآوردی از اثر قیمت در هزینه‌کردهای شهرداری در نتیجهٔ ساخت‌وسازهای پیشنهادی در پروژه‌های جدید خانه‌سازی در سه شهر ایالت نیویورک است. از آنجا که بسیاری از خدماتی که شهرداری ارائه می‌کند عمدتاً از طریق مالیات بر دارایی تأمین می‌شود، روشن است که تلاش برای تعیین این امر که آیا این پروژه‌ها موجب افزایش در هزینه‌کردها خواهند بود یا خیر در کانون توجه قرار دارد. رویکرد ما به این مسئله ایجاد مدل رگرسیونی مناسبی برای پیشگویی هزینه‌کردهای سرانه بر مبنای مجموعهٔ متداولی از پیشگوهای مربوط به آمارهای جمعیتی و مرتبط با درآمد است. این متغیرها را متعاقباً توصیف می‌کنیم، اما توجه به این مطلب حائز اهمیت است که استفاده از هزینه‌کردهای سرانه به‌عنوان متغیر پاسخ به ما اجازه می‌دهد که مقایسه‌های مستقیمی بین شهرهای مختلف صورت دهیم.

با این حال، باید همچنان به موضوع پیشگویی هزینه‌کردهای آتی بپردازیم زیرا این امر مبنای تصمیم‌گیری در این مورد خواهد بود که آیا جواز ساختمان برای پروژه‌های جدید خانه‌سازی صادر شود یا خیر. برای انجام‌دادن این مطالعهٔ موردی، توجه خود را به فرایند مدلبندی معطوف می‌کنیم. این کار در برگیرندهٔ گام‌های زیر است:

۱. داده‌هایی از همهٔ شهرداری‌های ایالت نیویورک دربارهٔ متغیرهای پاسخ و پیشگو برای سال ۱۹۹۲ گردآوری شد. داده‌های ۱۹۹۲ در `c73.dat` فراهم شده و برای به‌دست‌آوردن مدل رگرسیونی مناسبی برای پیشگویی هزینه‌کردهای سرانه به‌کار می‌روند.

۲. مقادیر پیش‌بینی برای متغیرهای پیشگو در سه شهر برای سال‌های

۲۰۰۵ و ۲۰۲۵ برآورد شدند. این پیش‌بینی‌ها در جدول ۶.۷ ارائه شده‌اند و در مدل رگرسیونی برای تولید هزینه‌کردهای آتی پیشگویی‌شده به‌کار گرفته شدند.

البته چالش ما عبور از نخستین گام است.

پیش از نگاه منصل‌تر به داده‌ها و متغیرها، تأکید می‌کنیم که باید سال‌های ۲۰۰۵ و ۲۰۲۵ را صرفاً به‌عنوان برچسب‌های مشاهدات جدیدی در نظر بگیریم که قصد داریم مقادیر پیشگویی را برای آن‌ها محاسبه کنیم. پیش‌بینی همهٔ تغییرات و رخدادهایی که طی ۳۰ سال بعدی در یک شهرداری اتفاق می‌افتند عملاً غیرممکن است. حقیقتاً سعی نداریم که هزینه‌کردهای آینده را پیش‌بینی کنیم، بلکه تنها به تغییرات مشخصی علاقه‌مندیم که محتملاً نتیجهٔ خانه‌سازی جدید در زمان حاضر خواهند بود. این اطلاعات است که آن را در مقادیر پیش‌بینی‌شدهٔ پیشگوها دخالت داده‌ایم؛ از هر چیز دیگری که مستقیماً تحت تأثیر پروژه‌های خانه‌سازی جدید نخواهند بود صرف‌نظر شده است.

داده‌ها

داده‌ها مشتمل بر یک پاسخ و شش پیشگوی مربوط به آمارهای جمعیتی‌اند. همچنین سه متغیر مربوط به خانه‌ها یعنی ایالت، شهرستان، و کد شناسایی (ID) موجودند. پاسخ هزینه‌کرد[1] است که به‌صورت سرانه برحسب دلار اندازه‌گیری می‌شود. پیشگوها عبارت‌اند از جمعیت، تراکم، درصد درون‌حکومتی[2]، ثروت سرانه، درآمد سرانه، و نرخ رشد جمعیت.

متغیر تراکم به‌صورت نسبت جمعیت به سطح تعریف می‌شود و همبستگی شدیدی با جمعیت دارد. از این امر چنین برمی‌آید که شاید لازم باشد که متغیر سطح را به مجموعه‌داده‌ها اضافه کنیم و یکی از دو متغیر دیگر را حذف کنیم. ثروت سرانه تنها اندازه‌ای برای ثروت مرتبط با ارزش واقعی مستغلات است و با درآمد همبسته است، اما این‌ها چیزهای مختلفی را اندازه‌گیری می‌کنند.

متغیر درصد درون‌حکومتی معرف درصد عایداتی است که از مبالغ اعطایی یا یارانه‌های ایالتی یا دولت فدرال حاصل می‌شود. این متغیر با ثروت و درآمد همبسته است، اما عامل‌های دیگری مانند خدمات عمومی را نیز که به‌وسیلهٔ شهرداری‌ها به دیگر مجامع می‌شود (مانند ایستگاه قطار رفت و برگشت بزرگی که در یک شهرداری کوچک واقع شده است) منعکس می‌کند.

1. Expenditure 2. Percent Intergovernmental

جدول ۴.۷ تعریف‌های متغیر

متغیر	تعریف
ID	کد شناسایی (برای جورسازی)
ST	کد ایالت
CD	کد شهر
EXPEN	هزینه‌کرد سرانه
WEALTH	ثروت سرانه
POP	جمعیت
PINTG	درصد درون‌حکومتی
DENS	تراکم
INCOME	درآمد
GROWR	نرخ رشد

مقصود از این بررسی پیشگویی هزینه‌کردهای سال‌های ۲۰۰۵ و ۲۰۲۵ برای شهرهایی است که در جدول‌های ۵.۷ و ۶.۷ فهرست شده‌اند.

روش‌شناسی

۲.۳.۷ تحلیل رگرسیونی

یک مدل رگرسیونی اغلب می‌تواند تقریب مناسبی برای مسئلهٔ پیشگویی در اختیار بگذارد. رویکرد کلی در برازش رگرسیونی در فصل ۳ (رگرسیون) توضیح داده شد. در آنجا نقش مهمی را که ابزارهای تشخیصی مانده‌ای در ارزیابی کیفیت و اعتبار مدل برازش‌یافته ایفا می‌کنند مورد بحث قرار دادیم. در این مطالعهٔ موردی، همچنین به ارزیابی کیفیت پیشگویی‌ها نیاز داریم. ما در زیر توجه خود را بر این مباحث معطوف می‌کنیم.

تبدیل‌ها

چند نمودار پراکنش و نمودار Q_Q برای توجیه نیاز به تبدیل داده‌ها کافی است. این نمودارها دم‌هایی بلند به سمت مقادیر بزرگ مثبت را برای اغلب متغیرها از جمله پاسخ نشان می‌دهند. این مقادیر بزرگ به‌عنوان نقاط اهرمی در رگرسیون عمل می‌کنند اما چون بخشی از دمی طولانی‌اند،

جدول ۵.۷ شهرهای مورد نظر (داده‌های سال ۱۹۹۲)

	وارویک	مونرو	تاکسیدو
ST	36	36	36
CO	33	33	33
EXPEN	237	159	926
WEALTH	78908	55067	155034
POP	16225	9338	2328
PINTG	24.7	8.8	6.1
DENS	170	599	52
INCOME	19044	16726	30610
ID	8730	5420	8400
GROWR	30.3	30.0	2.5

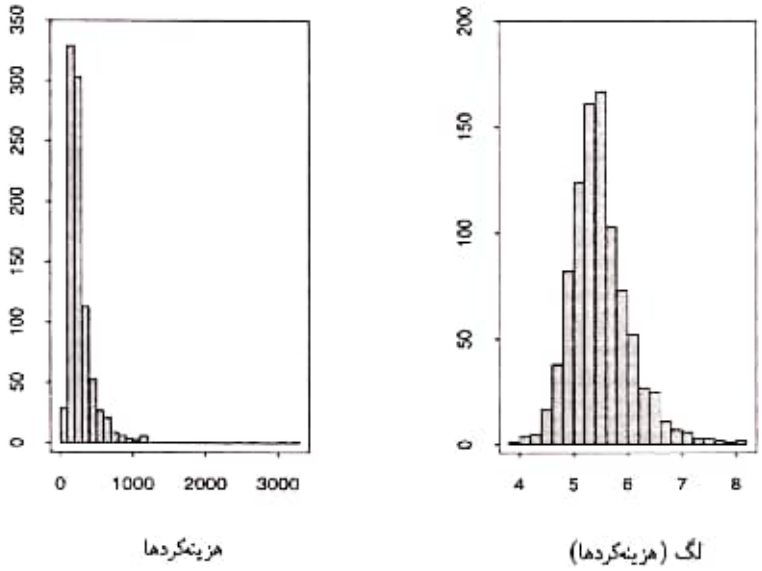
جدول ۶.۷ داده‌های پیش‌گویی‌شده برای شهرهای موردنظر

شهر	Year	POP	WEALTH	PINTG	DENS	INCOME	GROWR
وارویک	1992	16225	78908	24.7	170	19044	30.3
	2005	20442	85000	24.7	214	19500	35.0
	2025	31033	89000	26.0	325	20000	40.0
مونرو	1992	9338	55067	8.8	599	16726	30.0
	2005	10496	58000	8.8	695	17100	35.0
	2025	13913	60000	10.1	959	18000	35.0
تاکسیدو	1992	2328	155034	6.1	52	30610	2.5
	2005	10685	116000	6.1	249	28300	300.0
	2025	29246	115000	7.0	656	25000	100.0

اگر آن‌ها را حذف کنیم، حتی در صورت تکرار چندین‌بارهٔ فرایند، نقاط دیگری ظاهر خواهند شد. انجام‌دادن این کار شاید تمرین خوبی برای خواننده باشد.

از شکل ۲.۷ چنین برمی‌آید که یک تبدیل لگاریتمی ظاهراً می‌تواند به‌طرزی مناسب از پسِ داده‌های به‌شدت چولهٔ هزینه‌کرد برآید. گرچه مقادیر تبدیل یافته هنوز هم مختصری چولگی راست را نشان می‌دهند، تعبیر یک تبدیل لگاریتمی را آسانتر می‌توان برای مشتری توجیه کرد تا مثلاً یک تبدیل معکوس ریشهٔ دوم را.

تبدیل لگاریتمی و دیگر تبدیل‌های توانی زمانی که داده‌ها شامل صفر یا چند مقدار منفی باشند مناسب نیستند. یک امکان در چنین مواردی اضافه‌کردن مقداری ثابت و محاسبهٔ $\log(x+c)$ است. برای اعداد نامنفی اگر قرار دهیم $c = 1$، در این صورت تبدیل ۰ را به ۰ خواهد نگاشت. می‌توان از مقادیر بزرگتر c در صورت وجود چند مقدار منفی استفاده کرد و ممکن است مواردی



شکل ۲.۷ توزیع‌های اصلی و لگی تبدیل‌یافته.

وجود داشته باشند که در آن‌ها از داده‌ها مقداری مانند c را کسر کنیم زیرا داده‌ها شامل هیچ عدد کوچکی نیستند.

همچنین مواردی از این قبیل وجود دارند که متغیر دم‌هایی بسیار طولانی برای مقادیر مثبت و منفی از خود نشان می‌دهند. تبدیل‌های توانی رایج تنها برای مقادیر نامنفی x قابل تعریف‌اند. برای برطرف‌کردن این مشکل، از تبدیلی مانند $t(x)$ برای $x > 0$ و تبدیل دیگری مانند $-t(-x)$ برای $x < 0$ استفاده می‌کنیم، مادام که $t(x)$ و $-t(-x)$ یکنوا باشند. به‌عنوان مثال $\log(x+1)$ به‌ازای $x > 0$ و $-\log(-x+1)$ به‌ازای $x < 0$. این تبدیل صفر را در صفر نگاه می‌دارد و یکنواست.

نقاط دورافتاده

باید دورافتاده‌ها را تشخیص و مورد بررسی قرار داد تا معلوم شود که آیا خطاهایی واقعی‌اند یا مقادیر واقعی معتبری هستند. برای مقاصد مدلبندی سودمند است که آن‌ها را حذف کنیم یا زمانی که به‌شدت مدل خروجی را تحت تأثیر قرار می‌دهند به آن‌ها کمتر بها دهیم. یک مدل خطی برازش‌داده‌شده به روش کمترین توان‌های دوم خطا بدون حذف دورافتاده‌ها فرض‌های نرمال

استانداردبودن را دربارهٔ خطاها برآورده نخواهد کرد و خروجی مدل مرتبط با بازه‌های اطمینان و آماره‌های آزمون دربارهٔ مقادیر پارامترها نامعتبر خواهد شد.

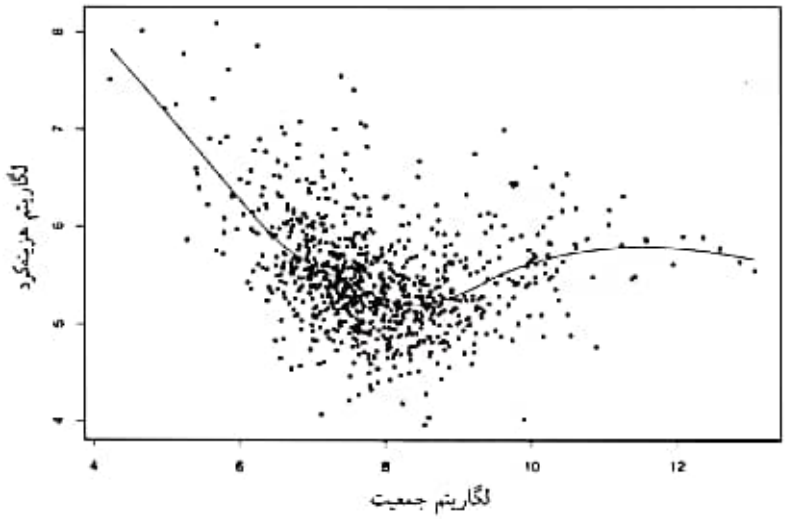
شیوهٔ رگرسیونی موجود در SAS یک روش‌شناسی اجرایی برای تشخیص دورافتاده‌ها به نام مباحث تشخیصی تأثیر دارد، اما فاقد هر نوع روش‌شناسی رگرسیونی استوار است. اگر یک یا چند دورافتاده پیدا شود، شیوه را باید با حذف دورافتاده‌ها تکرار کرد تا زمانی که دیگر هیچ دورافتاده‌ای پیدا نشود. دیگر نرم‌افزارها مانند S-PLUS شیوه‌های رگرسیونی استوار را منظور کرده‌اند که می‌توان از آن‌ها برای پرهیز از دورافتاده‌ها و مشاهدات تأثیرگذار استفاده کرد. یک قاعدهٔ عملی، اجرای رگرسیون با استفاده از هر دو روش کمترین توان‌های دوم و نیز روش رگرسیون استوار و مقایسهٔ برآوردهای پارامترها از دو روش است. اگر آن‌ها خیلی شبیه به هم باشند، آنگاه با استفاده از تحلیل استاندارد کار را پیش می‌بریم اما اگر این برآوردها کاملاً متفاوت باشند مجبوریم دورافتاده‌هایی را که مسئول این اختلاف‌اند بیابیم و سپس به پیش برویم.

رابطه‌های غیرخطی

گاهی به رابطه‌های غیرخطی برمی‌خوریم که برازش آن‌ها با استفاده از یک چندجمله‌ای درجهٔ پایین کار چندان آسانی نیست و مستلزم فنون رگرسیونی ناپارامتری پیچیده‌تری است. به‌علاوه، این روش‌شناسی غیرخطی محاسبهٔ بازه‌های اطمینان و Pـمقدارها را دشوار می‌کند و مدل‌ها اغلب فاقد تعبیر مناسبی هستند. مجدداً، یک قاعدهٔ سرانگشتی عملی عبارت است از مقایسهٔ R^2های حاصل از برازش‌دادن یک چندجمله‌ای درجهٔ پایین و یک مدل ناپارامتری غیرخطی نظیر آنچه از تابع gam() در S-PLUS به‌دست می‌آید.

وقتی مجموعه داده‌ها به قدرکافی بزرگ باشند، راهی ساده برای عبوراز این مشکل انتخاب یک زیرمجموعه حول نقطه‌ای است که می‌خواهیم پیشگویی را در آن انجام دهیم و سپس برازش‌دادن یک مدل خطی یا مدل چندجمله‌ای درجهٔ پایین.

مجموعه داده‌های تحت مطالعه، شامل مثال‌هایی از این رابطه‌های غیرخطی است. نمودار لگاریتم هزینه‌کرد در برابر لگاریتم جمعیت رابطه‌ای غیرخطی را نشان می‌دهد که نمی‌توان یک چندجمله‌ای درجهٔ ۲ یا ۳ به آن برازش داد. بنابراین، قابل توجیه است که زیرمجموعه‌ای از داده‌ها را با محدودکردن جمعیت به بازه‌ای حول آن مقداری که می‌خواهیم پیشگویی خود را در آن انجام دهیم اختیار کنیم.



شکل ۳.۷ لگاریتم (هزینه‌کرد) در برابر لگاریتم (جمعیت) با یک خم هموار برازش داده‌شده.

شکل ۳.۷ نشان می‌دهد که رابطهٔ بین لگاریتم هزینه‌کرد و لگاریتم جمعیت با مقادیر بالای لگاریتم هزینه‌کرد آغاز می‌شود و تا زمانی که به مینیمی حول جمعیتِ ۱۰۰۰۰ برسد به‌شدت کاهش می‌یابد و دوباره با آهنگ ملایم‌تری افزایش می‌یابد. جالب است بدانیم که آیا این الگو برای سایر وضعیت‌ها هم برقرار است یا خیر که در صورت مثبت‌بودن پاسخ از آن این گونه برداشت می‌شود که دامنه‌هایی از اندازهٔ جمعیت موجود است که هزینه‌کردهای شهری را بهینه می‌کند.

بازه‌های پیشگویی

لازم است که اثر پروژه‌های خانه‌سازی جدید را بر هزینه‌کردهای سرانهٔ این سه شهر برآورد کنیم. ویژگی خاص این مطالعه در تعبیر جملهٔ خطای مدل است. اگر مانده‌های سال به سال را از طریق مدلبندی هزینه‌کردهای سرانه با هم مقایسه کنیم، مشاهده می‌کنیم که بسیار همبسته‌اند. اقتصاددانان باور دارند که این امر ناشی از آن است که جملهٔ خطا عمدتاً ویژگی‌های آن شهر خاص را که بر اساس مدل توضیح داده نمی‌شود اندازه می‌گیرد. به‌علاوه، یک مؤلفهٔ خطای تصادفی نیز در بین است که نوسانات تصادفی مستقل سال به سال را اندازه‌گیری می‌کند. تفکیک این دو جمله بدون داده‌های طولانی‌مدت غیر ممکن است. برای پرهیز از این مشکل، گزینه‌ای دیگر عبارت از آن است که اثرهای پروژه‌های خانه‌سازی جدید را با محاسبهٔ تفاوت مقادیر پیشگویی‌شده با مقادیر فعلی پیشگو و مقادیر مربوط

به سال‌های ۲۰۰۵ و ۲۰۲۵ برآورد کنیم و، به جای پیشگویی‌های واقعی، تغییر را توصیف کنیم.

موضوع دوم آن است که دربارهٔ محاسبهٔ مقادیر پیشگویی‌شده چه باید کرد زیرا یک تبدیل لگاریتمی را بر داده‌ها اعمال کرده‌ایم. ما از نتیجه‌ای دربارهٔ میانگین توزیع لگ‌نرمال استفاده می‌کنیم که می‌گوید $E(e^X) = e^{\mu+\sigma^2/2}$. در حالت مورد نظر ما، این بدان معناست که، برای به‌دست‌آوردن مقادیر مورد انتظار در مقیاس اصلی، از فرمول بالا استفاده می‌کنیم و به جای μ مقدار لگاریتم هزینه‌کرد و به جای σ انحراف استاندارد، لگاریتم جمعیت را قرار می‌دهیم که هر دو از خروجی رگرسیونیِ تولیدشده به‌وسیلهٔ SAS یا S-PLUS قابل حصول‌اند.

همین نوع محاسبه وقتی تبدیل ریشهٔ دوم را هم به‌کار می‌بریم قابل اعمال است. برای محاسبهٔ مقدارهای پیشگویی‌شده برحسب مقیاس اصلی از فرمول $E(X^2) = \mu^2 + \sigma^2$ استفاده می‌کنیم.

این محاسبات را در صورت کوچک‌بودن اندازهٔ نمونه‌ها توصیه نمی‌کنیم زیرا از برآوردهای μ و σیی استفاده می‌کنیم که خود حاوی خطا هستند. فرض کنید که مدلی را به لگاریتم پاسخ برازش می‌دهیم و اینکه خطای استاندارد مقدار پیشگویی شده را بیش‌برآورد می‌کنیم. سپس آن را به توان دو می‌رسانیم و این را به نما اضافه می‌کنیم که منجر به بیش‌برآورد بسیار بزرگی برای پاسخ تبدیل‌یافته می‌شود. همین اتفاق در مورد بازه‌های اطمینان و بازه‌های پیشگویی استخراج‌شده از نمونه‌های کوچک نیز خواهد افتاد. پس از آنکه بازه‌ها را پس از تبدیلی لگاریتمی به نمای e می‌رسانیم، نتایج ممکن است به دلیل خطای برآوردها قابل اعتماد نباشند.

۳.۳.۷ تحلیل مقدماتی

با پیروی از ایده‌هایی که در بالا به‌اختصار بیان شد، یک تحلیل ممکن را برای داده‌ها پیشنهاد می‌کنیم:

- تبدیل‌ها را انتخاب کنید.
- زیرمجموعه‌ای از داده‌ها شامل ناحیه‌ای که پیشگویی‌ها را در آنجا انجام می‌دهیم برگزینید.
- جستجویی در مورد وجود دورافتاده‌ها انجام دهید و وجود هم‌خطی را وارسی کنید.
- بهترین مدل رگرسیونی را برازش دهید.
- مقدارهای پیشگویی‌شده و انحراف استانداردهای آن‌ها را برای نقاطی که پیشگویی‌ها را برای آن‌ها انجام‌می‌دهیم محاسبه کنید.
- پیشگویی‌ها را به مقیاس اصلی تبدیل کنید تا برآوردهای نهایی را تولید کنید.

برنامهٔ *SAS*

کد SAS زیر اسباب تحلیل مقدماتی برای داده‌های موجود در فایل `c73.dat` را فراهم می‌آورد.

```
   * In this data step we read the data and do the ;
   * preliminary transformations ;
   data a;
      infile 'c73.dat';
      input st co expen wealth pop pint
            dens income id grower ;
   * These are the transformations that were chosen ;
   *  after a few trials ;
      lex = log(expen);
      wealth = log(wealth);
      plop = log(pop);
      lens = log(dens);
      income = log(income);
      pint2 = pint**2;
      pint3 = pint**3;
      plop2 = plop**2;
      lens2 = lens**2;
      plop3 = plop**3;
      lens3 = lens**3;
   * Example of transformation when both the positive ;
   * and negative tails are very large ;
    if grower < 0 then lgrowr = - log(-grower);
    if grower > 0 then lgrowr =  log(grower);
    if _N_ ne 887;
    if _N_ ne 475;
* To check that the transformations work ;
proc univariate plot;
   var  st co expen wealth pop pint
        dens income id grower
        lex wealth plop lens income;
proc plot;
   plot  expen*(grower wealth pop dens income
                pint grower);
   plot lex*(lgrowr wealth plop lens income);
* Finally do the regression ;
proc reg;
   model lex = wealth plop plop2 plop3 pint
               pint2 pint3 lens lens2
               lens3 income lgrowr
               / method=stepwise ;
```

یادداشت‌های *S-PLUS*

اگر از S-PLUS استفاده می‌کنید، در اینجا به چند نکتهٔ اساسی دربارهٔ مدلبندی اشاره می‌شود.

lm ــ تابع اصلی برای مدل خطی

lmsreg ــ تابعی برای رگرسیون استوارکمترین توان‌های دوم میانه

gam ــ تابعی برای مدل‌های جمعی تعمیم‌یافته

مثال

```
# Read the data
ny <- read.table("c73.dat")
# Run the linear regression model
lm.ny <- lm(lex~lpop+ldens+lwealth+lincome+
                    pintg+growr,data=ny)
# Run the LMS robust regression model
rr.ny <- lmsreg(lex~lpop+ldens+lwealth+lincome+
                    pintg+growr,data=ny)
# Run the nonlinear regression model
gam.ny <- lm(lex~s(lpop)+s(ldens)+s(lwealth)+lincome+
                    pintg+growr,data=ny)
# Check the residual sum of squares from the summary output
#     to determine if the nonlinear model did much better.
summary(lm.ny)
summary(rr.ny)
summary(gam.ny)
```

شیءهایی نظیر lm. ny را می‌توان برای چاپ آماره‌های مدل به‌کار برد.

summary (lm. ny)	ضرایب، مقادیر آمارهٔ t، R^2 را چاپ می‌کند
anova (lm . ny)	جدول ANOVA را چاپ کنید

مدل‌ها

علامت "~ "مانند علامت مساوی است. این علامت، پاسخ را از پیشگوها تفکیک می‌کند. شما نمی‌توانید از "="استفاده کنید زیرا در زبان S-PLUS معنی دیگری دارد. می‌توانید جمله‌ای مانند

+pop^3 یا

+pop^2 + pop^3

را به فرمول مدل اضافه کنید. به R^2 نگاه کنید تا ببینید که آیا موجب اصلاح قابل ملاحظه‌ای می‌شود یا خیر.

اطلاعات بیشتر در مورد شی‌ها

شیٔ lm.ny را نیز می‌توان برای به‌دست‌آوردن مانده‌ها، مقادیر پیشگویی‌شده، یا نمودارها به‌کار برد:

resid(lm.ny)	مانده را بازمی‌گرداند.
predict(lm.ny)	مقادیر پیشگویی‌شده را بازمی‌گرداند.
predict(lm.ny,new)	مقادیر پیشگویی‌شده در نقاط جدید را بازمی‌گرداند.
plot(lm.ny)	نمودار مانده‌ها را در مقابل مقادیر پیشگویی‌شده رسم می‌کند.

۴.۳.۷ خلاصه

آشکار است که رابطهٔ بین هزینه‌کردها و متغیرهای پیشگو پیچیده است. در بهترین حالت، مدل رگرسیونی که به وجود می‌آوریم تقریب سودمندی برای مقاصد تحقیق مورد نظر در اختیار خواهد گذاشت. شکل ۳.۷ همچنین نشان می‌دهد که ممکن است لازم باشد که توجه خود را به برخی زیرمجموعه‌های داده‌ها محدود کنیم تا بتوانیم مدل رگرسیونی را به‌کار ببریم. از تابعی در S-PLUS موسوم به smooth.spline() برای افزودن اسپلاین هموارسازی به این نمودار پراکنش استفاده کرده‌ایم. مدل‌های پیچیده‌تری شامل اسپلاین‌ها را می‌توان با استفاده از تابع gam() مورد بررسی قرار داد، اما این کار را به عهدهٔ خواننده می‌گذاریم.

پرسش‌ها

پرسش اصلی مورد نظر در اینجا عبارت از محاسبهٔ پیشگویی‌های هزینه‌کردهای سرانه برای یکی از شهرداری‌های مفروض است. با پیروی از ایده‌هایی که در بالا به‌اختصار شرح داده شدند، یکی از راه‌های ممکن تحلیل داده‌ها به قرار زیر است.

۱. تبدیل‌ها را انتخاب کنید.

۲. زیرمجموعه‌ای از داده‌ها، شامل ناحیه‌ای که پیشگویی‌ها را انجام می‌دهیم، انتخاب کنید.

۳. جستجویی برای یافتن دورافتاده‌ها انجام دهید و وجود هم‌خطی را وارسی کنید.

۴. بهترین مدل رگرسیونی را برازش دهید.

۵. مقدار پیشگویی‌شده و خطاهای استاندارد را برای نقاطی که در آن‌ها پیش‌بینی‌ها را انجام می‌دهیم محاسبه کنید.

۶. پیشگویی‌ها را به مقیاس اصلی تبدیل کنید تا برآوردهای نهایی به‌دست آیند.

۴.۷ اندازه‌گیری زمان کیفیت

روش‌ها	تحلیل سری‌های زمانی مدل‌های ARIMA
داده‌ها	۴ سال داده‌های ماهانه ۲ سری زمانی

۱.۴.۷ تحلیل سری‌های زمانی

کیفیت یک محصول و سرویسی که به‌وسیلهٔ یک موسسهٔ کسب و کار ارائه می‌شود بسیار مهم است و بسیاری از شرکت‌ها یک مدیر تضمین کیفیت[1] تمام‌وقت، و شرکت‌های بزرگتر اغلب تیمی از پرسنل «QA» دارند. پدیدآوری کنترل کیفیت[2] به‌اختصار در فصل ۱ مورد بحث قرار گرفت و در حالی که روش‌های QC مبنای آماری طرح‌های تضمین کیفیت را می‌سازند، اصطلاح اخیر براهمیت ترکیب اصول مدیریتی با روش‌های آماری تأکید دارد.

پیش‌زمینه

در این مطالعهٔ موردی، شرکت محصولات متنوعی را تولید می‌کند و روش‌های کنترل کیفیت در سرتاسر فرایند تولید برای ارزیابی کیفیت مواد ورودی به‌کار گرفته می‌شوند. اغلب به این کار کنترل فرایند آماری (SPC[3]) اطلاق می‌شود و یکی از هدف‌های SPC آن است که شرکت را قادر سازد که هدف «بهسازی مستمر» محصولاتش را پیگیری کند.

یک شاخص کیفیت

شرکت برای خود یک شاخص کیفیت (IQ[4]) به وجود آورده به‌طوری که IQ=100 نشان‌دهندهٔ بالاترین مقدار ممکن است. نمرهٔ IQ این توانایی را برای شرکت فراهم می‌کند که امتیازهای کیفیت فردی را با هم ترکیب کند و آنچه را که ماکسیمم کیفیت را تشکیل می‌دهد بدون تغییر مقیاس IQ بازتعریف کند. بنابراین، می‌توان از نمره‌های IQ برای پیش‌بینی اهداف واقع‌بینانهٔ بهبود تولیداتی که به تغییرات فصلی فروش بستگی داشتند استفاده کرد.

1. Quality Assurance (QA) 2. Quality Control (QC) 3. statistical process control
4. index of quality

جدول ۷.۷ داده‌های IQ

ماه	تعداد دسته‌ها	نمرهٔ IQ	ماه	تعداد دسته‌ها	نمرهٔ IQ
Jan94	2339	86.63	Jan96	2971	89.25
Feb94	2275	84.60	Feb96	3083	90.54
Mar94	2881	87.04	Mar96	3504	89.89
Apr94	2780	87.19	Apr96	3580	90.28
May94	3227	87.91	May96	3855	89.46
Jun94	3291	87.99	Jun96	3894	89.42
Jul94	2944	88.09	Jul96	3772	89.28
Aug94	3163	88.25	Aug96	3705	89.17
Sep94	2770	87.62	Sep96	3364	90.42
Oct94	2827	87.43	Oct96	3341	90.46
Nov94	2392	86.74	Nov96	2680	88.63
Dec94	1973	84.86	Dec96	2418	89.74
Jan95	3006	87.44	Jan97	2963	90.48
Feb95	2924	87.77	Feb97	2890	89.76
Mar95	3592	88.09	Mar97	3455	90.20
Apr95	3460	88.53	Apr97	3747	90.68
May95	3807	88.11	May97	3685	90.19
Jun95	3753	88.59	Jun97	3672	89.78
Jul95	3648	88.67	Jul97	3865	89.72
Aug95	3698	88.87	Aug97	3729	90.78
Sep95	3166	89.92	Sep97	3205	90.32
Oct95	3159	88.93	Oct97	3158	90.64
Nov95	2545	87.17	Nov97	2552	90.97
Dec95	2208	89.07	Dec97	2135	90.23

داده‌ها

هدف این تحلیل به‌دست‌آوردن مدل سری زمانی خوبی برای مدلبندی امتیاز IQ است. داده‌های ماهانه در مورد امتیاز IQای چهار سال اخیر برای محصولی خاص در جدول ۷.۷ داده شده‌اند. شمارش BATCH با مقدار این محصول خاص تولیدشده به‌وسیلهٔ شرکت متناظر است. این مجموعه‌داده‌ها در فایل `c74.dat` نیز در دسترس است.

روش‌شناسی

فرض کنید که y_t معرف مقدار پاسخ مورد نظر در زمان t باشد. در اینجا واحد واقعی زمان، ماه است و ما علاقه‌مند به پیش‌بینی $y_t = \text{IQ}$ هستیم. در یک وضعیت رگرسیونی، می‌توانیم شاخص

زمان را با X و مقادیر پاسخ را با Y نشان می‌دهیم، و سپس از مدل رگرسیونی برازش‌داده‌شده برای به‌دست‌آوردن $\hat{y}_{t+1}$ استفاده کنیم. با این حال، در مدل رگرسیونی فرض می‌شود که مؤلفه‌های خطای مدل مستقل‌اند و این فرض معقولی برای فرایندی که در طول زمان در حال تطور است نیست. بنابراین، مدل‌های ردهٔ ARIMA را معرفی می‌کنیم.

۲.۴.۷ مدل‌های ARIMA

باکس و جنکینز (۱۹۷۰) ردهٔ مدل‌های اتورگرسیو جمع‌بستهٔ میانگین متحرک (ARIMA) را برای برازش به سری‌های زمانی عمومیت دادند. از رویکرد آن‌ها عموماً با عنوان روش باکس–جنکینز یاد می‌شود و مشتمل بر سه مرحلهٔ زیر است:

شناسایی از تفاضل‌گیری برای ماناکردن y_t استفاده می‌شود.

برآوردکردن یک مدل ARIMA(p,q) به سری تفاضل‌گیری‌شده برازش داده می‌شود.

پیش‌بینی از مدل ARIMA برای پیش‌بینی مقدارهای آتی y_t استفاده می‌شود.

شناسایی

قبل از آنکه بتوان همبستگی پیاپی را مدلبندی کرد، باید مشخصه‌های تعیینی سری زمانی از قبیل فصلی‌بودن و روندها مورد رسیدگی قرار گیرند. در عمل، می‌توان این مشخصه‌ها را اغلب با به‌کاربردن عملگرهای تفاضلی

$$\nabla y_t = y_t - y_{t-1} \qquad \nabla_s y_t = y_t - y_{t-s}$$

روی y_t حذف کرد تا آنکه سری حاصل، که آن را مثلاً w_t می‌نامیم، مانا به‌نظر برسد؛ به عبارت دیگر، میانگین و واریانس w_t تقریباً ثابت شوند. باید در بدو امر یک تفاضل‌گیری ساده بر سری اعمال می‌شود و نباید بیش از دو تفاضل‌گیری ساده ($\nabla^2 y_t = \nabla y_t - \nabla y_{t-1}$) را به یک سری زمانی اعمال کرد.[۱] در صورت نیاز، می‌توان یک تفاضل‌گیری فصلی را نیز اعمال کرد. به‌عنوان مثال، می‌توان از $s = 12$ برای حذف فصلی‌بودن سالانه در سری‌یی که ماهانه مشاهده شده استفاده کرد.

۱. تفاضل‌گیری بیش از حد ممکن است آنچه را که به ریشه‌های واحد موسوم است وارد بحث کند که موجب می‌شود ARIMA(p,q)ی برازش‌داده‌شده ناپایدار شود. این مشکل شبیه مسئلهٔ هم‌خطی‌بودن چندگانه در رگرسیون است.

برآوردکردن

با در دست‌داشتن w_t، اکنون می‌توان همبستگی پیاپی را با برازش‌دادن یک مدل ARIMA(p,q) به شکل

$$w_t + \sum_{j=1}^{p} \alpha_j w_{t-j} = \epsilon_t + \sum_{j=1}^{q} \beta_j \epsilon_{t-j} \qquad (۳.۷)$$

مدلبندی کرد که در آن ϵ_t معرف مؤلفه‌های تصادفی مجهول یا نوآوری‌هاست که معمولاً فرض می‌شود که مستقل و هم‌توزیع‌اند.

مرتبه‌های p و qی مدل ARMA را باید در ابتدا تعیین کرد و این کار را می‌توان با امتحان‌کردن همبستگی‌نگارهای خودهمبستگی (ACF) و خودهمبستگی جزئی (PACF) محقق کرد. این کار با استفاده از ملاک‌های اطلاع از قبیل AIC یا BIC تلفیق می‌شود که ما را به انتخاب مسک‌ترین مدل‌ها قادر می‌سازد. پس از تعیین مرتبه‌های p و q، می‌توان پارامترهای ARMA را برآورد کرد.

ACF عبارت از نمودار $\rho_k = Corr(y_{t-k}, y_k)$ در برابر k است. برای یک فرایند AR(p)، همبستگی بین y_t و y_{t-k} با افزایش زمان «تأخیر» k به‌طور نمایی کاهش می‌یابد. در مقابل، در یک فرایند MA(q)ی خالص، به‌ازای هر k، $k > p$، $\rho_k = 0$.

PACF خودهمبستگی جزئی بین y_t و y_{t-k} همبستگی **پس** از آن است که اثر y_{t-1}، y_{t-2}، ...، y_{t-k+1} با «رگرسیون کردن» خارج شده است. این همبستگی را با ϕ_k نشان می‌دهند و برای یک فرایند AR(p)، «ربط» بین y_t و y_{t-k} وقتی $k > p$ دیگر وجود ندارد ($\phi_k = 0$). در مقابل، یک MA(q) را می‌توان با یک فرایند AR(∞) نمایش داد به‌طوری که ϕ_k به‌طور نمایی کاهش می‌یابد. PACF نمودار ϕ_k در مقابل k است.

AIC/BIC ترکیب همبستگی نگارهای ACF و PACF می‌تواند در تعیین مرتبهٔ فرایندهای خالص AR یا MA بسیار مؤثر باشد. می‌توان آن‌ها را به‌عنوان یک ابزار تشخیصی برای مانده‌های یک مدل ARMAی برازش‌یافته نیز به‌کار گرفت. AIC را آکائیکه (۱۹۷۴) معرفی کرد که ملاک سودمندی برای مقایسهٔ مدل‌های مختلف سری زمانی در اختیار می‌گذارد. BIC معیار اطلاع دیگری است (آن را SBC نیز می‌نامند)، و هر دو معیار معمولاً به‌وسیلهٔ نرم‌افزارهای آماری با مدول‌های سری زمانی در دسترس ما گذاشته می‌شوند. این معیارها به‌صورت

$$\mathrm{AIC}(k) = \log \hat{\sigma}_k^2 + \frac{2k}{n} \qquad \mathrm{BIC}(k) = \log \hat{\sigma}_k^2 + \frac{k \log n}{n}$$

تعریف می‌شوند که در آن $\hat{\sigma}_k^2$ معرف واریانس ماندهٔ مربوط به برازش مدلی از مرتبهٔ k

است. جملهٔ دوم تاوان افزایش k است، لذا مقدار مینیمم‌کنندهٔ این ملاک‌ها برآوردی از مدسک‌ترین مدل در اختیار می‌گذارد.

پیش‌بینی

برای پیش‌بینی‌کردن، مدل «جمع‌بسته» را می‌توان از مدل برازش‌یافتهٔ ARMA و تفاضل‌گیری‌های به‌کاررفته در مرحلهٔ شناسایی بازسازی کرد تا یک مدل خطی برحسب سری زمانی اصلی y_t فراهم شود. به‌عنوان مثال، مدل $\text{ARIMA}(۱,۱,۱)$ به شکل زیر است:

$$y_t - y_{t-۱} + \alpha_۱(y_t - y_{t-۱}) = \epsilon_t + \beta_۱\epsilon_{t-۱}.$$

مدل‌های $ARIMA$ی فصلی

ردهٔ مدل‌های ARIMA می‌تواند با برازش‌دادن یک مدل $\text{ARMA}(P,Q)$ بر مشاهدات، با تفکیک دورهٔ فصلی، اثرهای فصلی را هم در خود بگنجاند. به عبارت دیگر، به‌جای j در (۳.۷) می‌توان s_j را قرار داد و w_tی سری حاصل را پس از تفاضل‌گیری فصلی (در صورت به‌کارگیری آن) نشان می‌دهد. مدل ARIMAی فصلی ضربی اغلب با $(P,D,Q)_s \times (p,d,q)$ نشان داده می‌شود که در آن

D = تعداد تفاضل‌گیری‌های فصلی

P و Q = مرتبه‌های مدل ARMAی فصلی

s = تأخیر فصلی

d = تعداد تفاضل‌گیری‌های ساده

p و q = مرتبه‌های مدل‌های ARMAی فصلی‌زدوده.

مانایی

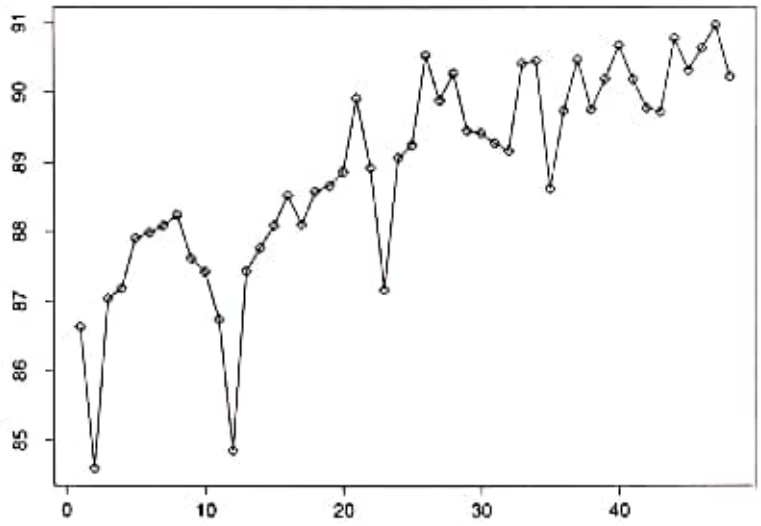
مدل‌های ARIMA اغلب می‌توانند نقطهٔ شروع سودمندی در اختیار بگذارند و برای مقاصد پروژهٔ در دست به قدرکافی دقیق باشند. با این حال، این مدل‌ها نیازمند آن هستند که بتوان سری اصلی را به یک سری مانا (w_tی بالا) فروکاست که در آن میانگین و واریانس ثابت‌اند. در حالت کلی، می‌توان با تفاضل‌گیری به مانایی در میانگین نائل شد، اما برای حصول واریانس ثابت، نیاز به یک تبدیل غیرخطی داریم.

۳.۴.۷ تحلیل مقدماتی

شکل ۴.۷ فصلی‌بودن و نیز روندی رو به بالا را نشان می‌دهد ــ که خوب است! این جنبه‌های ساختاری را باید حذف کرد و کد SAS زیر چند دستور عمومی برای مشخصات مدل ARIMA در PROC ARIMA ارائه می‌کند. ما همچنین استفاده از اطلاع‌رسان زمان "monyy5." را تشریح می‌کنیم که SAS را قادر می‌سازد که درایه‌های کاراکتریِ mon (مانند "jan94") را به‌درستی تعبیر و تفسیر کند.

```
filename in 'c74.dat' ;
data a ;                          * "monyy5." is similar to "$".   ;
  infile in ;                     * It tells SAS how to read the   ;
  input mon monyy5. count iq ;    * date variable "mon" (mon = xxx ;
  y = iq ;                        * and yy = 2 digit year)         ;
  format mon monyy5. ;   * use for printing purposes. otherwise   ;
                         * "jan94" is printed as a Julian date!   ;
proc arima ;
  * ------------------------------------------------------------ ;
  * Identifying an arima model                                   ;
  i var=y ;          * acf/pacf for time series y(t)             ;
                     * to detrend use (one of) the following     ;
                     *      instead                              ;
  i var=y(1) ;       * simple difference w(t) = y(t) - y(t-1)    ;
  i var=y(12) ;      * seasonal difference w(t) = y(t) - y(t-12) ;
  i var=y(1,12) ;    * simple difference giving w(t), then       ;
                     *   seasonal difference giving w2(t):       ;
                     *       w2(t) = w(t) - w(t-12)              ;
  * ------------------------------------------------------------ ;
  * Fitting an arima model                                       ;
  e p=(1) ;          * fit ar(1) to w(t)                         ;
  e p=(1) plot ;     * fit ar(1) and plot the acf/pacf of        ;
                     *      residuals                            ;
  e p=(1) q=(1) ;    * fit arma(1,1)                             ;
  e p=(1)(12) ;      * fit multiplicative ARIMA:                 ;
                     *      ar(1) x sar(1) (season lag=12)       ;
  * ------------------------------------------------------------ ;
  * Forecasting from the fitted arima model                      ;
  f out=r lead=1 back=0 id=mon ;* only forecast 1 step ahead.    ;
                               *    output to data=r             ;
  f out=r lead=12 id=mon ;     * forecast 12 steps ahead.       ;
                               *    include forecasts within    ;
                               *    series. output to data=r    ;
  * ------------------------------------------------------------ ;
```

مراحل شناسایی، برآوردکردن (برازش‌دادن)، و پیش‌بینی به‌طور طبیعی پی‌درپی انجام می‌شوند. بدین معنی که لازم‌است شیوهٔ proc arima را چندین بار اجرا کنیم که نخستین کار حصول



شکل ۴.۷ نمودار نمرهٔ IQای ماهانه.

مانایی از طریق تفاضل‌گیری یا روش‌های دیگر است. این کار را به‌عنوان تمرین به عهدهٔ خواننده گذاشته‌ایم و در عوض به‌اختصار رویکرد S-PLUS را ذکر می‌کنیم.

کد *S-PLUS*

در S-PLUS، پارامترهای یک مدل ARIMA به‌صورت یک list مشخص می‌شوند. در کد زیر، list اول، پارامترهای «درون‌فصلی» مدل ARIMA را به‌ترتیبِ `list=c(p,d,q)` با $p = ۱$، $d = ۱$، و $q = ۰$ مشخص می‌سازد. به‌عبارت دیگر، یک مدل AR(1) به ∇y_t برازش داده خواهد شد. list دوم، پارامترهای مدل ARIMAی «فصلی» را به‌ترتیبِ `list=c(P,D,Q)` با $P = ۱$ و دورهٔ فصلی مشخص‌شده با گزینهٔ«period=12» مشخص می‌کند. در این حالت، مدل ضربی

$$\nabla y_t - \alpha \nabla y_t - \Phi \nabla y_{t-۱۲} + \alpha\Phi \nabla y_{t-۱۳} = \epsilon_t$$

به کمک تابع `arima.mle()` برازش داده خواهد شد. مدل برازش‌داده‌شده را می‌توان با امتحان‌کردن ماندهٔ نمودارهای ACF و PACF که با تابع تشخیصی `arima.diag()` به‌دست می‌آید ارزیابی کرد.

```
z <- read.table("c74.dat")
y <- z$iq
m1 <- list(list(order=c(1,1,0),
           list(order=c(1,0,0),period=12)))
fit <- arima.mle(y, model=m1)

arima.diag(fit)
```

۴.۴.۷ خلاصه

روش‌های کنترل کیفیت و اصول مدیریتی، خود را به‌عنوان ابزارهای بسیار مؤثری برای شرکت‌های سازنده به ثبوت رسانیده‌اند. به‌عنوان مثال، نفوذ دمینگ[1] این فرصت را برای صنعت تولید خودروی ژاپن فراهم کرد که بخش عمده‌ای از بازار ایالات متحده را تصاحب کند. در این مثال، دیده‌ایم که بهسازی مستمر باید واقعیت فرایند تولید را نیز بازتاب دهد؛ بدان معنی که اثر تغییرات فصلی، در صورت وجود، باید در طرح‌های QA آمیخته شوند. از شکل ۴.۷ چنین به نظر می‌رسد که تغییرات فصلی در کیفیت طی سال آخر دیگر موجود نیست، هرچند که میزان تولید هنوز هم یک روند فصلی قوی دارد. شاید صرفاً در اختیارداشتن ابزارهایی مناسب برای نگاه به یک مسئله بتواند سبب حل و فصل آن شود!

پرسش‌ها

۱. از نمودار سری زمانی اولیه چنین برمی‌آید که به یک تبدیل برای پایدارسازی واریانس نیاز داریم. شما چه نوع تبدیلی را به‌کار می‌برید؟

۲. یک مدل ARIMA به داده‌های اولیه و تبدیل‌یافته برازش دهید. به چه نتیجه‌گیری‌هایی می‌رسید؟

1. Deming

۸

مطالعه‌های موردی از گروه III

چهار مطالعهٔ موردی ارائه‌شده در این فصل، پژوهش‌محورند و تحلیل آن‌ها نیازمند روش‌های چندمتغیره یا روش‌های آماری تخصصی است. برای به‌دست‌آوردن نتایج مناسب، ممکن است به چندین مرحله از تحلیل نیز نیاز پیدا کنیم. ممکن است «پاسخی» برای مسئلهٔ آماری لزوماً وجود نداشته باشد. مطالعه‌های موردی مورد نظر برحسب عنوان بخش در زیر فهرست شده‌اند.

۱.۸ داستان دو دزدک

مدل‌های آمیخته

۲.۸ ردیابی مواد منفجرهٔ پلاستیکی

تحلیل ممیزی

۳.۸ یک مطالعهٔ تحقیق بازار

تحلیل عاملی

۴.۸ فروش تجهیزات ارتوپدی

داده‌کاوی، تحلیل چندمرحله‌ای

۱.۸ داستان دو دزدک

روش‌ها	تحلیل واریانس اثرهای تصادفی مدل‌های آمیخته
داده‌ها	دو آزمایش پاسخ: سنجهٔ اجزای فعال ۲ ابزار نمونه‌گیری تحلیل کیفیت پس‌آمیزی

۱.۱.۸ تحلیل واریانس با اثرهای آمیخته

داروهای با نسخه و بدون نسخه محتوی ترکیبی از اجزای فعال و غیر فعال‌اند و دز دارو برحسب مقدار اجزای فعال موجود در هر قرص تعیین می‌شود. اطمینان از اینکه قرص‌ها حاوی دز درستی هستند در صنعت تولید دارو مسئله‌ای مهم است و، در این مطالعهٔ موردی، آزمایشی را که توسط یک شرکت داروسازی برای تحقیق در تغییرپذیری و اریبی نمونه‌گیری مربوط به تولید نوع خاصی قرص انجام شده است در نظر می‌گیریم.

شرح اجمالی مسئله

تولید قرص قرص‌ها با هم‌زدن اجزای فعال و غیر فعال در یک «V-مخلوط‌کن» تولید شدند و علت این نامگذاری آن است که این مخلوط‌کن به شکل یک **V**ی بزرگ است (نگاه کنید به شکل ۱.۸). عمل هم‌زدن با به‌چرخش‌درآوردن V-مخلوط‌کن در جهت عمودی انجام شد.

پس از اینکه مخلوط به‌طور کامل هم‌زده شد، پودر موجود از ته V-مخلوط‌کن خارج شد و با فشار به شکل قرص درآمد.

محتوای یکنواخت مهمترین شرط این فرایند تولید آن بود که قرص‌ها محتوای یکنواختی داشته باشند. این بدان معنی است که مقدار صحیحی از اجزای فعال در هر قرص حضور داشته باشد.

یکنواختی محتوای مخلوط داخل V-مخلوط‌کن را باید ارزیابی کرد.

نمونه‌گیری دزدکی از یک ابزار «دزدک» برای به‌دست‌آوردن نمونه‌هایی از جاهای مختلفی از داخل V-مخلوط‌کن استفاده شد. این ابزار اساساً عبارت بود از چوب بلندی که یک چَمچهٔ بسته در یک طرف آن قرار داشت و به‌وسیلهٔ یک ابزار مکانیکی به داخل مخلوط پودر فرو برده می‌شد. برای یک مکان معین در عمقی مناسب، درِ چمچه باز و نمونه‌ای گردآوری می‌شد.

برای فروبردن دزدک در مخلوطِ پودر نیروی زیادی لازم بود و مقایسهٔ دو نوع دزدک مورد توجه قرار داشت.

- دزدک دز واحد سه نمونهٔ دز واحد انفرادی از هر مکان گردآوری می‌کند.
- دزدک دز میانی نمونه‌ای بزرگ گردآوری می‌کند که از خود آن نمونه‌گیری می‌شود تا سه نمونهٔ دز واحد به‌دست آید.

شیوهٔ آزمایش

هدف این آزمایش این بود که تفاوت‌های اریبی و تغییرپذیری بین دو دزدک مطالعه و نتایج نمونه‌گیری‌های دزدکی با موارد مربوط به قرص‌ها مقایسه شوند. آزمایش به‌صورت زیر به اجرا درآمد.

۱. مخلوط را در V-مخلوط‌کن به مدت ۲۰ دقیقه هم بزنید.

۲. دزدک‌ها را به هم ببندید و از آن‌ها برای به‌دست‌آوردن نمونه‌هایی در شش مکان در داخل V-مخلوط‌کن استفاده کنید. طرحی از V-مخلوط‌کن و مکان‌های نمونه‌گیری در شکل ۱.۸ نشان داده شده است.

۳. پودر را از V-مخلوط‌کن تخلیه کنید و آن را با فشار به قرص تبدیل کنید. قرص‌ها را در ۳۰ ظرف استوانه‌ای‌شکل[1] قرار دهید.

۴. ۱۰ ظرف استوانه‌ای‌شکل را انتخاب کنید و سه قرص از هر یک از این ظرف‌ها به‌عنوان نمونه بردارید.

۵. همهٔ نمونه‌ها را بسنجید تا مقدار اجزای فعال در هر نمونه معین شود. مقدار سنجهٔ مشخص‌شده عبارت از ۳۵ میلی‌گرم در ۱۰۰ میلی‌گرم است.

در این مثال، به‌کاربردن فن تحلیل واریانس برای دو مجموعه از نتایج سنجه (دزدک و قرص) از روی این آزمایش بسیار پیچیده‌تر است زیرا این کار متضمن اثرهای تصادفی و ثابت است. وقتی هر دو نوع عامل حضور داشته باشند، این را یک مدل آمیختهٔ تحلیل واریانس می‌نامند.

1. Drum

داده‌ها

مجموعه‌داده‌های c81.thief.dat و c81.tablet.dat شامل نتایجی از این آزمایش است که در قالبی کمی فشرده‌تر به‌ترتیب در جدول‌های ۱.۸ و ۲.۸ ارائه شده‌اند. متغیرها و عامل‌های مورد نظر برای مقاصد تحلیل در جدول زیر به‌طور خلاصه بیان شده‌اند.

متغیر	سطوح	تعریف
Y	(میلی‌گرم)	مقدار سنجه در هر ۱۰۰ میلی‌گرم
METHOD	INTM	دزدک در متوسط
	UNIT	دزدک در واحد
LOC	۱، ۲، ... ۶	مکان نمونه‌گیری
REP	۱، ۲، ۳	تکرار در هر مکان
DRUM	$\mathrm{mod}(n, ۳)$	ظرف استوانه‌ای‌شکل به‌تصادف‌انتخاب‌شده
TABLET	۱، ۲، ۳	نمونهٔ‌قرص (در هر ظرف استوانه‌ای‌شکل)

مکان‌های نشان داده‌شده در شکل ۱.۸ نشان دهندهٔ مکان‌های «مطلوب» نمونه‌گیری برای دزدک‌ها هستند. در آزمایش واقعی، این مکان‌های «ثابت» در معرض مقداری تغییرپذیری بودند. نمونه‌های گردآوری‌شده از طریق دزدک‌ها را می‌توان، در داخل هر مکان، تصادفی تلقی کرد.

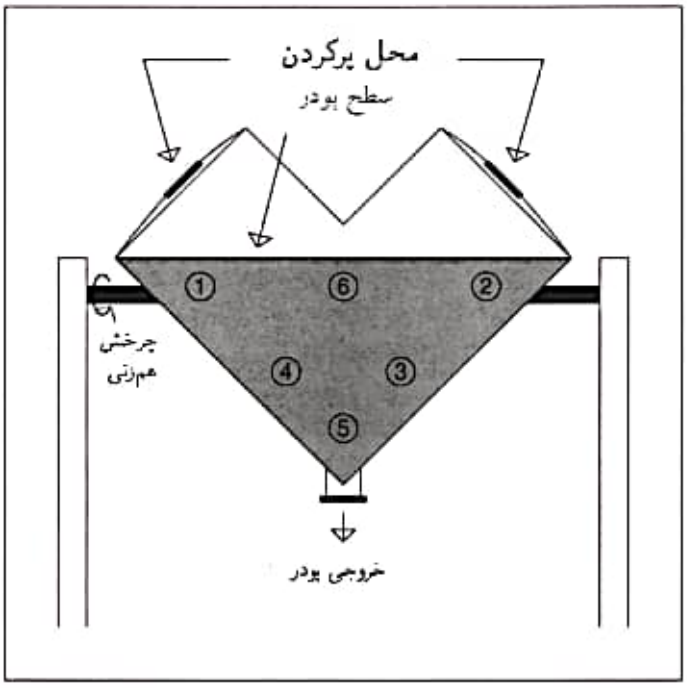
در آزمایش قرص‌ها، ترتیبی که طی آن‌ها ظرف‌های استوانه‌ای پر می‌شدند ثبت شد و این اطلاع در شیوهٔ انتخاب تصادفی دخالت داده شد. به‌خصوص، یک ظرف از هر یک از سه‌تایی‌های $\{۱, ۲, ۳\}$، $\{۴, ۵, ۶\}$، ... ، $\{۲۸, ۲۹, ۳۰\}$، به‌تصادف انتخاب شد. بنابراین، می‌توان از عامل DRUM برای آزمون‌کردن اثر «زمان» در داده‌های قرص‌ها استفاده کرد.

روش‌شناسی

برای کمک به ایجاد انگیزه نسبت به این بحث، مدل ANOVAی یک‌طرفهٔ متعادل را که در فصل ۳ معرفی کردیم (روش‌های استاندارد) به خاطر بیاورید:

$$y_{ij} = \mu + \tau_i + \epsilon_{ij} \qquad \begin{cases} i = ۱, ۲, ..., a \\ j = ۱, ۲, ..., n \end{cases} \qquad (۱.۸)$$

که در آن τ_i نشان‌دهندهٔ سطح یک عامل، مثلاً TRT، است و فرض می‌شود که ϵ_{ij}ها نوآوری‌های مستقل $N(۰, \sigma^۲)$ هستند. تفاوت عمدهٔ بین یک مدل اثرهای ثابت و یک مدل اثرهای تصادفی در تعبیر استنباط مرتبط با این مدل‌ها نهفته است.



شکل ۱.۸ طرح V-مخلوط‌کن و مکان‌های نمونه‌گیری.

اثر ثابت استنباط مشروط بر سطح‌های «ثابت» عامل TRT است. در این مثال، عامل METHOD آشکارا ثابت است زیرا هر استنباط وابسته به اثر METHOD به‌طور مشخص در مورد دو نوع دزدک به‌کار گرفته‌شده اعمال می‌شود.

اثر تصادفی استنباط دربارهٔ «جامعه»ای از سطوح TRT انجام می‌شود. عامل TRT در صورتی که τ_i به‌طور تصادفی از سطوح TRT انتخاب شده باشد، تصادفی در نظر گرفته می‌شود. بنابراین، τ_i برای مقاصد آزمون کردن، یک متغیر تصادفی $N(0, \sigma_\tau^2)$ (مستقل از ϵ_{ij}) فرض می‌شود. مدل (۱.۸) مدل مؤلفه‌های واریانس نیز نامیده می‌شود.

از آنجا که τ_iها به‌تصادف انتخاب می‌شوند، آزمون‌کردن اثرهای انفرادی بی‌معنی است و فرض‌های مورد نظر عبارت‌اند از $H_0 : \sigma_\tau^2 = 0$ در مقابل $H_1 : \sigma_\tau^2 > 0$. تحت H_0،

جدول ۱.۸ داده‌های دزدک

مکان	تکرار	روش سنجه	
		میانی	دزدک واحد
1	1	34.38	33.94
	2	34.87	34.72
	3	35.71	34.10
2	1	35.31	39.11
	2	37.59	37.51
	3	38.02	37.79
3	1	36.71	37.46
	2	36.56	34.12
	3	35.92	35.94
4	1	37.80	38.05
	2	37.41	34.82
	3	38.00	35.42
5	1	36.28	36.52
	2	36.63	38.60
	3	36.62	38.16
6	1	38.89	39.16
	2	39.80	32.77
	3	37.84	36.95

آمارهٔ آزمون

$$F_* = \frac{\mathrm{SS(TRT)}/(a-1)}{\mathrm{SSE}/(N-a)} = \frac{\mathrm{MS(TRT)}}{\mathrm{MSE}} \sim \mathcal{F}_{a-1,N-a}$$

دقیقاً به همان سبک حالت اثرهای ثابت ساخته می‌شود. با این حال، میانگین‌های توان‌های دوم مورد انتظار وابسته به مدل اثرهای تصادفی متفاوتند و برای ساختن برآوردگرهای مؤلفه‌های واریانس ضرورت دارند. برای مدل (۱.۸) می‌توان نشان داد که:

$$E[\mathrm{MS(TRT)}] = \sigma^2 + n\sigma_\tau^2 \qquad E[\mathrm{MSE}] = \sigma^2.$$

با برابر قراردادن میانگین‌های توان‌های دوم مشاهده‌شده و مورد انتظار، برآوردهای مؤلفه‌های واریانس به‌صورت زیر حاصل می‌شود:

$$\hat{\sigma}_\tau^2 = (\mathrm{MS(TRT)} - \mathrm{MSE})/n \qquad \hat{\sigma}^2 = \mathrm{MSE}.$$

جدول ۲.۸ داده‌های قرص

طرف استوانه‌ای شکل	قرص	سنجه	طرف استوانه‌ای شکل	قرص	سنجه
1	1	35.77	17	1	35.43
	2	39.44		2	33.80
	3	36.43		3	35.15
5	1	35.71	19	1	34.56
	2	37.08		2	35.33
	3	36.54		3	37.69
7	1	35.08	22	1	35.82
	2	34.25		2	35.67
	3	33.09		3	35.06
11	1	35.21	25	1	35.75
	2	34.36		2	37.32
	3	35.94		3	35.06
14	1	35.17	28	1	38.58
	2	36.54		2	36.63
	3	36.45		3	35.60

۲.۱.۸ تحلیل مدل آمیخته

مدل ANOVAی دوطرفهٔ

$$y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + \epsilon_{ijk} \qquad \begin{cases} i = ۱, ۲, ..., a \\ j = ۱, ۲, ..., b \\ k = ۱, ۲, ..., n \end{cases} \qquad (۲.۸)$$

را تحلیل واریانس مدل آمیخته می‌نامند هرگاه یکی از عامل‌ها ثابت و دیگری تصادفی باشد. در اینجا وضعیتی را در نظر می‌گیریم که در آن عامل A ثابت است، B تصادفی است و $n > ۱$، به‌طوری که جملهٔ اثر متقابل را بتوان برآورد کرد. حالتی که در آن A و B هر دو ثابت یا هر دو تصادفی‌اند، در فصل ۳ مورد بحث قرار گرفت.

تحلیل آماری

جملهٔ اثر متقابل AB را همواره اثری تصادفی فرض می‌کنیم، اما فرض‌های متفاوتی را می‌توان بر مؤلفه‌های تصادفی اعمال کرد. این امر به‌صورت‌های متفاوتی از مدل آمیخته منجر می‌شود:

A اثر ثابت: $\sum_{i=۱}^{a} \alpha_i = ۰$.

B اثر تصادفی: $\beta_j \sim N(0, \sigma_\beta^2)$.

AB اثر تصادفی:

مدل استاندارد: $\gamma_{ij} \sim N(0, (1-(1/a))\sigma_\gamma^2)$ و $\sum_{i=1}^{a} \gamma_{ij} = 0$.

مدل جایگزین: γ_{ij}ها ناهمبسته و دارای واریانس مشترک σ_γ^2 هستند.

تفاوت عمده در این است که در مدل استاندارد (که رایج‌تر است) اثرهای متقابل، در سطوح متفاوت عامل ثابت، ناهمبسته نیستند: به‌ازای $i \neq i'$، $Cov(\gamma_{ij}, \gamma_{i'j}) = -(1/a)\sigma_\gamma^2$. تحت مفروضات بالا، میانگین توان‌های دوم مورد انتظار، آماره‌های آزمون مناسب و برآوردهای مؤلفه‌های واریانس را به شرح زیر در اختیار می‌گذارند.

$H_0: \alpha_i = 0$ در برابر H_1: به‌ازای iای، $\alpha_i \neq 0$:

$$F_0 = \mathrm{MS(A)}/\mathrm{MS(AB)} \sim \mathcal{F}_{(a-1),(a-1)(b-1)}$$
$$E[\mathrm{MS(A)}] = \sigma^2 + n\sigma_\gamma^2 + \frac{bn}{n-1}\sum \alpha_i^2$$
$$\hat{\sigma}^2 = \mathrm{MSE}$$

$H_0: \sigma_\beta^2 = 0$ در برابر $H_1: \sigma_\beta^2 > 0$:

مدل استاندارد:

$$F_0 = \mathrm{MS(B)}/\mathrm{MSE} \sim \mathcal{F}_{(b-1),ab(n-1)}$$
$$E[\mathrm{MS(B)}] = \sigma^2 + an\sigma_\beta^2$$
$$\hat{\sigma}_\beta^2 = (\mathrm{MS(B)} - \mathrm{MSE})/an$$

مدل جایگزین:

$$F_0 = \mathrm{MS(B)}/\mathrm{MS(AB)} \sim \mathcal{F}_{(b-1),(a-1)(b-1)}$$
$$E[\mathrm{MS(B)}] = \sigma^2 + n\sigma_\gamma^2 + an\sigma_\beta^2$$
$$\hat{\sigma}_\beta^2 = (\mathrm{MS(B)} - \mathrm{MS(AB)})/an$$

$H_0: \sigma_\gamma^2 = 0$ در برابر $H_1: \sigma_\gamma^2 > 0$:

$$F_0 = \mathrm{MS(AB)}/\mathrm{MSE} \sim \mathcal{F}_{(b-1),ab(n-1)}$$
$$E[\mathrm{MS(AB)}] = \sigma^2 + n\sigma_\gamma^2$$
$$\hat{\sigma}_\gamma^2 = (\mathrm{MS(AB)} - \mathrm{MSE})/n$$

۳.۱.۸ تحلیل مقدماتی

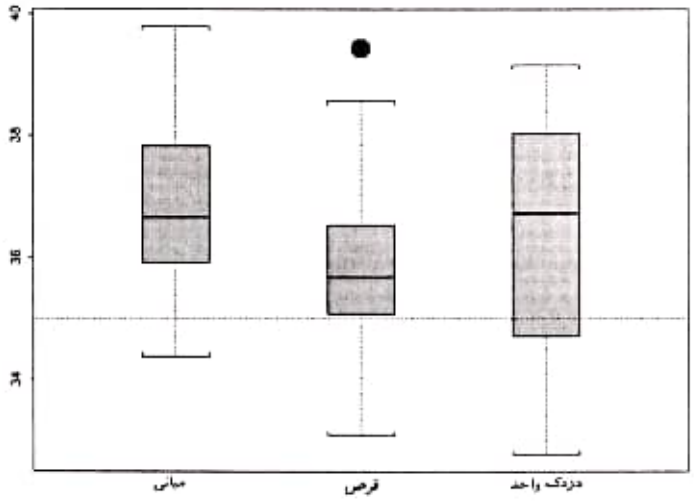
هدف این آزمایش عبارت از این بود که تفاوت‌ها در اریبی و تغییرپذیری بین دو دزدک را مطالعه و نتایج نمونه‌گیری دزدکی را با نتایج قرص‌ها مقایسه کنیم. سه مؤلفهٔ اصلی تحلیل مقدماتی ما به شرح زیرند:

- غربالگری مقدماتی داده‌ها
- مقایسهٔ دو دزدک
- نمونه‌های دزدک در برابر قرص

غربالگری داده‌ها

غربالگری مقدماتی داده‌ها دربرگیرندهٔ وارسی وجود دورافتاده‌ها و ارزیابی خاصیت‌های توزیعی مقادیر سنجه است. این مطلب از آن رو اهمیت دارد که قصد ما به‌کاربردن فن تحلیل واریانس است که نسبت به حضور دورافتاده‌ها و انحراف از نرمال‌بودن حساس است. البته، تنها با ۱۸ نمونه برای هر دزدک، لازم خواهد شد که در تفسیر جنبه‌های توزیعی، مقداری احتیاط به خرج دهیم.

از نمودار جعبه‌ای نمایش داده‌شده در شکل ۲.۸ به‌آسانی دیده می‌شود که همهٔ روش‌ها،



شکل ۲.۸ نمودارهای جعبه‌ای موازی مقادیرسنجه.

نسبت به مقدار سنجهٔ ۳۵ میلی‌گرمی مشخص‌شده، اریبی مثبت دارند. با این حال، تفاوت‌هایی در تغییرپذیری و توزیع مقادیر این سنجه بین این روش‌ها موجود است. به‌ویژه، مقدار سنجهٔ «متوسط» نمونهٔ قرص به میزان قابل توجهی کمتر است. جزء فعال تشکیل‌دهندهٔ آن کجا رفته است؟ آماره‌های خلاصه برای سه روش ارائه‌شده در جدول زیر نمایش داده شده‌اند.

	دزدک		قرصی
روش	UNIT	INTM	DRUM
N	18	18	30
Mean	36.40	36.91	35.79
Std Dev	1.98	1.40	1.36
Max	39.16	39.80	39.44
Q_3	38.08	37.88	36.54
Median	36.74	36.67	35.69
Q_1	34.57	35.87	35.08
Min	32.77	34.38	33.09

گرچه آزمون شاپیرو-ویلک و نمودارهای Q-Q با کران‌های اطمینان لیلیفور[1] (کانوور، ۱۹۸۰) بر هیچ‌گونه انحراف معنی‌دار از نرمال‌بودن برای هیچ‌کدام از روش‌ها دلالت نمی‌کنند، باید این نتیجه‌ها را به دلیل نمونه‌های کوچک دخیل در آن‌ها ناتمام تلقی کنیم. به همین نحو، بافت‌نگاری از داده‌های قرص نشان می‌دهد که «دورافتاده»های نشان داده‌شده در شکل ۲.۸ با روند دم‌راستِ کلفتی سازگاری دارند. این مشاهده نیز از $Q_3 + 2 * \text{IQR}$، که معمولاً به‌عنوان معیار (بالایی) محافظه‌کارانه‌تری برای تعیین دورافتاده‌های بالقوه به‌کار می‌رود، کوچکتر است.

مقایسهٔ دزدک‌ها

از آنجا که نتایج تحلیل مقدماتی نشان از هیچ تخلف آشکاری از فرض‌های زمینه‌ای شیوهٔ تحلیل واریانس نداشت، با استفاده از مدل تحلیل واریانس آمیخته برای مقایسهٔ دو دزدک پیش می‌رویم. چندین شیوهٔ SAS را می‌توان برای اجرای یک تحلیل آمیخته به‌کار برد که در کد برنامهٔ زیر ارائه شده‌اند.

1. Lilliefor

کد برنامهٔ *SAS*

```
*-------------------------------------------------------------;
  filename in1 'c81.thief.dat' ;  * input the two datasets  ;
  filename in2 'c81.tablet.dat' ;
  data a ;
    infile in1 ;
    input method $ loc rep ya ;
  data b ;
    infile in2 ;
    input method $ drum tablet yb ;
*-------------------------------------------------------------;
                                  * repeated measures ANOVA  ;
proc glm data=a ;                 * with "loc" as random     ;
  class method loc rep ;
  model ya = method loc(method) rep rep*method ;
  random loc(method) / test  ;

  test h=methda e=loc(method) ;
*-------------------------------------------------------------;
                                   * get variance components  ;
proc varcomp data=a method=ml ;    * (subsumed by proc mixed) ;
  class method loc rep ;
  model ya = method rep rep*method loc(method) / fixed=3 ;
*-------------------------------------------------------------;
                                   * set up the data for the  ;
proc sort data=a ;                 * multivariate version of  ;
     by method loc ;               * repeated measures ANOVA  ;
proc transpose out=a2(rename=(_1=y1 _2=y2 _3=y3)) ;
     by method loc ;
     id rep ;
proc glm data=a2 ;
  class methda loc ;
  model y1 y2 y3 = method ;
  repeated rep / summary ;
*-------------------------------------------------------------;
* Output from the "mixed" procedure is presented below.      ;
* ..............................                            ;
                                   * use the "mixed" procedure ;
proc mixed method=ml data=a ;      * which allows more general ;
  class method loc ;               * covariance structures to  ;
  model ya = method ;              * be specified. (SAS 6.08+) ;
  random loc ;
  repeated / subject=loc group=method ;
*-------------------------------------------------------------;
                                   * use the "mixed" procedure ;
proc mixed method=ml data=b ;      * to test for a time effect ;
  class drum ;                     * with respect to DRUM.     ;
  model yb = ;                     * ar(1) = autoregressive    ;
  random drum / s type=ar(1) ;     * covariance structure      ;
```

شیوهٔ proc mixed در نسخهٔ ۸ر۶ منتشرشدهٔ SAS در دسترس است و شیوهٔ proc varcomp را در خود دارد. این شیوه تعمیمی از شیوهٔ proc glm است و امکان برازش ردهٔ گسترده‌تری از مدل‌های آمیخته را میسر می‌سازد. همچنین دستور اندازه‌های «تکراری» انعطاف‌پذیرتری را در مقایسه با نسخه‌های تک‌متغیره یا چندمتغیره‌ای که در proc glm به‌کار گرفته شده‌اند در اختیار می‌گذارد.

در این مثال، LOC را اثر تصادفی اعلام و با سه سطح REP به‌عنوان اندازه‌های تکراری رفتار کرده‌ایم. خروجی proc mixed در زیر ارائه شده که هیچ اثر معنی‌دار METHOD را نشان نمی‌دهد. از این مطلب چنین نتیجه می‌شود که می‌توان از هر یک از دزدک‌ها برای نمونه‌گیری طی فرایند آمیختن و برای مقاصد عملی استفاده کرد، و معلوم شد کار با دزدک دز میانی آسان‌تر است. با این حال، تفاوت‌های معنی‌داری بین مکان‌های نمونه‌گیری وجود داشت که از آن چنین برمی‌آید که مخلوط پودر ممکن است محتوای یکنواختی نداشته باشد.

برای وارسی موضوع اریبی، می‌توانیم از شیوهٔ بوت‌استرپ (افرون و تیبشیرانی، ۱۹۹۳) استفاده کنیم. ایدهٔ بوت‌استرپ این است که از داده‌های مشاهده‌شده با جایگذاری بازنمونه‌گیری کنیم و از برآوردهای بازنمونه‌گیری، به عنوان مثال $\hat{\theta}_i$، برای ارزیابی تغییرپذیری $\hat{\theta}$ حول مقدار واقعی θ استفاده کنیم. بنابراین، اریبی $\hat{\theta}$ را می‌توان با گرفتن میانگین تفاضل‌های $\hat{\theta}_i - \hat{\theta}$ برآورد کرد.

خروجی *SAS* از *proc mixed*

The Mixed Procedure

Covariance Parameter Estimates

Cov Parm	Subject	Group	Estimate
loc			0.9977
Residual	loc	method intm	0.7068
Residual	loc	method unit	3.4001

Fit Statistics

Log Likelihood	-64.4
Akaike's Information Criterion	-67.4
Schwarz's Bayesian Criterion	-67.1
-2 Log Likelihood	128.8

Solution for Random Effects

Effect	loc	Estimate	Std Err Pred	DF	t Value	Pr > |t|
loc	1	-1.6389	0.5498	29	-2.98	0.0058
loc	2	0.2959	0.5498	29	0.54	0.5946
loc	3	-0.4341	0.5498	29	-0.79	0.4362
loc	4	0.5308	0.5498	29	0.97	0.3422
loc	5	-0.0792	0.5498	29	-0.14	0.8864
loc	6	1.3255	0.5498	29	2.41	0.0225

Type 3 Tests of Fixed Effects

Effect	Num DF	Den DF	F Value	Pr > F
method	1	29	1.14	0.2934

	UNIT vs. INTM	UNIT vs. DRUM	UNIT vs. DRUM
Pـمقدار مشاهده‌شده	۰ر۳۷۷۰	۰ر۲۱۳۵	۰ر۰۰۹۰
بوت استریپ میانگین	۰ر۳۹۳۲	۰ر۳۲۰۴	۰ر۰۵۸۶
اریبی	۰ر۰۱۶۲	۰ر۳۰۹۲	۰ر۰۴۹۸
خطای استاندارد	۰ر۲۹۵۸	۰ر۳۰۹۲	۰ر۱۲۴۴

در این مثال، از بوت‌استرپ برای ارزیابی قابلیت اعتماد آزمون t برای مقایسهٔ تفاوت‌های انفرادی بین سه روش استفاده کردیم. توجه کنید که برای مقایسه‌های همزمان، نابرابری بونفرونی[1] برآورد محافظه‌کارانه‌ای از آن سطح بحرانی که برای هر آزمون به‌کار می‌بریم در اختیار می‌گذارد: ۰ر۰۵/۳ = ۰ر۰۱۶۷. نتایج ارائه‌شده در جدول بالا بر مبنای ۱۰۰۰ نمونه است و از آن چنین برداشت می‌شود که تفاوت معنی‌داری که در ابتدا بین مقادیر سنجهٔ دزدک دز میانی و قرص مشاهده‌شده ممکن است لزوماً معتبر نباشد.

۴.۱.۸ خلاصه

ملاحظات عملی و محدودیت‌های زمانی اغلب اندازه و میزان آزمایش‌های تشخیصی را که می‌توان در بخش تولید انجام داد محدود می‌کند. برای تضمین اینکه آزمایش اهداف بررسی را برآورده می‌کند، برنامه‌ریزی دقیق و شیوه‌های نمونه‌گیری کارا از ضروریات است. باید خاطرنشان کنیم که گنجاندن اجزای مؤثر در قرص، بیش از سطحی مشخص، مشکلی برای مصرف‌کننده ایجاد نمی‌کند. با این حال، از منظر سازندهٔ دارو، به‌طور آشکار مطلوب آن است که به مقدار مشخص‌شده برسیم. نتایج این آزمایش را می‌توان بیشتر تحلیل کرد و در پرسش‌های زیر به برخی زمینه‌ها و رویکردهایی که خوانندهٔ علاقه‌مند ممکن است بخواهد دنبال کند اشاره کرده‌ایم.

پرسش‌ها

۱. آیا مقادیر سنجه عموماً خوش‌رفتارند؟ توجه کنید که وقتی با نمونه‌های دزدک به‌عنوان اندازه‌های تکراری رفتار می‌کنیم، لازم است مسئلهٔ همبستگی را در معیار خود برای تعیین دورافتاده‌ها منظور کنیم. پی ببرید که چه شیوه‌ها یا آزمون‌هایی در دسترس‌اند و آن‌ها را برای این داده‌ها به‌کار ببرید.

1. Bonferroni

۲. آیا هیچ شواهدی از اثر مکان در دست هست؟ تحلیل مقدماتی ما دلالت بر وجود آن دارد که این امر مورد علاقهٔ مدیریت تولید است. در گزارش خود به مدیریت چه توصیه‌هایی باید دربارهٔ این موضوع بکنیم؟

۳. آیا داده‌های قرص هیچ نشانه‌ای از اثر ظرف استوانه‌ای‌شکل یا زمان را در خود دارند؟ ما برای این منظور یک ساختار کوواریانس اتورگرسیو AR(1) را در آخرین شیوهٔ `proc mixed` به‌کار گرفتیم. آیا این کار معادل با استفاده از آزمون دوربین[1]ـواتسون[2] است؟

۴. آیا مقادیر نمونه‌گیری دزدکی با مقادیر قرص قابل مقایسه‌اند؟ یک رویکرد، در نظرگرفتن ضریب همبستگی توافقی است که لین (۱۹۸۹) ارائه کرده است؛ توافق بین دو نسخه از یک نمونهٔ یکسان را، با اندازه‌گیری تغییرات از خط 45° وارد به مبدأ، کمّی کنید.

۲.۸ ردیابی مواد منفجرهٔ پلاستیکی

منبع	مؤلفان
روش‌ها	تحلیل ممیزی
داده‌ها	۲۱ متغیر آخرین متغیر، پاسخ ۰/۱ است

۱.۲.۸ تشخیص الگو

اهمیت ردیابی مواد منفجرهٔ پلاستیکی بر اثر تراژدی پرواز ۱۰۳ شرکت پان آمریکن[3]، که بر روی لاکربی[4] اسکاتلند در ۲۱ دسامبر سال ۱۹۸۸ منفجر شد، آشکار گردید. یکی از مؤثرترین ابزارها برای ردیابی مواد منفجرهٔ پلاستیکی نوعی اسکنر اشعهٔ X است که نمایه‌ای از ترکیب شیمیایی ناحیه‌ای کوچک در چمدان را تولید می‌کند. اگر این نمایه الگویی شبیه به یکی از مواد منفجرهٔ شناخته‌شده نشان دهد، در این صورت چمدان به‌عنوان بمب رده‌بندی می‌شود.

تأکید این مطالعهٔ موردی بر قابلیت اعتماد ردیابی مواد منفجره بر مبنای یک نمونهٔ اولیه از دستگاه اشعهٔ X است. به این معنی که مواد منفجرهٔ پلاستیکی را تا چه حدی می‌توان ردیابی کرد؟ چه نوع روش رده‌بندی را باید به‌کار برد؟ قبل از ارائهٔ بحث، باید متذکر شویم که:

> تحلیل ارائه‌شده در اینجا به هیچ عنوان حاوی نتیجه‌گیری‌هایی دربارهٔ قابلیت اعتماد دستگاه‌های بازرسی واقعی، که در مراکز امنیتی نظیر فرودگاه‌ها به‌کار می‌روند، نیست.

1. Durbin 2. Watson 3. Pan American 4. Lockerbie

داده‌ها

برای اینکه به یک قاعدهٔ رده‌بندی برسیم، آزمایشی را ترتیب دادیم که در آن ۲۵۰۰ نمایه به‌دست آمد که از آن میان ۱۲۵۰ مورد متناظر با مواد منفجره بود و بقیهٔ ۱۲۵۰ مورد دیگر موادی نوعی بودند که در چمدان‌ها دیده می‌شوند. هر نمایه، برداری از ۲۰ عدد x_1، ...، x_{20} است که خلاصه‌شده‌ای از سیگنال جذب‌شده به‌وسیلهٔ مواد است. پاسخ y برابر ۱ است هرگاه چمدان حاوی بمب باشد و در غیر این صورت ۰ است. متغیرهای نمایه و پاسخ در مجموعه‌داده‌های c82.dat فراهم شده‌اند.

روش‌شناسی

مجموعه‌ای غنی از ابتکارات روش‌شناسی برای رده‌بندی در دست است. در اینجا خلاصهٔ کوتاهی را ارائه و تجربهٔ خود از روش‌های رده‌بندی را توصیف می‌کنیم.

۲.۲.۸ مروری کوتاه بر تحلیل ممیزی

فرض کنید که نمونه‌ای از k جامعه یا گروه‌های چندمتغیره داشته باشیم و x جدیدی را مشاهده می‌کنیم که جامعه یا گروه متناظر با آن نامعلوم است. می‌خواهیم یک قاعدهٔ رده‌بندی تعریف کنیم که x را به یکی از این گروه‌ها تخصیص دهد. در برخی موارد، شاید احتمال‌های پیشینی مانند π_iها را، که x به گروه i تعلق دارد، بتوانیم تخصیص دهیم. در چنین مواردی، رویکردی بیزی ترجیح داده می‌شود.

تابع‌های ممیزی خطی فیشر

فرض کنید که نمونه‌هایی از k جامعه در اختیار داریم و میانگین‌ها و ماتریس‌های کوواریانس نمونه‌ای $\bar{x}_1$، ...، $\bar{x}_k$ و S_1، ...، S_k را محاسبه می‌کنیم. فرض کنید ماتریس‌های کوواریانس جامعه یکسان‌اند. در این صورت، می‌توانیم ماتریس کوواریانس مشترک جامعه را به‌صورت $S_p = 1/(N-k)\sum_{i=1}^{k}(n_i-1)S_i$ محاسبه کنیم که در آن $N = \sum_{i=1}^{k} n_i$ و n_i اندازهٔ نمونه‌ای است که از جامعهٔ iام به‌دست آمده است. تابع‌های ممیزی خطی به شکل زیرند:

$$L_i(x) = \bar{x}_i' S_p^{-1} x - \bar{x}_i' S_p^{-1} \bar{x}_i = a_i' x + a_{i0}.$$

قاعدهٔ رده‌بندی این است که x به گروه iام تخصیص داده می‌شود هرگاه $L_i(x)$ ماکسیمم شود.

تابع‌های ممیزی درجهٔ دوم

فرض کنید که ماتریس‌های کوواریانس جامعه را برابر در نظر نگیریم. در این صورت، تابع ممیزی درجهٔ دوم زیر را ماکسیمم می‌کنیم.

$$Q_i(x) = -\log|S_i| - ۱/۲(x-\bar{x}_i)'S_i^{-۱}(x-\bar{x}_i)$$

از سایر جهات، این شیوه مانند تحلیل ممیزی خطی است.

وقتی احتمال‌های پیشین را به گروه‌ها تخصیص می‌دهیم، قاعدهٔ رده‌بندی به‌صورت $Q_i^*(x) = \log \pi_i + Q_i(x)$ درمی‌آید.

درخت‌های رده‌بندی

مقدمه‌ای بر درخت‌های رده‌بندی در مطالعهٔ موردی ۴.۸ ارائه شده است.

روش‌های رده‌بندی غیرخطی

استفاده از روش‌های رده‌بندی غیرخطی از قبیل تحلیل ممیزی انعطاف‌پذیر[1] (FDA) روز به روز با استقبال بیشتری مواجه می‌شود زیرا کامپیوترهای سریع در دسترس‌اند و اجرای بسیاری از آن‌ها در نرم‌افزارهای نوینی مانند S-PLUS یا Rتعبیه شده‌اند. شیوهٔ FDA مرکب از گام‌های زیر است.

۱. برازش غیر خطی. یک مدل غیرخطی با استفاده از نمایش عددی دودویی پاسخ‌ها به داده‌ها برازش دهید.

۲. گام ممیزی خطی. یک قاعدهٔ رده‌بندی ممیزی خطی به مقادیر برازش‌یافتهٔ گام اول به‌عنوان متغیر پیشگو و با همان متغیر پاسخ برازش دهید.

برای گام اول، از هر برآوردگر رگرسیونی ناپارامتری غیرخطی مانند مدل‌های جمعی تعمیم‌یافته، لووس[2]، رگرسیون جستجوی تصویری، یا مارس[3] استفاده کنید. برای یافتن مراجع بیشتر دربارهٔ FDA و مدل‌های غیرخطی نگاه کنید به هستی و همکاران (۱۹۹۵) و چمبرز و هستی (۱۹۹۲).

1. flexible discriminat analysis

۲. lowess (locally weighted scatter plot smoothing) یا هموارسازی موضعاً موزون پراکنش‌نگار از روش‌های «نوین» مدلبندی است که بر روش‌های کلاسیک مانند رگرسیون کمترین توان‌های دوم خطی و غیرخطی مبتنی است.ـم.)

۳. MARS (multivariate adaptive regression spline یا اسپلاین‌های رگرسیون سازوار چندمتغیره نوعی تحلیل رگرسیونی است که در سال ۱۹۹۱ توسط جروم فریدمن (Jerome Friedman) معرفی شده است.ـم.)

شبکه‌های عصبی

شبکه‌های عصبی ــ با نشان‌دادن اسلایدی دربارهٔ مغز انسان به ما این‌گونه گفته شد که ــ شبیه به مغز کار می‌کنند (و ما اضافه می‌کنیم که «اگر چنین باشد، در دردسری جدی قرار داریم»). در واقع امر، این شبکه‌ها معمولاً مدل‌هایی غیرخطی یا قواعد رده‌بندی به‌شدت بیش‌پارامتری شده‌اند. این مطلب خلاف تفکر آماری سنتی است که بیش‌پارامتری‌کردن را با بیش‌برازش یکی می‌داند.

با این حال، در بسیاری از مسائل رده‌بندی نوین، بیش‌پارامتری‌کردن ممکن است مزیت محسوب شود زیرا ممکن است داده‌ها برخی یا همهٔ مشخصه‌های زیر را داشته باشند.

۱. بین گروه‌ها همپوشانی وجود ندارد.

۲. رویهٔ تفکیک به‌شدت غیرخطی است.

۳. مقادیر عظیمی از داده‌ها موجودند به‌طوری که تفکیک، خوش‌تعریف است.

برای این نوع از مسائل، ریسک بیش‌برازش ممکن است ارزشمند باشد در صورتی که در عین حال برازش توانایی بازتاب‌دادن غیرخطی‌بودن‌های شدید را داشته باشد.

یک مدل پایه‌ای برای شبکه‌های عصبی تحت عنوان «شبکه‌های عصبی لایه‌مخفی یگانهٔ تغذیه رو به جلو»[1] ظاهر می‌شود که مرکب از یک لایهٔ ورودی، یک لایهٔ خروجی، و یک لایهٔ مخفی بین آن‌هاست. هر گره یک یا چند ورودی و یک خروجی دارد. مدل شبکهٔ عصبی به‌صورت زیر ساخته می‌شود:

۱. لایهٔ ورودی (یا نخستین لایه) مرکب از گره‌هایی به همان تعداد متغیرهای پیشگوست که برای برازش در اختیار داریم. خروجی هر گره، مقدار متناظر متغیر پیشگویی است که به آن تخصیص داده شده است.

۲. هر گره، خروجی‌های گره‌های لایهٔ پیش از خود را وارد می‌کند و خروجی آن، تابعی ثابت از ترکیب خطی است. این تابع، که گاهی «تابع انتقال» نامیده می‌شود، معمولاً یک تابع لوژستیک یا هر تابع هلالی‌شکل[2] برای مسائل رده‌بندی و تابع همانی برای مسائل رگرسیونی است.

۳. لایهٔ خروجی (یا آخرین لایه)، به همان تعداد پاسخ‌های موجود برای برازش، گره دارد.

فرایند برازاندن پارامترهای شبکهٔ عصبی از یک الگوریتم «پس‌برازیدن[3]» بسیار پیچیده استفاده می‌کند که همواره مقادیر بهینهٔ پارامتر را نمی‌یابد. اما از جنبه‌های خوب این مدل شبکهٔ

1. feed-forward single hidden layer neural nets 2. sigmoidal-shaped 3. backfitting

عصبی آن است که می‌توان آن را در قالب سخت‌افزارهایی در میکروچیپ‌ها تعبیه کرد که در کاربردهای مهندسی بسیار ارزشمند تلقی می‌شوند. برای آماردانان اشتباه خواهد بود که روش‌شناسی شبکهٔ عصبی را نادیده بگیرند زیرا در بسیاری از کاربردها نتایجی بسیار عالی تولید می‌کند. بر مشکلات مورد اشاره در بالا می‌توان با پژوهش‌های جدید و منابع محاسباتی عظیم‌تر فائق آمد.

k نزدیک‌ترین همسایه‌ها

روش‌های نزدیک‌ترین همسایه نیز برای حالت‌هایی که در آن شبکه‌های عصبی ترجیح داده می‌شوند سودمندند. این ایده کاملاً ساده است. لازم است که کار را با تعریفی از فاصله‌های بین‌نقطه‌ای $d(x_1, x_2)$ نظیر آن‌هایی که در مطالعهٔ موردی ۴.۸ به‌عنوان بخشی از روش‌شناسی تحلیل خوشه‌ای ذکر شد شروع کنیم. معمولاً فاصلهٔ ماهالانوبیس[1] را به‌کار می‌برند.

برای تعریف قاعدهٔ رده‌بندی برای x جدید، k نقطهٔ از همه نزدیک‌تر به x را برحسب فاصلهٔ مناسب x_1، ...، x_k پیدا می‌کنیم و p_i را که نسبت نقاط متعلق به گروه iام است محاسبه می‌کنیم. قاعدهٔ رده‌بندی این خواهد بود که x را به گروهی تخصیص دهیم که بیشترین نسبت را در درون مجموعهٔ نزدیکترین همسایه‌ها داشته باشد. اگر توزیع پیشینی داشته باشیم، قاعدهٔ رده‌بندی بیزی، حاصل‌ضرب $p_i \times \pi_i$ را ماکسیمم خواهد کرد.

یک راه ساده برای خلاصه‌کردن نتایج تحلیل رده‌بندی آن است که یک جدول دوطرفه از رده‌بندی پیشگویی‌شده به‌وسیلهٔ قاعدهٔ تصمیم در برابر رده‌بندی واقعی را به نمایش بگذاریم. از نسبت بدرده‌بندی‌شده‌ها برای ارزیابی شیوه‌ها استفاده می‌شود. در برخی از تنظیمات ممکن است بهتر باشد که داده‌ها را به مجموعه‌های آموزشی و آزمایشی تقسیم کنیم و از مجموعهٔ آزمایشی برای مقایسهٔ شیوه‌های رده‌بندی استفاده کنیم.

۳.۲.۸ تحلیل مقدماتی

در ابتدا داده‌ها را به‌تصادف بین مجموعهٔ آموزشی متشکل از ۱۰۰۰ مشاهده از هر گروه و یک مجموعهٔ آزمایشی متشکل از ۲۵۰ مشاهده برای هر گروه تقسیم می‌کنیم. مجموعهٔ آموزشی برای برآورد قاعده‌های رده‌بندی و مجموعهٔ آزمایشی برای مقایسهٔ نحوهٔ عمل قاعده‌ها به‌کار می‌روند.

ما تحلیل ممیزی خطی و درجهٔ دوم پایه‌ای را برازش دادیم. علاوه بر این مدل، FDA را با استفاده از رگرسیون جستجوی تصویر در مرحلهٔ برآورد تابع برازش دادیم.

1. Mahalanobis

جدول ۳.۸ نتایج تحلیل ممیزی خطی

شبکهٔ عصبی	تابع ممیزی			
	انعطاف‌پذیر	درجهٔ دوم	خطی	
۳ (۲)	۱ (۳)	۲ (۳)	۸ (۳۷)	هشدارهای غلط
۱ (۱)	۰ (۵)	۱ (۳)	۲ (۳)	بمب‌های کشف‌نشده
۴ (۳)	۱ (۸)	۳ (۶)	۱۰ (۳۸)	کل

سرانجام یک شبکهٔ عصبی تغذیه رو به جلو به کمک یک لایهٔ مخفی مرکب از ۱۰ گره (نصف پیشگوها) برازش داده شد. هم جستجوی تصویر و هم شبکه‌های عصبی را در S-PLUS با استفاده از تنظیمات پیش‌فرض برازش دادیم. این نتایج در جدول ۳.۸ نشان داده شده‌اند که در آن عدد داخل پرانتزها نتیجه را برای مجموعهٔ آموزشی به‌دست می‌دهد، در حالی که عدد اصلی نتیجه‌ای است که از مجموعهٔ آزمایشی به‌دست آمده است.

تفاوت بین برازش‌های خطی و درجهٔ دوم زیاد است. سه برازش غیرخطی نتایج مشابهی را به‌دست دادند، در حالی که شبکهٔ عصبی در مجموعهٔ آموزشی اندکی بهتر از مجموعهٔ آزمایشی عمل کرد.

۴.۲.۸ خلاصه

این مطالعهٔ موردی مثالی از یک مجموعه‌داده را تقریباً بدون هیچ خطای آماری تشریح می‌کند. روش‌شناسی‌های متعدد رقیب بسیار نزدیک به‌هم‌اند و برنده‌ای مطلق بین آن‌ها وجود ندارد. پرسش‌های زیر برخی از مباحثی را که لازم است یک گزارش به آن‌ها بپردازد مشخص می‌کند.

پرسش‌ها

۱. هدف عمده در اینجا انتخاب روشی است که در مورد داده‌ها بهتر از بقیه عمل می‌کند. تحلیل مقدماتی، بسیار اساسی است. ممکن است لازم باشد که پارامترهای روش‌های مختلف را با هم میزان کنیم تا عملکرد بهتر شود.

۲. آیا می‌توانید برای مقایسهٔ روش‌ها راه بهتری طرح‌ریزی کنید؟

۳. کدام‌یک بهترین نتایج را به‌دست می‌دهند؟

۳.۸ یک مطالعهٔ تحقیق بازار

روش‌ها	تحلیل عاملی
داده‌ها	۲۶ متغیر
	۵ ارزیاب برنامهٔ تلویزیونی
	۶ رستهٔ خرید

این مطالعه توسط ماریا درلیچ[1] از برنامهٔ MBA در دانشگاه راتگرز[2] ارائه شد و ایده‌های این گزارش از اوست. این مثال متناظر با وضعیتی کاملاً نوعی است که در آن داده‌های موجود قبلاً پردازش شده‌اند و این کار نتیجهٔ جدول‌بندی‌هایی است که از روی داده‌های خام انجام شده است. داده‌های خام در دسترس نیستند و تنها جدول‌ها را در اختیار داریم.

هدف این مطالعه، شناسایی بخش‌های مصرف‌کنندگان بود که نمایه‌های «خرید از روی کاتولوگ[3]» مشابهی با نمایه‌های بینندگان برنامه‌های محبوب تلویزیونی داشتند. هر نمایه مرکب از برداری شش‌عددی است که نشان‌دهندهٔ نسبت افرادی از زیرمجموعه‌ای خاص از جامعه است که محصولات را از هر یک از شش رسته خریداری کرده‌اند.

با این روش می‌توانیم الگوهای پیوند بین بخش‌هایی از جامعه و بینندگان برنامه‌هایی خاص را پیدا کنیم و سعی کنیم که نمایه‌های بینندگان مشخصی را برای محصولات خود ایجاد کنیم.

داده‌ها

داده‌های جدول ۵.۸ مرکب از نمایه‌های خرید ۲۹ زیرگروه جامعه است. این نتایج در `c83.dat` موجودند. متغیر توصیف زیرگروه[*Subgroup Description*] مشخصه‌های دقیق هر زیرگروه را ارائه می‌کند. الگوی خرید مرکب از شش اندازه‌گیری از تعداد افرادی از هر زیرگروه است که انواع معینی از کالاها را از روی کاتالوگ، طی ۱۲ ماه گذشته، به‌عنوان درصد کل جامعهٔ ایالات متحده در آن زیرگروه خریداری کرده‌اند. بنابراین، متغیرهای V_3 تا V_8 از مجموعه‌داده‌ها در جدول ۵.۸ نشان‌دهندهٔ درصدها هستند.

در عین حال، نمایه‌های مشابهی برای بینندگان پنج برنامهٔ تلویزیونی محاسبه شد. فرض زمینه‌ای در اینجا عبارت از این است که اگر نمایه‌های خرید بخشی از جامعه به نمایه‌های بینندگان

1. Maria Drelich 2. Rutgers

۳. purchasing from catalogue (خریدی که از روی کاتالوگ اقلام موجود در یک شرکت فروشندهٔ کالاهای مصرفی در منزل یا محل کار انجام و با پست ارسال می‌شوند. م.)

جدول ۴.۸ تعریف‌های متغیر

V_0	عدد ID
V_1	توصیف زیرگروه
V_2	کل جامعهٔ زیرگروه‌ها

درصد زیرگروه V_2 که خریدهای زیر را انجام داده‌اند

V_3	لباس
V_4	لوازم الکترونیکی
V_5	مبلمان منزل
V_6	لوازم خانگی
V_7	کالاهای ورزشی
V_8	اسباب و وسایل بازی

برنامه‌ای نزدیک باشد، می‌توانیم آن زیرگروه را به‌عنوان بخشی از بینندگان آن برنامه بدانیم. این مطلب ممکن است لزوماً درست نباشد، اما مانند بسیاری از مطالعه‌های تحقیق بازار، نتایج ممکن است پیوندهای جالبی در اختیار گذارد که از دیدگاه بازاریابی معنی و مفهوم داشته باشد.

روش‌شناسی

۱.۳.۸ مروری کوتاه بر تحلیل مؤلفه‌های اصلی

تحلیل مؤلفه‌های اصلی روشی برای کاهش بعد است. در زیر، مثال‌هایی از کاربردهای آماری ارائه می‌شود که در آن‌ها کاهش بعد اهمیت دارد.

- داده‌کاوی: این مبحث درکاربردهایی به خدمت گرفته می‌شود که به مجموعه‌داده‌هایی با متغیرها و حالت‌های بسیار زیاد سروکار دارند و یکی از اهداف اصلی آن کاهش داده‌هاست. مبحث مؤلفه‌های اصلی این توانایی را دارد که به‌عنوان ابزاری برای کاهش داده‌ها یا تقلیل تعداد متغیرها به‌کار گرفته شود. برای ملاحظهٔ جزئیاتی دربارهٔ داده‌کاوی نگاه کنید به مطالعهٔ موردی ۴.۸.
- تحلیل رگرسیونی: وقتی تعداد پیشگویی‌های q با درجه‌آزادی خطای n_ϵ قابل قیاس است. لازم است که تعداد متغیرها را به q^* با $q^* << n_\epsilon$ کاهش دهیم.

جدول ۵.۸ انواع کالاهای خریداری شده از روی کاتالوگ طی ۱۲ ماه گذشته برحسب متغیرهای جمعیت‌شناختی و برنامه‌های تلویزیونی به‌عنوان درصد گروه

V_0	V_1	V_2	V_3	V_4	V_5	V_6	V_7	V_8
1	18-24	23965	15.68	2.34	5.05	4.21	1.19	5.27
2	25-34	42832	22.57	3.98	7.99	5.74	4.20	8.99
3	35-44	39908	31.02	4.94	11.63	6.85	6.12	10.97
4	45-54	27327	31.10	5.72	9.43	8.39	4.99	6.54
5	55-64	21238	28.26	3.22	9.83	8.39	4.17	7.40
6	۶۵ به بالا	30552	24.36	3.09	7.27	4.66	1.43	4.07
7	فارغ‌التحصیلی دانشگاه	36463	36.34	5.97	10.16	6.39	6.24	10.41
8	با تحصیلات دانشگاهی	44294	29.08	4.55	10.92	6.86	4.37	8.28
9	فارغ‌التحصیل دبیرستان	66741	23.70	3.27	8.10	7.02	4.20	7.32
10	بدون مدرک دبیرستان	38324	15.14	2.75	5.77	4.30	2.38	4.56
11	کل مردان	88956	18.91	4.76	5.38	4.34	4.99	6.00
12	کل زنان	96866	31.93	3.30	11.74	8.09	3.60	9.04
13	مرد شاغل	65500	19.46	4.93	5.42	4.25	5.51	6.45
14	زن شاغل	55910	36.42	3.62	14.02	9.01	4.54	10.51
15	تمام‌وقت	110363	26.65	4.46	9.20	6.31	8.00	5.03
16	نیمه‌وقت	11047	33.51	2.98	10.74	7.78	5.47	11.51
17	بیکار	64412	22.73	3.37	7.41	6.01	2.76	6.20
18	مجرد	41284	16.86	3.60	5.17	4.21	3.60	3.75
19	متأهل	109023	29.94	4.51	10.65	7.23	5.12	9.61
20	بیوه/مترکه کرده/مسئله	35515	22.95	2.90	6.79	5.83	2.41	5.84
21	دانشجو	62342	29.22	4.76	10.75	6.40	5.65	12.37
22	بیش از ۷۵ هزار دلار	24165	35.59	4.42	11.09	8.03	5.00	11.19
23	بیش از ۶۰هزار دلار	40979	32.80	4.85	9.87	7.20	5.12	9.80
24	بیش از ۵۰ هزار دلار	57996	32.66	4.57	10.46	7.07	5.55	9.59
25	بیش از ۴۰ هزار دلار	80078	31.56	4.65	10.54	6.94	5.60	9.27
26	بیش از ۳۰ هزار دلار	106838	30.62	4.64	10.15	6.71	5.45	8.97
27	بیش از ۲۰ هزار دلار	30669	21.32	3.91	8.14	6.01	3.14	6.10
28	بیش از ۱۰ هزار دلار	29083	19.23	2.39	6.17	6.36	2.65	5.81
29	کمتر از ۱۰ هزار دلار	19232	15.12	2.96	5.35	4.37	1.91	4.92
30	نی ار	19640	31.65	4.22	10.65	7.22	4.97	9.58
31	فرندز	16000	23.92	3.52	7.54	5.86	4.50	5.64
32	فریزر	14840	31.03	3.98	10.51	8.25	4.66	8.51
33	جس	13550	19.14	3.77	6.49	4.97	4.24	5.28
34	الی مک بیل	10190	21.80	2.40	6.42	6.32	4.58	6.67

- MANOVA: وقتی تعداد پاسخ‌های p با درجه‌آزادی n_ϵ قابل قیاس است. لازم است که $p << n_\epsilon$.

فرض کنید که n بردار تصادفی $x_i = (x_{i1}, ..., x_{ip})'$، $i = 1, ..., n$، را مشاهده و فرض می‌کنیم که x_iها دارای میانگین صفر و ماتریس کوواریانس نمونه‌ای S_x باشند. فرض کنید که A تبدیل متعامدی باشد به‌طوری که $z_i = Ax_i$ و z_iها ناهمبسته‌اند. از آنجا که A متعامد است، $x_i'z_i = x_i'x_i$، بنابراین، z_iها دارای میانگین صفرند و ماتریس کوواریانس نمونه‌ای، ماتریس قطری

$S_z = AS_xA'$ است. سطرهای $\{a_i\}$ ماتریس A، ویژه‌بردارهای S_x نامیده می‌شوند و به عناصر قطری $\{\lambda_i\}$ ماتریس S_z، ویژه‌مقدارهای S_x اطلاق می‌شود.

ویژه‌بردارها مجموعه‌ای از متغیرهای جدید به نام «مؤلفه‌های اصلی» تولید می‌کنند که متناظر با دستگاه مختصات «طبیعی» داده‌ها هستند. ویژه‌مقدارها واریانس‌های متغیرهای مؤلفه‌های اصلی متناظر خود هستند و مؤلفه‌های اصلی معمولاً بنابر اندازهٔ ویژه‌مقدارهای خود، از بزرگترین به کوچکترین، مرتب می‌شوند. ویژه‌مقدارها اغلب به عنوان واریانس‌هایی که توسط مؤلفه‌های اصلی متناظر تبیین می‌شوند، توصیف می‌شوند. اولین مؤلفهٔ اصلی جهت بیشترین تغییرپذیری را نشان می‌دهد، دومین مؤلفهٔ اصلی جهت بیشترین تغییرپذیری در فضای عمود بر اولین مؤلفهٔ اصلی را نمایش می‌دهد، و به همین ترتیب الی آخر.

روش‌شناسی مؤلفه‌های اصلی زمانی به‌کار می‌رود که بسیاری از متغیرها در مجموعه‌داده‌ها زائدند؛ یعنی، هیچ اطلاع جدیدی بجز آن‌ها که دیگر متغیرها مبیّن آن‌ها هستند به‌دست نمی‌دهند. به‌لحاظ هندسی این بدان معنی است که داده‌ها تقریباً در زیرفضایی با بعد $k < p$، k، قرار دارند. اگر وضعیت بدین‌گونه باشد، انتظار داریم که k بزرگترین مؤلفهٔ اصلی تقریباً همهٔ تغییرپذیری در داده‌ها را تبیین کنند.

دو کمیت مورد توجه عبارت‌اند از $b_k = \lambda_k / \sum_{i=1}^{p} \lambda_i$ و $c_k = \sum_{i=1}^{k} b_i$ که در آن $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_p$. این دو کمیت به‌ترتیب نشان‌دهندهٔ نسبت واریانس تبیین‌شده به‌وسیلهٔ kامین مؤلفهٔ اصلی، و نسبت واریانس تبیین‌شده به‌وسیلهٔ نخستین k مؤلفهٔ اصلی هستند. ما سه معیار ساده را برای تعیین k پیشنهاد می‌کنیم.

- کمترین تعداد مؤلفه‌هایی که بیش از ۸۰٪ (یا هر عدد جادویی دیگر) واریانس را تبیین می‌کنند؛ یعنی، $c_k > ۰٫۸$.
- آن‌ها بیش از میانگین واریانس‌ها را تبیین می‌کنند؛ یعنی، $b_k > c_p/p$.
- رسم نمودار واریانس‌های عامل‌ها در برابر ترتیب، یعنی b_k در برابر k. آخرین جهش رو به پایین بزرگ، نقطهٔ برش را برای تعداد مؤلفه‌های اصلی نشان می‌دهد.

به طور ایده‌آل، این سه روش باید تعداد مؤلفه‌های اصلی واحدی را مشخص کنند. در صورت هر گونه تردید، بهتر است رویکرد محافظه‌کارانه‌ای را در پیش بگیریم.

۲.۳.۸ مروری کوتاه بر تحلیل عاملی

همان زمینه‌چینی بخش مؤلفه‌های اصلی را دنبال می‌کنیم. ایدهٔ تحلیل عاملی آن است که متغیرهای مشاهده‌شدهٔ $x_1, \ldots, x_p$ را به‌صورت ترکیب‌های خطّی چند کمیت نامعلوم موسوم به عامل‌های $f_1, \ldots, f_k$ که $(k < p)$ بیان کنیم. برای بحث زیر مهم است توجه کنیم که:

۱. عامل‌ها کمیت‌هایی نامعلوم‌اند.

۲. تعداد عامل‌ها نامعلوم است.

۳. برای سادگی، مدلی را در نظر می‌گیریم که نشان‌دهندهٔ متغیرهای مشاهده شدهٔ ما به‌عنوان ترکیب‌های خطی عامل‌های نامعلوم به‌اضافهٔ نوفه است.

در ارتباط با ایدهٔ تحلیل عاملی مناقشه‌های زیادی موجود است، به این دلیل که عامل‌ها را نمی‌توان اندازه گرفت؛ این عامل‌ها ممکن است موجود باشند یا نباشند. با این حال، در این کتاب، ما رویکرد عمل‌گرایانه‌ای را نسبت به این روش‌شناسی اتخاذ می‌کنیم و استفاده از تحلیل عاملی را هر زمان که سودمند باشد توصیه می‌کنیم.

برای اینکه تحلیل مؤلفه‌های اصلی را از تحلیل عاملی تمیز دهیم، توجه می‌کنیم که مؤلفه‌های اصلی ترکیب‌هایی خطی از متغیرها هستند اما متغیرها ترکیب‌هایی خطی از عامل‌ها هستند. مؤلفه‌های اصلی درصدد توضیح تغییرپذیری هستند، اما عامل‌ها درصدد بیان ساختار همبستگی‌اند.

مدل عاملی. مانند بخش قبل، فرض کنیم که x_iها مرکزی شده‌اند. مدلی برای تحلیل عاملی به‌صورت زیر است: $x = Af + \epsilon$، که در آن f برداری تصادفی با میانگین صفر و ماتریس کوواریانسی برابر با ماتریس واحد I_p است، خطاها ناهمبسته‌اند، و $Cov(f, \epsilon)$ ماتریس $k \times p$ی با درایه‌های صفر است.

بارهای عاملی. عناصر A بارهای عاملی نامیده می‌شوند که نمایش‌دهندهٔ کوواریانس‌های متغیرهایی با عامل‌های $a_{ij} = Cov(x_i, f_j)$ هستند. سطر iام A مختصات متغیرهای x_i را در مختصات عاملی می‌دهد.

همه‌داشت‌ها. ما واریانس متغیر را بین مؤلفه‌ای ناشی از عامل‌های مشترک (همه‌داشت) و متغیری خاصِ مؤلفه تجزیه می‌کنیم:

$$Var(x_i) = h_i^2 + \sigma_i^2 = \sum_{j=1}^{p} a_{ij}^2 + Var(\epsilon_i)$$

برآورد. برای تعیین تعداد عامل‌ها می‌توانیم از همان شیوهٔ سه‌طرفه که برای انتخاب تعداد مؤلفه‌های اصلی پیشنهاد شده، استفاده کنیم. این شیوه را در تحلیل مقدماتی در بخش بعدی تشریح می‌کنیم.

برای برآوردکردن پارامترها در مدل عاملی، می‌توانیم فرض کنیم که خطاها به‌صورت نرمال توزیع شده‌اند، که در این حالت، می‌توانیم A و f را به روش ماکسیمم درستنمایی برآورد کنیم.

اگر نخواهیم مدل نرمال را در نظر بگیریم، می‌توانیم عامل را با مؤلفه‌های اصلی تکراری برآورد کنیم که در آن، در هر مرحلهٔ تکرار، مدل عاملی را، با به‌کاربردن مؤلفه‌های اصلی، به مانده‌های تکرار قبلی برازش می‌دهیم.

دوران‌ها. یکی از مشخصه‌هایی که تحلیل عاملی را مناقشه‌آمیزتر می‌کند آن است که عامل‌های غیر قابل مشاهده، تحت تبدیل‌های متعامد، یکتا نیستند. فرض کنیم که H یک تبدیل نامتعامد باشد؛ در این‌صورت اگر $A^* = AH$ و $f^* = H'f$، آنگاه $Af = A^*f^*$ و A^* بارهای عاملی متناظر با عامل‌های تبدیل‌یافتهٔ f^* هستند. به این دلیل، عامل‌های اولیه معمولاً، به‌منظور یکی‌سازی هر متغیر با عاملی یگانه، دوران داده می‌شوند.

دوران‌های دستی. وقتی تنها دو یا سه عامل داشته باشیم، سودمند است که نمودار سطرهای A را به‌صورت نمودارهای ۲ یا ۳ بعدی رسم کنیم و تعیین کنیم که آیا دوران برای یکی‌سازی عامل‌ها با متغیرها یا گروه‌هایی از متغیرها سودمند است یا خیر. اگر چنین باشد، آسان است که زاویه‌های دوران را با چشم تعقیب کنیم و دوران را به‌طور دستی انجام دهیم.

دوران‌های واریماکس. دورانی که واریانس توان دوم بارهای هر عامل را ماکسیمم می‌کند دوران «واریماکس»[1] نامیده می‌شود. ایده آن است که هر عامل باری نزدیک به ۰ یا ۱ دارد (در ماتریس همبستگی).

۳.۳.۸ تحلیل مقدماتی

برای تعیین نمایهٔ خرید مشتری مشخص خود که برنامهٔ تلویزیونی خاصی را می‌بیند، کار را با اجرای یک تحلیل عاملی با استفاده از نرم‌افزار S-PLUS آغاز می‌کنیم.

روش انتخاب برای تحلیل عاملی استفاده از مؤلفه‌های اصلی و به دنبال آن یک دوران واریماکس است. ما دو عامل را برمی‌گزینیم زیرا سه معیاری که مطرح کردیم به دو عامل اشاره دارند:

- بیش از ۸۰٪ واریانس را تبیین می‌کنند.
- بیش از میانگین واریانس‌ها(۱۶٫۷٪) را تبیین می‌کنند.
- نموداری از واریانس‌های عاملی در برابر ترتیب نشان می‌دهد که آخرین جهش رو به پایین بزرگ از نقطهٔ ۲ به ۳ است.

1. varimax

جدول ۶.۸ نتایج تحلیل عامل

(الف) واریانس تبیین‌شده توسط عامل‌های دوران‌یافته

	عامل ۱	عامل ۲
بارهای SS	۲٫۹۴۰۳۴۳۳	۱٫۸۷۵۴۵۷۷
واریانس نسبتی	۰٫۴۹۰۰۵۷۲	۰٫۳۱۲۵۷۶۳
واریانس تجمعی	۰٫۴۹۰۰۵۷۲	۰٫۸۰۲۶۳۳۵

(ب) بارهای عاملی

بارها	V3	V4	V5	V6	V7	V8
عامل ۱	۰٫۸۵۷	۰٫۱۵۶	۰٫۹۲۱	۰٫۸۶۰	۰٫۳۳۱	۰٫۶۹۶
عامل ۲	۰٫۴۴۸	۰٫۷۵۳	۰٫۳۲۷	۰٫۱۱۱	۰٫۸۳۱	۰٫۵۴۶

از بارهای عاملی چنین برداشت می‌شود که عامل ۱ اساساً ترکیبی از فروش‌های البسه، مبلمان، لوازم خانگی، و اسباب‌بازی/ بازی کامپیوتری است، در حالی که عامل دوم ترکیبی از فروش‌های کالاهای الکترونیک و ورزشی است. شکل ۳.۸ تصویرهای متغیرهای اصلی را بر صفحهٔ تعریف‌شده با دو عامل نمایش می‌دهد. علاوه بر این، مشاهدات در مختصات عاملی نشان داده شده‌اند.

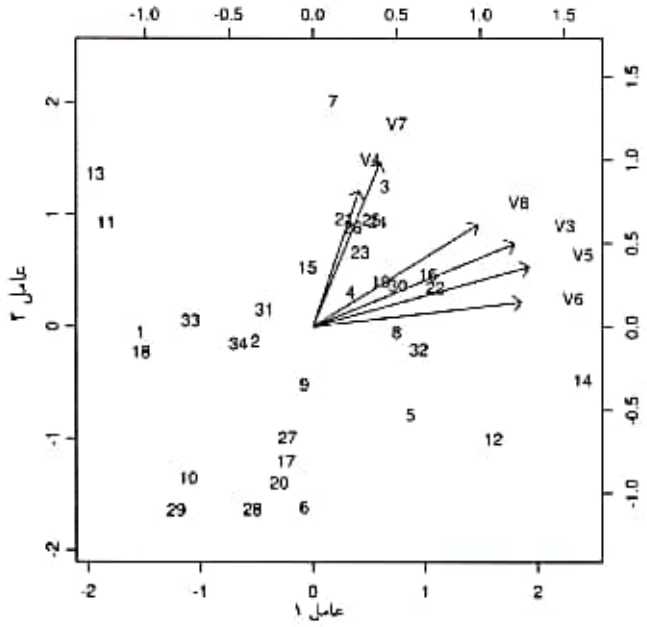
در مثال‌های ساده‌ای از این نوع، نیازی به اجرای تحلیل خوشه‌ای نیست، زیرا نمودار ۳.۸ آن‌قدر ساده هست که بتوان خوشه‌ها را به‌طور دستی استخراج کرد. در مجموعه‌داده‌های پیچیده‌تر شاید مناسب‌تر باشد که کار را با اجرای تحلیل خوشه‌ای پیش ببریم.

درگوشهٔ پایین از سمت چپ نمودارها، گروهی را مشاهده می‌کنیم که از ۱۰، ۲۰، ۲۸، ۲۹،... ساخته شده و متناظر با بینندگانی است که از دبیرستان فارغ‌التحصیل نشده‌اند، مطلّقه‌اند، متارکه کرده، یا بیوه‌اند، وگروه‌هایی که کمترین درآمد را دارند. این‌ها از لحاظ هر دو عامل در سطح پایینی قرار دارند. از سوی دیگر، گروه‌های ۱۱ و ۱۳ که متناظر با مردان شاغل‌اند، از لحاظ عامل ۲ بالا هستند ولی از لحاظ عامل ۱ پایین‌اند. در حالی که زنان شاغل و به‌طورکلی زنان (۱۲ و ۱۴) از لحاظ عامل ۱ بالا و از لحاظ عامل ۲ پایین‌اند. کل این موارد ممکن است کلیشه‌ای به نظر برسد، اما از اعداد چنین برمی‌آید.

برحسب برنامه‌های تلویزیونی، نتیجه‌گیری‌های اولیه به قرار زیرند:

بینندگان برنامه‌های ئی‌آر و فریزر[1] الگوهای خرید مشابهی دارند، درآمدهای متوسط به بالا دارند

۱. ER و Fraiser دو برنامهٔ تلویزیونی معروف آمریکایی‌اند ـ م.



شکل ۳.۸ بارهای عاملی و نقاط داده‌ای در پایهٔ عاملی.

(بالای چهل هزار دلار)، میانسال (۳۰ تا ۴۰ سال)، و صاحب فرزندند. بنابراین، طبق این تحلیل، نتیجه می‌گیریم که بسیاری از بینندگان تی‌آر و فریزر درآمدی متوسط به بالا دارند (بالای ۴۰ هزار دلار در سال)، میانسال‌اند و اغلب صاحب فرزندند.

مخاطبان برنامهٔ الی مک‌بیل[1] مرکب است از بینندگانی با درآمدی نسبتاً کمتر (۳۰ هزار دلار به بالا)، سنین بالاتر، اغلب متأهل و زن، و احتمالاً با تحصیلات دانشگاهی.

سرانجام، مخاطبان برنامهٔ جس[2] مرکب است از افراد جوان عمدتاً بین سنین ۲۵ تا ۳۴، شاغل تمام‌وقت، با تحصیلات دبیرستانی.

یک خوشهٔ بسیار قوی‌تر دیگر مرکب از زنان شاغل، فارغ‌التحصیلان دانشگاهی، و مردمی با درآمد بالای ۷۵۰۰۰ دلار هستند. الگوی خرید ازکاتولوگ برای این گروه قویتر از همه است. گرچه آن‌ها مستقیماً با هیچ‌یک از پنج برنامهٔ تلویزیونی وابستگی ندارند، تنها به‌طور متوسط با برنامه‌های تی‌آر، فرندز[3]، فریزر، و الی مک‌بیل ارتباط دارند اما با جس ارتباطی ندارند.

1. Ally McBeal 2. Jesse 3. Friends

گروه باقیمانده متشکل از جوان‌ترها (کمتر از ۲۴ سال)، یا از سالمندان (۶۵ سال به بالا)، بیکاران، مطلقه‌ها، یا متارکه‌کرده‌ها، با درآمد پایین است. آن‌ها از نمایه‌های بینندگان برای هیچ‌یک از برنامه‌های تلویزیونی بالا تبعیت نمی‌کنند. در طی کار، همچنان که برنامه‌های دیگری را اضافه می‌کنیم، تحلیل‌های مشابهی را برای هر یک از آن‌ها اجرا می‌کنیم تا نمایه‌های مناسبی از بینندگان به‌دست آوریم.

۴.۳.۸ خلاصه

این مطالعهٔ موردی برخی از روش‌شناسی‌های موجود برای مطالعه‌های تحقیقات بازار را تشریح می‌کند. یکی از مزیت‌های عمدهٔ تحلیل تحقیقات بازار، اکتشافی‌بودن آن است. در بسیاری از کاربردها مهم است که سریعاً اطلاعاتی به‌دست آوریم و لزومی ندارد که این یافته‌ها به کمک شیوه‌هایی رسمی مانند طرح آزمایش‌ها و آزمون فرض‌ها تأیید شوند. هدف آن است که اطلاعات جدید را سریع‌تر از رقبایتان، حتی با خطر درست‌نبودن آن‌ها در همهٔ موارد، به‌دست آورید. لطفاً پرسش‌های زیر را در گزارش تفصیلی خود مورد توجه قرار دهید.

پرسش‌ها

۱. کدام تبدیل‌ها، در صورت وجود، برای این داده‌ها مناسب‌اند؟

۲. چندین گزینه را برای برآورد و دوران در نظر بگیرید؛ آیا نتایج با هم سازگاری دارند؟

۳. نقدی بر تحلیل مقدماتی و تحلیل‌های جایگزین به‌عمل آورید. نتیجه‌ها را در ارتباط با اهداف این مطالعه تعبیر کنید.

۴.۸ فروش تجهیزات ارتوپدی

روش‌ها	داده‌کاوی تحلیل اکتشافی داده‌ها تحلیل ممیزی تحلیل رگرسیونی
داده‌ها	دو پاسخ ۸ پیشگوی جمعیت‌شناختی داده‌ها ۸ پیشگوی خاص برای ارتوپدی

۱.۴.۸ کاربردهای داده‌کاوی در تحقیقات بازار

در این مطالعهٔ موردی، مشتری ما شرکتی است که سازندهٔ بزرگ تجهیزات ارتوپدی در ایالات متحده، با مشتریان پایه‌ای است که مرکب از بیمارستان‌هایی تقریباً در کلیهٔ ۵۰ ایالت‌اند. محصولاتی که در این مطالعه مطرح‌اند از قطعات و تجهیزات ارتوپدی گرفته تا داروهایی را که در فرایند جراحی، توانبخشی، و ریکاوری تجویز می‌شوند شامل می‌شود.

از منظر فروش، شرکت ما باور دارد که بازار مرکب از سه گروه (یا بخش) از بیمارستان‌هاست:

۱. بیمارستان‌هایی که در آن‌ها فروش بالایی داریم و به وضعیت مطلوب بازار رسیده‌ایم.

۲. بیمارستان‌هایی که در آن‌ها فروش متوسطی داریم، اما به‌طور بالقوه هنوز امکان فروش بیشتری داریم.

۳. بیمارستان‌هایی که در آن‌ها حضور اندک یا هیچ داریم و سود بالقوه بالایی در آن‌ها موجود است.

از نظر فعالیتی، بیمارستان‌ها را به‌صورت زیر رده‌بندی کردیم:

۱. بیمارستان‌های عمومی کوچک که در آن‌ها فعالیت‌های ارتوپدی کمی صورت می‌گیرد. در این بیمارستان‌ها، اغلب عمل‌های ارتوپدی به بیمارستان‌های تخصصی‌تر ارجاع داده می‌شوند. این بیمارستان‌ها چندان یا اصلاً مورد توجه ما نیستند. زیرا فروش بالقوه در آن‌ها جزئی است.

۲. بیمارستان‌های عمومی بزرگ که اغلب دارای واحدهای سوانح یا توانبخشی‌اند یا برخی اعمال جراحی ارتوپدی را انجام می‌دهند. این‌ها لزوماً از لحاظ ارتوپدی خیلی تخصصی نیستند اما، به دلیل اندازه‌شان، به برخی تجهیزات برای این نوع عمل‌ها نیازمندند. آن‌ها به‌طور متوسط مورد توجه ما قرار دارند، به‌خصوص آن‌هایی که بزرگترند: به دلیل بزرگی‌شان می‌توانند نقش قابل توجهی در سهم شرکت ما از بازار داشته باشند.

۳. بیمارستان‌های تخصصی که شیوه‌های ارتوپدی زیادی را اجرا می‌کنند و واحدهایی دارند که به سوانح و توانبخشی اختصاص دارند. این بیمارستان‌ها گروه هدف اصلی ما هستند زیرا بخش بزرگی از بازار و رشد بازارند.

هدف این مطالعه، یافتن راه‌هایی برای افزایش فروش محصولات ما به بیمارستان‌ها در گروه‌های مطلوب‌تر است. ما این کار را با شناسایی بیمارستان‌هایی انجام می‌دهیم که مشتریان احتمالی ما هستند، اما به دلایل نامعلوم در حال حاضر کمتر از آنچه انتظار داریم از ما خرید می‌کنند. پس آن‌ها را یک به یک مورد مطالعه قرار می‌دهیم تا دلیل فروش کم را تعیین کنیم و اگر اعتقاد پیدا

جدول ۷.۸ گروه متغیرها

گروه‌ها	متغیرها
پاسخ‌ها	SALESY، SALESY12
جغرافیایی	HID، CITY، STATE،ZIP
عمومی	BEDS، OUT-V، ADM،SIR
مربوط به ارتوپدی	RBEDS، HIP95، KNEE95، TRAUMA،TH
	REHAB، HIP96، KNEE96، FEMUR96

کنیم که رشد بالقوه‌ای در فروش وجود دارد، برای بازاریابی سرمایه‌گذاری می‌کنیم و بنابراین تنها بر روی چند بیمارستان تمرکز خواهیم کرد که فکر می‌کنیم عایدی ما را ماکسیمم می‌کنند.

هدف دیگر این مطالعه ممکن است تعیین این باشد که آیا رده‌بندی بیمارستان‌ها در بالا واقع‌بینانه است یا راه‌های دیگری برای رده‌بندی بیمارستان‌ها موجود است که می‌توانیم از آن‌ها برای مقاصد بازاریابی خود استفاده کنیم.

داده‌ها

مجموعه‌داده‌های «c84.dat» شامل فهرست متغیرهایی است که بیشتر مورد توجه ما قرار دارند و بنابراین زیرمجموعه‌ای معقول از آن را انتخاب کرده‌ایم. هر مورد با بیمارستانی متناظر است و همهٔ بیمارستان‌های ایالات متحده در پایگاه داده‌ها موجودند. جدول ۷.۸ گروه‌هایی از متغیرها را توصیف می‌کند که به هر بیمارستان وابسته‌اند و جدول ۸.۸ تعریفی دقیق از هر متغیر به‌دست می‌دهد.

روش‌شناسی

۲.۴.۸ داده‌کاوی

این مورد، مثال جالبی از مبحثی است که به داده‌کاوی معروف شده است. این مورد را برای یک درس نیمسال دوم در سال ۱۹۹۳، بسیار پیش‌تر از زمانی که داده‌کاوی ابداع شده بود، تهیه کرده بودیم. بنابراین، می‌توانیم بگوییم که ما، پیش از ابداع رسمی داده‌کاوی، داده‌کاوی می‌کرده‌ایم. در این ادعا احتمالاً عدهٔ زیادی شریکند چون داده‌کاوی چیزی است که آماردانان و متخصصان علوم کامپیوتر طی مدتی طولانی به آن پرداخته‌اند. در نتیجه، داده‌کاوی از درون کاربردهای گوناگونی سر بر آورده است، بدون آنکه مخصوص علم آمار باشد. از دیدگاه کاربران، نرم‌افزارهای کاربردی برای داده‌کاوی (مانند Enterprise Miner، Diva، Spotfire) فهرستی از شیوه‌ها (ابزارهای

جدول ۸.۸ تعریف متغیرها

متغیر	تعریف
SALESY	فروش تجهیزات توانبخشی از اول ثانویه
SALES12	فروش تجهیزات توانبخشی برای ۱۲ ماه گذشته
ZIP	کد پستی ایالات متحده
HID	شمارهٔ شناسایی بیمارستان
CITY	نام شهر
STATE	نام ایالت
BEDS	تعداد تخت‌های بیمارستان
DUT-V	تعداد بیماران سرپایی
ADM	هزینه‌های مدیریتی (برحسب ۱۰۰۰ دلار)
SIR	درآمد حاصل از بیماران بستری
RBEDS	تعداد تخت‌های توانبخشی
PRHIP	تعداد عمل‌های جراحی استخوان مفصل ران (کل سال قبل)
PRKNEE	تعداد عمل‌های جراحی زانو (کل سال قبل)
RPKNEE	واحد سوانح، ۰=خیر، ۱=بلی
TRAUMA	واحد توانبخشی دارو، ۰=خیر، ۱=بلی
HIP	تعداد عمل‌های جراحی استخوان مفصل ران (کل سال جاری)
KNEE	تعداد عمل‌های جراحی زانو (کل سال جاری)
FEMUR	تعداد عمل‌های جراحی استخوان ران (کل سال جاری)

دیداری‌سازی داده‌ها، افرازبندی بازگشتی، تحلیل خوشه‌ای، شبکه‌های عصبی، و شاید برخی روش‌های هوش مصنوعی) در اختیار می‌گذارند که برای مجموعه‌داده‌هایی که به‌طور معقولی بزرگ‌اند (100K مشاهده×100 متغیر) به‌طرز مناسبی کارایی دارند. این شیوه‌ها لزوماً چندان نو نیستند اما توانایی‌های نرم‌افزارها برای مجموعه‌داده‌های بزرگ مناسب‌ترند. به همین دلیل، داده‌کاوی به وسیلهٔ متخصصان علوم کامپیوتر و مهندسان به پیش رانده شده است.

تحلیل داده‌ها در گذشته به‌عنوان کاری کارآگاهی جلوه داده می‌شد که هدف از آن یافتن ساختارهای جالب در داده‌ها بود. این شیوه، رویکردی خوب برای بسیاری از مسائل علمی است که در آن چنین ساختاری موجود است که باید یافته شود و نیز نوفه‌ای که چندان زیاد یا خیلی بدرفتار نیست. از سوی دیگر، در بسیاری از مسائل نوین، داده‌های عظیمی موجودند که در آن‌ها اطلاعات اندکی موجود است و تقریباً هیچ ساختاری در آن نیست.

هدف داده‌کاوی شناسایی کلوخه‌ها یعنی خوشه‌های کوچکی از مشاهدات در این داده‌هاست که حاوی اطلاعاتی غیر قابل انتظار و در عین حال به‌طور بالقوه ارزشمندند. تعریف ارزشمند عموماً از طریق مقدار پاسخ بزرگی از یک رستهٔ مشخص از پاسخ‌های کیفی آشکار می‌شود. الک‌کردن حجم بزرگی از داده‌ها که نوفه‌دار و بدرفتار است و ممکن است مقادیر گمشدهٔ فراوانی داشته باشد یا اینکه صرفاً نامربوط به هم باشند، چالش اصلی داده‌کاوی است. چالش دیگر ادغام پایگاه‌های داده‌های متفاوت به این امید است که اطلاعات خرده‌ریز در هر یک از آن‌ها شاید به مدد ترکیب با یکدیگر به شناسایی کلوخه‌ای در کل آن‌ها منجر شود. این کاری است که می‌توان در پایگاه داده‌های نوعی تحقیقات بازار، فروشگاه‌های مواد غذایی، خرده‌فروشی‌ها، اطلاعات مربوط به مکالمات تلفنی، غربالگری با ورودی بالا، با ریزآرایه‌های DNA، و در حوزهٔ آزمایه‌های بالینی یافت. در همهٔ این مسئله‌ها مقادیر بزرگی از داده‌ها موجودند که به‌طور نظام‌مند گردآوری می‌شوند؛ به‌عنوان مثال، همهٔ تماس‌های تلفنی که به‌وسیلهٔ شرکت‌ها انجام می‌شوند، یا همهٔ خریدهایی که در خرده‌فروشی‌ها صورت می‌گیرد. داده‌ها در انبارهای داده‌ها ذخیره می‌شوند و برای تحلیل در دسترس قرار دارند[۱]

ابزارهای استاندارد جستار از پایگاه داده‌ها توسط متخصصان علوم کامپیوتر به‌منظور بازیافت زیرمجموعه‌هایی با ویژگی‌های مشخص به وجود آمده‌اند، و از سوی دیگر، آماردانان ابزارهایی برای یافتن ویژگی‌های جالب داده‌ها ابداع کرده‌اند. آنچه واقعاً در اینجا می‌خواهیم راهی برای تلفیق هر دوی این ایده‌ها با یکدیگر است. چگونه می‌توان جستارهای هوشمندی به وجود آورد؟ چگونه می‌توان زیرمجموعه‌های جالبی را برگزید که در آن معنی «جالب» به‌طور دقیق در جستار پایگاه داده‌ها تعریف شده باشد. این‌ها از نوع سؤال‌هایی هستند که داده‌کاوی سعی در پاسخ به آن‌ها دارد. رویکردی ساختاریافته به داده‌کاوی ممکن است به‌ترتیب شامل موارد زیر باشد:

۱. کاهش بعد (متغیرها).
   - مؤلفه‌های اصلی.
   - تحلیل عاملی.
۲. بخش‌بندی و گزینش داده‌ها.
   - تحلیل خوشه‌ای.

۱. البته در ازای پرداخت وجه! پایگاه‌های داده‌ها به کالاهای روز به روز پرقیمت‌تری بدل شده‌اند که دلیل دیگری برای آن است که چرا داده‌کاوی به‌سرعت رشد یافته است.

- روش‌های درختی.
- شبکه‌های عصبی.

۳. تحلیل داده‌ای بخش‌های جالب.

این مطلب تا حد زیادی دستورکاری برای فروکاهی داده‌هاست که در آن اغلب داده‌های فاقد اطلاع به‌طور نظام‌مند حذف می‌شوند و به شخص اجازه می‌دهد که تنها بر چند بخش جالب تمرکز کند.

می‌توانیم داده‌های مورد کاوش را به‌عنوان آرایه‌ای از موردها برحسب متغیرها تلقی کنیم. متغیرها ممکن است عددی، رسته‌ای، یا حتی ساختارهای داده‌ای ساده‌ای باشند؛ بنابراین، چنین زمینه‌چینی‌یی کاملاً کلی است. می‌توانیم کار را با انتخاب مجموعه‌ای از متغیرها آغاز کنیم که ممکن است در ارتباط با اهداف ما اطلاع‌بخش باشند. معمولاً کاهش بعد بیشتری لازم است. می‌توان به این امر با استفاده از ابزارهای آماری استاندارد برای فروکاهی متغیرها مانند تحلیل مؤلفه‌های اصلی و تحلیل عاملی نائل شد.

فروکاهی موردها را می‌توان با استفاده از تحلیل خوشه‌ای برای افراز مجموعه‌داده‌ها به خوشه‌ها یا بخش‌های همگن به سرانجام رساند. راهی دیگر برای اجرای بخش‌بندی استفاده از یک روش درختی مانند CART یا روش‌های مشابه افرازبندی بازگشتی است (برای توصیفی از این روش‌ها، زیربخش بعدی را ببینید). ایدهٔ بیان‌شده در آنجا کمی متفاوت است زیرا مستلزم منظورکردن متغیرهای پاسخ است. بنابراین، ممکن است مجموعه‌ای از خوشه‌های طبیعی از آن عاید نشود، اما بخش‌بندی کاراتری برحسب پاسخ‌ها حاصل شود. از سویی دیگر، ممکن است مجموعه‌ای از بخش‌های طبیعی در جامعه موجود نباشد و روش درختی پاسخ‌های مرتبط‌تری خواهد داد. در بهترین شرایط، نتایج حاصل از هر دو شیوه مشابه خواهند بود، اما زیاد هم روی آن حساب نکنید.

به محض آنکه زیرمجموعه‌های به‌طور بالقوه جالب شناسایی شوند، می‌توان آن‌ها را مورد بازبینی قرار داد یا آماره‌هایی خلاصه را محاسبه کرد، تحلیل‌های کافی را انجام داد، و قابلیت تفسیر علمی را ارزیابی کرد تا تصمیم بگیریم که کدام بخش‌ها سزاوار تحقیق بیشترند. این کار ممکن است عملی ساده از قبیل محاسبهٔ متوسط فروش هر بخش باشد یا آنکه نیاز به مدلبندی پیچیده‌تری داشته باشیم.

از اینجا به بعد تحلیل زیرمجموعه‌های کوچک و استخراج قطعه‌های نهایی اطلاعات مربوط که به مسئله وابسته‌اند آغاز می‌شوند. در مثالی که اینجا آورده‌ایم، یک راهبرد خوب، اجرای تحلیل

رگرسیونی روی بخش‌ها و استخراج آن بیمارستان‌هایی است که میزان فروش به آن‌ها از میزان مورد انتظار بسیار پایین‌تر است. این کار در مقولهٔ تحلیل داده‌های بخش زیر قرار می‌گیرد.

۳.۴.۸ مروری کوتاه بر تحلیل خوشه‌ای

یک مجموعه‌داده در اختیار داریم و می‌خواهیم داده‌ها را در k گروه طبیعی مجزا گروه‌بندی کنیم. رویکردهای متعددی بر تحلیل خوشه‌ای موجودند و این مطلب در اجراهای نرم‌افزاری تحلیل خوشه‌ای در بسته‌های آماری معمول به چشم می‌خورد که ۱۰ یا ۱۵ روش مختلف را اجرا می‌کنند. گرچه ایده‌های تحلیل خوشه‌ای کاملاً شهودی‌اند، رسمیت‌بخشیدن به ایدهٔ «خوشه» در یک معنای یکتای کلی به‌طرز شگفت‌آوری دشوار است. دو روش متداول برای اجرای تحلیل خوشه‌ای موجودند.

خوشه‌بندی سلسله‌مراتبی این روش نیاز به تعریف فاصله‌های بین‌نقطه‌ای و فاصله‌های بین‌خوشه‌ای دارد. فاصلهٔ بین‌نقطه‌ای معمولاً فاصلهٔ اقلیدسی در نظر گرفته می‌شود:

$$d_E(x_1, x_2) = \sqrt{\sum_{j=1}^{p} (x_{1j} - x_{2j})^2}.$$

گاهی ممکن است فاصلهٔ منهتن[1] را به‌کار ببریم:

$$d_M(x_1, x_2) = \sum_{j=1}^{p} |x_{1j} - x_{2j}|.$$

فاصلهٔ بین‌خوشه‌ای بین دو خوشه به‌عنوان تابعی از فاصله‌های بین‌نقطه‌ای بین جفت‌هایی از نقاط تعریف می‌شود که در آن هر نقطه از خوشه‌ای متفاوت می‌آید. تعریف‌های متداول فاصله‌های بین‌خوشه‌ای عبارت‌اند از:

تک‌اتصالی:	فاصلهٔ بین نزدیکترین دو نقطه
اتصال کامل:	فاصلهٔ بین دورترین دو نقطه
اتصال متوسط:	متوسط فاصله‌های بین هر جفت از نقاط
وارد[2]:	تغییر R^2.

یک درخت سلسله‌مراتبی را با شروع از یک خوشه در هر نقطهٔ نمونه‌ای بنا می‌کنیم و، در هر مرحلهٔ درخت، دو نزدیکترین خوشه را به هم وصل می‌کنیم تا خوشهٔ جدیدی تشکیل شود.

به محض اینکه کار ساختن درخت را تمام کنیم، پرسش بدل به این می‌شود که «چند خوشه را انتخاب کنیم؟». یک راه برای تعیین این تعداد، وارسی درخت سلسله‌مراتبی و یافتن نقطه‌ای

1. Manhattan 2. Ward

قابل قبول برای تفکیک خوشه‌هاست. همچنین می‌توانیم تابع ملاک را برای تعداد متفاوت خوشه‌ها رسم کنیم و جهش‌های به‌طور نامعمول بزرگ را بیابیم.

روش‌های گرانیگاهی الگوریتم Kـمیانگین .

K بذردانه[1] انتخاب و داده‌ها بین K خوشه توزیع می‌شوند. سپس خوشه‌ها به‌آرامی و با استفاده از برخی معیارها مانند R^2 بهینه‌سازی می‌شوند. در هر مرحلهٔ الگوریتم، یک نقطه به خوشه‌ای منتقل می‌شود که تابع ملاک را بهینه خواهد کرد. این کار تا زمانی تکرار می‌شود که همگرایی رخ دهد. پیکربندی نهایی تا حدی به پیکربندی آغازین وابسته است، بنابراین مهم است که آغازی خوب صورت گیرد. یک امکان این است که روش واژد اجرا و از نتیجهٔ آن به عنوان پیکربندی آغازین برای kـمیانگین استفاده شود.

۴.۴.۸ افرازبندی و رده‌بندی بازگشتی و درخت‌های رگرسیونی

افرازبندی بازگشتی و روش‌های مبتنی بر درخت مانند CART[2]راه‌های دیگری برای اجرای بخش‌بندی یک مجموعه‌داده‌های چندمتغیره هستند که شامل یک متغیر پاسخ و چندین پیشگوست. متغیر پاسخ و نیز پیشگوها ممکن است رسته‌ای یا عددی باشند.

گرچه این روش‌ها از این نظرکه درختی را با بهینه‌سازی برخی معیارها روی همهٔ جداسازی‌های ممکن گسترش می‌دهند بسیار مشابه یکدیگر هستند، تفاوت‌هایی نیز موجود است. روش‌های افرازبندی بازگشتی از یک قاعدهٔ توقف مبتنی بر Pـمقدار آمارهٔ آزمونی مانند χ^2 یا F استفاده می‌کنند. روش‌های درختی، درخت را آن‌قدر گسترش می‌دهند که اندازهٔ گره‌ها از عددی معین کوچکتر شوند و سپس به هرس‌کردن آن اقدام می‌کنند. هرس کردن برای اجتناب از بیش برازش ضروری است.

۵.۴.۸ تحلیل مقدماتی

ما مطالعهٔ خود را به زیرمجموعه‌ای از ایالت‌های مربوط به حدود ۳۰۰ الی ۵۰۰ بیمارستان متمرکز می‌کنیم و آن‌هایی را پیدا می‌کنیم که مصرف بالایی از چنان تجهیزاتی دارند اما در آن‌ها فروش ما پایین است. هدف آن است که به گروهی برگزیده برسیم که فکر می‌کنیم در آن‌ها تلاش‌های ما مثمر ثمر خواهند بود.

نام‌های متغیرها در جدول ۸.۸ داده شده‌اند. در شکل ۴.۸ شبکه‌ای از یک طرحواره را برای

1. Seed 2. Classification And Regression Trees



شکل ۴.۸ نموداری برای تحلیل داده‌کاوی.

تحلیل این داده‌ها ارائه می‌کنیم. نقطهٔ تمرکز این نمودار بر مراحل داده‌کاوی و فروکاهی داده‌هاست و نه بر تحلیل آماری استاندارد. کد SAS برای اجرای این تحلیل مقدماتی را آورده‌ایم.

بخش ۱. اجزای بازار خود را انتخاب کنید.

فروکاهی داده‌ها. زیرمجموعهٔ حالت‌های مربوط را انتخاب کنید. ما ایالت یا گروهی از ایالت‌ها را برای بررسی انتخاب می‌کنیم، اما مایلیم که حداقل ۳۰۰ بیمارستان را در ایالت یا گروه ایالت‌ها داشته باشیم.

نمودار هر متغیر تنها را رسم کنید و تصمیم بگیرید که آیا تبدیلی مناسب است یا خیر. متغیرها را به دو گروه تفکیک کنید: پاسخ‌ها و آمارهای جمعیت‌شناختی.
متغیرهای آمارهای جمعیت‌شناختی برای تقسیم فهرست بیمارستان‌ها (همهٔ مشتریان ممکن = بازار) به زیرمجموعه‌هایی که ما بخش‌های بازار خواهیم نامید به‌کار می‌روند.

فروکاهی متغیرها. اگر متغیرهای آمارهای جمعیت‌شناختی زیادی داشته باشید، یک تحلیل مؤلفه‌های اصلی یا تحلیل عاملی را به‌کار برید تا

متغیرهای آمارهای جمعیت‌شناختی را به چند مؤلفه یا عامل تخصیص دهید.

بخش‌بندی بازار: از تحلیل خوشه‌ای یا روش‌های درختی برای یافتن بخش‌های بازار یا خوشه‌ها استفاده کنید.

فروکاهی داده‌ها: روی خوشه‌های جالب تمرکز کنید. به محض اینکه خوشه‌ها انتخاب شدند، آماره‌های خلاصه را برای هر خوشه بررسی و سعی کنید که مضمون آن‌ها را توصیف کنید. تفسیر در این مرحله خیلی مهم است.

زیرمجموعهٔ نهایی: مجموعه‌داده‌های کوچک. خوشه(ها)یی را انتخاب کنید که با هدف‌ها توافق داشته باشند. در این مطالعه به دنبال بخش‌هایی (بیمارستان‌هایی) با فروش همه‌جانبهٔ بالا هستیم که در آن‌ها فروش شرکت پایین است. این کار بیمارستان‌هایی را مشخص خواهد کرد که تعداد کمی بیمار دارند اما به تجهیزات ما نیاز دارند و بنابراین برای مقاصد بازاریابی خود به این بخش‌ها علاقه‌ای نداریم.

بخش ۲. برآورد فروش‌های بالقوه

برای پایان‌دادن به تحلیل، یک تحلیل رگرسیونی برای هر یک از بخش‌بندی‌های انتخاب‌شده انجام می‌دهیم. توجه کنید که چون بخش‌بندی‌ها بسیار همگن‌اند، ممکن است انتظار مقادیر کوچک R^2 را داشته باشیم؛ بنابراین به مقادیر کوچک R^2 توجه نکنید.

بیمارستان‌هایی با مانده‌های منفی بزرگ، آن‌هایی‌اند که فروش‌های پایینی دارند ولی از ویژگی‌های آن‌ها چنین برداشت می‌شود که پایین‌تر از فروش بالقوهٔ خود قرار دارند (مقادیر پیشگویی‌شده را به‌عنوان فروش‌های بالقوه به‌کار ببرید). برای به پایان رساندن کار، فهرستی از بیمارستان‌های موجود در بخش‌بندی را که می‌توان فروش را در آن‌ها بهبود بخشید تهیه می‌کنیم و برآوردی از عایدی‌های بالقوه ارائه می‌کنیم.

این مدل تحلیل، تنها یکی از راه‌های گوناگون تحلیل این داده‌هاست. به آن به‌عنوان فهرستی از پیشنهادها فکر کنید که می‌توان برای این یا آن مجموعه‌داده‌ها به‌کار برد. برای سهولت، در اینجا از مجموعه‌داده‌های خیلی بزرگی استفاده نکرده‌ایم اما گام‌های بالا به همان نسبت که اندازهٔ مجموعه‌داده‌ها بزرگتر می‌شوند مهم و مهم‌تر می‌شوند.

کد SAS زیر را می‌توان برای اجرای محاسبات توصیف‌شده در بالا به‌کار برد. گام تحلیل عاملی حذف شد و تحلیل خوشه‌ایِ اجراشده به روش وارد بود.

برنامهٔ SAS برای تحلیل فروش تجهیزات ارتوپدی

```
options ps=55 ls=78;

data ortho;
infile 'c84.dat' delimiter=',';
input zip $ hid $ city $ state $ beds rbeds outv adm sir
      salesy sales12 hip95 knee95 th trauma rehab hip96
      knee96 femur96;

/* 1. Select subset of interest */
if state eq 'fl' or state eq 'ga';
array x {12} beds rbeds prhip prknee hip knee femur
             outv adm sir salesy sales12;

/* 2. Transform the variables */
do i=1 to 7 ;
   x{i} = sqrt(x{i}) ;
end;
do i=8 to 12 ;
   x{i} = log(1+x{i}) ;
end;
run;
/* 3. Factor analysis step is omitted */

/* 4. Data segmentation using cluster analysis */
proc cluster method=ward;
   var beds rbeds outv adm sir prhip knee th trauma
       rehab hip knee femur;
   copy beds rbeds outv adm sir prhip prknee th trauma
       rehab hip knee femur sales12 salesy ;

/* Select 8 clusters or market segments. Assign the
   cluster variable to the cases */
proc tree noprint ncl=8 out=txclust;
   copy beds rbeds outv adm sir prhip prknee th trauma
        rehab hip knee femur sales12 salesy ;
run;

/* 5. Data reduction. Examine the composition of the
      clusters and pick the most interesting ones. */
```

```
proc sort data=txclust; by cluster;
proc means data=txclust;
   var beds rbeds outv adm sir prhip prknee th trauma
       rehab hip knee femur sales12 salesy ;
   by cluster;
   output out = c
   mean = mbeds mrbeds moutv madm msir mprhip mprknee
          mth mtrauma mrehab mhip mknee mfemur
          msales12 msalesy ;
proc print;
   var cluster mbeds mrbeds moutv madm msir mprhip
       mprknee mth mtrauma mrehab mhip mknee mfemur
       msales12 msalesy ;
run;

/* 6. Final subset. Do the standard analysis of the small
      sample */
```

در مثال داده‌کاوی بالا، می‌توانیم خوشه‌ها را به‌دقت امتحان کنیم زیرا تنها تعداد متوسطی از مشاهدات را داریم. با این حال، در مسائل داده‌کاوی با مجموعه‌داده‌های بزرگ، شاید امتحان‌کردن اغلب خوشه‌ها به‌طور دقیق ممکن نباشد و تنها می‌توان تعدادی اندک را که جالب به نظر می‌رسند به‌دقت امتحان کرد. بنابراین، شاید هدف داده‌کاوی تحلیل «همهٔ» داده‌ها نباشد، بلکه تمرکز بر زیرمجموعه‌های جالب باشد.

۶.۴.۸ خلاصه

تحلیل بالا مخصوص پرسش‌هایی است که در این مطالعه ارائه شد، اما می‌توان از آن به‌عنوان بخش اساسی بسیاری از مسائل داده‌کاوی استفاده کرد. پرسش‌های زیر شاید خاص مسئلهٔ حاضر باشند، اما می‌توان آن‌ها را با جرح و تعدیل در وضعیت‌های بسیار دیگری به‌کار برد.

پرسش‌ها

۱. متغیرهایی را پیدا کنید که لازم است تبدیل شوند و تبدیل‌های متناظر را پیدا کنید.

۲. امکان استفاده از مؤلفه‌های اصلی را برای تلخیص متغیرها کاوش کنید. آیا تعبیری از چند مؤلفهٔ اصلی اول موجود است؟

۳. نمودار مؤلفه‌های اصلی عمده را رسم کنید. آیا خوشه‌های قابل رؤیتی وجود دارند؟

۴. یک تحلیل خوشه‌ای با بهترین مؤلفه‌های اصلی و با متغیرهای خام اجرا کنید. این تحلیل‌ها را با هم مقایسه و تعبیر کنید. آیا لازم است که پاسخ را در تحلیل خوشه‌ای منظور کنید؟

۵. به محض اینکه افرازی رضایت‌بخش در دست داشته باشید، چند خوشهٔ جالب را انتخاب و داده‌ها را با استفاده از تحلیل رگرسیونی یا هر روش آماری دیگری تحلیل کنید. به خاطر داشته باشید که هدف در اینجا یافتن بیمارستان‌هایی است که از ویژگی‌های آن‌ها فروش‌های بالایی برمی‌آید اما فروش‌های مشاهده‌شده پایین‌اند.

۹

چند مطالعهٔ موردی دیگر

حالا نوبت شماست! شما را آن یکی مشتری به‌عنوان مشاور ی به‌شدت توصیه کرده است. پس چند پروژه‌ای را که ما روی آن کار می‌کنیم، ملاحظه کنید... می‌شود هفتهٔ بعد برای گرفتن نتایج به شما مراجعه کرد؟

۱.۹ بهبود آموزش

۲.۹ نمونه‌گیری تصادفی؟

۳.۹ چپ یا راست؟

۴.۹ برانگیختن حس در اسب‌ها

۵.۹ سرخ‌موی بلندقد

۶.۹ انتقام بنتلی

۷.۹ هر چه می‌خواهید می‌پوشید؟

۸.۹ مطالعه‌ای در مورد ایدز

۱.۹ بهبود آموزش

توصیف

هدف این مطالعه، تحقیق دراین‌باره بود که آیا گنجاندن فعالیت‌هایی متناسب با ترجیح‌های سبک یادگیری دانش‌آموزان شواهدی مبنی بر بهبود «عملکرد» در مقایسه با روش‌های آموزش سنتی (کتاب درسی) فراهم می‌آورد یا خیر. ابزار سیاههٔ آموزشی (LSI)[۱] معیاری برای سنجش قدرت فرد نسبت به توانایی‌های حس لامسه (ساختن با دست)، جنبشی (ایفای نقش)، حس شنوایی و بینایی آن‌هاست. فردی که نمرهٔ ۶۰ یا بیشتر در یکی از این مقوله‌ها دریافت می‌کند دارای ترجیحی قوی برای یک سبک یادگیری، که توانایی ادراکی خاصی از آن‌ها را مورد توجه قرار می‌دهد، تلقی می‌شود. از یک دیدگاه آموزشی، از این امر چنین برداشت می‌شود که اگر ترجیح سبک یادگیری دانش‌آموزان در روند آموزشی در نظر گرفته شود، آن‌ها اطلاعات جدید را بسیار آسان‌تر جذب و نگهداری خواهند کرد.

طرح

جمعاً به ۵۹ دانش‌آموز ابتدایی (۹ تا ۱۰ ساله) در سه کلاس [CLASS] به مدت سه هفته سه واحد [UNIT] پیاپی علوم تدریس شد. پیش از آغاز آزمایش، ترجیح‌های سبک یادگیری دانش‌آموزان با استفاده از ابزار LSI ارزشیابی شد. برای جلوگیری از اریبی، طی آزمایش در مورد ترجیح‌های سبک یادگیری دانش‌آموزان به آن‌ها اطلاعی داده نشد.

برای ثبت دانش پیشین دانش‌آموزان، پیش‌آزمونی [PRE] مرکب از پرسش‌های چند گزینه‌ای در شروع هر واحد به آن‌ها داده شد. پس از پایان هر واحد، دانش‌آموزان، به کمک یک پس‌آزمون [PST] چندگزینه‌ای، یک مقیاس افتراقی معنایی[۲] (SDS) یا آزمون «نحوهٔ نگرش» (دانش‌آموزان نسبت به یادگیری علوم چه احساسی داشتند)، و یک آزمون یادگیری عالی[۳] [HLT] که مستلزم ارائهٔ راه‌حلی برای یک مسئله بود ارزشیابی شدند. حداکثر نمره برای PRE و PST ۱۰۰٪ بود. برای SDS، نمرهٔ ۶۰ نشان‌دهندهٔ بالاترین نمرهٔ ممکن برای نحوهٔ نگرش (مثبت) بود. آزمون HLT با مقیاس ۱، ۲، ۳، ۴ نمره داده می‌شد که ۴ نشان‌دهندهٔ راه‌حل کامل و درست بود و ۱، ۲، ۳ به راه‌حل‌های جزئاً درست، مطابق با میزان درستی آن، اختصاص داشت.

۱. ابزار LSI (Learning Style Inventory) در فصل چهار نیز مورد بحث قرار گرفته است.

2. semantic differential scale(SDS) 3. higher-learning test

کلاس دیگری به‌عنوان گروه کنترل (GROUP=C) برای هر واحد انتخاب شد. به این کلاس تنها با استفاده از روش‌های سنتی تدریس شد. دو کلاس دیگر به‌عنوان گروه‌های آزمایشی (GROUP=E) تعیین شدند که در آن‌ها به تدریس سنتی، فعالیت‌های لمسی و جنبشی برای دانش‌آموزان افزوده شده بود. فعالیت‌های شنیداری و دیداری در این آزمایش مد نظر نبودند. تخصیص‌های گروه کنترل به‌صورت زیر بودند.

```
UNIT = 1 : C = CLASS 3
UNIT = 2 : C = CLASS 1
UNIT = 3 : C = CLASS 2
```

داده‌ها

نتایج این آزمایش در مجموعه‌داده‌های c91.dat گنجانده شده‌اند که دارای فرمت زیر است. متغیرها در جدول ۱.۹ تعریف شده‌اند.

			UNIT=1				UNIT=2				UNIT=3				PREF			
ID	CLASS	SEX	P1	F1	S1	H1	P2	F2	S2	H2	P3	F3	S3	H3	T	K	A	V
s01	2	M	70	100	60	4	48	88	52	3	72	64	60	2	61	44	57	55
s02	2	F	80	85	60	2	50	76	60	3	70	72	60	1	35	44	57	34
:				:				:				:				:		
s59	3	F	78	78	58	3	18	70	60	3	26	90	60	4	61	51	48	50

تکلیف

هدف از این پروژه ارزشیابی معنی‌داری آماری چهار فرض زیر بود که در فرم راهنما بیان شده بودند.

H_1 اختلافی بین PST برحسب GROUP وجود خواهد داشت.

H_2 اختلافی بین HLT برحسب GROUP وجود خواهد داشت.

H_3 اختلافی بین SDS برحسب GROUP وجود خواهد داشت.

H_4 اثر متقابلی بین PST و PREF برحسب GROUP وجود خواهد داشت.

۲.۹ نمونه‌گیری تصادفی؟

توصیف

طرح‌های نمونه‌گیری عموماً از طرف نهادهای حکومتی ایالتی و فدرال برای نظارت بر ادعاهای جبران زیان بکار گرفته می‌شوند که از طرف شرکت‌هایی که خدماتی برای برنامه‌های یارانه‌ای دولت، نظیر مراقبت‌های بهداشتی و کمک‌های اسکان به دریافت‌کنندگان واجد

جدول ۱.۹ تعریف‌های متغیر

متغیر	توصیف
ID	کد شناسایی تخصیص‌داده‌شده به دانش‌آموز
CLASS	گروه کلاسی: ۱، ۲، ۳ $(n_1 = 21, n_2 = n_3 = 19)$
UNIT	تدریس واحد درسی: ۱، ۲، ۳
GROUP	C = کنترل: روش آموزش سنتی E = آزمایشی: فعالیت‌های T و K را در بر دارد.

متغیرهای ارزشیابی — کدگذاری ستونی

PRE	پیش‌آزمون (برای هر UNIT)	P1	P2	P3
PST	پس‌آزمون	F1	F2	F3
SDS	نمرهٔ نگرش	S1	S2	S3
HLT	آزمون یادگیریِ عالی	H1	H2	H3

ترجیح‌های سبک آموزشی

PREF	T = لمسی	A = شنیداری
	K = جنبشی	V = دیداری

شرایط با هزینه‌های کاهش‌یافته، ارائه می‌کنند به دولت تسلیم می‌شود. در این تمرین مطالعهٔ موردی، یک شرکت نتایج بررسی یک نهاد دولتی را از ادعاهای او بر این مبنا مورد مناقشه قرار داده است که روش نمونه‌گیری به‌کار رفته از طرف این نهاد نامناسب بوده است.

بازبینی

گزارش ارائه‌شده به‌وسیلهٔ نهاد مذکور حاکی از این است که یک نمونهٔ تصادفی آماری از ۴۰۰۰۰ ادعای نوع I و نوع IIای تسلیم‌شده به‌وسیلهٔ شرکت اینکس[1] در سال پیش منجر به انتخاب $n_1 = 400$ ادعای نوع I و $n_2 = 350$ ادعای نوع II برای بازبینی شد. جبران خسارت برای ادعاهای نوع II عموماً بزرگتر از ادعاهای نوع I هستند. بر مبنای بازبینی نهاد از این ادعاهای نمونه‌ای، ۵۰٪ ادعاهای نوع I و ۳۰٪ ادعاهای نوع II غیرواجد شرایط برای جبران خسارت تحت شرایط مندرج در برنامهٔ یارانه‌ای یاپریون[2] تشخیص داده شدند. چون این ادعاها به‌طور اتوماتیک

1. Inexs 2. Yapereven

پرداخت می‌شدند، این نهاد، به جای اینکه یک رسیدگی کامل از هر عمل مالی را انجام دهد، برآوردی از بیش‌پرداخت بر مبنای نتایج نمونه‌ای را معین کرد.

فرجام‌خواهی

شرکت اینکس تلاش کرد که شیوهٔ نمونه‌گیری آماری را، که از قرار معلوم به وسیلهٔ نهاد مزبور به‌کار رفته بود، مانندسازی کند، و نتایج کاملاً متفاوتی به‌دست آورد. برای حذف تغییرپذیری ذاتی وابسته به شیوه‌های نمونه‌گیری، کل پایگاه داده‌های ادعاها مورد تحلیل قرار گرفت تا یک نمایهٔ واقعی از «جامعه»ی دریافت‌کنندگان نوع I و نوع II، به صورتی که توسط نهاد ناظر بر برنامهٔ پایریون تعریف شده بود، به‌دست آید.

نمایهٔ پایگاه داده‌ها

۴۰۰۰۰	تعداد کل ادعاها	
۸۰۰۰	تعداد کل دریافت‌کننده‌ها	
نسبت	کل	نوع
۰٫۸۷۵	۳۵۰۰۰	I
۰٫۱۲۵	۵۰۰۰	II

تکلیف

۱. یک برآورد محافظه‌کارانه برای احتمال به‌دست‌آوردن یک نمونهٔ تصادفی آماری مرکب از ۴۰۰ ادعای نوع I، به‌طوری که در بازبینی نهاد گزارش شده است، محاسبه کنید.

۲. وقتی شرکت اینکس با این نتایج نهاد روبه‌رو شد، نهاد اظهار داشت که یک نمونهٔ تصادفی طبقه‌بندی‌شده را برای برآوردکردن نسبت کلی ادعاهای واجد شرایط به کار گرفته است. آیا نتایج گزارش‌شده به‌وسیلهٔ نهاد با روش نمونه‌گیری سازگاری دارد؟ آیا این امر اثری بر برآورد بیش‌پرداخت دارد؟

۳.۹ چپ یا راست؟

توصیف

داده‌ها در این تمرین مطالعهٔ موردی طی چندین سال گردآوری شده‌اند و مرکب‌اند از اندازه‌گیری‌های مختلفی از «تیغه»هایی که به‌عنوان ابزارهای برش به‌وسیلهٔ انسان‌های اولیه در عصر چهارم زمین‌شناسی[1] به‌کار می‌رفتند. این دست‌ساخته‌ها از مکان‌های زیادی گردآوری

۱. یعنی، تیغه‌هایی که ۲ میلیون سال قبل به‌کار می‌رفتند.

شدند و یکی از هدف‌های اصلی این پروژه تلاش برای تعیین برخی مشخصه‌های مشترک بین مشاهدات بود. علاوه بر آن، از مطالعات قبلی چنین برآمده بود که برخی الگوهای تیغه‌ای با «دست‌ساخت‌بودن» سازگاری دارند و بنابراین انسان‌های اولیه شاید توانایی‌های استدلال شناختی را داشته‌اند. این نتیجه‌گیری البته محل مناقشه است، اما به این پرسش جالب منجر می‌شود که آیا ترجیح دست از سوابق باستان‌شناسی قابل تشخیص است یا خیر.

داده‌ها

اندازه‌گیری‌های تیغه‌ها از چندین مکان در مجموعه‌داده‌های c93.dat درج شده‌اند. اغلب متغیرهای این مجموعه‌داده گسسته‌اند و متغیر اصلی برای تشخیص ترجیح دست FT (نوع تیغه[1]) است. به‌ویژه، سطوح FT که با ۲٫۲ و ۲٫۵ کدبندی شده‌اند، به‌عنوان ویژگی‌های مرتبط با تیغه‌های تراش‌داده‌شده برای راست‌دستان تلقی می‌شوند و سطوح ۲٫۳ و ۲٫۵ با تیغه‌های تراش‌داده‌شده برای چپ‌دستان مرتبط‌اند. فهرستی کامل از متغیرها در جدول ۲.۹ داده شده است. این متغیرها به همان ترتیبی ارائه شده‌اند که ستون‌وار در مجموعه‌داده‌های c93.dat ظاهر می‌شوند.

تکلیف

آیا داده‌ها شواهدی از ترجیح دست را نشان می‌دهند؟

۴.۹ برانگیختن حس در اسب‌ها

توصیف

آزمایشی بر روی هفت اسب، برای ارزیابی اثر دو تیمار بر سطوح گلوکز و انسولین در خون اسب‌ها به مرور زمان، انجام شد. هر دو متغیر پاسخ در هر یک از ۱۲ لحظهٔ زمانی که در بازه‌های هم‌فاصله قرار داشتند اندازه‌گیری شدند. یک طرح متقاطع به‌کار گرفته شد که با شمای زیر مطابقت داشت.

	اسب						
آزمایه	h1	h2	h3	h4	h5	h6	h7
1	a	a	a	a	b	b	b
2	b	b	b	b	a	a	a

تیمار= a b

پاسخ=گلوکز، انسولین

زمان=۱، ۲،...، ۱۲

1. flake type

جدول ۲.۹ تعریف‌های متغیر وکد

FN	مشخص‌کنندهٔ مکان	
ID	مشخص‌کنندهٔ دست‌ساخته	
L، W، T	طول، عرض، کلفتی (میلی‌متر)	
MD	ماکسیمم ابعاد: max (L، W، T)	
FS	شکل تیغه	۱، ۲، ...، ۷
FT	شکل تیغه	۲ر۲، ۲ر۵ (راست) ۳ر۲، ۳ر۵ (چپ) ۱، ۱ر۲، ...، ۹ (سایرین)
PC	مشخصهٔ طرح	۱، ۲، ...، ۶
PA ،PW ،PL	طول، پهنا، زاویهٔ طرح (میلی‌متر یا درجه)	
TQ	فن	۱، ۲
RM	مادهٔ خام	۱، ۲، ...، ۹۹
FA	تیغه یا قطعهٔ گوشه‌دار	۱، ۲، ۳
CN	وضعیت تیغه	۱، ۲، ...، ۶
H	لولا	۱=بلی، ۰=خیر
X	پوسته	۱=بلی، ۰=خیر
R	لبه	۱=بلی، ۰=خیر
E	زدودگی لولا	۱=بلی، ۰=خیر
B	خردشدگی خلفی	۱=بلی، ۰=خیر
A	دگرگونی گرمایی	۱=بلی، ۰=خیر

تکلیف

هدف این تمرین مطالعهٔ موردی ایجاد یک برنامهٔ SAS یا S-PLUS بود که تحلیل‌های توصیف‌شده در زیر را انجام دهد. چون داده‌ها در دسترس نیستند، لازم است که یک مجموعه‌داده‌های ظاهری برای مقاصد آزمون ایجاد کنیم.

۱. نخستین گام ایجاد یک متغیر جدید «seq» است که نشان‌دهندهٔ دنبالهٔ `treatment + trial` است:

```
seq = 1 if (trial = 1 & treatment = a)
seq = 2 if (trial = 1 & treatment = b)
seq = 2 if (trial = 2 & treatment = a)
seq = 1 if (trial = 2 & treatment = b)
```

۲. آماره‌های خلاصهٔ زیر را در ارتباط با هر خم پاسخ محاسبه و شیوهٔ تحلیل واریانس مناسب را برای آزمون‌کردن یک اثر «تیمار» معنی‌دار به‌کار ببرید. توضیح دهید که چگونه باید خروجی ANOVA را تفسیر کرد و در صورتی که اثر seq یا trial معنی‌دار باشد نتایج چه معنایی دارند.

- سطح زیر منحنی
- مقدار قلهٔ ماکسیمم منحنی
- زمان رسیدن به قلهٔ ماکسیمم.

۳. «زمان» را در تحلیل دخالت دهید. توجه کنید که می‌توان با زمان به‌صورت یک عامل یا یک متغیر کمکی رفتار کرد.

۵.۹ سرخ‌موی بلندقد

یک مورد قضایی تبعیض

بنا به درخواست رئیس گروه صنعتی اره‌وان[1]، از ما تقاضا شد که تحلیل ارائه‌شده در یک گزارش تهیه‌شده از طرف خواهان در رابطه با مورد قضایی زیر را بازبینی کنیم.

تحلیل رابطهٔ آماری بین تصمیم به موافقت با افزایش حقوق و رنگ مو یا قد

در قضیهٔ

الیاس آفن[2] بر علیه گروه صنعتی اره‌وان

خلاصهٔ گزارش

در این گزارش، سه عامل در خصوص کارمندانی در اره‌وان، که متقاضی افزایش حقوق بودند، مورد تحلیل قرار گرفت. این عامل‌ها عبارت بودند از:

افزایش حقوق [RAISE] موافقت شد یا نشد

1. Erewhon 2. Ilias Offen

رنگ [COLOR] سرخ‌مو یا غیر سرخ‌مو
قد [HEIGHT] قد کارمند (کوتاه یا بلند)

قد [HEIGHT] های کارمندان در صورتی که قد واقعی آن‌ها کمتر از ۱۸۰ سانتی‌متر بود، به عنوان «کوتاه‌قد» رسته‌بندی شد؛ در غیر این صورت، به‌عنوان «بلندقد» رسته‌بندی شد. تحلیل آماری عامل‌ها به دو بخش تفکیک شد و در هر حالت، آزمون دقیق فیشر به‌کار گرفته شد.

RAISE (A) در برابر COLOR

RAISE (B) در برابر HEIGHT.

نتایج

جدول زیر در گزارش، نتایج به‌دست آمده از تحلیل آماری بالا را تلخیص می‌کند. برای هر تحلیل، یک نتیجهٔ از لحاظ آماری معنی‌دار به‌دست آمد.

(A) درصد و تعدادی که برحسب COLOR افزایش حقوق نگرفتند.

واحد SD	$P-$ مقدار	سرخ‌مو	غیر سرخ‌مو
۲٫۰۸	۰٫۰۳۸	٪۶۶٫۷ (۴ از ۶)	٪۱۸٫۲ (۴ از ۲۲)

(B) درصد و تعدادی که برحسب HEIGHT افزایش حقوق نگرفتند.

واحد SD	$P-$ مقدار	بلند قد	کوتاه قد
۲٫۱۸	۰٫۰۲۹	٪۴۲٫۱ (۸ از ۱۹)	٪۰٫۰ (۰ از ۹)

نتیجه‌گیری گزارش

بر مبنای نتایج بالا، نویسندهٔ گزارش اظهار می‌دارد که:

«... داده‌هایی که به من ارائه شد یک احتمال معنی‌دار آماری برای این امر نشان می‌دهد که رفتار با سرخ‌موها با غیر سرخ‌موها و رفتار با کارمندان بلندقد با رفتار با کارمندان کوتاه‌قد در تصمیم به موافقت با افزایش حقوق توسط رئیس گروه صنعتی اره‌وان متفاوت بوده است.»

تکلیف

بازبینی شما باید به شکل گزارشی تفصیلی نوشته شود: «پاسخ به تحلیل ارائه‌شده در گزارش که دغدغه‌هایی را که باید! در رابطه با تحلیل و نتیجه‌گیری‌های ارائه‌شده در گزارش خواهان داشته باشید مورد توجه قرار می‌دهد. گرچه این وضعیت کاملاً ساختگی است، مطمئناً عامل‌های «واقعی» زیادی وجود دارند که می‌توان آن‌ها را با تصمیم به موافقت با افزایش حقوق به کارمندان مرتبط دانست. این عامل‌ها نیز باید در بازبینی شما در نظر گرفته شوند.

۶.۹ انتقام بنتلی

ادامهٔ مطالعهٔ موردی ۳.۶ ...

در مطالعهٔ موردی ۳.۶، ادعای تولیدگری را بررسی کردیم مبنی بر اینکه نصب‌کردن دستگاه او بر روی موتور به‌طور مستمر موجب کاهش در انتشار هیدروکربن و مونوکسید و افزایش خروجی دیوکسیدکربن خواهد شد. در این تمرین مطالعهٔ موردی، مجموعهٔ دومی از آزمایش‌ها را ارائه می‌کنیم که با استفاده از یکی از خودروهای بنتلی که در اختیار شرکت استتکان قرار داشت صورت گرفت.

آزمایش بنتلی

- خودرو تحت چهار وضعیت جاده‌ای مختلف در هر یک از ۱۴ روز متفاوت تحت آزمایش قرار گرفت. وضعیت‌های جاده‌ای عبارت بودند از:
  - وضعیت ۱: بار متوسط موتور در سرعت ۵۵ مایل در ساعت.
  - وضعیت ۲: بار سنگین موتور در سرعت ۳۵ مایل در ساعت.
  - وضعیت ۳: بار سبک موتور در سرعت ۵۵ مایل در ساعت.
  - وضعیت ۴: بالارفتن از تپه در سرعت ۱۴۵ مایل در هر ساعت.
- در هر یک از روزهای آزمون، شرایط آزمایشی سه بار تکرار شدند.
- دستگاه بعد از ۶ روز آزمایش بر روی موتور قرار گرفت.
- آلاینده‌های هیدروکربن (ppm)، دیوکسیدکربن، و مونوکسیدکربن (٪ حجم) با استفاده از ابزاری اندازه‌گیری شد که پس از اندازه‌گیری‌های روزانه مجدداً تنظیم می‌شد.

داده‌ها

برخی اندازه‌گیری‌ها، که به دلیل در دست ساختمان بودن جاده انجام نشده بودند، در مجموعه‌داده‌های c96.dat، که دربرگیرندهٔ نتایج آزمایش است، با "A" نشان داده می‌شوند. دراید‌های هر ستون با عنوان ستون‌ها مشخص شده‌اند. روزهای آزمایش پشت سر هم نبودند و ماشین تقریباً ۶۰۰ تا ۴۰۰۰ مایل بین آزمون‌ها راه برده شد. آزمون‌ها در ماه دسامبر شروع شدند و در ماه ژوئیهٔ سال بعد به اتمام رسیدند. همچنین باید توجه کنیم که کمیت آلاینده‌های اندازه‌گیری‌شده، در اغلب موارد، از دقت ابزار اندازه‌گیری به‌کار رفته کمتر بود.

تکلیف

۱. بر مبنای داده‌های گردآوری‌شده از آزمون‌ها، با استفاده از خودروی تولیدگر و تحلیل شما از آزمایش بالا، آیا شواهدی دال بر تأیید ادعای تولیدگر در دست هست؟

۲. تغییرپذیری در اندازه‌گیری‌های آلاینده‌ها ممکن است به همان نسبت، معلول چندین عامل دیگر باشد که در این آزمایش‌ها کنترل نشده‌اند. شما چطور آزمایش را طراحی می‌کردید؟

۷.۹ هر چه می‌خواهید می‌پوشید؟

توصیف

تمرکز اصلی این مطالعه عبارت از امتحان‌کردن این مطلب بود که آیا تعامل بین دانش‌آموزان و معلم، بسته به نوع لباسی که دانش‌آموز پوشیده بود، متفاوت بود؟ آزمایش مشاهده‌ای، مشتمل بر این بود که مشتری چندین کلاس را که به‌وسیلهٔ یک معلم درس داده می‌شد مورد مشاهده قرار دهد و تعداد تعامل‌هایی را که هر دانش‌آموز با این معلم دارد بشمارد. یک تعامل بسته به نحوهٔ پاسخ یا واکنش معلم به شاگرد، از سوی مشتری «مثبت» یا «منفی» رده‌بندی می‌شد. تعداد تعامل‌ها، بنابر جنسیت یا نوع لباس دانش‌آموز، متقاطع‌رده‌بندی[۱] شده بود.

داده‌ها

مجموعه‌داده‌های c97.dat نتایج این آزمایش مشاهده‌ای را دربردارد. در c97.dat، برای هر دانش‌آموز دو سطر در نظر گرفته شد و نخست تعداد عامل‌های مثبت فهرست‌بندی شد. مقادیر گمشده با یک نقطه "." نشان داده شدند که ممکن است به‌عنوان ۰ تعبیر شود. سطوح هر متغیر در جدول زیر، به ترتیبی که در مجموعه‌داده‌های c97.dat ظاهر شده‌اند، تعریف شده‌اند.

۱. crossclassification؛ رده‌بندی بر اساس بیش از یک صفت کیفی در آن واحد. م.

تعریف‌های متغیر و سطوح

Clothing	نوع لباس پوشیده‌شده توسط دانش‌آموز
Unisex	لباس هم دخترانه هم پسرانه
Std	لباس تک‌جنسیتی استاندارد
Other	لباس نامعمول
Gender	جنسیت دانش‌آموز (پسر/دختر)
Count	تعداد تعاملهای مشاهده شده
Interaction	نوع تعامل (از منظر معلم)
PosInt	تعامل مثبت بین دانش‌آموز و معلم
NegInt	تعامل منفی بین دانش‌آموز و معلم

تکلیف

این داده‌های «شمارشی» را چگونه تحلیل کنیم؟ یک رویکرد ساده در نظرگرفتن اختلاف بین تعداد تعامل‌های مثبت و منفی و ملاحظهٔ این مطلب است که آیا اثری ناشی از جنسیت یا نوع لباس وجود دارد یا خیر؟ به روش دیگر، می‌توان از یک رویکرد چندمتغیره استفاده کرد که در آن تعامل‌های مثبت و منفی نشان دهندهٔ دو پاسخ مجزا هستند.

۸.۹ مطالعه‌ای در مورد ایدز

توصیف

اندازه‌گیری‌های «شمارش سلولی» سلول‌هایی خاص، وسیله‌ای مؤثر برای نظارت بر بیمارانی است که مبتلا به ویروس ایدز هستند، یا بیماری‌هایی مانند سرطان یا هپاتیت دارند. برای کسی که HIVی مثبت است، دو ابزار تشخیصی مهم عبارت‌اند از شمارش سلول‌های CD4 و CD8 آن‌ها. CD4 سلول‌های سفید خون هستند که ویروس ایدز از آن‌ها به عنوان میزبان برای تکثیر خود استفاده می‌کند. بنابراین، شمارش سلول‌های CD4 یک نشانگر مهم از وضعیت سیستم ایمنی فرد در اختیار می‌گذارد:

زیر ۲۰۰ کاملاً مبتلا به ایدز

۲۰۰ تا ۵۰۰ وضعیت بینابینی

بالای ۵۰۰ سیستم ایمنی با کارکرد سالم

سلول‌های CD8، با کشتن سلول‌هایی که بدن آن‌ها را خارجی تشخیص می‌دهد، به سرکوب عفونت ناشی از ویروس کمک می‌کند. متأسفانه، خود ویروس ایدز با ساکن‌شدن در درون سلول CD4 از این تشخیص می‌گریزد. با این حال، شمارش سلول‌های CD8، معیاری از توانایی فرد برای مبارزه با سایر عفونت‌ها در اختیار می‌گذارد. معیار دیگر «بار» ویروسی که یک شخص ناقل آن‌هاست شمارش RNA است. RNA رشتهٔ واحدی از DNA است که ویروس ایدز از آن برای بازتکثیر خود استفاده می‌کند.

مطالعه

هدف از این مطالعه دریافتن این مطلب است که آیا سه معیارِ شمارش‌های CD4، CD8، و RNA توانایی تمیز بین دو گروه از زوج‌های ناهمسان[1] (DP: تنها یکی از دو نفر HIVی مثبت است) و همسان[2] (CP: هر دو HIVی مثبت‌اند) را دارند؛ تنها یکی از دو عضو هر زوج در این مطالعه منظور شد و عضو مبتلا در گروه DP اندازه‌گیری شد. این کار، همگروهی همگن‌تری را فراهم آورد و، برای حذف اثرهای اختلاط، معتادان به مواد مخدر و زوج‌های دارای بیش از یک همسر از مطالعه بیرون گذاشته شدند.

تکلیف

نتایج این مطالعه در مجموعه‌داده‌های c98.dat در اختیار گذاشته شده‌اند که عنوان ستون‌ها توضیح‌های آشکاری دارند. در این تحلیل، تبدیلی از متغیرهای پیشگو شاید کمک‌کننده باشد. نتایج تحلیل خود را برای یک مشتری، که با روش‌هایی که به‌کار برده‌اید، آشنا نیست، تفسیر کنید.

1. discordinant 2. concordinant

پیوست الف

منابع

مهارتی پراهمیت که هر مشاور آماری باید در خود بپروراند، توانایی یافتن اطلاعات و منابع است. این اطلاعات و منابع اغلب مرجعی برای شیوهٔ آماری خاصی خواهند بود که مشاور با آن ناآشناست (یا هرگز قبلاً به‌کار نبرده است). با این حال، ممکن است کاملاً بی‌ارتباط با هر موضوع آماری باشد. بهترین راه ایجاد یک سامانهٔ حسابداری چیست؟ پایگاه داده‌های مشتری در یک فرمت یا یک برنامهٔ کاربردی است که مشاور به آن دسترسی ندارد، یا فایل (های) شامل داده‌ها به‌هم‌ریخته به نظر می‌رسد. در این صورت، مشاور آماری باید با چه کسی برای گرفتن کمک مشورت کند؟

الف.۱ مراجع

برای آماردان، اینترنت به منبعی مهم تبدیل شده است. برای مشاور آماری اینترنت از ضروریات شده است. با این حال، اطلاعاتی که در اینترنت وجود دارد کامل نیست؛ مشاور هنوز هم باید به منابع کاغذی از قبیل کتاب متکی باشد. بنابراین، فهرست زیر از مجله‌های علمی، منبعی سودمند برای مشاوران است.

مجله‌های علمی

JRSS مجلهٔ انجمن سلطنتی آمار[1]، چهار سری منتشر می‌کند: A، B، C، و D. سری A نقد کتاب خوبی دارد و به آمار در مباحث اجتماعی می‌پردازد. سری B مقاله‌هایی دربارهٔ روش‌شناسی آماری منتشر می‌کند. سری C در واقع «آمار کاربردی» نامیده می‌شود. بنابراین، عنوان آن گویاست. این مجله اغلب دربردارندهٔ مقالات مطالعه‌های موردی است. سری D («آماردان»)[2] مقاله‌هایی در زمینهٔ عام دارد.

ASA انجمن آمار آمریکا[3] چندین مجله از جمله *JASA*، مجلهٔ انجمن آمار آمریکا[4]، را منتشر می‌کند که آمیزه‌ای از نظریه و کاربردها و نیز نقد کتاب است. نظیر اغلب مجلات علمی، کیفیت و تنوع مقالات به سردبیر بستگی دارد. آماردان آمریکایی[5] مقاله‌هایی برای عموم، و نیز نقد نرم‌افزارها (و کتب) را چاپ می‌کند. تکنومتریکس[6]، که مشترکاً با انجمن تضمین کیفیت آمریکا[7] چاپ می‌شود، بر استفاده از آمار و کاربردها در حوزه‌های علوم و مهندسی تأکید دارد. دو مجلهٔ علمی قابل ذکر، که مقاله‌هایی مورد نظر مشاوران آماری چاپ می‌کنند، عبارت‌اند از: مجلهٔ آمار بازرگانی و اقتصادی[8]، و مجلهٔ گرافیک‌های محاسباتی و آماری[9] (که مشترکا با مؤسسهٔ آمار ریاضی[10] (IMS) و بنیاد واسط[11] چاپ می‌شود.)

IMS عمدتاً مقاله‌های نظری در سالنامهٔ آمار[12] و سالنامهٔ احتمال[13] منتشر می‌کند. اخبار کنفرانس‌ها، فرصت‌های شغلی، و سایر خبرها در بولتن IMS[14] در اختیار گذاشته می‌شود. مقاله‌های مورد نظر مشاوران آماری را به احتمال زیاد می‌توان در آماردان یافت.

دیگر مجله‌های علمی گفتن اینکه تعداد زیادی مجلهٔ علمی دیگر وجود دارند که ارزش خواندن دارند و برخی از آن‌ها در محدودهٔ سنتی آمار قرار ندارند، تا حدی سطحی خواهد

---

1. The Journal of the Royal Statistical Society 2. The Statistician 3. The American Statistical Association 4. Journal of the American Statistical Association 5. The American Statistician 6. Technometrics 7. American Society for Quality Assurance 8. Journal of Business and Economic Statistics 9. Journal of Computational and Statistical Graphics 10. The Institute of Mathematical Statistics 11. Interface Foundation 12. Annals of Statistics 13. Annals of Probability 14. IMS Bulletin

بود. به‌عنوان مثال، موارد زیر را ذکر می‌کنیم: بیومتریکا[1]، بیومتریکس[2]، آمار در پزشکی[3]، مجلهٔ بازاریابی[4]، مجلهٔ بین‌المللی پیش‌بینی[5]، مجلهٔ *IEEE*ی کنترل خودکار[6].

... وقت رفتن به اینترنت فرارسیده است.

اینترنت اینترنت به منبعی مهم برای همه تبدیل شده است. اینترنت در بدو امر برای ارتش آمریکا (آرپانت[7]) طراحی شد و سپس به دانشگاه‌ها راه یافت و امروزه به منبع عمومی مهمی بدل شده است. از بین سایت‌های مختلفی که می‌توانیم برای آمار ذکر کنیم، یکی از سایت‌هایی که خوب پشتیبانی می‌شود استت‌لیب[8] است. این سایت منبعی غنی از اطلاعات است.

`lib.stat.cmu.edu` این همان استت‌لیب است که توسط دانشگاه کارنگی ملون[9] پشتیبانی می‌شود. به‌طوری‌که در بالا گفتیم، این سایت اطلاعاتی غنی دارد که مورد توجه مشاوران آماری خواهد بود. در زیر فهرست کوتاهی از برخی منابع مندرج در استت‌لیب ارائه می‌شود.

- Other Sites: شامل نشانی‌های وب‌سایت‌های بین‌المللی است.
- DESIGNS: مجموعه‌ای از طرح‌ها و برنامه‌ها و الگوریتم‌ها برای ایجاد طرح‌هایی برای آزمایش‌های آماری.
- Directory of People: فهرست نشانی‌ها و آدرس‌های ای‌میل آماردان‌ها.
- DOS/Windows: Dos و Windows که تنها از طریق FTP و WWW قابل دسترسی‌اند.
- General Archive: شامل نرم‌افزارهای گوناگونی است که به زبان FORTRAN، C، و Lisp نوشته شده‌اند، برخی از آن‌ها سیستم‌های آماری کاملی هستند، و بقیه اقلام متفرقه‌اند.
- Genstat: نرم‌افزارها و ماکروها[10] برای زبان Genstat.
- Glim: ماکروها و نرم‌افزارهای مربوط به بستهٔ GLIM.

1. Biometrika 2. Biometrics 3. Statistics in Medicine 4. Journal of Marketing 5. International Journal of Forecasting 6. IEEE Journal of Automatic Control 7. Arpanet 8. StatLib 9. Carnegie Mellon 10. macros

- JASA Software: این مجموعه شامل برنامه‌های مربوط به مقاله‌هایی است که در مجلهٔ انجمن آمار آمریکا چاپ می‌شوند.
- JCGS: بایگانیِ jcgs دربردارندهٔ بایگاه داده‌ها و نرم‌افزارها و چکیده‌های چاپ‌شده در مجلهٔ گرافیک‌های محاسباتی و آماری است.
- MacAnova: سیستم آماری MacAnova، برای Mac، Windows، و UNIX.
- Maps: یک نقشهٔ دنیا در فایل tar فشرده با حجم بیش از ۴ مگابایت. تنها از طریق FTP (از مد *binary* استفاده کنید) قابل دسترسی است.
- Meetings: تقویم‌ها و برنامه‌های انواع همایش‌های داخلی و بین‌المللی.
- Minitab: ماکروهای آمار صنعتی Minitab و ماکروهای گروه کاربران Minitab.
- Multivariate: انواع روتین[1]ها برای تحلیل چندمتغیره و رده‌بندی. برای دیگر فنون چندمتغیره، مجموعهٔ عمومی را ببینید.
- PoliSci Data: داده‌هایی از مجله‌ها و مؤلفان علوم سیاسی.
- R: 'GNU S' ــ زبان و محیطی برای محاسبات و گرافیک آماری.
- S Archive: نرم‌افزار و افزونه[2]هایی برای زبان S (S-PLUS). بیش از ۱۳۰ بستهٔ جداگانه مشتمل بر بسیاری ایده‌های جدید آماری.
- Xlispstat : سیستم Luke Tierney's Xlispstat برای سیستم‌های UNIX.

stat.fsu.edu : کتابخانهٔ مجازی شبکهٔ جهان‌گستر[3] آمار که توسط گروه آمار دانشگاه فلوریدا پشتیبانی می‌شود.

1. routine 2. extension 3. World Wide Web Virtual Library

سایت‌های دولتی ایالات متحده و دیگر کشورها وب‌سایت‌هایی را پشتیبانی می‌کنند که آمارهای «دولتی» را از بسیاری منظرهای اجتماعی در اختیار می‌گذارند. برای اطلاع از نهادهای دولتی ایالات متحده به www.fedstats.gov بروید. فهرستی از چند سایت دولتی بین‌المللی در زیر ارائه شده است.

دفتر آمار استرالیا	www.abs.gov.au
بخش آمار ـ هنگ‌کنگ	www.info.gov.hk
دفتر آمار ـ اندونزی	www.bps.go.id
موسسهٔ دانمارکی علوم کشاورزی	www.jbs.agrsci.dk
بخش آمار ـ اردن	www.dos.gov.jo
بخش آمار ـ مالزی	www.statistics.gov.my
مؤسسهٔ آمار هند	www.isical.ac.in
مؤسسهٔ ملی آمار ـ اسپانیا	www.ine.es
ادارهٔ مرکزی آمار ایرلند	www.cso.ie
مؤسسهٔ ملی آمار ایتالیا	www.istat.it
دفتر آمار ژاپن	www.stat.gov.jp
مؤسسهٔ ملی آمار ـ بولیوی	www.ine.gov.bo
ادارهٔ ملی آمار ـ کرهٔ جنوبی	www.nso.gov.kr
ادارهٔ ملی آمار ـ تایلند	www.nectec.or.th
فدراسیون روسیه دربارهٔ آمار	www.gks.ru
آمار کانادا	www.statcan.ca
آمار دانمارک	www.dst.dk
بخش آمار ـ لیتوانی	www.std.lt
آمار گرینلند	www.statgreen.gl
آمار هلند	www.cbs.nl
آمار سنگاپور	www.singstat.gov.sg

نرم‌افزار مشاوران آماری نیاز به نرم‌افزارهای آماری دارند. برای یافتن آن‌ها چه جایی بهتر از اینترنت؟ در اینجا فهرستی از برخی سایت‌های نرم‌افزارهای آماری موجود است.

شرکت نرم‌افزاری CyTEL	www.cytel.com
EaSt ،LogXact ،StatXact	
شرکت دیتا دسکریپشن[1]	www.datadesk.com
Activstats ،Vizion ،Data Desk	
Statgraphics	www.statgraphics.com
شرکت Mathsoft ــ **S-PLUS**	www.mathsoft.com
شرکت Minitab	www.minitab.com
گروه الگوریتم‌های عددی	www.nag.co.uk
Genstat،GLIM ،NAG Libraries	
مؤسسهٔ **SAS**	www.sas.com
SPSS	www.spss.com

الف.۲ مجموعه‌داده‌های مطالعه‌های موردی قسمت II

به مجموعه‌داده‌های استفاده‌شده در مطالعه‌های موردی قسمت II، می‌توان از طریق وب‌سایت اشپرینگر-فرلاگ[2] یا وب‌سایت مؤلفان [این کتاب] دسترسی یافت:

```
www . springer-ny.com
www . rci.rutgers.edu/~ cabrera
www . csam.montclair.edu/~ mcdougal
```

مجموعه‌داده‌ها مطابق با زیربخش‌هایی که در آن‌ها ظاهر می‌شوند نامگذاری شده‌اند (مثلاً، «c64.dat» مجموعه‌داده‌های مطالعهٔ موردی ۴.۶ است). اطلاعات بیشتر دربارهٔ کد SAS و S-PLUS در اختیار داده شده‌اند.

مؤلفان همچنین اطلاعات بیشتری از قبیل توصیهٔ خاص ESL که در این کتاب نیامده در اختیار گذاشته‌اند.

الف.۳ درس مشاورهٔ آماری

الف.۱.۳ توصیف درس

هدف ایجاد درسی جدید در سطح کارشناسی ارشد دربارهٔ مشاورهٔ آماری. مسائل واقعی آماری و علمی، که در تعاملات نوعی بین آماردانان و دانشمندان به عرصه می‌آیند، به دانشجویان

1. Data Description 2. Springer-Verlag

عرضه می‌شود. اجزای درس بر محور مطالعه‌های موردی که به‌وسیلهٔ سخنرانان مدعو ارائه می‌شود متمرکز است.

مخاطبان درس دانشجویان ارشد سال دومی. دانشجو باید با محاسبات و روش‌شناسی آماری کاربردی آشنا باشد.

پیش‌نیازها به سطح مهارت دانشجویان و/یا نظر مدرس بستگی دارد. فهرستی ممکن از پیش‌نیازها عبارت‌اند از:

سطح ۱ برای دانشجویان ارشد سال اولی، تأکید درس ممکن است بر جنبه‌های ارتباطی مشاورهٔ آماری معطوف باشد. تحلیل مطالعه‌های موردی لزوماً به سطح معلومات آماری که در سطح دورهٔ کارشناسی (سال‌های آخر) مؤسسه ارائه شده محدود می‌شود.

- تحلیل اکتشافی داده‌ها.
- یک درس محاسباتی که در آن دانشجویان پایهٔ محاسبات و تحلیل داده‌ها را فرامی‌گیرند.
- ANOVAی ساده، رگرسیون.

سطح ۲ برای درسی که برای سال دوم طراحی شده، روش‌شناسی آماری را می‌توان گسترش داد تا مطالب زیر را شامل شود:

- طرح آزمایش‌ها.
- رگرسیون و GLM.
- درس روش‌شناسی یا نظریه. مبانی محاسبات آماری.
- محاسبات آماری: مبانی SAS، S-PLUS (یا معادل آن‌ها).

سطح ۳ برای دانشجویان دکترای مرحلهٔ پژوهشی، می‌توان این درس را مورد استفاده قرار داد تا اطمینان حاصل شود که این دانشجویان مهارت‌های کافی یافته‌اند که بعد از فارغ‌التحصیلی پیشنهادهای جالب شغلی دریافت دارند. به‌ویژه، مهارت‌های ارتباطی خوب حائز اهمیت‌اند.

- تحلیل اکتشافی داده‌ها (شاید ... یک مرور؟).
- GLM، چندمتغیره، سری زمانی، و مدل‌های رسته‌ای.
- محاسبات آماری با SAS و S-PLUS.

اندازهٔ کلاس حداکثر ۱۲ تا ۱۵ نفر در هر کلاس که به سه تا پنج تیم مشاوره تقسیم شده‌اند. ما متوجه شده‌ایم که دانشجویان اغلب ترجیح می‌دهند که به‌صورت دو نفری کار کنند اما این‌کار آشکارا هدف تعامل‌کردن به‌عنوان جزئی از یک تیم را از بین می‌برد. قویاً مدرسان را تشویق می‌کنیم که اندازهٔ تیم‌ها را در حداقل سه نفر نگه دارند.

تکالیف برای هر مطالعهٔ موردی، تکلیفی همراه با مجموعه‌داده‌های کامل به تیم‌های مشاوره واگذار می‌شود، و آن‌ها باید گزارشی تولید کنند و در اغلب موارد ارائه‌کردی داشته باشند. برای هر مطالعهٔ موردی، تیم‌ها سرتیمی دارند که ارائه‌کرد و نوشتن گزارش بر عهدهٔ اوست. مسئولیت سرتیم‌بودن برای مطالعه‌های موردی مختلف چرخشی خواهد بود.

مطالعه‌های موردی شناسایی مطالعه‌های موردی مناسب از اهمیت فراوانی برخوردار است زیرا مطالعه‌های موردی قلب این درس‌اند. ما مصمم شده‌ایم که مطالعه‌های موردی از گروه‌های زیر را انتخاب کنیم.

گروه I مطالعه‌های موردی ساده‌ای که جواب‌های آن‌ها معلوم‌اند، اما مباحث آماری و علمی جالب در آن‌ها قابل بحث است. مطالعه‌های موردی گروه I در یک جلسهٔ درسی یک‌ساعته ارائه خواهد شد و پس از آن بحثی صورت خواهد گرفت. متعاقب این مباحثه، به هر تیم پروژه‌ای واگذار خواهد شد. داده‌های همهٔ مطالعه‌های موردی باید در دسترس دانشجو باشد.

دانشجویان در هفتهٔ بعد از آن گزارش‌های تیمی خود را ارائه خواهند کرد.

گروه II مطالعه‌های موردی دشوارتری که مسائل آماری موجود در آن‌ها عموماً خوش‌تعریف‌اند، اما وسعت دیدی فراخ‌تر از مطالعه‌های موردی گروه I دارند. ممکن است لازم باشد که چندین راه‌حل مورد ارزیابی قرار گیرند. دشواری مسئله ممکن است تا حدی ناشی از اندازه یا فرمت مجموعه‌داده‌ها باشد.

مطالعه‌های موردی گروه II در فرمتی مشابه با گروه I ارائه خواهند شد.

نتایج اولیه در هفتهٔ بعد از آن ارائه می‌شوند. بحث و پرسش باید مورد توجه باشد. ارائه‌کرد نهایی در هفتهٔ دوم.

گروه III مطالعه‌های موردی پژوهش‌محوری که در آن‌ها پی‌بردن به مسئلهٔ آماری دشوار است و دانشجویان باید زیاد فکر کنند و پاسخ ممکن است معلوم باشد یا نباشد. ممکن است چندین مرحلهٔ تحلیل برای رسیدن به نتایج مناسب ضروری باشد. ممکن است لزوماً یک «پاسخ» به مسئلهٔ آماری وجود نداشته باشد.

هفتهٔ بعد از آن، دانشجویان گزارش‌های اول را ارائه و بحث موضوعات اصلی را ادامه خواهند داد و به اهداف نهایی دست خواهند یافت.
در هفتهٔ دوم یا سوم ارائه‌کردهای نهایی را انجام خواهند داد.

الف.۳.۲ فهرست مباحث برحسب هفته

هفتهٔ ۱ آشنایی با درس و سازماندهی آن. دانشجویان تیم‌های مشاوره‌ای را تشکیل می‌دهند که در آماده‌سازی و ارائهٔ چندین گزارش شامل تحلیل مطالعه‌های موردی با هم کار خواهند کرد.
تاریخ علم و نقش آمار. آشنایی با روش علمی. محیط‌های مشاورهٔ آماری. نقش آماردان در یک محیط علمی.
بخش کتاب درسی: فصل ۱.
سخنران: مدرس درس.

هفتهٔ ۲ برقراری ارتباط با پژوهشگران دیگر حوزه‌ها. گزارش‌نویسی.
بخش کتاب درسی: فصل ۲.
سخنران: مدرس درس.

هفتهٔ ۳ جنبه‌های روش‌شناختی. مروری بر روش‌های آماری که در این درس به‌کار خواهند رفت.
مروری بر ابزارهای محاسباتی و نرم‌افزارهای آماری از قبیل SAS و S-PLUS که در اختیار دانشجویان است.
بخش کتاب درسی: فصل ۳، پیوست ب.
سخنران: مدرس درس.

هفتهٔ ۴ مطالعهٔ موردی ۱ (از گروه I)
آشنایی با مطالعه‌های موردی. توصیف شیوه‌ای که در مورد مطالعه‌های موردی در پیش گرفته خواهد شد. فرمت گزارش‌نویسی و ارائه‌کرد توسط دانشجویان.
ارائه‌کرد مطالعهٔ موردی ۱ توسط سخنران مدعو. بحث.
واگذاری پروژه به تیم‌های دانشجویی.
بخش کتاب درسی: فصل ۴.
سخنران: مدرس درس و سخنران مدعو برای مطالعهٔ موردی ۱.

هفتهٔ ۵ مطالعهٔ موردی ۲ (از گروه I)
ارائه‌کرد توسط تیم‌های مشاوره از گزارش‌های مطالعهٔ موردی ۱ به‌صورت ضبط ویدئویی.

ارائه مطالعهٔ موردی ۲ توسط سخنران مدعو. بحث.
مرور نوار ویدئویی. بحث.
واگذاری پروژه‌ها به تیم‌های دانشجویی.
بخش کتاب درسی: مرور فصل‌های ۲، ۳، ۴.
سخنران: سخنران مدعو برای مطالعهٔ موردی ۲.

هفتهٔ ۶ مطالعهٔ موردی ۳ (از گروه I)
ارائه‌کردهایی توسط تیم‌های مشاوره از گزارش‌های مطالعهٔ موردی ۲. بحث
واگذاری پروژه‌های مطالعهٔ موردی ۳ به تیم‌های دانشجویی.
مرور گزارش‌های تحویل‌شده و ارائه‌کردها. بحث مهارت‌های ارتباطی.
بخش کتاب درسی: فصل ۶.
سخنران: مدرس درس.

هفته ۷ مطالعهٔ موردی ۴ (از گروه II)
ارائه‌کردهایی توسط تیم‌های مشاوره از گزارش‌های مطالعهٔ موردی ۳.
ارائهٔ مطالعهٔ موردی ۴ توسط سخنران مدعو. بحث.
واگذاری پروژه‌ها به تیم‌های دانشجویی.
مقاله‌های پژوهشی و گزارش‌هایی از دیگر مطالعه‌های موردی واگذارشده به دانشجویان.
بحث دربارهٔ مطالب خوانده شده در دیگر مطالعه‌ها.
بخش کتاب درسی: پیوست الف.
سخنران: سخنران مدعو برای مطالعهٔ موردی ۴.

هفته ۸ مطالعهٔ موردی ۴ (از گروه II).
گزارش پیشرفت کار توسط تیم‌های دانشجویی. ارائه‌کرد مقدماتی تحلیل اولیهٔ آن‌ها برای مطالعهٔ موردی ۴.
بحث، ارزیابی پیشرفت، و بازنگری اهداف پروژه‌ها. جمع‌آوری گزارش‌های مربوط به مقاله‌های پژوهشی.
سخنران: مدرس درس.

هفتهٔ ۹ مطالعهٔ موردی ۵ (از گروه II).
ارائه‌کردها توسط تیم‌های مشاور از گزارش‌های مطالعهٔ موردی ۴.
ارائه‌کردی از مطالعهٔ موردی ۵. بحث مطالعهٔ موردی ۵. مرور گزارش‌های مقاله‌ها.

بخش کتاب درسی: فصل ۷.
سخنران: مدرس درس.

هفته ۱۰ گزارش پیشرفت کار توسط تیم‌های مشاوره. ارائه‌کردهای مقدماتی از تحلیل اولیهٔ مطالعهٔ موردی ۵.
بحث، ارزیابی پیشرفت، و بازنگری اهداف پروژه‌ها.
بخش کتاب درسی: فصل ۷.
سخنران: مدرس درس.

هفتهٔ ۱۱ مطالعهٔ موردی ۶ (از گروه III).
ارائه‌کردهایی توسط تیم‌های دانشجویی از گزارش‌های مطالعهٔ موردی ۵.
دومین ضبط ویدئویی.
ارائه‌کرد سخنران مدعو از مطالعهٔ موردی ۶. بحث.
مرور ویدئو. انتقاد از خود دربارهٔ ارائه‌کردها. واگذاری پروژه‌ها به تیم‌های دانشجویی.
بخش کتاب درسی: فصل ۸.
سخنران: سخنران مدعو برای مطالعهٔ موردی ۶.

هفتهٔ ۱۲ مطالعهٔ موردی ۶ (از گروه III).
گزارش پیشرفت کار توسط تیم‌های مشاوره. ارائه‌کرد مقدماتی تحلیل اولیهٔ آن‌ها از مطالعهٔ موردی ۶.
بحث، ارزیابی پیشرفت، و بازنگری اهداف پروژه‌ها.
بخش کتاب درسی: فصل ۸.
سخنران: مدرس درس.

هفته ۱۳ مطالعهٔ موردی ۷ (مشکل نامعلوم)
ارائه‌کرد توسط تیم‌های مشاوره از گزارش‌های مطالعهٔ موردی ۶.
ارائه‌کردی از مطالعهٔ موردی ۷. مهلت: یک هفته برای تغییر نظر. بحث توسط دانشجویان اداره می‌شود. واگذاری پروژه‌ها.
بخش کتاب درسی: فصل ۹.
سخنران: مدرس درس.

هفتهٔ ۱۴ مطالعهٔ موردی ۷.
ارائه‌کردهای گزارش‌های نهایی برای مطالعهٔ موردی ۷.

تحویل‌دادن گزارش‌ها برای ارزیابی نهایی. پیتزا سفارش دهید.
بحث آزاد.
سخنران: مدرس درس.

هفتهٔ ۱۵ پس از هفتهٔ ۱۴ (اختیاری). تجربه‌های زندگی واقعی. کارآموزی محدود یا پروژه‌های مشاوره‌ای.

الف ۳.۳ فهرست مراجع

عمومی: کتاب‌های درسی که می‌توان به‌عنوان مکمل این کتاب به‌کاربرد.

1 J. Derr (2000), *Statistical Consulting: A Guide to Effective Communication*, Duxbury Press, Pacific Grove, CA.

2 C. Chatfield (1995), *Problem Solving. A Statistician's Guide*, (2nd ed.) Chapman & Hall, London.

3 D. J. Hand and B. S. Everitt (Eds.) (1987), *The Statistical Consultant in Action*, Cambridge University Press, Cambridge.

4 J. Tanur (Ed.) (1988), *Statistics: A Guide to the Unknown*, (3rd ed.), Duxbury Press, Belmont, CA.

5 S. Conrad (Ed.) (1989), *Assignments in Applied Statistics*, John Wiley, New York.

6 D.R. Cox and E.J. Snell (1981), *Applied Statistics, Principles and Examples*. Chapman & Hall, London.

آمار زیستی: کتاب‌های درسی که به‌طور خاص به کاربردهای آمار زیستی می‌پردازند.

7 R. G. Miller, B. Efron, B. W. Brown, and L. E. Moses (Eds.) (1980), *Biostatistics Casebook*, John Wiley, New York.

آمار داروسازی: کتاب‌های داروسازی که به‌طور خاص به کاربردهای آمار داروسازی می‌پردازند.

8 K. E. Peace (Ed.) (1988), *Biopharmaceutical Statistics for Drug Development*, Marcel Dekker, New York.

داده‌ها: برخی منابع داده‌ها از مرجع زیر در دسترس‌اند:

9 D. F. Andrews and A. M. Herzberg (1985), *Data: A Collection of Problems from Many Fields for the Student and Research Worker*, Springer-Verlag, New York.

پیوست ب

نرم‌افزارهای آماری

ب.۱ SAS

SAS یک بستهٔ قدرتمند تحلیل آماری است که به‌طور گسترده‌ای توسط شرکت‌های بزرگ (و نه چندان بزرگ) بسیاری به‌خصوص شرکت‌های داروسازی به‌کار برده می‌شود. با این حال، SAS ساختار زبانی نسبتاً خاصی دارد که می‌تواند دستکاری داده‌ها را تا حدودی با ماجراجوئی توأم کند که در نتیجه به افراد کم‌جرأت توصیه نمی‌شود!

ب.۱.۱ تجهیزات SAS

در اجرای هر برنامهٔ SAS، با فایل‌های متعددی سروکار پیدا می‌کنیم. اگر SAS را به‌صورت تعاملی اجرا کنید، آنگاه هر یک از فایل‌های زیر در پنجرهٔ جداگانه‌ای ظاهر خواهند شد.

prog.sas: فایل «program» که توسط کاربر ایجاد شده و شامل دستور SAS است که، اگر به‌صورت موفقیت‌آمیزی پردازش شود، فایل خروجی `prog.lst` را ایجاد می‌کند. در اغلب سیستم‌ها پسوند «.sas» لازم است، اما نام «prog» را خودمان انتخاب کرده‌ایم.

prog.lst: فایل «output» که ممکن است دربردارندهٔ **همهٔ** نتایج مورد نظر ما نباشد، زیرا جایی خطایی رخ داده است. برای یافتن مکان خطا، فایل prog.log را بازبینی کنید.

prog.log: فایل «error». این فایل، در واقع، فهرست‌نامهٔ مواردی است که با اجرای هر دستور SAS در فایل برنامه prog.sas رخ داده‌اند. همیشه باید آن را در مرحلهٔ ورود داده‌ها بازبینی کنیم، زیرا SAS خواندن فایل داده‌ها را حتی در صورت مواجه‌شدن با خطا ادامه می‌دهد.

مثالی از برنامهٔ SAS

مثال زیر برخی از دستورهای پایه‌ای SAS را نشان می‌دهد که همگی آن‌ها در فایل prog.sas گنجانده خواهد شد.

```
options ls=78 ;
filename in1 'example.dat';
data a;
      infile in1;
      input x y ;
* proc print ;
data b ;
      set a ;
      if y > 10 ;
proc means ;
run ;
endsas ;
```

توجه کنید که ساختار یک برنامهٔ SAS به‌واقع کاملاً ساده است. این برنامه شامل دو گام اساسی است که در هر چندبارکه لازم باشد توسط کاربر برای اجرای تحلیل مورد نظر به‌کار گرفته می‌شوند:

گام DATA: پردازش و اجرای عملیات روی داده‌ها.

گام PROC: اِعمال شیوه‌ای مشخص روی داده‌ها.

این گام‌ها باید با شرایط مشخصی سازگاری داشته باشند و هر دستور باید مطابق با نحو SAS باشد. همان‌گونه که در مثال بالا ملاحظه می‌شود، جنبه‌های زیر در هر برنامهٔ SAS عمومیت دارند.

نحو: کلیهٔ دستورها باید به یک نقطه‌ویرگول ; ختم شوند.

تذکرها: متن یا دستورهایی که بین * <text> ; ظاهر می‌شوند به‌عنوان تذکر تلقی می‌شوند و SAS آن‌ها را در طول اجرا نادیده می‌گیرد.

گزینه‌ها: می‌توانیم گزینه‌هایی را برای درکنترل‌داشتن خروجی prog.lst مشخص کنیم. در اینجا، ls=78 خروجی را به ۷۸ کاراکتر در هر سطر محدود می‌کند.

filename: به SAS می‌گوید که یک فایل داده‌های خارجی را در کجا پیدا کند.

endsas: به SAS می‌گوید که اجرا را متوقف کند. اگر در حال اجرای تعاملی SAS هستید، از این دستور استفاده نکنید زیرا در این صورت برنامهٔ SAS به کلی بسته خواهد شد (به جای آن از دستور نهایی ; run استفاده کنید).

حال می‌توانیم تحلیل خاصی را که توسط این برنامه اجرا می‌شود توصیف کنیم.

---

۱. example.dat یک فایل دادهٔ خارجی است که شامل دو ستون دادهٔ عددی است که حداقل با یک فاصلهٔ سفید (یعنی، فضای خالی) از هم جدا شده‌اند.

۲. به این فایل داده‌ها یک *fileref*[ارجاع به فایل] تخصیص داده می‌شود که "in1" نامیده می‌شود. نام واقعی در اختیار کاربر است و در نتیجه می‌توان با استفاده از دستورهای filename متوالی به بیش از یک فایل خارجی با استفاده از *fileref*های متفاوت دسترسی پیدا کرد.

۳. ; data a یک مجموعه‌داده‌های «کاری» به نام "a" (به انتخاب کاربر) ایجاد می‌کند که چگونگی پردازش اطلاعات موجود در example.dat توسط SAS را نشان می‌دهد.

۴. ; infile in1، SAS را وامی‌دارد که به فایل خروجی دسترسی پیدا کند. اگر این‌کار موفقیت آمیز بود، آنگاه

۵. ; input x y به SAS می‌گوید چگونه example.dat را بخواند و چه نام‌هایی به متغیرهای ستون‌ها تخصیص دهد.
توجه: اگر ستونی شامل یک دادهٔ کاراکتری باشد، آنگاه باید نماد $ را بلافاصله بعد از نام متغیر به‌کار برد؛ برای مثال، ; input x $ y نشان می‌دهد که ستون x دربردارندهٔ مقادیر کاراکتری است.

۶. SAS دستور ; proc print * را نادیده خواهد گرفت. اگر * برداشته شود، این شیوه، مندرجات جدیدترین مجموعه‌داده‌های SAS ایجادشده (در این مورد "a") را در فایل prog.lst چاپ خواهد کرد.

۷. دستور ; data b ; set a ; if y > 10 یک مجموعه‌داده‌های جدید SAS به‌نام "b" ایجاد می‌کند که با مندرجات "a" آغاز می‌شود (دستور ; set a)، اما صرفاً

آن سطرهایی از "a" را برمی‌گزینند که به‌ازای آن‌ها $y > ۱۰$. مقادیر x متناظر با این مقادیر y نیز در این مجموعه‌داده‌های "b"ی SAS گنجانده خواهند شد.

۸. دستور ; proc means شیوهٔ MEANS را در مورد جدیدترین مجموعه‌داده‌های ایجادشدهٔ SAS، در این مورد "b"، به‌کار می‌بندد. برای به‌کارگیری این شیوه بر "a"، به دستور ; proc means data=a نیاز داریم. دستورها و گزینه‌های بیشتری وجود دارند که می‌توان برای این شیوه مشخص کرد. در شکل پیش‌فرض، به صورتی که در اینجا ارائه شده است، SAS میانگین، مینیمم، ماکسیمم، تعداد مشاهدات، و انحراف استاندارد همهٔ متغیرهای عددی موجود در "b" را محاسبه و نتایج را در prog.lstچاپ خواهد کرد. SAS نمایش نتایج در یک فرمت مناسب را به عهده می‌گیرد.

ب.۱.۲ جزئیات در مرحلهٔ DATA

در مثال بالا، خواندن example.dat در SAS آسان است، زیرا این داده‌ها در فرمت ستونی با یک یا چند فاصلهٔ سفید بین ستون‌ها است. در این صورت روشن است که میزان زمان و تلاشی که لازم است صرفِ دستور ; input کنیم، در صورتی که فایل دادهٔ خارجی قبلاً در فرمت ستونی بوده باشد، به حداقل خواهد رسید. این مطلب حتی در صورتی که ستون‌ها شامل مقادیر کاراکتری خاص، تاریخ، یا زمان باشند نیز صادق است، زیرا SAS دامنهٔ گسترده‌ای از تعدیل‌گر[1]ها را برای پرداختن به این ساختارها فراهم می‌کند. در نتیجه:

همواره سعی کنید فایل‌های داده‌ای خارجی را در ساختار ستونی آماده کنید.

البته، ما همواره آن‌قدر خوش‌اقبال نیستیم و در این صورت زمان و تلاشی که لازم است صرف کنیم ممکن است به میزان چشمگیری افزایش یابد. با این حال، به شرط آنکه فایل داده‌ها فرمتی عادی داشته باشند، باید قادر باشیم که از فرمت‌کردن مجدد فایل داده‌ای به‌طور مستقیم اجتناب کنیم. مثال‌های زیر برخی از وضعیت‌هایی را نشان می‌دهند که در آن‌ها دستور پیچیده‌تری برای ورود داده‌ها مورد نیاز است.

مثال ب.۱ دنباله‌داربودن [Trailing @@]

```
input trt $ x1 x2 x3 y  @@;
```

1. modifier

از یک سطر، بیش از یک مشاهده (دنباله) بخوانید. در اینجا trt یک گروه تیمار (کاراکتری) است، x1 x2 x3 سه متغیر طرح‌اند، و y متغیر پاسخ است. فرمت فایل داده‌ها به شکل زیر خواهد بود:

```
placebo 0 1 0  23.5 drugA 0 1 1 13.4 drugB 1 0 0 11.4
placebo 1 0 1  33.1 drugA 1 0 0 12.7 drugB 1 1 1 16.9
...
```

مثال ب.۲ فرمت ثابت

```
input x 4-7 @15 y  ;
```

مقادیر X را از ستون‌های ۴ تا ۷ بخوانید سپس به ستون ۱۵ بروید تا مقادیر Y متناظر را بخوانید.

مثال ب.۳ ورودی‌های چندگانه

```
if(mod(_N_,11) = 1) then
   do ;
      input group $ ;
      retain group ;
   end ;
else
   do ;
      input  qlabel :$15. answer ;
      output ;
   end ;
```

برای خواندن داده‌های بررسی، از شمارندهٔ داخلی مشاهدات، یعنی -N- (که توسط SAS تعبیه شده)، استفاده کنید که این داده‌ها مشتمل بر ۱۰ سؤال‌اند که در آن، پیش از پاسخ‌های (عددی)، answer برچسب (سؤالی) qlabel : $15 آمده که می‌تواند تا ۱۵ کاراکتر طول داشته باشد. اگرچه یک پاسخگو با متغیری مانند group رده‌بندی می‌شود، این مدخل صرفاً در ابتدای هر مجموعه از پاسخ‌ها ظاهر می‌شود و این دلیل نیاز به استفاده از دستور ; retain group است.

مثال ب.۴ هر چیزی را بخوانید.

```
input tmp $ @@ ;
```

هر داده را به‌عنوان یک مقدار کاراکتری بخوانید! نکته در اینجا این است که گاهی خواندن داده‌های داده‌ها به‌عنوان متغیرهای کاراکتری و سپس تبدیل آن‌ها به فرمت مناسب آسان‌تر است.

خطاهای SAS

در اغلب موارد، خطاهای نحوی را می‌توان به‌سرعت با بازبینی فایل prog.log رفع و رجوع کرد، زیرا SAS پردازش DATA یا PROC Step را در جایی که خطا رخ داده است متوقف می‌کند. فایل prog.log همچنین منبع مهمی برای مشخص‌کردن جایی است که یک فایل داده‌ای به‌درستی خوانده نشده است. خطاهایی که حل و فصل آن‌ها دشوار است، معمولاً در جایی رخ می‌دهند که کاربر می‌خواهد نوعی دستکاری روی داده‌ها انجام دهد، اما SAS به تولید نتایج غیر منتظره ادامه می‌دهد. این امر ممکن است به‌خصوص زمانی دردسرآفرین باشد که کاربر می‌خواهد به عناصر سطری خاصی از یک متغیر دسترسی یابد یا روی آن عملیاتی انجام دهد. در SAS، هیچ راهی برای دسترسی مستقیم به «iامین عنصر»، مثلاً $X[i]$، از متغیری مانند X وجود ندارد. استفادهٔ ماهرانهٔ توأم از چندین DATA Step ممکن است برای زیرمجموعه‌ای از سطرهای یک متغیر مورد نیاز باشد.

ب.۳.۱ شیوه‌های SAS

گام PROC اساساً عبارت از اِعمال یک شیوهٔ SAS از پیش‌فراهم‌شده بر روی یک مجموعه‌داده‌های SAS است. به‌عبارت دیگر، ما باید از گزینه‌های موجود درون یک شیوهٔ SAS استفاده کنیم تا یک تحلیل خاص را اجرا کنیم. خوشبختانه، دامنهٔ گزینه‌ها گسترده است و شیوه‌های عامی مانند GLM و CATMOD موجودند که تحلیل‌های چندین شیوهٔ خاص را تکرار می‌کنند. برای مثال، می‌توان از هر دوی GLM و ANOVA برای تحلیل طرح‌های متعادل استفاده کرد، اما زمانی که طرح نامتعادل باشد به GLM نیاز داریم. در زیر خلاصه‌ای از برخی شیوه‌ها و گزینه‌های مفید SAS آمده است.

EDA

FREQ جدول‌های پیشایندی چندگانه و آماره‌های مربوط به آن‌ها

```
proc freq  ;
tables a a*b / exact norow nocol nopercent ;
```

یک جدول فراوانی از متغیر A و یک جدول پیشایندی از $A \times B$ را با محاسبهٔ آزمون دقیق فیشر تولید کنید. چاپ سطر، ستون، و درصدهای خانه‌ها را متوقف کنید. این درصدها به‌طور پیش‌فرض چاپ خواهند شد.

MEANS میانگین‌ها، انحراف استانداردها، و دیگر آماره‌های خلاصه برای متغیرهای عددی.

```
proc means n mean std stderr t prt ;
var x1-x5 ;
output out=mdata mean=m1-m5 ;
```

آماره‌های خلاصهٔ مشخص‌شده را محاسبه و آزمون tی $H_0 : \mu_1 = 0$ را برای پنج متغیر عددی X_i، $i = 1, 2, 3, 4, 5$، اجرا کنید. توجه کنید که می‌توان از x1-x5 برای نمایش دنبالهٔ x1 x2 x3 x4 x5 استفاده کرد. دستور آخر پنج میانگین نمونه‌ای را در مجموعه داده‌های SAS mdata خروجی می‌دهد.

سایر: CHART، PLOT، PRINT، TIMEPLOT.

آماره‌های ساده

UNIVARIATE خلاصهٔ مفصل‌تری از متغیرهای عددی همراه با آزمون نرمال‌بودن، دیگر ویژگی‌های توزیعی، و نمودارهای تشخیصی. این نمودارها عموماً چندان سودمند نیستند.

```
proc univariate ;
vars x1-x5 ;
by group ;
```

خلاصهٔ تک‌متغیرهٔ X_1، ... X_5 در هر سطح متغیر group.

TTEST آزمون tی دونمونه‌ای. یک متغیر دوسطحی از نوع class مورد نیاز است. برای اجرای یک آزمون tی جفتی، نخست به ایجاد تفاضل‌ها (در یک گام DATA) و سپس استفاده از proc means ذکرشده در بالا نیاز است.

```
proc ttest ;
class group ;
var x1-x5 ;
```

CORR آلفای کرونباخ (فصل ۳، همبستگی) را می‌توان با استفاده از proc corr همان‌طور که در زیر نشان داده شده به‌دست آورد. دستور with <variable list>; اختیاری است. در صورت استفاده‌نکردن از آن، همهٔ همبستگی‌های جفتی با دستور var <vartiable list> ; تولید می‌شود.

```
proc corr data=a alpha ;
var x1-x5 ;
with y1-y4 ;
```

سایر: SUMMARY، TABULATE.

رگرسیون

REG یک مدل رگرسیونی خطی برازش دهید. همهٔ رگرسورهای گنجانده‌شده درمدل را باید از پیش تعریف کرد (مثلاً، $X_{12} = X_1 * X_2$).

```
proc reg ;
model y = x1 x2 x12 / all ;
output out=rdata p=yhat r=resid ;
```

گزینهٔ all مستلزم چاپ آماره‌هایی اضافی است. خروجی آماره‌های مشخص‌شده برای یک مجموعه‌دادهٔ SAS از طریق keyword = name در دسترس است. در اینجا نام مجموعه‌داده‌ها rdata است، و p= r= کلیدواژه‌هایی هستند که مقادیر پیش‌بینی‌شده و مانده‌های مدل را به‌ترتیب در متغیرهای yhat و resid به عنوان خروجی به‌دست می‌دهند.

RSREG یک رگرسیون رویهٔ پاسخ برازش می‌دهد و یک تحلیل کانونی (ویژه‌بردار) از رویه انجام می‌دهد. دستور مدل زیر معادل با مدل رگرسیونی مشخص‌شدهٔ بالاست.

```
proc rsreg ;
model y = x1 x2 ;
```

CALIS, NLIN, ORTHOREG, TRANSREG.

سایر: CALIS، NLIN، ORTHOREG، TRANSREG.

طرح آزمایشی ANOVA برای طرح‌های متعادل، ANOVA کارآمدتر از GLM است اما اگر برای تحلیل‌های بعدی نیاز به آماره‌های خاصی داریم، ممکن است ملزم به استفاده از دستور GLM باشیم. همچنین وقتی اثرهای تصادفی و آزمون‌های F غیر استاندارد دخیل می‌شوند، کاربر ملزم به انجام یک تحلیل مناسب با استفاده از ANOVA است.

```
proc anova ;
class group trt ;
model y = group trt group*trt ;
means group trt / t lines ;
```

ANOVAی دوطرفه با اثر متقابل. مجدداً، کاربر مسئولیت تفسیر درست خروجی را بر عهده دارد. مقایسه‌های چندگانه با استفاده از آزمون‌های tی جفتی با تفاوت‌های معنی‌دار، با گزینهٔ lines نشان داده می‌شود.

GLM از این شیوه می‌توان برای برازش مدل‌های خطی کلی و بنابراین در اصل تکرار تحلیل

موارد خاص همچون رگرسیون (REG) و طرح‌های متعادل (ANOVA) استفاده کرد. با این حال، GLM ویژگی‌های خاصی دارد که در این شیوه‌های اختصاصی در دسترس نیست و به‌طور خاص هنگام سروکارداشتن با داده‌های نامتعادل طرحریزی شده است.

```
proc glm ;
class group trt ;
model y = group trt group*trt ;
lsmeans group trt ;
means group trt ;
random group group*trt / test ;
test h=group e=group*trt ;
output out=gdata r=resid ;
```

در این مثال، group به‌عنوان یک اثر تصادفی در نظر گرفته می‌شود (GLM به‌طور خودکار اثرهای متقابل را تصادفی در نظر نمی‌گیرد) و lsmeans معادلی برای means فراهم می‌کند، با این فرض که طرح متعادل در نظر گرفته شده بود. دستور test(که در ANOVA نیز موجود است) استفاده از یک جملهٔ خطای مناسب را برای آزمودن group میسر می‌سازد. خاطرنشان می‌کنیم که آزمون‌های F مناسب برای یک اثر تصادفی را می‌توان از طریق option / test در دستور random نیز انجام داد. در نهایت، دستور output این امکان را فراهم می‌کند که بتوان مانده‌های مدل را بعداً آزمون کرد.

سایر: MIXED، NESTED، NPARIWAY، PLAN، VARCOMP، GLMMOD، LATTICE، MULTTEST.

داده‌های رسته‌ای LOGISTIC وقتی متغیر پاسخ دودویی یا ترتیبی است، یک مدل رگرسیونی خطی مبتنی بر تابع لوجیت$g(p) = p/(1-p)$ که در آن $p = P[Y = 1|X]$ ممکن است مناسب باشد. به عبارتی، مدلی به فرم $g(p) = \alpha + \beta' X$.

```
proc logistic outest=beta ;
model y = x1 x2 / ctable details ;
output out=ldata p=phat ;
```

فرمت دستور model مشابه با REG است (بدان معنی که همهٔ متغیرهای تبیینی مشخص‌شده در دستور model را باید از پیش تعریف کرد). گزینه‌های ctable details مخصوص مدل لوژستیک‌اند.

CATMOD این شیوه معادلی همه‌منظوره برای GLM به‌منظور تحلیل داده‌های رسته‌ای

است. با استفاده از آن می‌توان گستره‌ای از مدل‌های رسته‌ای به‌انضمام مدل‌های لوژستیک (بالا) و لگ‌خطی را برازش داد. مثال زیر یک مدل لگ‌خطی به جدول پیشایندی مشخص‌شده توسط $A \times B$ برازش می‌دهد.

```
proc catmod ;
response marginals ;
model a*b = _response_  ;
```

CORRESP, GENMOD, PROBIT.

سایر: CORRESP، GENMOD، PROBIT.

روش‌های چندمتغیره CLUSTER، TREE، ACECLUS، FASTCLUS، MODECLUS، DISCRIM، CANDISC، PRINCOMP، CANCORR، FACTOR، CALIS، MDS.

طیف بسیار وسیعی از شیوه‌های چندمتغیره موجودند که توسط SAS ارائه می‌شوند و راهنمای SAS/STAT (SAS 1990) به‌وضوح بهترین مکان برای یادگیری جزئیات خاص و نحو مرتبط با این شیوه‌هاست (در مطالعهٔ موردی ۴.۸، نحوهٔ استفاده از proc cluster تشریح شده است). برخی از خوانندگانی که با SAS ناآشنا هستند ممکن است متوجه شده باشند که هیچ شیوهٔ MANOVA فهرست نشده است. این از آن روست که مدل‌های MANOVA در شیوهٔ GLM (یا ANOVA و MIXED) تعبیه شده‌اند. تشخیص تحلیل MANOVA تا حدودی مرموز است، بنابراین اهمیت دارد که نحو را به‌دقت دنبال کنید. همچنین باید خاطرنشان کنیم که SAS گرایش به تولید «انبوهی» از خروجی از شیوه‌های چندمتغیره دارد. خود را برای خواندن زیاد آماده کنید!

سری زمانی نیاز به نرم‌افزار *SAS/ETS* دارد.

ARIMA رویکرد باکس–جنکینز به برازش مدل‌های ARIMA(p, d, q) (به‌انضمام صورت فصلی) در شیوهٔ ARIMA تعبیه شده است. گام‌های تفاضل‌گیری، برآورد پارامترها، و پیش‌بینی جداگانه مشخص شده‌اند و بنابراین معمولاً عبور از چندین مرحله برای به‌دست‌آوردن مدلی مناسب مورد نیاز است.

```
proc arima ;
i var=y(1,12) ;
e p=(1)(12) q=2 plot;
f out=adata lead=12 id=t ;
```

مثال بالا مدل چندمتغیرهٔ فصلی را، که با $(۱,۱,۲) \times (۱,۱,۰)_{۱۲}$ نشان داده

می‌شود، برازش می‌دهد و از این مدل برای پیش‌بینی ۱۲ دورهٔ زمانی بعد استفاده می‌کند. خاطرنشان می‌کنیم که q=2 صورت مختصرشدهٔ q=(1,2) است و گزینهٔ plot همبستگی‌نگارها و آزمون کلی‌نگر مانده‌ها را تولید می‌کند.

سایر: SPECTRA، STATESPACE، X12.

سیستم‌های اضافی همان‌طورکه در بالا مشخص شد، شیوه‌های تحلیل سری‌های زمانی در مدول نرم‌افزاری SAS/STAT گنجانده نشده است. بنابراین، برای استفاده از proc arima، باید مدول SAS/ETS روی سیستم کاربر نصب شود. با وجود اینکه محصولات اختصاصی SAS بسیار دیگری نیز در دسترس‌اند، ما توجه خود را به مدول‌های QC و GRAPH معطوف می‌کنیم.

QC نرم‌افزار SAS/QC گستره‌ای از شیوه‌های کنترل کیفیت فراهم می‌کند که برای یک مشاور آماری، که روی پروژه‌هایی کار می‌کند که در آن‌ها تضمین کیفیت نقش مهمی بازی می‌کند، سودمند خواهد بود. در زیر مثالی از یک نمودار گردشی[1] شوهارت $\overline{X}$-R بر روی مجموعه‌داده‌های spray ارائه می‌شود که در آن batch نشان‌دهندهٔ نمونه‌ای است که از اندازه‌گیری پاسخ ingred به‌دست آمده و tests=1 to 8 چندین آزمون تشخیصی نمودار گردشی را اجرا می‌کند.

```
proc shewhart history=spray graphics;
xrchart ingred*batch /
   tests=1 to 8
   tableall
   zonelabels;
```

GRAPH نمودارهای با کیفیت ارائه‌کرد بخش مهمی از یک بستهٔ نرم‌افزاری آماری باارزش هستند و سیستم SAS/GRAPH این قابلیت را فراهم می‌کند. در حالی که یقیناً می‌توان از طریق شیوه‌های «G» (مثلاً GPLOT، GCONTOUR، GCHART، G3D) نمودارهایی با وضوح بالا به‌دست آورد، به عقیدهٔ ما انعطاف‌پذیری و سهولت فراهم‌شده توسط نرم‌افزار S-PLUS (که در بخش بعدی مورد بحث قرار می‌گیرد) قابل مقایسه با این شیوه‌ها نیست. همراه با نمودارهای با وضوح بالا باید درایور گرافیکی خاصی مشخص شود و در مثال زیر، دو حالت ممکن با dev=XCOLOR (مناسب استفادهٔ تعاملی SAS روی یک ایستگاه کاری یونیکس) و dev=PS (که یک فایل PostScript تولید می‌کند)

1. run chart

مشخص می‌شود. البته در عمل فقط باید از یک دستور goptions استفاده کرد. باقیماندهٔ مثال، یک نمودار پراکنش با خط رگرسیونی خطی ساده و حدود اطمینان ۹۵٪ را روی نمودار رسم می‌کند. می‌توان از عنوان‌ها، فونت‌ها،رنگ‌ها،و نشانه‌های مختلفی برای بهبود نمودار استفاده کرد.

```
goptions dev=XCOLOR ; * <-- for interactive use ;
                      *     on UNIX workstation ;
filename gout1 '~/gexample.ps' ;
goptions gsfname=gout1 gsfmode=append dev=PS ;

filename in1 '~/datafilename' ;

data a ;
    infile in1 ;
    input x1-x4 ;
    rename x4=y ;

title1 f=xswiss 'Example of GPLOT for SLR' ;
title2 'SAS Graphics' ;
footnote '95% Confidence Interval Overlaid' ;

proc gplot data=a ;
      plot y*x3 ;
symbol1 c=red i=rlclm v=diamond ;
run;
```

ASSIST، INSIGHT. علاوه بر راهنماها و متون منتشر شده توسط SAS، یک سیستم تعاملی help موجود است که جزئیاتی دربارهٔ همهٔ شیوه‌های SAS فراهم می‌کند. این نرم‌افزار رابط با کاربر در مدول SAS/ASSIST در دسترس است. در صورت تعبیهٔ آن، دستیابی به SAS ASSIST از منوی آبشاری «Options» میسر خواهد بود. در نهایت، باید خاطرنشان کنیم که یک نسخهٔ منویی از SAS موسوم به SAS INSIGHT موجود است.

ب.۴.۱ جزئیات بیشتری از SAS

مشخص‌سازی اثرها

SAS از نمادگذاری زیر برای مشخص‌سازی شیوه‌های تحلیل واریانس اثرهای متقاطع و آشیانی استفاده می‌کند.

متقاطع A*B هر سطح A در هر سطح B مشاهده می‌شود.

آشیانی B(A) سطوح B در هر سطح A آشیانی می‌شود.

نمادگذاری میله‌ای A|B نمادگذاری مختصرشده‌ای برای A + B +AB است.

مجموعه‌داده‌های دائمی SAS

برای مجموعه‌داده‌های کوچک تا متوسط،خواندن یک فایل دادهٔ خارجی در هر بار اجرای برنامهٔ prog.sas محتملاً مورد توجه زیادی است. با این حال، برای مجموعه‌داده‌های بزرگ، این کار نسبتاً ناکارآمد است و باید یک مجموعه‌دادهٔ دائمی SAS از طریق دستور libname ساخته شود. مثال زیر این کار را برای اجرای SAS تحت یونیکس تشریح می‌کند.

```
options ls=78 ;
libname scp '~/sasuser/job1' ;
filename in1 'example.dat' ;
data a;
   infile in1 ;
   input x y ;
data scp.xydat ;
   set a ;
run ;
endsas ;
```

دایرکتوری ~/sasuser باید قبلاً توسط SAS ساخته شده باشد و ما می‌توانیم زیردایرکتوری‌هایی مانند job1/ را برای پروژه‌های جداگانه ایجاد کنیم. *libref* موسوم به scp، SAS را به آن زیردایرکتوری خاص ارجاع می‌دهد و بنابراین در دومین گام DATA یک مجموعه‌دادهٔ دائمی SAS موسوم به xydat.ssd01 در آنجا تولید می‌شود (پسوند آن توسط SAS ایجاد می‌شود). از آنجا که این مجموعه‌دادهٔ دائمی دربردارندهٔ همهٔ اطلاعات مورد نیاز SAS است، برای تحلیل کافی است که به آن فایل دسترسی پیدا کنیم. به عبارتی، فایل برنامهٔ prog2.sas زیر را می‌توان اکنون مستقیماً برای تحلیل داده‌ها به‌کار برد.

```
options ls=78 ;
libname scp '~/sasuser/job1' ;
data a ;
   set scp.xydat ;
proc means ;          * The catch is that we need to know ;
   var x y ;          * what variables exist in "xydat".  ;
endsas ;              * Use "proc contents" to find out.  ;
```

انتقال فایل‌های SAS

انتقال مجموعه‌داده‌های دائمی SAS بین کامپیوترها ممکن است کاراتر از شروع از داده‌های خام باشد. اغلب، قبل از ایجاد یک مجموعه‌دادهٔ دائمی SAS، مقدار معینی پاکسازی داده‌ها انجام شده است و بازتولید داده‌های خام ارزشمند یا حتی معقول نخواهد بود.

دربرخی موارد، ممکن است بتوان مجموعه‌داده‌های SAS را مستقیماً از طریق یک شبکهٔ ارتباطی انتقال داد. در موارد دیگر، مجموعه‌داده‌های SAS را باید به فرمتی موسوم به فرمت انتقال[1] یا به شکل Export بازنویسی کرد. بسته به سیستم عامل، فرمت‌های انتقال را می‌توان از طریق گام DATA یا یکی از شیوه‌های PROC COPY، XCOPY، یا CPORT با گزینه‌های مناسبی همچون export ایجاد کرد. برای خواندن یک مجموعه‌دادهٔ SAS در فرمت انتقال، از یکی از دستورهای بالا همراه با گزینهٔ import یا در صورت لزوم از PROC CIMPORT استفاده کنید. با مجموعه‌داده‌های بزرگ SAS، ممکن است لازم باشد نخست صورت Export را از حالت فشرده خارج کرد.

ماکروهای SAS

SAS یک قابلیت MACRO فراهم می‌کند که ما را قادر به ساختن برنامه‌هایی می‌کند که به طور خاص برای انجام‌دادن یک شیوه یا تحلیلی طراحی می‌شود که اغلب مورد استفادهٔ ما قرار می‌گیرد. نکتهٔ کلیدی در اینجا این است که یک ماکروی SAS باید به‌دفعات مورد استفاده باشد تا تلاش ما ارزش داشته باشد. این حالت برای موردی در محیط‌های مشاوره‌ای پژوهشی یا تجاری صادق است که با مجموعه‌داده‌های مشابه سروکار دارند، اما این مسئله در مورد مشتریانی که تنها یک بار از برنامه استفاده خواهند کرد صدق نمی‌کند. ما در این کتاب از هیچ ماکروی SAS استفاده نکرده‌ایم.

ب.۲ S-PLUS

S-PLUS از زبان برنامه‌نویسی S که به همت بکر و چمبرز (۱۹۸۴) و بکر و همکاران (۱۹۸۸) در آزمایشگاه‌های AT&T ایجاد شد نشأت گرفت. بنابراین، با اینکه S-PLUS کاربرد عملی آماری بیشتری از S دارد، رویکرد زمینه‌ای آن برنامه‌محور باقی مانده است. یکی از عمده‌ترین نقاط قوت S-PLUS این است که می‌توان از آن به‌طریقی کارآمد در یک محیط تعاملی استفاده کرد. افزون بر آن، S-PLUS قابلیت‌های گرافیکی بسیار خوبی در اختیار کاربر قرار می‌دهد. با این حال، بهای این انعطاف‌پذیری این است که کاربر برای انجام یک تحلیل آماری کامل باید به میزان زیادی «برنامه‌نویسی» انجام دهد.

1. transport format

ب.۲.۱ مقدمات S-PLUS

S-PLUS یک برنامهٔ کاربردی سطح بالای شی‌مگرا است و به محض اینکه کاربر درک خوبی از برخی فرمان‌های پایه‌ای S-PLUS به‌دست آورد، ساختن توابع مورد نظر وی سرراست خواهد بود و محیط غنی و توسعه‌پذیری برای تحلیل داده‌ها گشوده می‌شود. به عبارتی، کاربر اسیر محدودیت‌های تابع‌های خودساختهٔ S-PLUS نمی‌شود، اما از آنجا که گردایهٔ جامعی از این تابع‌های S-PLUS موجود است، کاربر درگیر مواجهه با الگوریتم‌های درونی (کد C) زمینه‌ای این شیوه‌ها نمی‌شود. پیش از توجه به هنر ساخت تابع‌های خود، نخست باید دو فرمان مهم را بشناسیم.

شروع و پایان S-PLUS

برای فراخوانی S-PLUS به کمک یک دستور یونیکس[۱]:،تایپ کنید

```
Splus
```

اگر اولین باری است که از آن استفاده می‌کنیم، دایرکتوری خاص "/.Data" ایجاد خواهد شد و پیامی که دقیقاً همان را می‌گوید ظاهر می‌شود. سپس علامت > ظاهر می‌شود. اکنون در حالت تعاملی هستیم و می‌توانیم به‌طرز مناسب دستورها را به فرم فرمان‌های S-PLUS تایپ کنیم. برای خارج‌شدن از S-PLUS، دستور ()q را کنار علامت > تایپ کنید:

```
> q()
```

البته، احتمالاً مایلیم کاری بیشتر از صرفاً شروع یا پایان‌دادن به یک اجرای S-PLUS انجام دهیم! بنابراین، با پوشش‌دادن برخی فرمان‌های مقدماتی S-PLUS با ارجاع به مجموعه‌داده‌های زیر آغاز می‌کنیم که فرض می‌شود در قالب فایل ex.dat روی دایرکتوری اصلی یونیکس ما قرار دارد. فایل ex.dat دربردارندهٔ درایه‌های زیر (به‌انضمام برچسب‌های ستونی) است:

مجموعه داده‌های x.dat

فایل ex.dat شامل درایه‌های زیر است (مشتمل بر برچسب‌های ستونی):

name	class	height	weight	size
clm	1	8	10	104
gag	2	1	3	11
gam	2	9	11	120
grc	1	2	4	12
mra	1	6	11	127
rmm	3	6	6	33
spm	3	6	7	52
tcs	3	10	12	150

۱. یونیکس نسبت به بزرگ یا کوچک بودن حروف حساس است، بنابراین Splus، splus یکی نیست.

ورود داده‌ها

پس از فراخوانی S-PLUS، اولین کار این است که مجموعه‌داده‌ها را در S-PLUS بخوانیم و متغیری را به آن تخصیص دهیم. به یاد داشته باشید که فرمان‌های زیر در کنار علامت < نوشته می‌شوند (خود علامت < را تایپ نکنید).

```
> x <- read.table("ex.dat", header=T, row.names=NULL)

    #   the assignment symbol "<-" consists of the
    #  "less than" key followed by the "dash" key

> print(x)
     name  class  height  weight  size
      clm      1       8      10   104
      gag      2       1       3    11
      gam      2       9      11   120
      grc      1       2       4    12
      mra      1       6      11   127
      rmm      3       6       6    33
      spm      3       6       7    52
      tcs      3      10      12   150
> q()
```

گرچه این جلسهٔ S-PLUS خیلی مختصر (!) است، برخی از ویژگی‌های عمومی یک جلسهٔ S-PLUS را تشریح می‌کند.

تخصیص چندین راه برای تخصیص یک مقدار به دیگری وجود دارد:

-> برای تخصیص یک متغیر به یک کمیت بکار برده می‌شود.

= فقط درون تابع‌ها برای جایگذاری مقداری برای گزینه‌ها بکار می‌رود.

تذکرها در S-PLUS، قبل از یک تذکر از علامت # استفاده می‌شود.

رشته‌ها رشته‌های کاراکتری مانند اسم یک فایل "ex.dat" باید درون علامت نقل قول قرار بگیرند.

منطقی‌ها [1] علامت منطقی T مخفف TRUE است و ارزش ۱ را دارد. به‌طور مشابه، F=FALSE و ارزش ۰ را دارد.

درونی‌ها [2] NULL مقدار خاصی است که در اصل یا هیچ حجمی از حافظه را به یک متغیر

1. Logicals 2. Intrinsics

اختصاص نمی‌دهد و یا یک تنظیمات شناسه مانند row names را به یک تنظیمات پیش‌فرض اختصاص می‌دهد.

تابع‌ها بسیاری از تابع‌های S-PLUS تنظیمات پیش‌فرضی را برگزینه‌ها اعمال می‌کنند که کاربر می‌تواند آن‌ها را مجدداً تنظیم کند: read.table("ex.dat", header=T). برخی تابع‌ها هیچ گزینه‌ای ندارند، اما (تقریباً) همهٔ آن‌ها یک شناسه دارند که کاربر باید آن را پر کند: print(x). یک استثنا q() است.

Data./ x به‌طور دائمی در دایرکتوری Data./ ذخیره خواهد شد.

ب.۲.۲ تنظیمات S-PLUS

S-PLUS شاید بیش از آنکه یک بستهٔ نرم‌افزاری باشد، یک محیط است. یک کاربر نوعی S-PLUS به‌طور تعاملی در این محیط کار می‌کند، داده‌ها را می‌خواند و می‌نویسد، متغیرها و تابع‌ها را ایجاد می‌کند، و نیز نتایج را از یک شیوهٔ آماری استخراج می‌کند. به همهٔ این‌ها شیء اطلاق می‌شود و باید به‌طور دائمی در Data./ تحت نام یکسان ذخیره شوند. بنابراین، در مثال بالا x به عنوان یک «فایل» و با محتوایی ذخیره‌شده به فرمت S-PLUS در Data./ موجود خواهد بود. از آنجا که Data./ ممکن است به‌سرعت توسط همهٔ اشیائی که ساخته‌ایم در هم و برهم شود، نخستین گام این است که پروژه‌های خود را از هم جدا کنیم.

در یونیکس، می‌توانیم به‌سادگی از طریق فرمان mkdir یک زیردایرکتوری بسازیم، بنابراین همهٔ کاری که باید انجام دهیم این است که تایپ کنیم: mkdir .Data(نقطهٔ "." ضروری است و به یاد داشته باشید که یونیکس نسبت به حروف بزرگ وکوچک حساس است). این دستور، .DATA/ را در دایرکتوری جاری، مثلاً project1/، ایجاد می‌کند و هرگاه که S-PLUS در project1/ فراخوانی شود، اشیای S-PLUS در project1/.Data/ ذخیره خواهند شد. اگر هیچ زیردایرکتوری Data./ وجود نداشته باشد، S-PLUS از Data./ در دایرکتوری بالاتر ما ($HOME) استفاده می‌کند. در S-PLUS خانگی، اشیا در زیردایرکتوری DATA_ ذخیره می‌شوند.

تابع‌ها(شیء‌ها)ی خودساختهٔ S-PLUS در یک دایرکتوری سیستمی، مانند /usr/local/splus که توسط اپراتور سیستم تنظیم می‌شود، قرار می‌گیرد. دایرکتوری سیستمی برای S-PLUS دربردارندهٔ چندین زیردایرکتوری است که هرگاه S-PLUS را از دایرکتوری جاری خود فراخوانی می‌کنیم «ضمیمه می‌شود». برای دریافتن اینکه این‌ها چیستند، S-PLUS را اجرا و تایپ کنید: search(). روی سیستم خود به‌دست می‌آوریم:

دایرکتوری‌های S-PLUS

```
> search()
 [1] "/home/faculty/mcdougal/.Data"
 [2] "/opt/local/ss/splus/splus/.Functions"
 [3] "/opt/local/ss/splus/stat/.Functions"
 [4] "/opt/local/ss/splus/s/.Functions"
 [5] "/opt/local/ss/splus/s/.Datasets"
 [6] "/opt/local/ss/splus/stat/.Datasets"
 [7] "/opt/local/ss/splus/splus/.Datasets"
 [8] "/opt/local/ss/splus/library/trellis/.Data"
```

«موقعیت‌ها»ی . . . [2] [1] ترتیبی را که S-PLUS بر طبق آن به جستجوی یک شیء می‌پردازد نشان می‌دهد. بنابراین، دایرکتوری Data./ ی ما نخستین جایی است که S-PLUS باید در آن به جستجوی یک شیء بپردازد. خاطرنشان می‌کنیم که ما «می‌توانیم» تابع print() خود را بسازیم ــ ایدهٔ خوبی نیست ــ[1] که روی تابع print() خودساختهٔ S-PLUS نوشته نخواهد شد. برای مشاهدهٔ اینکه چه توابعی در دایرکتوری تابع‌های S-PLUS گنجانده شده، تایپ کنید: ls(pos=2) (آماده باشید ــ فقط تعداد بسیار کمی وجود دارد!). یک فرهنگ اصطلاحات فنی تابع‌های S-PLUS در پایان این بخش آمده است. برای جزئیاتی دربارهٔ نحوهٔ استفاده و گزینه‌های متعدد موجود، از تابع راهنمای برخط help() در S-PLUS استفاده کنید. خاطرنشان می‌کنیم که گردایهٔ جامعی از مستندات S و S-PLUS از طریق http://lib.stat.cmu.edu در دسترس است.

پنجرهٔ پایانه‌ای S-PLUS (یونیکس)

در یونیکس، S-PLUS به‌طور نرمال از یک پنجرهٔ پایانه‌ای فراخوانی می‌شود و بنابراین به محض اینکه خروجی حاصل از S-PLUS روی این صفحه چاپ می‌شود، سطرهایی که تایپ می‌کنیم سرانجام ناپدید می‌شوند. مشابه با فایل prog.sas در SAS، می‌توانیم یک فایل برنامه نیز از فرمان‌های S-PLUS بسازیم که می‌توان آن را با استفاده از تابع source ("prog.spl") ارسال کرد (پسوند ".spl" برای سهولت در اینجا به‌کار رفته و نیازی به آن نیست). هر چیزی در فایل prog.spl درست همان‌طور که گویی ما سطرها را مستقیماً تایپ کرده‌ایم توسط S-PLUS اجرا خواهد شد و خطاها را نیز دربردارد. بنابراین یک راه خوب برای اجرای یک جلسهٔ S-PLUS از درون یک پنجرهٔ محافظ Emacs با پنجرهٔ محافظ دیگری شامل فایل «منبع[2]» prog.spl است.

۱. مشکل این است که نسخه‌ای از print() در Data./ ذخیره می‌شود و بنابراین دیگر نمی‌توان از نسخهٔ خودساختهٔ این تابع توسط S-PLUS، استفاده کرد.

2. source

پنجرهٔ گرافیکی S-PLUS (UNIX)

برای مشاهدهٔ گرافیک‌های با وضوح بالا، لازم است که یک دستگاه گرافیک را مشخص کنیم. چندین دستگاه برای استفادهٔ تعاملی و نیز دستگاه `postscript()` برای ذخیرهٔ گرافیک‌ها در یک فایل موجودند. دستگاه `printer()` در ترکیب با فرمان `show()` را می‌توان برای تولید گرافیک‌های متنی (که معمولاً بی‌ارزش‌اند) بکار برد. برخی از انواع دستگاه‌های گرافیکی تعاملی عبارت‌اند از:

```
X11() motif() openlook() hp2468() tek4105()
```

اکنون همهٔ گرافیک‌های با وضوح بالا در پنجرهٔ گرافیکی ظاهر خواهند شد که کنترل‌هایی برای چاپ و دیگر امکانات دارند.

ب.۳.۲ فرمان‌های مقدماتی S-PLUS

جلسهٔ S-PLUS زیر برخی از فرمان‌های مقدماتی را برای کار با یک فایل داده‌ای عددی ساده موسوم به `mat.dat` که متشکل از دو ستون پنج‌تایی است تشریح می‌کند. همهٔ فرمان‌های S-PLUSی که در اینجا تایپ شدند در واقع فرمان‌های استاندارد S هستند. فایل داده‌های `"mat.dat"` که توسط S-PLUS خوانده می‌شود در زیر نشان داده می‌شود. خاطرنشان می‌کنیم که دو ستون از اعداد در این فایل در فرمت شسته‌رفته‌ای نیستند، اما حداقل یک فاصله بین ستون مدخل‌ها در هر سطر وجود دارد. فایل را همواره باید با فشاردادن کلید RETURN (کاراکتر نامرئی `^M`) پایان داد.

```
unix.prompt% more mat.dat
   2 4
   -1  7
   0 4
    11 2
   1 1
```

حال S-PLUS را فراخوانی کنید و سعی کنید با استفاده از تابع `scan()`، فایل `"mat.dat"` را بخوانید.

ب.۴.۲ استفادهٔ کارآمد از S-PLUS

وقتی از S-PLUS خارج می‌شویم، تنها اشیائی که طی جلسه ایجاد کردیم حفظ و در دایرکتوری `Data/` ذخیره می‌شوند. جایی باید دایرکتوری `Data/` را با حذف متغیرهایی مانند `tmp`، `i`، `j`، ... که نیازی به آن‌ها نداریم پاکسازی کنیم. خاطرنشان می‌کنیم که هر یک از این متغیرها نیز یک فایل‌اند و بنابراین به فضای مینیمم معینی (مانند ۲۰۴۸ بایت) نیاز دارند که می‌توانند به‌سرعت سهمیهٔ مجاز ما را مصرف کنند.

```
unix.prompt% S-PLUS
> scan("mat.dat")
 [1]   2   4  -1   7   0   4  11   2   1   1
> # not quite right ... hence
>
> y_matrix(scan("mat.dat"), ncol=2,byrow=T)
> # data is now in variable "y"
> # typing y just outputs its contents.
> y
     [,1] [,2]
[1,]    2    4
[2,]   -1    7
[3,]    0    4
[4,]   11    2
[5,]    1    1
>                  # Now we can try some simple manipulations
> y[,1]                     # subsetting 1st column
[1]   2  -1   0  11   1
> y[1,]                     # subsetting 1st row
[1] 2 4
> a_rev(y[-2,1])    # assign variable "a" to the 1st column,
                    # but exclude the 2nd element, and then
                    # reverse the order of those 4 elements

> a                 # show contents of "a" ... it worked
[1]   1  11   0  2

> y[,1]+y[1,]       # add 1st col and 1st row ?? (the lengths
[1]  4  3  2 15  3  # are different!) S-PLUS does it, but
                    # provides a warning message ...
Warning messages:
  Length of longer object is not a multiple of the length of
  the shorter object in: y[, 1] + y[1,  ]

 # S-PLUS does the above vector addition operation by its own
 # special rules which were partly violated. Hence a warning.

> solve(y)              # S-PLUS won't do illegal operations
Error in solve.qr(a): matrix inverse only for square matrices
Dumped

> stem(y)  # do a stemplot of matrix y. Although y is matrix,
           # the S-PLUS function "stem" will treat y as a
           # vector here. Not all functions will do this.
```

```
N = 10   Median = 2
Quartiles = 1, 4

Decimal point is at the colon

  -1 : 0
  -0 :
   0 : z          # S-PLUS uses "z" for 0 on the 0 stem only
   1 : 00
   2 : 00
   3 :
   4 : 00
   5 :
   6 :
   7 : 0

High:  11          # S-PLUS puts outliers on a separate line

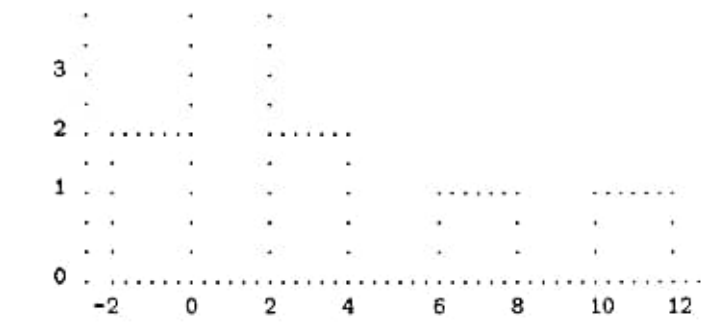
> args(stem)             # get the arguments of "stem" These
                         # can be used to modify the default
Stem-and-Leaf Display    # display. For full documentation
                         # use help(stem)
USAGE:
       stem(x, nl=0, scale=1000,
            twodig=FALSE, fence=2, head=TRUE, depth=FALSE)

> hist(y)                #  now try a histogram ... error?

Error in barplot(counts, width = breaks, histo = T..: Device
not active
Dumped
Error in barplot  # hist() actually uses barplot() to draw
                  # a histogram. However, barplot() requires
                  # a graphics device to be specified first.

                     # we will use the printer() device, but
> printer(60,20)     # make the plotting region smaller than
                     # the default settings: 80,64
> hist(y)            # hist can now be done, however device
> show()             # printer() needs show() to display

                     # result is a rather low quality display
```




```
> postscript("g.hist.ps")   # to get high quality printout,
> hist(y)                   # use postscript() device which
> q()                       # creates the PostScript file
                            # "g.hist.ps" on exit from S-PLUS
```

دایرکتوری Data./

هر شیء، تابع، یا متغیری که تولید می‌کنیم (به‌استثنای متغیرهایی که داخل یک تابع می‌سازیم) به‌طور خودکار در یک فایل خاص در Data./ ذخیره می‌شوند. پس از یک دورهٔ زمانی، تعداد شیءهای واقع در Data./ ممکن است بسیار بزرگ شود. همچنین فایل Audit. که در اصل هر حرکت کلیدی را که طی یک جلسهٔ S-PLUS انجام می‌دهیم ثبت می‌کند، ممکن است بسیار بزرگ شود. از آنجا که هیچ‌کس از آن‌هایی که تاکنون می‌شناسیم از فایل Audit. استفاده نمی‌کند، شما باید آن را از طریق rm Data/Audit. با با اختصاص متغیرهای محیطی مناسبی که اندازه‌ای برابر یا صفر دارند حذف کنید.

S-PLUS همهٔ اشیائی را که ما تولید می‌کنیم ذخیره می‌کند، اما بدون هیچ هشداری! به ما اجازه می‌دهد که روی هر شیئی در Data./ فایل دیگری را ذخیره کنیم. بنابراین، باید سعی کنیم نام‌های مناسبی برای اشیائی که نمی‌خواهیم روی آن‌ها فایل دیگری نوشته شود انتخاب کنیم. مزیت این کار این است که وقتی یک مجموعه‌داده را خواندیم و آن را به یک متغیر خوش‌نام، مثلاً a1.org.dat اختصاص دادیم، آنگاه لازم نیست دوباره داده‌ها را بخوانیم. صرفاً تخصیص y <- a1.org.dat را انجام می‌دهیم و به تحلیل با استفاده y (که تایپ‌کردن آن ساده‌تر است) به پیش می‌رویم.

انتخاب یک نامگذاری مناسب می‌تواند بسیار سودمند باشد. یک رویکرد این است که پسوند (یا پیشوند)ی به اشیائی که می‌خواهیم از آن‌ها به سبکی نظام‌مند نگهداری کنیم بیفزاییم. به‌عنوان مثال، اگر aname نشان‌دهندهٔ شیئی باشد که می‌خواهیم آن را نگه داریم،از این‌ها استفاده می‌کنیم:

شیء قالب داده‌ای یا ماتریس داده‌ای	d.aname.dat
تابع داده‌ای برای پردازش شیء قالب داده‌ها	d.aname.fun
تابع برای خروجی نتیجه	f.aname.out
تابع عمومی	f.aname.fun
تابع برای اجرای یک شیوه	f.aname.run
تابع گرافیکی برای تولید نمودار	g.aname.fun

فایل برنامهٔ S-PLUS

در ایجاد یک برنامهٔ S-PLUS سه قسمت عمده دخیل‌اند:

- ورود/خروج داده‌ها
- ایجاد تابع‌هایی برای تحلیل داده‌ها
- ایجاد گرافیک‌های ارائه‌کرد

برنامهٔ S-PLUS این یک فایل، مثلاً prog.spl[1] است که مثلاً با استفاده از Emacs تولید شده که دربردارندهٔ فهرستی از دستورات S-PLUS است که طبق اهداف کاربر نوشته شده است. با استفاده از تابع S-PLUS

```
> source("prog.spl")
```

همهٔ دستورات موجود در فایل prog.spl درست همان‌طور اجرا خواهند شد که گویی هر سطر را جداگانه تایپ کرده‌ایم. به‌وضوح، اگر اشتباهی رخ دهد، S-PLUS آن را بیان خواهد کرد و معمولاً پردازش بقیهٔ دستورات موجود در فایل را متوقف خواهد کرد.

باگ‌زدایی باگ‌زدایی ساده متشکل از اجرای برنامهٔ S-PLUS و سپس استفاده از Emacs برای ویرایش جایی است که خطا در آن به وقوع پیوسته است. چندین بار تکرار این کار سرانجام کل برنامهٔ S-PLUS را باگ‌زدایی خواهد کرد. روشی بهتر آن است که برنامهٔ S-PLUS را به گردایه‌ای از تابع‌های S-PLUS بشکنیم تا بتوان آن‌ها را به‌تنهایی باگ‌زدایی کرد. تابع‌ها باید به شکل کلی نوشته شوند به‌قسمی که بتوان از خروجی یک تابع به‌راحتی به‌عنوان ورودی تابعی دیگر استفاده کرد.

خطاها خطاهای ناشی از source("prog.spl") < به‌راحتی رفع می‌شوند، اما یافتن آن‌ها دشوار است، زیرا خطای نحوی واقعی ممکن است مدت‌ها پیش از اتمام اجرای

۱. برنامهٔ S-PLUS خود را به‌صورت «prog.s» نامگذاری نکنید. Emacs ممکن است با «.s*» به‌عنوان یک فایل زبان اسمبلی رفتارکند و ممکن است اتفاق عجیبی بیفتد.

S-PLUS اتفاق افتاده باشد. مثال‌هایی از چیزهایی که باید وارسی کرد عبارت‌اند از:

```
ym = mean(y)
sqrt(mean(ym)
{{ [[ ( ) ]] }}
# ym <- mean(y)
"missing }"
```

تخصیص مستقیم توسط دستور مقابل انجام می شود: - یا - >
")" موجود نیست
اطمینان حاصل کنید که () [] {} متقارن اند
تذکرها: همه چیز پس از # نادیده گرفته می شوند
S-PLUS منبع سازی نمی‌کند؟
یک RETURN پس از آخرین "}" قرار دهید

رفع خطاهای برنامه‌نویسی (بعد از اینکه خطاهای منبع مرتفع شدند) دشوارتر است زیرا آن‌ها معمولاً دلالت بر یک عمل غیر مجاز یا محاسبه‌ای نامناسب دارند.

تعمیم افزودن گزینه‌هایی به تابع‌هایی که تولید می‌کنیم این تابع‌ها را برای استفاده انعطاف‌پذیرتر خواهد ساخت. به‌طور مشابه، باید در تابع‌های خود از «کدنویسی سخت» پرهیز کنیم. به عبارتی، `x[-length(x)]` به `x[-65]` ترجیح داده می‌شود زیرا طول x همیشه برابر با ۶۵ نیست.

تحلیل اغلب تابع‌های برنامه‌ای که ایجاد می‌کنیم احتمالاً به‌منظور دستکاری داده‌ها و فرایندهای گرافیکی است. تابع‌های S-PLUS موجود برای تحلیل داده‌ها و مدلبندی معمولاً برای استفادهٔ تعاملی مناسب‌تر است زیرا در هرگام از یک تحلیل فقط تعدیل‌هایی جزئی صورت می‌گیرد. تابع‌هایی برای اجرای یک تحلیل کامل را می‌توان همان‌طور که تحلیل را پیش می‌بریم ایجاد کنیم. در پایان می‌توان آن را اجرا کرد تا همهٔ خروجی‌های لازم برای گزارش را به‌دست آورد.

خروجی به محض اینکه همهٔ بخش‌های برنامهٔ S-PLUS باگ‌زدایی شدند، بخش خروجی وارسی می‌شود. متأسفانه، خروجی حاصل از شیوه‌های S-PLUS فرمتی مینیمال دارند و احتمالاً باید بعداً برای ویرایش یک فایل خروجی بسازیم. در برخی موارد، ارزشمند است که تابع‌هایی ایجاد کنیم که کارهای فرمت‌بندی خاصی را انجام می‌دهند، اما این کار زمان‌بر است و ممکن است آسان‌تر این باشد که صرفاً نتایج S-PLUS را کپی و در فایل خروجی بچسبانیم و آن را ویرایش کنیم.

گرافیک‌ها برای گرافیک‌های با کیفیت ارائه‌کرد،ایجاد یک تابع گرافیکی جداگانه در فایل برنامه‌ای دیگر، رفع اشتباهات آن و تغییر نمودار را ساده‌تر خواهد ساخت. برای تولید نمودار نهایی و ذخیرهٔ تغییراتی که مبتنی بر نمایش نشان‌داده‌شده در پنجرهٔ گرافیکی استفاده می‌شود، از یک مرحلهٔ باگ‌زدایی مشابه با آنچه در بالا ذکر شد استفاده خواهد شد. نسخهٔ نهایی را نیز باید در فرمت PostScript ذخیره کرد و بنابراین می‌توان در صورت نیاز بعداً آن را مجدداً چاپ کرد.

ایجاد تابع‌های S-PLUS

تاکنون تنها از تابع‌های خودساختهٔ S-PLUS استفاده کرده‌ایم. بااین حال، قدرتمندترین مشخصهٔ S-PLUS این است که می‌توانیم تابع‌های خودمان را تولید کنیم. دو مثال ارائه می‌شوند.

مثال ب.۵ یک تابع ساده برای خواندن داده‌ها

نخستین تابع S-PLUS ما read.dat() نام دارد و کد S-PLUS زیر همهٔ آن چیزی است که در فایل prog.spl ما گنجانده شده است. می‌توان این تابع را مستقیماً در S-PLUS ایجاد کرد، اما هر اشتباهی به معنی تایپ دوبارهٔ کل تابع است.

```
read.dat <- function(file="",nc=1)
{                          # first check whether a file
   if(!missing(file)) {    #     has been specified
      if(nc == 1)          # if so, default is to assume
          y_scan(file)     #     input as a vector
      else
          y_matrix(scan(file), ncol=nc, byrow=T)
      return(y)            # return the result. Here
   }                       # "y" is a temporary variable
   else {                  # and is NOT saved in /.Data
      cat("No input file specified\n")
   }
```

خاطرنشان می‌کنیم که وقتی تنها یک دستور در دستور if() and or else پردازش می‌شود، نیازی به { } نیست. تابع cat() ما را قادر می‌سازد که پیغام‌های سودمندی را روی صفحه چاپ کنیم. در اینجا نتیجهٔ اجرای این برنامه در S-PLUS آمده است.

```
> source("prog.spl")
> read.dat()
No input file specified
> read.dat("mat.dat")
 [1]   2   4  -1   7   0   4  11   2   1   1
> read.dat("mat.dat",nc=2)
     [,1] [,2]
[1,]    2    4
[2,]   -1    7
[3,]    0    4
[4,]   11    2
[5,]    1    1
> read.dat("mat.dat",nc=5)
     [,1] [,2] [,3] [,4] [,5]
[1,]    2    4   -1    7    0
[2,]    4   11    2    1    1
> y_read.dat("mat.dat",nc=2)
```

به یاد داشته باشید که باید برای استفاده از دستور ورود داده‌ها در S-PLUS یک متغیر را تخصیص دهیم و این y در Data./ ذخیره خواهد شد. تابع read.dat() کامل نیست (3 یا nc = -1را امتحان کنید)، و تا حدودی در نمایش فرمان S-PLUS read.table() زائد است (چرخ را دوباره اختراع نمی‌کنیم). بااین حال، مشخصه‌ها (و مشکلات) ابتدایی مرتبط با ایجاد تابع‌های دلخواهمان را تشریح می‌کند.

مثال پ.۶ یک تابع گرافیکی ساده

یک مشکل با گرافیک‌ها این است که هر نموداری که می‌خواهیم بدون استثنا با آخرین نمودار متفاوت است و این مسئله، نوشتن یک تابع رسم نمودار را غیر عملی می‌کند (گزینه‌های بسیار زیادی موجودند). با این حال، یک کالبد ساده که ممکن است سودمند باشد در زیر می‌آید.

```
gxy.fun_function(x,y,tk=1,skip=0,psfile="g.slr.ps",msg="")
{
    switch(tk,              # the switch variable is tk
       ,                    # default is to specify no
       printer(70,25),      # graphics device (option tk=1)
       tek4105(),           # since we may have already
       X11(),               # specified one.
       postscript(psfile) )
#-------------------------------------------------------------
# Any graphical display of interest to the user can go here.

    plot(x,y,pch="*")           # do a scatter plot
    if(skip == 1)               # if skip=1, add SLR fit
        abline(lsfit(x,y)$coef)
    if(!missing(msg))
        title(msg)       # put a title on plot if supplied
}
```

ایدهٔ اصلی این تابع این است که می‌توان آن را با استفاده از دستگاه‌های گرافیکی مختلف اجرا کرد. به‌عنوان مثال، قراردادن tk=5 بااستفاده از نام فایل پیش‌فرض g.slr.ps یا قراردادن "psfile="new_name در سطر گزینه‌ها، یک فایل PostScript ایجاد می‌کند. حتی در این مورد ساده می‌توان ملاحظه کرد که تعداد بی‌شماری گزینه مانند onefile=T وجود دارد که می‌توان اضافه کرد. می‌توان این را در سطر گزینه‌های gxy.fun تعبیه کرد تا کاربر قادر به تولید یک فایل PostScript جداگانه برای هر نمودار باشد (با این فرض که نمودارهای بیشتری به این تابع افزوده شده باشد).

چاپ خروجی S-PLUS

خروجی گرافیکی S-PLUS را می‌توان در فایل‌های PostScript قرار داد. خروجی داده‌ها مبحث کمی پیچیده‌تری است زیرا نمی‌خواهیم کل یک جلسهٔ S-PLUS را در فایل انبار کنیم. به عبارتی، تنها باید بخش‌های مرتبطی از تحلیل را نگه داشت. ساده‌ترین روش این است که S-PLUS را در Emacs اجرا کنیم و بخش‌های مرتبط جلسهٔ S-PLUS را کپی کنیم و در یک فایل خروجی مانند "al.out" بچسبانیم و al.out را چاپ کنیم. با این حال، در مورد مجموعه‌داده‌های بزرگ، این کار ممکن است کسل‌کننده باشد و باید مستقیماً خروجی را به فایل‌های خارجی ببریم. این کار را می‌توان از طریق تابع‌های S-PLUS "write.table(file="dat.out", x)" یا write(file="dat.out", x) انجام داد. روشی کندتر این است که از تابع cat() استفاده کنیم:

```
write(file="dat.out", x)
use the cat() function:

   for(i in 1:nrow(y)) {
      cat(file="dat.out", y[i,], append=T)
      cat(file="dat.out","\n",append=T)
   }
```

ب.۵.۲ شیوه‌های آماری S-PLUS

Data Frameها

data frameهای S-PLUS در مطالعهٔ موردی ۴.۶ مورد بحث قرارگرفتند و به کاربر اجازه می‌دهند که از شیوه‌های مدلبندی آماری به روال طبیعی‌تری استفاده کند. ما نخست برخی از مشخصه‌های ابتدایی یک data frame را به کمک فایل «mat.dat» که قبلاً معرفی شد تشریح می‌کنیم.

```
> a_read.table("mat.dat")
> a
   V1 V2   # read.table() works out the number of columns
1   2  4   # for itself and labels them as V1 V2
2  -1  7
3   0  4
4  11  2
5   1  1
> a[1,]    # subsetting is as before
  V1 V2
```

```
1  2  4
> mode(a)     # However "a" is now a dataframe object ...
[1] "list"    # hence V1 is  a "list" component of "a"
> a$V1                    # ==> a$V1 = a[,1]
[1]   2  -1   0  11   1
> a[,1]
[1]   2  -1   0  11   1

> summary(a)  # provides summary statistics for
              # dataframe objects

       V1              V2
 Min.   : -1.0   Min.   :1.0
 1st Qu.:  0.0   1st Qu.:2.0
 Median :  1.0   Median :4.0
 Mean   :  2.6   Mean   :3.6
 3rd Qu.:  2.0   3rd Qu.:4.0
 Max.   : 11.0   Max.   :7.0
```

برای تشریح نحوهٔ استفاده از S-PLUS برای برازش رگرسیون لوژستیک، مجموعه‌داده‌های زیر را در نظر بگیرید.

```
> d <- example.dat
> d  # print the dataframe object "example.dat"
     # read-in via "read.table()". To do this
     # using the "scan()" is more complicated.

Age Number Start    Y
 2    3      5    abs
12    3     25    pres
22    5      4    abs
 4    4     23    pres
15    5     31    abs
61    4     45    pres

> attach(d)   # a key feature of dataframes is that
              # the variables in "d" can now be
              # accessed directly. Thus, we can use
              # the glm() function to fit the logistic
              # regression model. Model statements in
              # S-PLUS are defined by the "~" character.

d.model <- glm(Y ~ Age + Start + Number,
               family = binomial, data = d )
```

حال تابع‌های زیر می‌توانند مؤلفه‌های متعددی از شیء `glm` با عنوان `"d.model"` را استخراج کنند.

```
residuals()  # residuals
fitted()     # fitted values
predict()    # predicted values
coef()       # regression coefficents
deviance()   # the model deviance
```

برای تشریح مفصل‌تر d.mode;، از این دستور استفاده کنید:

```
summary(d.model)
```

برای مقایسهٔ مدل‌های مختلف می‌توانیم متغیرها یا مشاهدات را به یک مدل اضافه یا از آن حذف کنیم:

تحلیل جدول‌های انحرافه:

```
d2.model <- update(d.model, ~ . - Age)
       # drops the Age variable

d3.model <- update(d.model, subset= -79)
       # drops the 79th observation
```

ب.۲.۶ فرهنگ اصطلاحات S-PLUS

برخی از فرمان‌ها و مشخصه‌های S-PLUS را فهرست می‌کنیم. برای اطلاعات بیشتر، از فرمان help() در S-PLUS استفاده کنید یا مراجع S و S-PLUS را ببینید.

توجه: همهٔ فرمان‌های تابعی S-PLUS برای فراخوانی یک فرمان نیازمند "()"یا (شناسهٔ ۱، شناسهٔ۲،...) هستند: length(x) طول بردار x را به‌دست می‌دهد. اگر فقط تایپ کنیم: length، در این صورت S-PLUS کد منبع تابع length را می‌دهد (که ممکن است سودمند باشد یا نباشد).

```
> length
function(x)
.Internal(length(x), "S_extract", T, 6)
```

کمک بر خط از طریق فرمان help() S-PLUS یا args() قابل دسترسی است.

```
> help(cmd)   full documentation on {\tt cmd}
> args(cmd)   brief documentation on {\tt cmd}
```

در برخی موارد، نیاز به گیومهٔ دوتایی است (مثلاً، help("[[")).

تابع‌های پایه‌ای S-PLUS

عملگرها

حساب	`+ - / * ^ ** %% %/`
مقایسه	`< <= > >= == !=`
تخصیص $x-۲$ یا $x<-۲$ (هر دو قرار می‌دهند $x=۲$)	`_ -> <- assign`
منطقی‌ها: و، یا، نه	`& | !`
داده‌های زیرمجموعه‌ای: چهار عنصر اول $x[۱:۴]$	`[] : seq`

صفت داده‌ها

بردارهای عددی سطر و ستون ماتریس	`col row`
طول و حالت بردار	`length mode`
بُعد بردار	`ncol nrow dim`
منطقی شناسهٔ تابع را بازمی‌گرداند	`missing`
صفت شیء S-PLUS را باز می‌گرداند	`attr attributes`
نوع حالت شیء S-PLUS را باز می‌گرداند	`mode`
نام‌های شیء ماتریس را بازگرداند	`names dimnames`

دایرکتوری داده‌ها

دایرکتوری را ضمیمه یا حذف کن	`attach detach`
فایل‌ها (متغیرها) را در دایرکتوری فهرست کن	`ls object`
متغیرها را حذف کن	`rm`
تنظیمات کنترل S-PLUS جاری را باز می‌گرداند	`options`
دایرکتوری‌های S-PLUS را فهرست کن	`search`
یک کتابخانه S-PLUS ضمیمه کن	`library`

دستکاری داده‌ها

بردار $x<-c(۱,۰,-۱)$ را تولید کن	`c`
با ترکیب ستون‌ها یا سطرها ماتریس تولید کن	`cbind rbind`
بردار را مرتب، تکرار، وارون یا جور کن	`order rep rev sort`
رشته‌های کارکتر را به‌هم بچسبان	`aste`
شیء فهرستی بردار خردشده را برحسب گروه بازمی‌گرداند	`split`

ساختارهای داده‌ای

matrix	ماتریس تولید کن
array	آرایهٔ چندبعدی تولید کن
list	فهرستی از شی‌ءهای S-PLUS تولید کن
ts	سری زمانی تولید کن
data.frame	یک شیء قالب داده‌ای تولید کن

تهیه مستندات

help args	مستندسازی کمک مختصر یا کامل

نمایش‌های گرافیکی

stem	نمودار ساقه و برگ
printer	نوع دستگاه (رزولوشن پایین)
show	گرافیک‌ها را نشان بده (تنها برای دستگاه ()printer)
X11 postscript	انواع دستگاه‌ها (رزولوشن بالا)ـضروری برای:
boxplot hist	بافت نگار نمودار جعبه‌ای
plot	(x,y) را رسم کن [با گزینه‌های زیاد]
tsplot	رسم نمودار سری زمانی
qqplot	نمودار Q-Q
contour	نمودار خط تراز

افزونه‌های گرافیکی

par	فهرست گستردهٔ گزینه‌های گرافیکی
abline axis box	:
lines points	خط‌ها، نقطه‌ها، عنوان‌ها، و نظایر آن‌ها را به نمودارها اضافه کن
arrows segments	:
legend title	:
text mtext symbols	:

فایل‌های I/O

print	متغیری را بر صفحهٔ نمایش پایانه‌ها چاپ کن
scan	یک فایل داده‌ای خارجی را بخوان
read.table	فایل داده‌ها را به‌صورت data.frame بخوان
write write.table	شیء یا قالب داده‌ها را در یک فایل خارجی بنویس
source	فایل فرمان S-PLUS با منبع خارجی
cat	خروجی ساده در صفحهٔ نمایش یا فایل خارجی
unix	یک فرمان UNIX اجرا کن (unix("ls"
!	ترک کار هسته‌ای UNIX: !"ls"

جبر خطی

%*% t	ضرب، ترانهادهٔ ماتریس
apply	اعمال تابع به سطرها یا ستون‌های ماتریس
diag	عنصرهای قطری ماتریس
eigen	ویژه‌مقدارها و ویژه‌بردارهای ماتریس

عملگرهای منطقی

all any	برای دستورهای شرطی بکار می‌رود
!= < <= == > >=	if((a&&b)||c){...}
if ifelse else	"if (a *and* b)"
&& || !	عملگرهای منطقی که در دستورهای اگر-در غیر این‌صورت بکار می‌رود

حلقه‌سازی و تکرار

apply sapply tapply	تابع را در مورد شیء S-PLUS بکار ببر
for(i in start:finish)	{... S-PLUS commands ... }
	حلقه‌ها در S-PLUS ممکن است بسیار کند باشد
	در صورت امکان از عمل‌های برداری استفاده کنید

عمل‌های ریاضی

abs ceiling floor trunc cos sin diff exp log log10
max min mean median order range rank round sort
sqrt sum prod var cor

ماتریس‌ها

cbind rbind col row diag dim dimnames ncol nrow

data.matrix	قالب داده‌ای را به ماتریس عددی تبدیل کن
dist	ماتریس فاصلهٔ جفتی
scale	ستون‌های ماتریس را مرکزی و مقیاس‌بندی می‌کند

توزیع‌های احتمال و اعداد تصادفی

*norm: p q r	کمیت‌های نرمال تولید کن:
pnorm	pnorm(2) = P[Z < 2]
qnorm	qnorm(0.5) ⇒ z st. P[Z<z]=0.5
rnorm	rnorm(10) ⇒ 10 quantiles from N(0,1)
*chisq *t	خی دو، توزیع t
*exp *unif	نمایی، یکنواخت
*cauchy *f	کوشی، توزیع F
qqplot qqnorm	نمودارهای Q-Q
sample	نمونه‌ای از یک بردار به‌دست آور

فنون استوار/مقاوم

mad mean median sabl smooth twoway

عمل‌های آماری

```
hist min max mean median prod sum range stem var cor
```

سری زمانی

```
diff lag sabl sablplot smooth ts tsplot
```

شیوه‌های آماری S-PLUS

تابع‌های عددی

`mean median`	میانگین و میانهٔ (پیراستهٔ) نمونه‌ای
`quantile`	چندک‌ها؛ پیش‌فرض خلاصهٔ ۵ عددی است
`var`	σ^2 یا $\hat{\sigma}$. برای به‌دست آوردن s از `sqrt (var(x))` استفاده کنید
`summary`	همهٔ موارد بالا را در اختیار می‌گذارد

آزمون‌های کلاسیک

`t.test`	آزمون tی تک‌نمونه‌ای، دونمونه‌ای، و جفتی
`cor.test`	آزمون همبستگی ($\rho = 0$)
`chisq.test`	آزمون خی‌دو: GOF یا جدول‌های پیشایندی
`fisher.test`	آزمون دقیق فیشر
`var.test`	آزمون F برای واریانس‌ها
`wilcox.test`	آزمون رشد علامت‌دار ویلکاکسون
`kruskal.test`	آزمون کروسکال‌والیس
`binom.test`	آزمون علامت
`prop.test`	آزمون z برای نسبت‌ها

رگرسیون

`lm`	یک مدل خطی برازش کن
`lsfit`	رگرسیون کمترین توان‌های دوم
`ls.summary`	خلاصهٔ رگرسیون کمترین توان‌های دوم
`l1fit`	رگرسیون $\mathcal{L}_1$
`rbiwt rreg`	رگرسیون استوار
`step`	انتخاب مدل قدم‌به‌قدم

انووا

`aov raov`	انووای با اثرهای ثابت یا تصادفی را محاسبه کنید

مدل‌های خطی تعمیم‌یافته

`glm`	مدل خطی تعمیم‌یافته
`gam`	مدل جمعی تعمیم‌یافته
`lm lme`	مدل خطی ساده یا پیچیده
`varcomp`	تحلیل مؤلفه‌های واریانس

گرافیک‌ها

trellis	تابع‌های نمودار تعمیم‌یافته

چندمتغیره

cancor	تحلیل همبستگی متعارف
hclust plclust	خوشه‌بندی سلسله مراتبی، نمودار درختی
kmeans	خوشه‌بندی k- میانگین
pairs	نموداری پراکنش جفتی
prcomp princomp	تحلیل مؤلفه‌های اصلی
tree	یک شیء درختی ایجاد می‌کند

فنون استوار/مقاوم

loess lowess	خم‌های اسپلاین

عمل‌های آماری

summary	خلاصهٔ شیء S-PLUS

سری زمانی

acf	همبستگی نگارها
ar arima.mle	مدل ar یا arima برازش کن
fft spectrum	برآورد طیف

مرجع‌های 'S و S-PLUS

1 Ripley, B.D. (1994) *Introductory Guide to S-Plus.* PostScript file available from StatLib: **/S/sguide2.ps**

```
FTP:   ftp  lib.stat.cmu.edu  [ then cd to S ]
Web:   http://lib.stat.cmu.edu/S/
```

2 Venebles, W.N. and Ripley, B.D. (1994) *Modern Applied Statistics with S-Plus.* Springer-Verlag, New York.

3 Chambers, J.M. and Hastie, T.J. (Editors) (1993) *Statistical Models in S.* Chapman & Hall, New York.

4 Becker, R.A., Chambers, J.M. and Wilks, A.R. (1988) *The New S Language.* Wiley, New York.

پیوست ج

افزودهٔ آماری

به‌طور ایده‌آل، این نوع پیوست‌ها دربردارندهٔ همهٔ جزئیات آماری‌یی است که یک مشاور نیاز دارد. متأسفیم ... نه در این کتاب! درواقع، بسیار غیرعملی و تا حدودی زائد است زیرا قطعاً منابع جامع دیگری در دسترس‌اند (مثلاً کاتس و جانسون (۱۹۸۲)). در عوض، چندین جدول راگنجانده‌ایم که دربردارندهٔ نوعی اطلاعات‌اند که ما (مؤلفان) هرگز نمی‌توانیم جزئیات دقیق آن‌ها را خاطرنشان کنیم. حال، حداقل، اطلاعات را یک‌جا داریم؛ در اینجا صورت مختصری از جدول‌های ارائه‌شده در این پیوست آمده است.

ج. ۱ توزیع‌های تک‌متغیره

جدول ج. ۱ توزیع‌های گسسته

جدول‌های ج. ۲، ج. ۳، ج. ۴ توزیع‌های پیوسته

ج. ۲ توزیع‌های چندمتغیره

ج. ۳ آزمون‌های استاندارد

جدول‌های ج. ۵، ج. ۶ آزمون‌های یک و دونمونه‌ای استاندارد

جدول‌های ج. ۷، ج. ۸ آزمون‌های ناپارامتری استاندارد

ج. ۴ اندازهٔ نمونه

جدول ج. ۹ اندازه‌های نمونهٔ مورد نیاز برای توان ۸۰٪

جدول ج. ۱۰ طرح‌های مرکب مرکزی

ج.۱ توزیع‌های تک‌متغیره

جدول ج.۱ توزیع‌های گسسته

توزیع گسسته	تابع احتمال		میانگین μ	واریانس σ^2
دوجمله‌ای	$\binom{n}{x}p^x q^{n-x}$	$0<p<1$ $x=0,1,\ldots,n$	np	npq $q=1-p$
پواسون	$e^{-\lambda}\frac{\lambda^x}{x!}$	$\lambda>0$ $x=0,1,\ldots$	λ	λ
هندسی[1]	pq^x	$x=0,1,\ldots$	$\frac{q}{p}$	$\frac{q}{p^2}$
	pq^{y-1}	$y=1,2,\ldots$	$\frac{1}{p}$	$\frac{q}{p^2}$
پاسکال یا[1]	$\binom{r+x-1}{x}p^r q^x$	$r>0$ $x=0,1,\ldots$	$\frac{rq}{p}$	$\frac{rq}{p^2}$
دوجمله‌ای منفی	$\binom{y-1}{r-1}p^r q^{y-r}$	$y=r,r+1,\ldots$	$\frac{r}{p}$	$\frac{rq}{p^2}$
فوق هندسی	$\frac{\binom{K}{x}\binom{N-K}{n-x}}{\binom{N}{n}}$	$p^*=\frac{K}{N}$	np^*	$np^*q^*(1-f)$
	$x=\max\{0,n-N+K\},\ldots,\min\{K,n\}$			$q^*=1-p^*$ $f=\frac{n-1}{N-1}$
چندجمله‌ای[2]	$\frac{n!}{\prod_{i=1}^k n_i!}\prod_{i=1}^k p_i^{n_i}$		np	$n(\mathrm{diag}(p)-pp')$
	$n=\sum n_i$	$\sum p_i=1$		$p=(p_1,\ldots,p_k)'$

۱. توزیع‌های هندسی و پاسکال را می‌توان برحسب X = تعداد شکست‌ها، یا Y = تعداد آزمایش‌ها تعریف کرد.
۲. توزیع چندجمله‌ای، توزیعی چندمتغیره است.

جدول ج.۲ توزیع‌های پیوستهٔ I

توزیع های تعریف شده بر $\mathcal{R}$ $(-\infty < X < \infty)$

توزیع	تابع احتمال†		میانگین	واریانس
نرمال[۱]	$X \sim N(\mu, \sigma^2)$ $Z \sim N(0,1)$	$\sigma>0$	μ 0	σ^2 1
توزیع t[۲]	$T \sim t_\nu$ $T = \frac{Z}{\sqrt{W/\nu}}$	$\nu>0$ $Z\sim N(0,1)$ $W\sim\chi^2_\nu$	$\nu>1$ اگر 0	$\nu>2$ اگر $\frac{\nu}{\nu-2}$
کوشی	t_1 $\frac{1}{\pi\beta\left[1+(\frac{x-\theta}{\beta})\right]^2}$	$\theta=0$ $\beta=1$ وقتی $\beta>0$	وجود ندارد	وجود ندارد
لاپلاس یا یا نمایی دوگانه	$\frac{\lambda}{2}e^{-\lambda\lvert x-\theta\rvert}$	$\lambda>0$	θ	$\frac{2}{\lambda^2}$
لوژستیک	$\left[1+e^{-(\frac{x-\theta}{\beta})}\right]^{-1}$	$\beta>0$	θ	$\frac{\beta^2\pi^2}{3}$
گامبل یا مقدار کرانگین	$e^{-e^{-(x-\theta)/\beta}}$	$\beta>0$ $\gamma\approx0.577216$	$\theta+\beta\gamma$ [ثابت اویلر]	$\frac{\beta^2\pi^2}{6}$

† . نشان‌دهندهٔ توزیع ترجیهی است: $\mathrm{P}[X \le x]$

۲. تابع چگالی‌احتمال نرمال (گاوسی): $f(x) = \frac{1}{\sqrt{2\pi\sigma^2}}e^{-(x-\mu)^2/2\sigma^2}$

۳. تابع چگالی‌احتمال توزیع t : $f(x) = \frac{\Gamma((\nu+1)/2)}{\Gamma(\nu/2)\sqrt{\nu\pi}}\left[1+\frac{x^2}{\nu}\right]^{-(\nu+1)/2}$

جدول ج.۳ توزیع‌های پیوستهٔ II

توزیع‌های تعریف‌شده بر $\mathcal{R}^+(X > 0)$

توزیع	تابع احتمال		میانگین	واریانس†
خی‌دو[1]	$W \sim \chi^2_\nu$ $\chi^2_n = \sum^n Z_i^2$	$\nu > 0$ $z_i \sim N(0,1)$	ν	2ν
توزیع F[2]	$F \sim \mathcal{F}_{\nu_1,\nu_2}$ $F = \frac{W_1/\nu_1}{W_2/\nu_2}$	$\nu_i > 0$ $W_i \sim \chi^2_{\nu_i}$	اگر $\nu_2 > 2$ $\frac{\nu_2}{\nu_2-2}$	اگر $\nu_2 > 4$ $\frac{2\nu_2^2(\nu_1+\nu_2-2)}{\nu_1(\nu_2-2)^2(\nu_2-4)}$
نمایی	$\lambda e^{-\lambda x}$	$\lambda > 0$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
گاما	$\frac{\lambda^a}{\Gamma(a)} x^{a-1} e^{-\lambda x}$	$\lambda > 0$ $a > 0$	$\frac{a}{\lambda}$	$\frac{a}{\lambda^2}$
وایبل	$a\lambda x^{a-1} e^{-\lambda x^a}$	$\lambda > 0$ $a > 0$	$\frac{\Gamma(1+a^{-1})}{\lambda^{1/a}}$	$\mu_2 = \frac{\Gamma(1+2a^{-1})}{\lambda^{2/a}}$
لگ‌نرمال[3]	$Y = e^X$	$x \sim N(\mu, \sigma^2)$	$e^{\mu+(\sigma^2/2)}$	$\mu_2 = e^{2\mu+2\sigma^2}$

† برای وایبل ولگ نرمال: $\sigma^2 = \mu_2 - \mu^2$

۱. $W \sim \text{Gamma}\left(a = \frac{\nu}{2}, \lambda = \frac{1}{2}\right)$

۲. $f(x) = \frac{\Gamma((\nu_1+\nu_2)/2)}{\Gamma(\nu_1/2)\Gamma(\nu_2/2)} \left(\frac{\nu_1}{\nu_2}\right)^{\nu_1/2} x^{(\nu_1-2)/2} \left[1 + x\left(\frac{\nu_1}{\nu_2}\right)\right]^{-(\nu_1+\nu_2)/2}$

۳. $f(x) = \frac{1}{y\sqrt{2\pi\sigma^2}} e^{-(\log_e y - \mu)^2/2\sigma^2}$

جدول ج.۴ توزیع‌های پیوسته III

توزیع‌های تعریف‌شده روی $\chi \subset \mathcal{R}$

توزیع	تابع احتمال		میانگین	واریانس
یکنواخت	$\frac{1}{b-a}$	$x \in [a,b]$	$\frac{1}{2}(a+b)$	$\frac{1}{12}(b-a)^2$
بتا	$\frac{x^{a-1}(1-x)^{b-1}}{B(a,b)}$	$x \in [0,1]$ $a > 0, b > 0$	$\frac{a}{a+b}$	$\frac{ab}{(a+b+1)(a+b)^2}$
پارتو	$\frac{\theta x_0^\theta}{x^{\theta+1}}$	$x > x_0 > 0$ $\theta > 0$	اگر $\theta > 1$ $\frac{\theta x_0}{\theta-1}$	اگر $\theta > 2$ $\frac{\theta x_0^2}{(\theta-1)^2(\theta-2)}$

ج.۲ توزیع‌های چندمتغیره

نرمال چندمتغیره

فرض کنید Y یک بردار تصادفی d-بعدی با تابع چگالی احتمال توأم

$$f(y) = (۲\pi)^{-d/۲}|\Sigma|^{-۱/۲}\exp[-\frac{۱}{۲}(y-\mu)'\Sigma^{-۱}(y-\mu)]$$

باشد که در آن $\Sigma > ۰$. در این صورت، Y دارای یک توزیع نرمال چندمتغیره (ناتکین) است که به‌صورت $Y \sim N_d(\mu, \Sigma)$ نشان داده می‌شود. نتیجه می‌شود که $E[Y] = \mu$ و $Var[Y] = \Sigma$. تعریفی جایگزین به‌صورت زیر است:

Y نرمال چندمتغیره است اگر به‌ازای هر $\ell \in \mathcal{R}^d$، $\ell'Y$ نرمال تک‌متغیره باشد.

ویشارت[1]

فرض کنید $Y_۱$، $Y_۲$، ...، Y_m دارای توزیع $N_d(۰, \Sigma)$ و مستقل باشند. در این صورت، $W = \Sigma_{i=۱}^{M} Y_i Y_i'$ دارای توزیع ویشارت با m درجه‌آزادی و ماتریس واریانس $d \times d$ ی Σ است و به‌صورت $W \sim W_d(m, \Sigma)$ نشان داده می‌شود. نتیجه می‌شود که به‌ازای هر $\ell \in \mathcal{R}^d$، $\ell'W\ell \sim \ell'\Sigma\ell\chi^۲_m$. با این حال، این شکل درجهٔ دوم از ویژگی خطی نرمال چندمتغیره، تابع چگالی توأم $d(d+۱)/۲$ عنصر مجزای W را به‌طور یکتا مشخص نمی‌کند. به‌منظور تکمیل:

$$f(w) = \frac{|W|^{(m-d-۱)/۲}\exp[-\frac{۱}{۲}\mathrm{tr}(\sum^{-۱}W)]}{۲^{md/۲}|\sum|^{m/۲}\pi^{d(d-۱)/۴}\prod_{j=۱}^{d}\Gamma(\frac{۱}{۲}(m+۱-j))}.$$

$T^۲$ ی هتلینگ

یک صورت چندمتغیره از آمارهٔ t است. فرض کنید $Y \sim N_d(\mu, \sum)$ و $W \sim W_d(m, \sum)$ و مستقل از یکدیگر باشند. در این صورت

$$T^۲ = m(y-\mu)'W^{-۱}(y-\mu) \sim T^۲_{d,m}$$

که در آن $T^۲_{d,m}(m-d+۱)/md \equiv \mathcal{F}_{d,m-d+۱}$. وقتی $d = ۱$، $T^۲ \sim \mathcal{F}_{۱,m} \equiv (t_m)^۲$ که در آن t_m نشان‌دهندهٔ توزیع t با m درجه‌آزادی است.

1. Wishart

MANOVA

آزمون F در شیوهٔ تحلیل واریانس بر مبنای نسبت‌هایی به صورت $T = H/E$ و $V = H/(H+E)$ است که در آن‌ها $H \sim \sigma^2 \chi^2_{m_H}$ نشان‌دهندهٔ مجموع توان‌های دوم «تحت فرض صفر» و مستقل از $E \sim \sigma^2 \chi^2_{m_E}$ است که مجموع توان‌های دوم «خطا» را نشان می‌دهد. چگالی‌های این نسبت‌ها عبارت‌اند از $T \sim (m_H/m_E)\mathcal{F}_{m_H,m_E}$ و $V \sim Beta(m_H/2, m_E/2)$ که دارای رابطهٔ توزیعی $1 - Beta(a,b) \sim Beta(b,a) \sim (1 + (a/b)\mathcal{F}_{2a,2b})^{-1}$ هستند.

این نتایج را می‌توان با درنظرگرفتن $E \sim W_d(m_E, \sum)$ و $H \sim W_d(m_H, \sum)$ به‌عنوان ماتریس‌های ویشارت مستقل با $m_E, m_H \geq d$ به حالت چندمتغیره تعمیم داد. تجزیه‌های ویژه‌مقداری زیر با این امر مرتبط‌اند.

$$|HE^{-1}| = \prod_{j=1}^{k} \psi_j = \prod_{j=1}^{k} \theta_j/(1-\theta_j)$$

$$|H(E+H)^{-1}| = \prod_{j=1}^{k} \theta_j$$

که در آن $k = \min(d, m_H)$ تعداد ویژه‌مقدارهای ناصفر است. در توصیف آماره‌هایی که در زیر ارائه می‌شوند، ν_1 و ν_2 نشان‌دهندهٔ کمیت‌های زیرند.

$$\nu_1 = \frac{1}{2}(|m_H - d| - 1) \qquad \nu_2 = \frac{1}{2}(m_E - d - 1)$$

آمارهٔ **U** این آماره را لاندای ویلک نیز می‌نامند.

$$U \equiv \Lambda = |E|/|E+H| = \prod_{j=1}^{k}(1-\theta_j) \sim U_{d,m_H,m_E}.$$

به‌ازای $k = 1, 2$، توزیع دقیق آمارهٔ U به‌صورت

$$(k=1) \qquad \frac{1-U}{U}\frac{\nu_2+1}{\nu_1+1} \sim \mathcal{F}_{2\nu_1+2,2\nu_2+1}$$

$$(k=2) \qquad \frac{1-U^{1/2}}{U^{1/2}}\frac{2\nu_2+2}{2\nu_1+3} \sim \mathcal{F}_{2\nu_1+6,2\nu_2+2}$$

ارائه می‌شود. به‌ازای $k > 2$، می‌توان از تقریب‌های زیر استفاده کرد.
برای m_Eی بزرگ:

$$-f \log U \sim \chi^2_{dm_H}$$

که در آن $f = m_E - (d - m_H + ۱)/۲$

$$\frac{۱ - U^{۱/t}}{U^{۱/t}} \frac{ft - g}{dm_H} \sim \mathcal{F}_{dm_H, ft-g}$$

که در آن

$$t = \max\left(۱, \sqrt{(d^۲ m_H^۲ - ۴)/(d^۲ + m_H^۲ - ۵)}\right)$$

و $g = dm_H - ۲/۲$

اثر پیلایی $V = \mathrm{tr}[H(E+H)^{-۱}]$. تقریب:

$$\frac{۲\nu_۲ + k + ۱}{۲\nu_۱ + k + ۱} \frac{V}{k - V} \mathcal{F}_{k(۲\nu_۱ + k + ۱), k(۲\nu_۲ + k + ۱)}$$

اثر لالی‌ـهتلینگ تعمیم آمارهٔ $T^۲$ نیز نامیده می‌شود.

$$T_g^۲ = m_E \mathrm{tr}[HE^{-۱}]$$

تقریب: $(۱/c)\mathrm{tr}[HE^{-۱}] \sim \mathcal{F}_{a,b}$ که در آن

$$a = dm_H, b = ۴ + \frac{(a+۲)(m_E - d - ۳)(m_E - d)}{(m_E + m_H - d - ۱)(m_E - ۱)} \text{ و } c = \frac{a(b-۲)}{b(m_E - d - ۱)}.$$

ریشهٔ ماکسیمم روی [۱] جداولی برای ریشهٔ ماکسیمم روی به‌ازای $\theta_{\max} = \lambda_{\max}(H(E+H)^{-۱})$ و $\psi_{\max} = \lambda_{\max}(HE^{-۱})$ موجودند.

وقتی $k = ۱$، $\theta_{\max} \sim Beta(\nu_۱ + ۱, \nu_۲ + ۱)$. بنابراین،

$$\frac{۲(\nu_۲ + ۱)}{۲(\nu_۱ + ۱)} \frac{\theta_{\max}}{۱ - \theta_{\max}} = \frac{۲(\nu_۲ + ۱)}{۲(\nu_۱ + ۱)} \psi_{\max} \sim \mathcal{F}_{۲(\nu_۱ + ۱)/۲(\nu_۲ + ۱)}.$$

تقریب: $\psi_{\max}(m_E - r + m_H)/r$، که در آن $r = \max(d, m_H)$ کران بالایی برای $\mathcal{F}_{r, m_E - r + m_H}$ است که کران پایینی برای سطح معنی‌داری به‌دست می‌دهد.

حالت $m_H = ۱$ وقتی $m_H = ۱$، آماره‌های بالا در

$$U = ۱ - \theta_{\max} \quad V = \theta_{\max} \quad T_g^۲ = m_E \frac{\theta_{\max}}{۱ - \theta_{\max}} \sim T_{d, m_E}^۲$$

صدق می‌کنند. بنابراین،

$$\frac{m_E - d + ۱}{d} \frac{\theta_{\max}}{۱ - \theta_{\max}} \sim \mathcal{F}_{d, m_E - d + ۱}$$

1. Roy

آزمون کروی‌بودن فرض کنید X_1، X_2، ...، X_n متغیرهای هم‌توزیع و مستقل $N_d(\mu, \sum)$ باشند. آزمون کروی‌بودن عبارت است از $H_0 : \sum = \lambda I_d$ در مقابل $H_A : \sum \neq \lambda I_d$.
آزمون کروی‌بودن معادل با آزمودن این است که همهٔ ویژه‌مقدارهای $\sum$ برابرند:

$$1 = \frac{\text{میانگین هندسی } \lambda_j}{\text{میانگین حسابی } \lambda_j} = \frac{(\prod_j \lambda_j)^{1/d}}{\frac{1}{d}\sum_j \lambda_j} = \frac{|\sum|^{1/d}}{\mathrm{tr}[\sum/d]}.$$

با قراردادن برآورد ماکسیمم درستنمایی $\sum$ به‌جای آن، آمارهٔ نسبت درستنمایی

$$V = |\mathrm{SS}_X|/(\frac{1}{d}\mathrm{tr}[\mathrm{SS}_X])^d$$

به‌دست می‌آید که در آن $\mathrm{SS}_X = \sum_{i=1}^n (X_i - \overline{X})(X_i - \overline{X})'$. این آزمون H_0 را به‌ازای مقادیر کوچک V رد می‌کند. برای جزئیات تقریب‌های بزرگ‌نمونه‌ای این آزمون، سبر (۱۹۸۴) و موثیرهد (۱۹۸۲) را ببینید.

ج.۳ آزمونهای استاندارد

جدول ج.۵ آزمون‌های تک‌نمونه‌ای استاندارد

ج.۵-۱	t-آزمون: $H_0 : \mu = \mu_0$		
	آمارهٔ آزمون	توزیع H_0	فرض‌ها و شرط‌ها
	$t_0 = \dfrac{\overline{x} - \mu_0}{s/\sqrt{n}}$	t_{n-1} $t_0 \approx N(0,1)$	فرایند نرمال $n > 50$
ج.۵-۲	z—آزمون: $H_0 : \mu = \mu_0$، σ^2 معلوم		
	$z_0 = \dfrac{\overline{x} - \mu_0}{\sigma/\sqrt{n}}$	$N(0,1)$ $z_0 \xrightarrow{d} N(0,1)$	فرایند نرمال $n > 30$
ج.۵-۳	آزمون t‌ی جفتی: $H_0 : \mu_A - \mu_B = \delta_0$		
	$t_0 = \dfrac{\overline{d} - \delta_0}{s_d/\sqrt{n}}$	t_{n-1}	آزمایش قبل و بعد تفاضل‌های نرمال $d_i = x_{i,\text{after}} - x_{i,\text{before}}$
ج.۵-۴	نسبت: $H_0 : p = p_0$		
	$v_0 = \sum_{i=1}^{n} x_i$	$Bin(n, p_0)$	آزمایش دوجمله‌ای $x_i \sim Bin(1, p_0)$
	$z_0 = \dfrac{v_0 - np_0}{\sqrt{np_0q_0}}$	$z_0 \xrightarrow{d} N(0,1)$	$np_0, nq_0 \geq 5$ $(q_0 = 1 - p_0)$
	$w_0 = z_0^2$	$w_0 \sim \chi_1^2$	$np_0, nq_0 \geq 5$
ج.۵-۵	واریانس: $H_0 : \sigma^2 = \sigma_0^2$		
	$w_0 = \dfrac{(n-1)s^2}{\sigma_0^2}$	χ_{n-1}^2	فرایند نرمال
ج.۵-۶	آزمون نرمال بودن ویلکسشاپیرو: $H_0 : X \sim N(\mu, \sigma^2)$		
	W[§]	جدولبندی شده[۱]	دوری از نرمال که توسط Wای کوچک آشکار می‌شود آزمون KS که روی $n>50$ به‌کار می‌رود ن. ک [ج، ۷-۲]

§ $W = \left\{ \sum_{i=1}^{[n/2]} a_{n-i+1}^{(n)} (x_{(n-i+1)} - x_{(i)}) \right\}^2 / \sum_{i}^{n} (x_{(i)} - \overline{x})^2$

۱. ن. ک. میر (۱۹۸۴، پیوست‌های د۷، د۸)

جدول ج.۶ آزمون‌های دونمونه‌ای استاندارد

ج.۶-۱	آزمون t: $H_o : \mu_x = \mu_y$ یا $(\sigma_x = \sigma_y)$		
	آمارهٔ آزمون	توزیع H_o	فرض‌ها و شرط‌ها
	$t_o = \dfrac{\overline{x} - \overline{y}}{s_p\sqrt{\frac{1}{n_x} + \frac{1}{n_y}}}$	$t_{n_x+n_y-2}$	نرمال‌های مستقل $s_p^2 = \dfrac{(n_x-1)\sigma_x^2+(n_y-1)\sigma_y^2}{n_x+n_y-2}$
ج.۶-۲	آزمون اسمیت–ساترث ویت: $H_o : \mu_x = \mu_y$ و $\sigma_x \neq \sigma_y$		
	$t_o = \dfrac{\overline{x} - \overline{y}}{\sqrt{\frac{s_x^2}{n_x} + \frac{s_y^2}{n_y}}}$	$t_o \approx t_\nu$	نرمال‌های مستقل $\nu \geq \min(n_x, n_y) - 1$ [§]
		$t_o \approx N(0,1)$	$n_x, n_y > 50$
ج.۶-۳	آزمون z: $H_o : \mu_x = \mu_y$ و σ_x^2 و σ_y^2 معلوم		
	$z_o = \dfrac{\overline{x} - \overline{y}}{\sqrt{\frac{\sigma_x^2}{n_x} + \frac{\sigma_y^2}{n_y}}}$	$N(0,1)$ $z_o \xrightarrow{d} N(0,1)$	نرمال‌های مستقل $n_x, n_y > 30$
ج.۶-۴	نسبت‌ها: $H_o : p_x = p_y$ $\Leftrightarrow$ آزمون دقیق فیشر (2×2)		
	$v_o = \sum_{i=1}^{n_x} x_i$	فوق‌هندسی	دوجمله‌ای‌های مستقل $x_i, y_j \sim Bin(1,p)$ *see Table C.1*
	$z_o = \dfrac{\hat{p}_x - \hat{p}_y}{\sqrt{\overline{p}\,\overline{q}\left(\frac{1}{n_x} + \frac{1}{n_y}\right)}}$	$z_o \xrightarrow{d} N(0,1)$	n_x, n_y large (see [C.5-4]) $\hat{p}_{x,y} = \sum x_i/n_x \ , \ \sum y_j/n_y$ $\overline{p} = \frac{n_x\hat{p}_x + n_y\hat{p}_y}{n_x+n_y}$
	$z_o' = \dfrac{\psi(\hat{p}_y) - \psi(\hat{p}_x)}{\sqrt{1/4n}}$	$z_o' \approx N(0,1)$	$\psi(p) = \sin^{-1}(\sqrt{p})$ (arcsine)
	$w_o = z_o^2$	$w_o \sim \chi_1^2$	ک. جدول ج.۳
ج.۶-۵	واریانس: $H_o : \sigma_x^2 = \sigma_y^2$		
	$f_o = s_x^2/s_y^2$	$\mathcal{F}_{n_x-1,n_y-1}$	نرمال‌های مستقل
ج.۶-۶	همبستگی: $H_o : \rho = \rho_1$		
	$z_o = \dfrac{V(r) - V(\rho_o)}{s_V}$	$z_o \xrightarrow{d} N(0,1)$	نرمال دومتغیره $V(r) = \frac{1}{2}\log\left(\frac{1+r}{1-r}\right)$ $s_V = 1/\sqrt{n-3}$
	$t_o = \dfrac{r\sqrt{n-2}}{\sqrt{1-r^2}}$	t_{n-2}	$\rho_o = 0$

§. تقریب داده‌ها: $\nu = \left(\frac{s_x^2}{n_x} + \frac{s_y^2}{n_y}\right)^2 \Big/ \left[\frac{(s_x^2/n_x)^2}{n_x-1} + \frac{(s_y^2/n_y)^2}{n_y-1}\right]$.

جدول ج.۷ آزمون‌های ناپارامتری I

ج-۷-۱	آزمون علامت : $H_o : \eta = \eta_o \Leftrightarrow H_o : p = p_o = P[X < \eta_o]$ آزمون میانه : $H_o : M = \eta_o$ ($Median$) $\Leftrightarrow H_o : p = p_o = 0.5$	
آمارهٔ آزمون	توزیع H_o	فرض‌ها و شرط‌ها
$v_o = \sum_{i=1}^{n} v_i^+$	$Bin(n, p_o)$	$v_i^+ = \mathcal{I}[x_i > \eta_o]$§ For $n\{p_o, q_o\} \geq 5$ see [C.5-4]
ج-۷-۲	آزمون رتبه علامت‌دار ویلکاکسون : $H_o : M = \eta_o$	
$S_o = \sum_{i=1}^{n} v_i^+ R_i$	$S_o \sim S_n$	توزیع متقارن $R_i = Rank\big[\lvert x_i - \eta_o \rvert\big]$ v_i^+ مانند بالا جدول‌هایی برای $S_{n,\alpha} (n \leq 20)$
$z_o = \dfrac{S_o - S_m}{\sqrt{S_v}}$	$z_o \xrightarrow{d} N(0,1)$	$n > 20$ $S_m = n(n+1)/2$ $S_v = n(n+1)(2n+1)/24$
ج-۷-۳	آزمون کولموگوروف‌اسمیرنوف (KS) : $H_o : X \sim F_o$	
$D_o = \sup_x \lvert\Delta(x)\rvert$	$D_o \sim \mathcal{D}_n$	$\Delta(x) = F_n(x) - F_o(x)$ $F_n(x) = n^{-1} \sum \mathcal{I}[x_i \leq x]$ جدول‌هایی برای $\mathcal{D}_{n,\alpha}$
	$D_o \xrightarrow{d} H_1$§§	$n > 5$ $x_1 = D_o\big[\sqrt{n} + 0.12 + \frac{0.11}{\sqrt{n}}\big]$ $P[D > D_o] \approx 1 - H_1(x_1)$
$V_o = D_o^+ + D_o^-$	$V_o \xrightarrow{d} H_2$†	آمار: کویپر‡ $D_o^{\pm} = \sup_x \pm\Delta(x)$ $x_2 = V_o\big[\sqrt{n} + 0.155 + \frac{0.24}{\sqrt{n}}\big]$ $P[V > V_o] \approx 1 - H_2(x_2)$

§ . تابع نشانگر: $\mathcal{I}(A) = 1$ اگر A رخ دهد، 0 در غیر این‌صورت

§§ . $H_1(x) = 1 - 2\sum_{j=1}^{\infty} (-1)^j e^{-2j^2x^2}$.

† . $H_2(x) = 1 - 2\sum_{j=1}^{\infty} (4j^2x^2 - 1) e^{-2j^2x^2}$.

‡ . ن. ک. به پرز و همکاران (۱۹۹۲؛ بخش ۳.۱۴).

جدول ج.۸ آزمون‌های ناپارامتری II

ج.۸-۱	KS آزمون: $H_o : F = G \quad (X \sim F \text{ and } Y \sim G)$		
	امارهٔ آزمون	توزیع H_i	فرض‌ها و شرط‌ها
	$D_o = \sup_x \lvert\Delta^*(x)\rvert$	$D_o \xrightarrow{d} H_1$	$\Delta^*(x) = F_n(x) - G_n(x)$ $n_e = n_x n_y/(n_x + n_y) > 5$
	$V_o = D_o^+ + D_o^-$	$V_o \xrightarrow{d} H_2$	Use [C.7-3] with $n = n_e$

ج.۸-۲	آزمون گردش‌ها: $H_o : F = G$ or $H_o : \{X_i\} -$ نمونه تصادفی		
	$A_o = \#\{\text{گردش‌ها}\}$	$A_o \sim A_{n_x,n_y}$	آزمون تصادفی بودن §
	$z_o = \dfrac{A_o - A_m}{\sqrt{A_v}}$	$z_o \xrightarrow{d} N(0,1)$	$n_x, n_y > 10$ $A_m = \frac{2n_x n_y}{n_x + n_y} + 1$ $A_v = \frac{(A_m-1)(A_m-2)}{n_x+n_y-1}$

ج.۸-۳	آزمون میانه: $H_o : F = G \Rightarrow H_o : \eta_X = \eta_Y$ (آزمون میانه)		
	$M_o = \#[x_i < M]$	فوق هندسی	$M = \text{med}\{x_i, y_j\}$ †

ج.۸-۴	آزمون U ای من-ویتنی: $H_o : F = G \Rightarrow H_o : \eta_X = \eta_Y$ آزمون مجموع رتبه‌ای ویلکاکسون(§): $S = U + n_x(n_x + 1/2)$		
	$U_o = \sum_i^{n_x} \sum_j^{n_y} u_{ij}$	$U_o \sim \mathcal{U}_{n_x,n_y}$	$u_{ij} = \mathcal{I}[x_i < y_j]$ جدول‌هایی برای $\mathcal{U}_{n_x,n_y}a$
	$S_o = \sum_i^{n_x} R_i^X$	$S_o \sim \mathcal{S}_{n_x,n_y}$	$\mathcal{S}_{n_x,n_y}a$ $R_i^X =$ Rank of x_i in $\{x_i, y_j\}$
	$z_o = \dfrac{U_o - U_m}{\sqrt{U_v}}$	$z_o \xrightarrow{d} N(0,1)$	$n_x, n_y > 10$ $U_m = n_x n_y/2$ $U_v = n_x n_y(n_x + n_y + 1)/12$

ج.۸-۵	آزمون کروسکال-والیس: H_o هیچ اثر تیمار		
			مدل های اثر ثابت :
	$K_o = \sum_i^a g_{n_i}(R_{ij})$	$K_o \approx \chi^2_{a-1}$ ‡	$y_{ij} = \mu_i + \epsilon_{ij} \quad \epsilon_{ij} \sim (0, \sigma^2)$ $R_{ij} = \text{Rank}(y_{ij})$ $n_i > 5, \ i = 1, \ldots, a$

ج.۸-۶	همبستگی اسپیرمن: $\rho = 0 : H_o$		
	$t_o = \dfrac{r_s\sqrt{n-2}}{\sqrt{1-r_s^2}}$	$t_o \approx t_{n-2}$	$r_s = 1 - 6\sum_{i=1}^n d_i^2/n(n^2-1)$ $d_i = \text{Rank}(x_i) - \text{Rank}(y_i)$

§ $P[A = a] = \binom{n_x-1}{\lfloor a/2 \rfloor}\binom{n_y-1}{\lfloor (a+1)/2 \rfloor} + \binom{n_x-1}{\lfloor (a+1)/2 \rfloor}\binom{n_y-1}{\lfloor a/2 \rfloor} \Big/ \binom{n_x+n_y}{n_x}$.

† $P[M = a] = \binom{n_x}{a}\binom{n_y}{\lfloor (n_x+n_y)/2 \rfloor - a} \Big/ \binom{n_x+n_y}{\lfloor (n_x+n_y)/2 \rfloor}$.

‡ $K_o = R_v^{-1}\left[\sum_{i=1}^a n_i^{-1}\left(\sum_{j=1}^{n_i} R_{ij}\right)^2 - n(n+1)^2/4\right]$,

که در آن $n = \sum_{i=1}^a n_i$

و $R_v = (n-1)^{-1}\left[\sum_{i=1}^a \sum_{j=1}^{n_i} R_{ij}^2 - n(n+1)^2/4\right]$.

ج.۴ اندازهٔ نمونه

برای آزمون‌ها و شیوه‌های آماری فهرست‌شده در جدول ج.۹، اندازه‌های نمونه‌ای، ۸۰٪ از توان را در سطح معنی‌داری و اندازهٔ اثر مفروض تأمین خواهند کرد. به‌عنوان مثال، برای اینکه یک پژوهشگر ۸۰٪ اطمینان داشته باشد که یک همبستگی ناصفر حداقل برابر با ۰٫۳۰ را با استفاده از آزمون t (ج.۶ـ۶ را ببینید) با $\alpha = ۰٫۰۵$ مشخص خواهد کرد، لازم است که نمونه‌های X و Y حداقل ۸۵ مشاهده داشته باشند. نکات دیگری که لازم است ذکر شوند عبارت‌اند از:

- برای آزمون ANOVAی یک‌طرفه، اندازهٔ نمونه به مینیمم تعداد مشاهدات مورد نیاز برای هر سطح تیمار اشاره دارد. بنابراین، اندازهٔ نمونهٔ کل مورد نیاز برابر با $N = I \times n$ است که در آن n درایهٔ ارائه‌شده در جدول است.
- درکل، اندازه‌های اثر ارائه‌شده در جدول به‌عنوان «متوسط» و «بزرگ» در نظر گرفته می‌شوند (مقدار واقعی اندازهٔ اثر با مقدار شاخص d مشخص می‌شود). بنابراین، این اندازه‌های نمونه را می‌توان به‌عنوان نمایندهٔ حداقل تعداد مشاهداتی انگاشت که پژوهشگر می‌تواند به‌کار گیرد، و هنوز جای مطالعه‌ای معنی‌دار از منظر آماری را دارد.

جدول ج.۱۰ دربردارندهٔ تخصیص‌های نقاط عاملی، محوری، و مرکزی برای طرح‌های مرکب مرکزی با بیش از هفت عامل است. برای مقدار مفروضی از عامل‌ها، تفاوت عمدهٔ بین دقت یکنواخت، طرح‌های متعامد، و بلوکی متعامد در تعداد نقاط مرکزی و مقیاس‌بندی محوری مورد نیاز برای متعامدسازی اثرها است. نمادگذاری [8,8,0] اشاره به یک مؤلفهٔ خاص طرح (عاملی، محوری، یا مرکزی) دارد و برابر با تعداد اجراهای تخصیص داده‌شده به هر بلوک (برای آن مؤلفهٔ طرح) دارد. این مطالب را برای طرح‌های بلوکی با ۷ و ۶ $= k$ عامل به‌طور خلاصه به‌صورت $B_{۱-۸} = ۸$ نشان می‌دهند (بدین معنی که هر یک از بلوک‌های ۱ تا ۸، ۸ مشاهده دارند). شرط‌های لازم و کافی برای ساختن یک طرح مرکب مرکزی در b بلوک متعامد این است که هر بلوک یک طرح مرتبهٔ اول متعامد، مانند X_j، تشکیل می‌دهد و $\text{trace}(X_j'X_j)/\sum_j \text{trace}(X_j'X_j) = n_j/n$ که در آن n_j تعداد اجراهای ژامین بلوک است.

جدول ج.۹ اندازه‌های نمونهٔ لازم برای ۸۰٪ توان

آزمون	شاخص برای ES	معنی‌داری: اندازه اثر (ES)	سطح α: 5%	1%
آزمون t: $H_o: \mu_1 = \mu_2$	$d = \lvert\mu_1 - \mu_2\rvert/\sigma$	0.50	64	95
		0.80	26	38
همبستگی: $H_o: \rho = 0$	$d = r$	0.30	85	125
		0.50	28	41
نسبت: $H_o: p = 0.5$	$d = p - 0.5$	0.15	85	127
		0.25	30	44
$H_o: p_1 = p_2$	$d = \psi(p_1) - \psi(p_2)$	0.50	63	93
آزمون z	$\psi = $ ن.ک (ج.۶-۴)	0.80	25	36

H_o: خی دو $p_{o1} = p_i$ ، $i = 1, 2, \ldots, k$			سطح α: 5%	1%
شاخص	ES	DF		
$d = \sqrt{\sum_{i=1}^{k} \frac{(p_{a1} - p_{o1})^2}{p_{o1}}}$	0.50	1	26	38
		2	39	56
		3	44	62
		4	48	67
		5	51	71
		6	54	75

رگرسیون: $H_o: B_1 = \cdots = B_k = 0$						
شاخص	سطح α	ES	k=2	k=3	k=4	k=5
$d = \frac{R^2}{1 - R^2}$	5%	0.15	67	76	84	91
		0.35	30	34	38	42
	1%	0.15	97	108	118	126
		0.35	45	50	55	59

انووا (یک‌طرفه): $H_o: \mu_1 = \cdots = \mu_I$						
شاخص	سطح α	ES	I=3	I=4	I=5	I=6
$d = \frac{\sigma_{TRT}}{\sigma_E}$	5%	0.25	52	45	39	35
		0.40	26	21	18	16
	1%	0.25	76	63	55	49
		0.40	30	25	22	20

جدول ج.۱۰ طرح‌های مرکب مرکزی بلوکی دقت یکنواخت (U. P) و متعامد (OR)

عاملی : $n_f = 2^k$ or 2^{k-1}

نقاط محوری : $n_a = 2k$

نقاط مرکزی : n_o

مقیاس‌بندی محوری

عامل‌ها		نوع طرح			اثرهای متعامد: α_{rot}		
k	n_s	U.P.	OR.	بلوک	دوران‌پذیر : $\alpha_{rot} = \sqrt[4]{n_f}$		
2	n_f	4	4	[4, 0]		1.267103	(up)
	n_a	4	4	[0, 4]	α_{or}	1.414214	(or)
	n_o	5	8	[3, 3]		1.414214	(bk)
	n_{runs}	**13**	**16**	**14**	α_{rot}	**1.414214**	
3	n_f	8	8	[4, 4, 0]		1.524649	(up)
	n_a	6	6	[0, 0, 6]	α_{or}	1.668032	(or)
	n_o	7	9	[2, 2, 2]		1.632993	(bk)
	n_{runs}	**21**	**23**	**20**	α_{rot}	**1.681793**	
4	n_f	16	16	[8, 8, 0]		1.770742	(up)
	n_a	8	8	[0, 0, 8]	α_{or}	2.0	(or)
	n_o	7	12	[2, 2, 2]		2.0	(bk)
	n_{runs}	**31**	**36**	**30**	α_{rot}	**2.0**	
5 ($\frac{1}{2}$rep)	n_f	16	16	[16, 0]		1.820359	(up)
	n_a	10	10	[0,10]	α_{or}	2.0	(or)
	n_o	6	12	[6, 1]		2.0	(bk)
	n_{runs}	**32**	**36**	**33**	α_{rot}	**2.0**	
5	n_f	32	32	[8, 8, 8, 8, 0]		2.096683	(up)
	n_a	10	10	[0, 0, 0, 0,10]	α_{or}	2.353860	(or)
	n_o	10	16	[2, 2, 2, 2, 4]		2.366432	(bk)
	n_{runs}	**52**	**58**	**54**	α_{rot}	**2.378414**	
6 ($\frac{1}{2}$rep)	n_f	32	32	[16,16, 0]		2.142723	(up)
	n_a	12	12	[0, 0,10]	α_{or}	2.353860	(or)
	n_o	9	14	[4, 4, 2]		2.366432	(bk)
	n_{runs}	**53**	**58**	**54**	α_{rot}	**2.378414**	
6	n_f	64	64	$B_{1-8} = 8$		2.481445	(up)
	n_a	12	12	$B_9 = 12$	α_{or}	2.828427	(or)
	n_o	15	24	B_{1-8} =1 B_9 =2		2.828427	(bk)
	n_{runs}	**91**	**100**	**90**	α_{rot}	**2.828427**	
7 ($\frac{1}{2}$rep)	n_f	64	64	$B_{1-8} = 8$		2.523223	(up)
	n_a	14	14	$B_9 = 14$	α_{or}	2.828427	(or)
	n_o	14	22	B_{1-8} =1 B_9 =4		2.828427	(bk)
	n_{runs}	**92**	**100**	**90**	α_{rot}	**2.828427**	

مراجع

[1] Adcock, C. (1997). Sample size determination: a review. *The Statistician*, **46**(2), 261–283.

[2] Agresti, A. (1990). *Categorical Data Analysis*. Wiley, New York.

[3] Akaike, H. (1974). Stochastic theory of minimal realization. *IEEE Trans. Automatic Control*, **AC-19**, 667–674.

[4] Andersen, B. (1990). *Methodological Errors in Medical Research*. Blackwell Scientific, Oxford.

[5] Anderson, T. (1971). *The Statistical Analysis of Time Series*. Wiley, New York.

[6] Anderson, V. and McLean, R. (1974). *Design of Experiments: A Realistic Approach*. Marcel Dekker, New York.

[7] Andrews, D. and Herzberg, A. (1985). *Data: A Collection of Problems from Many Fields for the Student and Research Worker*. Springer-Verlag, New York.

[8] Barnett, V. (1982). *Comparative Statistical Inference*. 2nd edition. Wiley, Chichester.

[9] Becker, R. and Chambers, J. (1984). *S: An Interactive Environment for Data Analysis and Graphics*. Wadsworth and Brooks/Cole, Pacific Grove, CA.

[10] Becker, R., Chambers, J., and Wilks, A. (1988). *The New S Language — A Programming Environment for Data Analysis and Graphics.* Wadsworth and Brooks/Cole, Pacific Grove, CA.

[11] Belsley, D., Kuh, E., and Welsch, R. (1980). *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity.* Wiley, New York.

[12] Boen, J. and Zahn, D. (1982). *The Human Side of Statistical Consulting.* Lifetime Learning Publications, Belmont, CA.

[13] Box, G. (1957). Evolutionary operation: A method for increasing industrial productivity. *Applied Statistics*, **6**, 81–101.

[14] Box, G. and Cox, D. (1964). An analysis of transformations. *JRSS(B)*, **26**, 211–252.

[15] Box, G. and Draper, N. (1969). *Evolutionary Operation.* Wiley, New York.

[16] Box, G., Hunter, W., and Hunter, J. (1978). *Statistics for Experimenters.* Wiley, New York.

[17] Box, G. and Jenkins, G. (1970). *Time Series Analysis: Forecasting and Control.* Holden-Day, San Francisco.

[18] Box, G., Jenkins, G., and Reinsel, G. (1994). *Time Series Analysis: Forecasting and Control.* 3rd edition. Prentice Hall, Englewood Cliffs, NJ.

[19] Box, J. (1978). *R. A. Fisher, The Life of a Scientist.* Wiley, New York.

[20] Brieman, L., Friedman, J., Olshen, R., and Stone, C. (1984). *Classification and Regression Trees.* Wadsworth International Group, Belmont, CA.

[21] Brockwell, P. and Davis, R. (1991). *Time Series: Theory and Methods.* 2nd edition. Springer-Verlag, New York.

[22] Carr, D. (1998). Presenting spatially-indexed summary statistics using linked micromap plots. Presented at the Data Visualization Conference, Drew University, New Jersey.

[23] Chambers, J. and Hastie, T., Editors (1993). *Statistical Models in S.* Wadsworth & Brooks/Cole, Pacific Grove, CA.

[24] Chatfield, C. (1995). *Problem Solving: A Statistician's Guide.* 2nd edition. Chapman & Hall, London.

[25] Cleveland, W. (1985). *The Elements of Graphing Data*. Wadsworth, Inc., Belmont, CA.

[26] Cleveland, W. (1993). *Visualizing Data*. Hobart Press, Summit, NJ.

[27] Cochran, W. (1957). Analysis of covariance: Its nature and uses. *Biometrics*, **13**, 261–281.

[28] Cochran, W. (1977). *Sampling Techniques*. 3rd edition. Wiley, New York. First published in 1953.

[29] Cochran, W. (1983). *Planning and Analysis of Observational Studies*. Wiley, New York.

[30] Cochran, W. and Cox, D. (1957). *Experimental Designs*. 2nd edition. Wiley, New York.

[31] Cohen, J. (1988). *Statistial Power Analysis for the Behavioral Sciences*. 2nd edition. Erbaum, Hillsdale, NJ.

[32] Cohen, J. (1992). A Power Primer. *Psych. Bulletin*, **112**(2), 155–159.

[33] Collett, D. (1991). *Modelling Binary Data*. Chapman & Hall, New York.

[34] Conover, W. (1980). *Practical Nonparametric Statistics*. 2nd edition. Wiley, New York.

[35] Conrad, S., Editor (1989). *Assignments in Applied Statistics*. Wiley, New York.

[36] Cox, D. (1958). *The Planning of Experiments*. Wiley, New York.

[37] Cox, D. and Snell, E. (1981). *Applied Statistics, Principles and Examples*. Chapman & Hall, London.

[38] Cronbach, L. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, **16**(3), 297–334.

[39] Crowder, M. and Hand, D. (1990). *Analysis of Repeated Measures*. Chapman & Hall, New York.

[40] Dempster, A., Laird, N., and Rubin, D. (1977). Maximum likelihood for incomplete data via the EM algorithm. *JRSS(B)*, **39**, 1–22.

[41] Derr, J. (2000). *Statistical Consulting: A Guide to Effective Communication*. Duxbury Press, Pacific Grove, CA.

[42] Desu, M. and Raghavarao, D. (1990). *Sample Size Methodology*. Academic Press, Boston.

[43] Dillman, D. (1978). *Mail and Telephone Surveys: The Total Design Method.* Wiley, New York.

[44] Dixon, W. and Kronmal, R. (1965). The choice of origin and scale for graphs. *J. of the Assoc. for Computing Machinery*, **12**, 259–261.

[45] Dobson, A. (1990). *An Introduction to Generalized Linear Models.* 2nd edition. Chapman & Hall, London.

[46] Dodge, H. and Romig, H. (1959). *Sampling Inspection Tables: Single and Double Sampling.* Wiley, New York.

[47] Draper, N. and Smith, H. (1981). *Applied Regression Analysis.* Wiley, New York.

[48] Duncan, D. (1955). Multiple range and multiple F-tests. *Biometrics*, **11**, 1–42.

[49] Dunckel, J. and Parnham, E. (1993). *Effective Speaking for Business Success: Making Confident Presentations, Using Audio-Visuals, and More.* 2nd edition. Self-Counsel Press, Bellingham, WA.

[50] Efron, B. and Tibshirani, R. (1993). *An Introduction to the Bootstrap.* Chapman & Hall, San Francisco.

[51] Everitt, B. (1995). The analysis of repeated measures: A practical review with examples. *The Statistician*, **44**, 113–135.

[52] Fienberg, S. (1989). *The Evolving Role of Statistical Assessments as Evidence in the Courts.* Springer-Verlag, New York.

[53] Fink, A. (1995). *The Survey Handbook.* Sage, Thousand Oaks, CA.

[54] Finklestein, M. and Levin, B. (1990). *Statistics for Lawyers.* Springer-Verlag, New York.

[55] Finney, D. (1982). The questioning statistician. *Statistics in Medicine*, **1**, 5–13.

[56] Fisher, R. (1925). *Statistical Methods for Research Workers.* Oliver and Boyd, Edinburgh. 14th edition, 1970.

[57] Fleiss, J. (1986). *The Design and Analysis of Clinical Experiments.* Wiley, New York.

[58] Frank, I. and Friedman, J. (1993). A statistical view of some chemometrics regression tools. *Technometrics*, **35**, 109–148.

[59] Freedman, D. and Diaconis, P. (1981). On the maximum deviation between the histogram and the underlying density. *Z. Wahrsch. Verw. Gebiete*, **57**, 453–476.

[60] Friedman, L., Furberg, C., and Demets, D. (1985). *Fundamentals of Clinical Trials.* 2nd edition. PSG Publishing, Littleton, MA.

[61] Fuller, W. (1987). *Measurement Error Models.* Wiley, New York.

[62] Gail, M. and Gart, J. (1973). The determination of sample size for use with the exact conditional test in 2×2 comparative trials. *Biometrics*, **29**, 441–448.

[63] Gnanadesikan, R. (1997). *Methods for Statistical Data Analysis of Multivariate Observations.* Wiley, New York.

[64] Gopen, G. and Swan, J. (1990). The science of scientific writing. *American Scientist*, **78**, 550–558. Also available from the JCGS Website: `www.amstat.org`.

[65] Greenfield, A. (1979). Statisticians in industrial research: The role and training of the industrial consultant. *The Statistician*, **28**, 71–82.

[66] Hacker, D. (1998). *A Writer's Reference.* 4th edition. St. Martins Press, New York.

[67] Hahn, G. and Hoerl, R. (1998). Key challenges for statisticians in business and industry. *Technometrics*, **40**, 195–200.

[68] Hamilton, C. and Parker, C. (1993). *Communicating for Results: A Guide for Business and the Professions.* Wadsworth, Inc., Belmont, CA.

[69] Hand, D. (1994). Deconstructing statistical questions. *JRSS(A)*, **157**, 317–356.

[70] Hand, D. J. and Everitt, B. S., Editors (1987). *The Statistical Consultant in Action.* Cambridge University Press, Cambridge.

[71] Hastie, T., Tibshirani, R., and Buja, A. (1995). Flexible discriminant and mixture models. In Kay, J. and Titterington, D., Editors, *Neural Networks and Statistics.* Oxford University Press, Edinburgh. Web page: `http://www-stat.stanford.edu/~hastie/Papers/`.

[72] Hinkley, D. (1977). On quick choice of power transformation. *Applied Statistics*, **26**, 67–69.

[73] Hoaglin, D., Mosteller, F., and Tukey, J., Editors (1983). *Understanding Robust and Exploratory Data Analysis.* Wiley, New York.

[74] Hoaglin, D., Mosteller, F., and Tukey, J., Editors (1985). *Exploring Data Tables: Trends and Shapes.* Wiley, New York.

[75] Hoaglin, D., Mosteller, F., and Tukey, J., Editors (1991). *Fundamentals of Exploratory Analysis of Variance*. Wiley, New York.

[76] Hoerl, R., Hooper, J., Jacobs, P., and Lucas, J. (1993). Skills for industrial statisticians to survive and prosper in the emerging quality environment. *The American Statistician*, **47**, 280–292.

[77] Hurlbert, S. (1984). Pseudoreplication and the design of ecological field experiments. *Ecological Monographs*, **54**, 187–211.

[78] Johnson, R. and Wichern, D. (1998). *Applied Multivariate Statistical Analysis*. 4th edition. Prentice Hall, Upper Saddle River, NJ.

[79] Joiner, B. (1982). Practicing statistics, or, what they forgot to say in the classroom. In Rustagi, J. and Wolfe, D., Editors. *Teaching of Statistics and Statistical Consulting*. Academic Press, New York.

[80] Kaye, D. and Zeisel, H. (1997). *Statistics for Lawyers*. Springer-Verlag, New York.

[81] Kettenring, J. (1995). What industry needs. *The American Statistician*, **49**, 2–4.

[82] Kettenring, J. (1997). Message to students: Will you get a job in industry? *AmStat News*, **240**, 9–10.

[83] Kimball, A. (1957). Errors of the third kind in statistical consulting. *JASA*, **52**, 133–142.

[84] Kirk, R. (1991). Statistical consulting in a university: Dealing with people and other challenges. *The American Statistician*, **45**, 28–34.

[85] Kotz, S. and Johnson, N., Editors (1982). *Encyclopedia of Statistical Sciences*. Wiley, New York. (in 9 volumes).

[86] Kraemer, H. and Thiemann, S. (1987). *How Many Subjects? Statistical Power Analysis in Research*. Sage, Newbury Park, CA.

[87] Lin, L. (1989). A concordance correlation coefficient to evaluate reproducibility. *Biometrics*, **45**, 255–269.

[88] Little, R. and Rubin, D. (1987). *Statistical Analysis with Missing Data*. Wiley, New York.

[89] Lurie, W. (1958). The impertinent questioner: The scientist's guide to the statistician's mind. *American Scientist*, **46**, 57–61.

[90] Maindonald, J. (1992). Statistical design, analysis, and presentation issues. *New Zealand Journal of Agricultural Research*, **35**, 121–141.

[91] Manchester, L., Field, C., and McDougall, A. (1999). Regression for overdetermined systems. A fisheries example. *The Canadian Journal of Statistics*, **27**, 25–39.

[92] Mardia, K., Kent, J., and Bibby, J. (1979). *Multivariate Analysis*. Academic Press, New York.

[93] Marquardt, D. (1979). Statistical consulting in industry. *The American Statistician*, **33**, 72–76.

[94] Mason, S. (1962). *A History of the Sciences*. MacMillian, New York.

[95] MathSoft (1997). *S-Plus Trellis Graphics User's Manual*. MathSoft Inc, Seattle, WA.

[96] MathSoft (2000). *S-PLUS 6.0 Guide to Statistics (for UNIX)*. MathSoft Inc, Seattle, WA. Data Analysis Division.

[97] McCullagh, P. and Nelder, J. (1989). *Generalized Linear Models*. 2nd edition. Chapman & Hall, London.

[98] McDougall, A. J. and Cook, D. (1994a). *Exploring Time Series Using Interactive Graphics*. Video. ASA Statistical Graphics Video Library.

[99] McDougall, A. J. and Cook, D. (1994b). XQz: An X application for Interactive Exploratory Time Series Analysis. *Contributed software*. StatLib: `lib.stat.cmu.edu`.

[100] Meinert, C. (1986). *Clinical Trials: Design, Conduct and Analysis*. Oxford University Press, New York.

[101] Miller, R., Efron, B., Brown, B., and Moses, L., Editors (1980). *Biostatistics Casebook*. Wiley, New York.

[102] Montgomery, D. (1997). *Design and Analysis of Experiments*. 4th edition. Wiley, New York.

[103] Moore, D. and McCabe, G. (1998). *Introduction to the Practice of Statistics*. 3rd edition. Freeman, New York. (2nd ed., 1993).

[104] Morgan, B. (1992). *Analysis of Quantal Response Data*. Chapman & Hall, New York.

[105] Muirhead, R. J. (1982). *Aspects of Multivariate Statistical Theory*. Wiley, New York.

[106] Myers, R. and Montgomery, D. (1995). *Response Surface Methodology: Process and Product Optimization Using Designed Experiments*. Wiley, New York.

[107] Newton, J. and Schmiediche, H. (1993). Graphical analysis for time series. In *Proc. of the 27th Symp. on the Interface: Computing Science and Statistics*, Amsterdam. Elsevier.

[108] Ott, L. (1993). *An introduction to statistical methods and data analysis*. PWS-Kent, Boston.

[109] Peace, K., Editor (1988). *Biopharmaceutical Statistics for Drug Development*. Marcel Dekker, New York.

[110] Pearson, K. (1900). On a criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. *Philos. Mag.*, **50**(Series 5), 157–175. Reprinted in *Karl Pearson's Early Statistical Papers*, ed. by E.S. Pearson (1948). Cambridge Universty Press.

[111] Plewis, I. (1985). *Analysing Change: Measurement and Explanation using Longitudinal Data*. Wiley, Chichester.

[112] Pocock, S. (1983). *Clinical Trials: A Practical Approach*. Wiley, Chichester.

[113] Press, W., Teukolsky, S., Vetterling, W., and Flannery, B. (1992). *Numerical Recipes in C: The art of scientific computing*. 2nd edition. Cambridge University Press, New York.

[114] Priestley, M. (1981). *Spectral Analysis and Time Series*. Academic Press, London.

[115] Radelet, M. (1981). Racial characteristics and the imposition of the death penalty. *Amer. Sociolog. Review*, **46**, 918–927.

[116] Rencher, A. (1995). *Methods of Multivariate Analysis*. Wiley, New York.

[117] Ripley, B. (1994). Introductory guide to S-Plus. PostScript file available from StatLib: `http://lib.stat.cmu.edu/S/sguide2.ps`.

[118] Ross, N. (1995). What government needs. *The American Statistician*, **49**, 7–9.

[119] Rousseeuw, P. and Leroy, A. (1987). *Robust Regression and Outlier Detection*. Wiley, New York.

[120] Rustagi, J. and Wolfe, D., Editors (1982). *Teaching of Statistics and Statistical Consulting*. Academic Press, New York.

[121] SAS Institute Inc. (1990). *SAS/STAT User's Guide, Version 6*. SAS Institute Inc., Cary, NC.

[122] SAS Institute Inc. (2000). *JMP, Version 4*. SAS Institute Inc., Cary, NC.

[123] Scheffé, H. (1959). *The Analysis of Variance*. Wiley, New York.

[124] Scott, D. (1979). On optimal and data-based histograms. *Biometrika*, **66**, 605–610.

[125] Searle, S. (1987). *Linear Models for Unbalanced Data*. Wiley, New York.

[126] Seber, G. (1977). *Linear Regression Analysis*. Wiley, New York.

[127] Seber, G. (1984). *Multivariate Observations*. Wiley, New York.

[128] Seber, G. and Wild, C. (1989). *Nonlinear Regression*. Wiley, New York.

[129] Selvin, S. (1996). *Statistical Analysis of Epidemiologic Data*. Oxford University Press, Oxford.

[130] Shao, J. and Tu, D. (1995). *The Jackknife and Bootstrap*. Wiley, New York.

[131] Shewhart, W. (1931). *Economic Control of Quality of Manufactured Products*. D. Van Nostrand, Princeton, NJ.

[132] Snedecor, G. (1937). *Statistical Methods*. Iowa State Press, Ames, IA.

[133] Sprent, P. (1993). *Applied Nonparametric Statistical Methods*. 2nd edition. Chapman & Hall, London.

[134] Stigler, S. (1986). *The History of Statistics: The Measurement of Uncertainty Before 1900*. Harvard University Press, Cambridge, MA.

[135] Stoffer, D. S., Tyler, D. E., and McDougall, A. J. (1993). Spectral analysis for categorical time series: Scaling and the spectral envelope. *Biometrika*, **80**, 611–622.

[136] Student (1908). The probable error of a mean. *Biometrika*, **6**, 1–25. Reprinted on pp. 11–34 in *"Student's" Collected Papers*. Edited by E.S. Pearson and John Wishart with a Foreword by Launce McMullen, Cambridge University Press for the Biometrika Trustees, 1942.

[137] Swayne, D., Cook, D., and Buja, A. (1991). XGobi: Interactive dynamic graphics in the X Window System with a link to S. In *ASA Proc. of the Section on Statistical Graphics*, pages 1–8.

[138] Tanur, J., Editor (1988). *Statistics: A Guide to the Unknown*. 3rd edition. Duxbury Press, Belmont, CA.

[139] Thisted, R. (1988). *Elements of Statistical Computing*. Chapman & Hall, New York.

[140] Thompson, S. (1992). *Sampling*. Wiley, New York.

[141] Tierney, L. (1991). *LispStat: An Object-Orientated Environment for Statistical Computing and Dynamic Graphics*. Wiley, New York.

[142] Tufte, E. (1983). *The Visual Display of Quantitative Information*. Graphics Press, Cheshire, Connecticut.

[143] Tufte, E. (1997). *Envisioning Information*. Graphics Press, Cheshire, Connecticut.

[144] Tufte, E. (1999). *Visual Explanations: Images and Quantities, Evidence and Narrative*. Graphics Press, Cheshire, Connecticut.

[145] Tukey, J. (1953). The problem of multiple comparisons. Unpublished Notes, Princeton University.

[146] Tukey, J. (1977). *Exploratory Data Analysis*. Addison-Wesley, Reading, MA. Originally published in 1970.

[147] Tweedie, R. (1998). Consulting: real problems, real interactions, real outcomes. *Statistical Science*, **13**, 1–29.

[148] Venables, W. and Ripley, B. (1994). *Modern Applied Statistics with S-Plus*. Springer-Verlag, New York.

[149] Williford, W., Krol, W., Bingham, S., Collins, J., and Weiss, D. (1995). The multicenter clinical trials coordinating center statistician: More than just a consultant. *The American Statistician*, **49**, 221–225.

[150] Yeatts, D. and Hyten, C. (1998). *High-Performing Self-Managed Work Teams: A Comparison of Theory and Practice*. Sage, Thousand Oaks, CA.