دانشگاه گیلان
۱۳۵۳_1974

چاپ اول

صد پرسش و پاسخ درباره آمار در روانشناسی

تألیف:
نیل جی. سالکایند
ترجمه:
دکتر رضا سلطانی شال
استادیار دانشکده ادبیات و علوم انسانی دانشگاه گیلان

اداره چاپ و انتشارات دانشگاه گیلان

صد پرسش و پاسخ درباره آمار در روانشناسی

100 Questions (and Answers) About Statistics

By:
Neil J. Salkind
Translated by:
Reza Soltani Shal, Ph. D

University of Guilan Press

تألیف: نیل جی. سالکایند ترجمه: دکتر رضا سلطانی شال

9 786001 532344

ISBN: 978-600-153-234-4

صد پرسش و پاسخ درباره آمار در روانشناسی

تألیف:

نیل جی. سالکایند

ترجمه:

دکتر رضا سلطانی شال

استادیار دانشکده ادبیات و علوم انسانی دانشگاه گیلان

انتشارات دانشگاه گیلان

۱۴۰۰

۱۳۵۳-1974

شابک: ۴ ـ ۲۳۴ ـ ۱۵۳ ـ ۶۰۰ ـ ۹۷۸

سرشناسه	: سالکایند، نیل ج.، ۱۹۴۷ - م. Salkind, Neil J.
عنوان و نام پدیدآور	: صد پرسش و پاسخ درباره آمار در روانشناسی/جی.نیل سالکایند ؛ ترجمه رضا سلطانی‌شال ؛ ویراستار علمی: عباس ابوالقاسمی ؛ ویراستار ادبی: فرشته گلچین.
مشخصات نشر	: رشت: دانشگاه گیلان، ۱۳۹۹.
مشخصات ظاهری	: ۱۸۵ص.: مصور،جدول.نمودار؛ ۲۴×۱۷ س‌م.
شابک	: 978-600-153234-4
وضعیت فهرست نویسی	: فیپا
یادداشت	: عنوان اصلی : ۱۰۰ questions (and answers) about statistics ,[2015].
یادداشت	: کتابنامه.
موضوع	: آمار -- مطالب گونه‌گون
موضوع	: Statistics -- Miscellanea
شناسه افزوده	: سلطانی شال، رضا، ۱۳۶۴ -، مترجم
شناسه افزوده	: ابوالقاسمی، عباس، ۱۳۴۴ -، ویراستار
شناسه افزوده	: دانشگاه گیلان
رده بندی کنگره	: QA۲۷۶/۱۲
رده بندی دیویی	: ۵۱۹/۵
شماره کتابشناسی ملی	: ۷۳۶۹۳۰۸

اداره چاپ و انتشارات دانشگاه گیلان

نام کتـاب : صد پرسش و پاسخ درباره آمار در روانشناسی
مؤلف : نیل. جی سالکایند
مترجم : دکتر رضا سلطانی شال
ویراستار علمی : دکتر عباس ابوالقاسمی
ویراستار ادبی : فرشته گلچین راد
نوبت چـاپ : اول، ۱۴۰۰
ناشـر : انتشـارات دانشـگاه گیـلان
شـمارگان : ۱۰۰۰ نسخه

* هر گونه چاپ و تکثیر صرفاً در اختیار انتشارات دانشگاه گیلان است.*

زندگی فقط با دو چیز زیباست: یاد گرفتن ریاضیات، یاد دادن ریاضیات

سیمون دنیس پواسون (۲۱ ژوئن ۱۷۸۱، لوآره –۲۵ آوریل ۱۸۴۰)

فهرست

درباره نویسنده ۶
مقدمه مترجم ۷
بخش ۱: چرا آمار ۹
پرسش ۱: آمار چیست و مطالعه آن چرا مهم است؟ ۱۰
پرسش ۲: علم آمار چگونه شروع شد؟ ۱۱
پرسش ۳: آمار توصیفی چیست و چگونه استفاده می‌شود؟ ۱۳
پرسش ۴: آمار استنباطی چیست و چگونه استفاده می‌شود؟ ۱۴
پرسش ۵: من نمیخواهم متخصص آماری شوم، پس چرا باید دوره آمار بگذرانم؟ ۱۶
پرسش ۶: از کدام بسته‌های نرم‌افزاری آماری می‌توانم استفاده کنم؟ SPSS یا اِکسل و یا سایر بسته‌های نرم‌افزاری؟ ۱۷
بخش ۲: درک مقیاس‌های گرایش مرکزی ۱۹
پرسش ۷: شاخص‌های گرایش مرکزی کدام‌اند و چرا به‌کاربرده می‌شوند؟ ۲۰
پرسش ۸: آیا مثالی برای چگونگی کاربست شاخص‌های گرایش مرکزی وجود دارد؟ ۲۲
پرسش ۹: میانگین چیست و چگونه محاسبه می‌شود؟ ۲۳
پرسش ۱۰: آیا مثالی برای چگونگی کاربست میانگین وجود دارد؟ ۲۴
پرسش ۱۱: میانه، انحراف چارکی و رتبه درصدی چیست؟ ۲۶
پرسش ۱۲: آیا مثالی برای چگونگی کاربست میانه وجود دارد؟ ۲۷
پرسش ۱۳: نما چیست و چگونه محاسبه می‌شود؟ ۲۸
۱۴: آیا مثالی برای چگونگی کاربست نما وجود دارد؟ ۲۹
پرسش ۱۵: چگونه میتوانم درباره اینکه از کدام شاخص گرایش مرکزی استفاده کنم تصمیم بگیرم؟ ۳۰
پرسش ۱۶: چگونه نرم‌افزار اکسل برای محاسبه شاخص‌های گرایش مرکزی به‌کاربرده می‌شود؟ ۳۱
بخش ۳: درک شاخص‌های پراکندگی ۳۳
پرسش ۱۷: شاخصهای پراکندگی چیست‌اند و چرا به‌کاربرده می‌شوند؟ ۳۴
پرسش ۱۸: آیا مثالی برای چگونگی کاربست شاخص پراکندگی وجود دارد؟ ۳۶
پرسش ۱۹: دامنه تغییرات چیست و چگونه محاسبه می‌شود؟ ۳۷
پرسش ۲۰: آیا مثالی برای چگونگی کاربست دامنه تغییرات وجود دارد؟ ۳۸
پرسش ۲۱: انحراف استاندارد چیست و چگونه محاسبه می‌شود؟ ۳۹
پرسش ۲۲: چرا از برآورد بدون تورش N-1 بجای برآورد تورش دار N در محاسبه واریانس و انحراف استاندارد استفاده می‌کنیم؟ ۴۱
پرسش ۲۳: آیا مثالی برای چگونگی کاربست انحراف استاندارد وجود دارد؟ ۴۳
پرسش ۲۴: برای چگونگی کاربست واریانس چه مثالی وجود دارد؟ ۴۴
پرسش ۲۵: چگونه نرم‌افزار اِکسل برای محاسبه شاخص‌های پراکندگی به‌کاربرده می‌شود؟ ۴۵
بخش ۴: نمایش داده ۴۶
پرسش ۲۶: آیا واقعاً یک تصویر یا نمودار ارزش هزاران واژه را دارد؟ چرا داده‌ها را نمایه می‌کنند؟ ۴۷
پرسش ۲۷: برای نمایه‌سازی کارآمد داده‌ها چه دستورالعمل‌هایی وجود دارد؟ ۴۹
پرسش ۲۸: جداول توزیع فراوانی و توزیع فراوانی تراکمی چیستند و چگونه محاسبه می‌شوند؟ ۵۰
پرسش ۲۹: نمودار هیستوگرام چیست و چگونه می‌توان آن را ترسیم کرد؟ ۵۲
پرسش ۳۰: چگونه می‌توان از طریق نرم‌افزار اِکسل یک نمودار هیستوگرام ترسیم کرد؟ ۵۴
پرسش ۳۱: نمودار اجایو یا تراکمی چیست و چگونه می‌توان آن را ترسیم کرد؟ ۵۷

پرسش ۳۲: نمودار ستونی چیست و چگونه می‌توان آن را ترسیم کرد؟ ۵۹
پرسش ۳۳: نمودار میله‌ای چیست و چگونه می‌توان آن را ترسیم کرد؟ ۶۰
پرسش ۳۴: نمودار خطی چیست و چگونه می‌توان آن را ترسیم کرد؟ ۶۱
پرسش ۳۵: نمودار دایره‌ای چیست و چگونه می‌توان آن را ترسیم کرد؟ ۶۲
پرسش ۳۶: نمودار پراکندگی چیست و چگونه می‌توان آن را ترسیم کرد؟ ۶۴
پرسش ۳۷: چگونه می‌توانم نمودارهایی را که در اِکسل رسم کرده‌ام ویرایش کنم؟ ۶۶
پرسش ۳۸: چگونه می‌توانم یک نمودار را در سایر اسنادی که دارم ترکیب کنم؟ ۶۸
پرسش ۳۹: چه هنگامی می‌بایست از یک نمودار استفاده کنم؟ نوع انتخاب نمودار به چه عواملی بستگی دارد؟ ۶۹
بخش ۵: درک روابط ۷۱
پرسش ۴۰: ضریب همبستگی چیست و چگونه به کار می‌رود؟ ۷۲
پرسش ۴۱: آیا مثالی برای چگونگی کاربست ضریب همبستگی وجود دارد؟ ۷۳
پرسش ۴۲: انواع شدت ضرایب همبستگی کدامها هستند؟ ۷۴
پرسش ۴۳: نمودارهای پراکندگی چگونه به فهم ضرایب همبستگی کمک می‌کنند؟ ۷۶
پرسش ۴۴: چگونه می‌توانم ضریب همبستگی پیرسون را محاسبه کنم؟ ۷۸
پرسش ۴۵: چگونه می‌توانم از طریق اِکسل ضریب همبستگی را محاسبه کنم؟ ۷۹
پرسش ۴۶: مستقیم‌ترین مسیر برای تفسیر ارزش‌های یک ضریب همبستگی چیست؟ ۸۱
پرسش ۴۷: ضریب تعیین چیست و چگونه محاسبه می‌شود؟ ۸۲
پرسش ۴۸: مؤلفه‌های مهمی که درباره فهم و کاربست ضرایب همبستگی باید به یاد داشته باشم، کدام‌اند؟ ۸۳
پرسش ۴۹: چگونه می‌توانم از یک ماتریس برای نمایش چند ضریب همبستگی استفاده کنم؟ ۸۴
پرسش ۵۰: سایر شاخص‌های همبستگی کدام ها هستند و چگونه به کار می‌روند؟ ۸۶
بخش ۶: درک اندازه‌گیری و اهمیت آن ۸۷
پرسش ۵۱: چرا اندازه‌گیری، یک موضوع مهم در میان متخصصین آمار است؟ ۸۸
پرسش ۵۲: سطوح اندازه‌گیری کدام‌اند و چرا اهمیت دارند؟ ۸۹
پرسش ۵۳: پایایی چیست و چه کاربردی دارد؟ ۹۱
پرسش ۵۴: انواع مختلف پایایی چیست و کاربرد آن‌ها کجاست ؟ ۹۳
پرسش ۵۵: چگونه می‌توان اعتماد (پایایی) یک آزمون را افزایش می‌یابد ؟ ۹۴
پرسش ۵۶: روایی چیست و چه کاربردی دارد؟ ۹۶
پرسش ۵۷: انواع مختلف اعتبار (روایی) چیست و چگونه محاسبه می‌شوند ؟ ۹۸
پرسش ۵۸: چگونه می‌توان اعتبار (روایی) یک آزمون را افزایش می‌یابد؟ ۹۹
پرسش ۵۹: ارتباط بین پایایی و روایی چیست؟ ۱۰۰
بخش ۷: درک نقش فرضیه در علم آمار ۱۰۲
پرسش ۶۰: فرضیه چیست و چرا در پژوهش علمی اهمیت دارد؟ ۱۰۳
پرسش ۶۱: ویژگی‌های یک فرضیه خوب چیست؟ ۱۰۵
پرسش ۶۲: یک نمونه و یک جامعه چه تفاوت‌هایی باهم دارند؟ ۱۰۶
پرسش ۶۳: فرضیه صفر چیست و چگونه به کار می‌رود؟ ۱۰۷
پرسش ۶۴: فرضیه تحقیق چیست و چگونه به کار می‌رود؟ ۱۱۰
پرسش ۶۵: فرضیه‌های صفر و تحقیق چه تفاوت‌هایی با یکدیگر دارند؟ ۱۱۲
بخش ۸: درک منحنی بهنجار و احتمال ۱۱۳
پرسش ۶۶: چرا در مطالعه آمار، مفهوم احتمال مهم است؟ ۱۱۴

پرسش ۶۷: منحنی بهنجار یا زنگوله‌ای چیست؟ ۱۱۶
پرسش ۶۸: کجی و کشیدگی چیست و توزیع‌های نمرات چگونه ازنظر این ویژگی متفاوت‌اند؟ ۱۱۹
پرسش ۶۹: قضیه حد مرکزی چیست و چرا مهم است؟ ۱۲۲
پرسش ۷۰: نمره استاندارد Z چیست و چگونه محاسبه می‌شود؟ ۱۲۵
پرسش ۷۱: چگونه می‌توانم از طریق اِکسل نمرات استاندارد Z را محاسبه کنم؟ ۱۲۸
پرسش ۷۲: بین نمرات استاندارد Z و منحنی بهنجار چه ارتباطی وجود دارد؟ ۱۲۹
پرسش ۷۳: نمرات استاندارد Z چگونه با آزمون فرضیه ارتباط دارد؟ ۱۳۱
بخش ۹: درک مفهوم معناداری ۱۳۲
پرسش ۷۴: استنباط آماری چه کاربردی دارد؟ ۱۳۳
پرسش ۷۵: مفهوم معناداری چیست و چرا مهم است؟ ۱۳۴
پرسش ۷۶: خطای نوع ۱ و ۲ چیست؟ ۱۳۶
پرسش ۷۷: مراحل کاربرد یک آزمون آماری برای آزمون یک فرضیه پژوهشی چیست؟ ۱۳۹
پرسش ۷۸: تفاوت بین معناداری آماری و معناداری منطقی چیست؟ ۱۴۰
پرسش ۷۹: جعبه‌ابزار اِکسل چیست و چگونه می‌توانم از آن برای آزمون‌های آمار استنباطی استفاده کنم؟ ۱۴۲
پرسش ۸۰: یک تابع اِکسل چیست و چگونه می‌توانم از آن برای آزمون‌های آمار استنباطی استفاده کنم؟ ۱۴۴
بخش ۱۰: درک تفاوت‌های بین گروهی ۱۴۶
پرسش ۸۱: چگونه بدانم که از کدام آزمون آماری استفاده کنم؟ ۱۴۷
پرسش ۸۲: آزمون T میانگین‌های مستقل چیست و یک مثال برای چگونگی کاربست آن کدام است؟ ۱۴۹
پرسش ۸۳: چگونه می‌توانم از طریق اِکسل به آزمون تفاوت بین میانگینهای مستقل بپردازم؟ ۱۵۱
پرسش ۸۴: آزمون T بین میانگین‌های وابسته چیست و یک مثال برای چگونگی کاربست آن کدام است؟ ۱۵۳
پرسش ۸۵: چگونه می‌توانم از طریق اِکسل به آزمون تفاوت بین میانگینهای وابسته بپردازم؟ ۱۵۵
پرسش ۸۶: تحلیل واریانس ساده چیست و یک مثال برای چگونگی کاربست آن کدام است؟ ۱۵۷
پرسش ۸۷: چگونه می‌توانم از طریق اِکسل تحلیل واریانس ساده را محاسبه کنم؟ ۱۵۹
پرسش ۸۸: تحلیل واریانس عاملی چیست و یک مثال برای چگونگی کاربست آن کدام است؟ ۱۶۲
پرسش ۸۹: چگونه می‌توانم از طریق اِکسل تحلیل واریانس عاملی را محاسبه کنم؟ ۱۶۳
پرسش ۹۰: چگونه میتوانم از آزمون‌های ناپارامتریک برای آزمون معناداری استفاده کنم؟ ۱۶۶
پرسش ۹۱: اندازه اثر چیست و چرا مهم است؟ ۱۶۸
بخش ۱۱: نگاهی به روابط بین متغیرها ۱۶۹
پرسش ۹۲: روابط بین متغیرها و معناداری آن‌ها چگونه محاسبه می‌شوند؟ ۱۷۰
پرسش ۹۳: چگونه می‌توانم از طریق اِکسل به بررسی معناداری ضریب همبستگی بپردازم؟ ۱۷۱
پرسش ۹۴: رگرسیون ساده چیست و یک مثال برای چگونگی کاربست آن کدام است؟ ۱۷۴
پرسش ۹۵: چگونه می‌توانم از طریق اِکسل معادله رگرسیون ساده را محاسبه کنم؟ ۱۷۸
بخش ۱۲: روشهای آماری دیگر ۱۸۱
پرسش ۹۶: تحلیل واریانس چند متغیره چیست و چه کاربردی دارد؟ ۱۸۲
پرسش ۹۷: تحلیل کوواریانس چیست و چه کاربردی دارد؟ ۱۸۳
پرسش ۹۸: تحلیل واریانس با اندازه‌گیری مکرر چیست و چه کاربردی دارد؟ ۱۸۴
پرسش ۹۹: رگرسیون چندگانه چیست و چه کاربردی دارد؟ ۱۸۵
پرسش ۱۰۰: تحلیل عاملی چیست و چه کاربردی دارد؟ ۱۸۶

درباره نویسنده

پروفسور نیل جی سالکایند[1]، دکترای خود را از دانشگاه مریلند[2] ایالات‌متحده در رشته روانشناسی تحولی اخذ کرد و بعد از ۳۵ سال تدریس در گروه روانشناسی تربیتی[3] در دانشگاه کانزاس[4]، هم‌اکنون یک استاد بازنشسته است. پروفسور سالکایند، در زمینه روانشناسی شناختی، تحول شناختی کودکان و خانواده‌درمانی پژوهش و تدریس کرده است و فوق‌دکترای خود را در حیطه سبک‌های شناختی و اختلال بیش فعالی کودکان در دانشگاه کارولینای شمالی[5] کسب کرده است. وی حدود ۱۵۰ مقاله و سمینار علمی و حدود ۱۰۰ کتاب و متن علمی به دنیای علم ارائه کرده‌اند. او علاقه‌مند به نوشتن درباره آمار و روش‌های تحقیق به شیوه‌های آموزنده، غیرترسناک و در عین حال بسیار جدی است. وی مؤلف کتاب آمار برای افرادی که فکر می‌کنند از آمار متنفرند (ویرایش پنجم)[6]؛ کتاب آمار برای افرادی که فکر می‌کنند از آمار متنفرند (ویرایش سوم اکسل ۲۰۱۰)[7]؛ کتاب آمار اکسل-راهنمای آسان (ویرایش دوم)[8]؛ کتاب صد پرسش و پاسخ درباره روش‌های تحقیق[9] و ویراستار مجموعه دائره‌المعارف طرح تحقیق[10] بوده است.

---

1. Neil J. Salkind
2 . The University of Maryland
3. Department of Educational Psychology
4 . The University of Kansas
5 . The North Carolina State University
6. Statistics for People Who (Think They) Hate Statistics
7. Statistics for People Who (Think They) Hate Statistics (the Excel Edition)
8. Excel Statistics; A Quick Guide, Second Edition
9 . 100 Questions (and Answers) About Research Methods
10. Encyclopedia of Research Design

مقدمه مترجم

اجازه دهید مقدمه خود را با جمله زیبای دانشمند و مخترع سرشناس ایتالیایی در سده‌های ۱۶ و ۱۷ میلادی آغاز کنم:

"کتابِ بزرگِ طبیعت را با علائم ریاضی نگاشته‌اند"

گالیلئو گالیله (۱۵ فوریه ۱۵۶۴–۸ ژانویه ۱۶۴۲)

کتاب صد پرسش و پاسخ آماری در روانشناسی، کتاب بسیار ارزشمندی است چراکه حاصل تجربیات چهل ساله پروفسور سالکایند در تدریس آمار و روش تحقیق در دانشگاه‌های معتبر دنیا نظیر کانزاس، مریلند و کارولینای شمالی بوده است. پروفسور سالکایند، پرسش‌های دانشجویان خود در حیطه آمار و روش تحقیق در دوره کارشناسی، کارشناسی ارشد، دکتری و پسادکترا را طبقه‌بندی نموده و در قالب سه کتاب تألیف نموده است. این کتاب، اولین مجموعه از این سه کتاب است. من نیز در تجربه چندین ساله خود در تدریس آمار و روش تحقیق در سطوح مختلف تحصیلی در رشته‌های روانشناسی و مشاوره، با پرسش‌های کاملاً مشابهی از جانب دانشجویان مواجه بوده و هستم. لذا با توجه به اشتراک موضوع، محتوا و نیز با توجه به تجربیات ارزشمند پروفسور سالکایند و همچنین متن بسیار روشن و فصیح ایشان، ترجمه این اثر گران‌سنگ را سودمند یافته و بر آن همت گماشتم، چرا که به فرموده امام علی علیه السلام، بر خردمند لازم است که رأی خردمندان را به رأی و نظر خود بیفزاید و دانش حکیمان را به دانش خود اضافه کند[۱].

این کتاب، پاسخ‌های جامع، موجز و مختصری به پرسش‌های مطرح‌شده داده است و برای کسب اطلاعات کامل‌تر می‌بایست به سایر منابع موجود رجوع شود، بنابراین به‌عنوان مدرّس دروس آمار و روش تحقیق، مطالعه این کتاب را به‌تمامی دانشجویان رشته های علوم انسانی بویژه روانشناسی و علوم تربیتی در سطوح تحصیلی کارشناسی و کارشناسی ارشد پیشنهاد می‌کنم. امیدوارم ترجمه این کتاب ارزشمند گام ارزشمندی در راستای دانش‌افزایی دانشجویان علوم انسانی و علوم تربیتی در حیطه آمار باشد.

دکتر رضا سلطانی شال
استادیار گروه روانشناسی دانشگاه گیلان

۱ . امام علی علیه السلام : حقٌّ عَلَی العاقِلِ أن یُضِیفَ إلی رَأیِهِ رَأیَ العُقَلاءِ ، و یَضُمَّ إلی عِلمِهِ عُلومَ الحُکَماءِ. غرر الحکم، ۴۹۲۰

بخش ۱: چرا آمار؟

پرسش ۱: آمار چیست و مطالعه آن چرا اهمیت دارد؟

ما در دنیایی زندگی می‌کنیم که به‌صورت پیش‌رونده‌ای با داده[1] تسخیر شده است. دنیای پر از داده! چه کسی فکر می‌کرد که تحلیل داده‌ها، همان‌قدر که قبلاً برای مؤسسات آموزشی، پژوهشی و بسیاری از مشاغل اهمیت داشت. امروزه برای تیم‌های ورزشی حرفه‌ای اهمیت پیدا کند. با افزایش پدیده‌ها، مردم تمایل پیدا کرده‌اند که اصول علم آمار را نیز برای بررسی الگوهای موجود در حرفه خود و همچنین پیش‌بینی پدیده‌ها به کار ببرند. علم آمار شامل مجموعه‌ای از ابزارهایی است که به ما کمک می‌کنند تا نتایج پیچیده را بهتر درک کنیم، تصمیم کارآمدتری بگیریم و با احتمال خطای کمتر و با اطمینان بیشتر، احتمال وقوع آن‌ها را پیش‌بینی و کنترل کنیم.

علم آمار شامل توصیف، سازمان‌دهی، تحلیل و تفسیر اطلاعات کیفی[2] است. این اطلاعات ممکن است مجموعه‌ای از نمرات یک آزمون، ترجیح نوع خاصی از اتومبیل یا چگونگی امتیازگیری یک گروه بسکتبال نسبت به سایر گروه‌ها، باشد. شما می‌توانید به این مسائل و بسیاری از مسائل دیگری که با تحلیل داده سروکار دارند با کاربست ابزارهای متنوعی که در این کتاب یاد خواهید گرفت، پاسخ دهید.

مطالعه علم آمار به چند دلیل اهمیت دارد. در کاربردی‌ترین پاسخ می‌توان گفت که علم آمار به ما کمک می‌کند تا بر اساس اطلاعاتی که ممکن است تفسیرشان بسیار آسان و یا بسیار سخت باشد، تصمیم‌گیری کنیم. این سودمندی را می‌توان در چند مثال ساده به‌روشنی نشان داد. برای مثال، آیا شاخص میانگین درس آمار دانشجویان یک کلاس، شاخص سودمندتر و کارآمدتری نسبت به نمرات تک تک دانشجویان آن کلاس در آن درس نیست؟ آیا شاخص میانگین ارزیابی کیفیت خدمات دریافتی گروهی از مشتریان، شاخص سودمندتر و کارآمدتری نسبت به نمرات ارزیابی مجزا هر یک از آن‌ها در پاسخ به بیست پرسش مختلف نیست؟

این دو مثال اشاره می‌کنند که توانایی جمع‌آوری، توصیف و تحلیل داده‌ها به ما کمک خواهد کرد تا تصمیم‌های بهتری اتخاذ کنیم که مبتنی بر شواهد دقیق‌تر، منسجم‌تر و کارآمدتر خواهد بود. علم آمار ابزاری است که به ما اجازه می‌دهد دریابیم که مجموعه داده‌های نامنظم قبلی چه معنایی می‌توانستند داشته باشند.

بنابراین، علم آمار، برای تمامی دانش‌پژوهانی که تحقیق انجام می‌دهند، برای تصمیماتی که سیاستمداران ملّی و منطقه‌ای و بین المللی اتخاذ می‌کنند و حتی برای هرکسی که در حرفه خود اطلاعات موجود را تحلیل می‌کند و برای تمام کسانی که برای رسیدن به مقصود مطلوب خود دست به تحلیل اطلاعات می‌زند، ضروری و حیاتی است.

برای پرسش‌های بیشتر، پرسش‌های ۲، ۳ و ۴ را مطالعه کنید.

---

1 . data

2 . quantitative information

پرسش ۲: علم آمار چگونه شروع شد؟

مطالعه علم آمار بسیار فراتر از جمع‌آوری و تحلیل داده‌هاست و البته بیشتر درباره جمع‌آوری و کاربرد اطلاعات در جهت اتخاذ تصمیم‌های مهم است. احتمالاً مردم هرگز نگران این نمی‌شوند که چه مقدار از چه چیزهایی دارند (مثلاً "چقدر غذا تا رسیدن ما به منزل باقی‌مانده است؟" یا "تا زمستان چند روز باقی‌مانده است؟"). درواقع، مردم هرگز نگران این نمی‌شوند که این مقدارها و تعدادها چگونه بر نتایج معینی (مثلاً بر سلامتی) اثرگذار خواهند بود. امّا هنگامی که پای اتخاذ تصمیم‌های مهم به میان بیاید، بتدریج بر نگرانی ها افزوده شده و نیاز به یک تحلیل و تصمیم دقیق احساس می شود. از زمان آغاز علم آمار تاکنون، اعداد همواره با نتایج خاصی همراه و همگام بوده‌اند. اگر دانش‌آموزی در مدرسه عالی عمل می‌کرد و نمرات خوبی کسب می‌کرد، احتمال اینکه در سطوح تحصیلی و رده‌های بالاتر به موفقیت دست می‌یافت بسیار بالا بود. اگر دانش‌آموزی آموزش خوبی دریافت می‌کرد، به‌احتمال بسیار زیاد، شغل مناسبی بعد از پایان دوره تحصیلات در انتظارش بود. بعلاوه، به گذشته نه‌چندان دور که بنگریم، افرادی را می‌توان به یاد آورد که تا حد زیادی شبیه به جمعیت‌شناس[1] های امروزی بودند و کارشان مطالعه جمعیت‌ها و ویژگی‌های شخصیتی‌شان بود و به شمارش و بررسی توزیع جمعیت از منظر مکان زیست، نوع حرفه و نوع تفریحات آن‌ها می‌پرداختند. تمامی این‌ها توسط ریاضیدان‌ها انجام می‌شد، امّا رشته-هایی نظیر زیست‌شناسی و بعدها روانشناسی برای درک و فهم آنچه مورد مشاهده قرار می‌دادند تحت‌فشار قرار داشتند. چرا که اطلاعات زیادی داشتند امّا در طبقه‌بندی، تحلیل و کاربست آن‌ها قدرت کافی نداشتند. در این زمان بود که علم آمار متولد شد. احتمالاً مُبدع اصلی این تولد، کارهای علمی فرانسیس گالتون[2] بود. وی، پسرعموی چارلز داروین[3] (متولد اوایل قرن ۱۹) بود. گالتون مُبدع شیوه بسیار معروفی به نام ضریب همبستگی[4] بود که به بررسی رابطه بین دو متغیر[5] می‌پرداخت. علاقه شخص گالتون مطالعه هوش در خانواده‌ها بود. کار او منجر به شکل‌گیری یک چارچوب برای مقایسه این‌گونه روابط در میان اعضای خانواده شد. بعد از گالتون، جهان شاهد تحولات جدید بود. در حقیقت، در جامعه پیچیده و به‌سرعت در حال رشد، مجموعه اطلاعات درهم‌تنیده‌ای در دسترس قرارگرفته بود که نیازمند روش‌های آماری برای تحلیل و تفسیر بود. در این میان، برخی از افراد متخصص نظیر کارل پیرسون[6] (ریاضیدان) و رونالد آیلمر فیشر[7] (مهندس کشاورزی) شروع به کاربست مفاهیم رشته‌های تخصصی خود در ابعاد مختلف رفتار انسان کردند. از طرف دیگر، با ظهور کامپیوترهای شخصی در حدود ۴۰ سال قبل، فنون آماری بسیار قدرتمندتری در دسترس تمامی کسانی قرار گرفت که علاقه‌مند به بررسی الگوها و مطالعه

---

1 . demographer
2 . Sir Francis Galton, FRS (/ˈfrɑːnsɪs ˈgɔːltən/; 16 February 1822ت 17 January 1911)
3 . Charles Robert Darwin (/ˈdɑːrwɪn/ 12 February 1809ت 19 April 1882)
4 . correlation coefficient
5 . variables
6 . Karl Pearson FRS (/ˈpɪərsən/ 27 March 1857ت 27 April 1936)
7 . Ronald Aylmer Fisher FRS (17 February 1890ت 29 July 1962)

روند مجموعه داده‌های بزرگ بودند. این مجموعه داده‌ها، از مهم‌ترین بخش‌های علم آمار عصر مدرن است. این مسیر به‌جایی رسیده است که امروزه، حتی دانشکده‌ها و متخصصان تیم‌های ورزشی نیز از این رویکرد برای شناسایی آنچه سودمند یا ناکارآمد است، بهره می‌برند. دانش‌پژوه گرامی، شما می‌توانید برای دستیابی به اطلاعات بیشتر درباره تاریخچه علم آمار به وب‌سایت دانشکدهٔ آنسلِم قدیس[1] به آدرس اینترنتی[2] مراجعه کنید.

برای پرسش‌های بیشتر، پرسش‌های ۱، ۳ و ۴ را مطالعه کنید.

1 . Saint Anselm College

2 .www.anselm.edu/homepage/jpitocch/biostatshist.html

پرسش ۳: آمار توصیفی چیست و چگونه استفاده می‌شود؟

آمار توصیفی[1] به‌منظور سازمان‌دهی و توصیف ویژگی‌های یک مجموعه از داده‌ها که توزیع داده‌ها نیز نام دارد، به‌کار برده می‌شود. آمار توصیفی اولین روشی است که برای اکتشاف داده‌ها به‌کار برده می‌شود و اطلاعات مهمی درباره شاخص‌های مهم مجموعه داده نظیر پراکندگی مشاهدات در مجموعه داده در اختیار ما قرار می‌دهد. دو دسته عمومی از آمار توصیفی وجود دارد.

دسته اول شامل آن آماره‌هایی است که به ارزیابی گرایش‌های مرکزی[2] نظیر میانگین[3]، میانه[4] و نما[5] می‌پردازند. این آماره‌ها، شاخص‌های مرکزی نام دارند. تمامی این شاخص‌های مرکزی، به‌نوعی نشان‌دهنده "حد وسط" یک مجموعه داده هستند. هرکدام از این آماره‌ها می‌توانند نمایانگر بهترین نقطه‌ای باشند که نشان‌دهنده یک مجموعه داده است. برای مثال، اگر ما علاقه‌مند به فروش همبرگر در یک رستوران محلی هستیم، اولین پرسشی که باید بپرسیم این است که میانگین فروش هفتگی همبرگر در این مکان چقدر است؟ یا بیشترین فراوانی فروش اتومبیل مربوط به کدام مدل است؟

دسته دوم آمار توصیفی، شامل روش‌هایی است که بیان‌گر تغییرپذیری یا پراکندگی و چگونگی انتشار داده‌ها در یک مجموعه داده است. دامنه تغییرات[6]، انحراف استاندارد[7] و واریانس[8] از این دسته آماره‌ها محسوب می‌شوند. این آماره‌ها، شاخص‌های پراکندگی نام دارند. شاخص‌های پراکندگی به ما می‌گویند که مشاهدات یک مجموعه داده تا چه میزان از یکدیگر متفاوت هستند. برای مثال، اگر بخواهیم بدانیم که مهارت‌های خواندن یک گروه از کودکان، چقدر مشابه با هم و یا متفاوت از هم است، می‌بایست به انحراف استاندارد نمرات آزمون خواندن‌شان نگاهی بیاندازیم. هر چه انحراف استاندارد کمتر و کوچک باشد، نشان‌دهنده آن است که تفاوت اندکی بین کودکان در مهارت‌های خواندن وجود دارد و در واقع انحراف کمی از میانگین دارند و این یعنی تشابه زیادی بین‌شان وجود دارد.

ترکیب این دو دسته از آماره‌های توصیفی که به شاخص‌ها مرکزی و شاخص‌های پراکندگی معروف‌اند، به ما تصویر مناسبی از ویژگی‌های مجموعه داده‌ها و چگونگی تفاوت آن‌ها از سایر مجموعه داده‌ها می‌دهد و همان‌گونه که در ادامه کتاب خواهید دید، این دو دسته از شاخص‌ها، از بنیادی‌ترین مؤلفه‌های عملیات پیچیده آماری نظیر ارزیابی معناداری تفاوت بین میانگین دو گروه از مشاهدات محسوب می‌شوند.

برای پرسش‌های بیشتر، پرسش‌های ۴، ۷ و ۱۶ را مطالعه کنید.

---

1 . descriptive statistics
2 . central tendency
3 . mean
4 . median
5 . mode
6 . range
7 . standard deviation
8 . variance

پرسش ۴: آمار استنباطی چیست و چگونه استفاده می‌شود؟

آمار استنباطی[1] به ما اجازه می‌دهد که یافته‌های یک گروه بزرگ‌تر که جامعه[2] نام دارد را از داده‌های یک گروه بسیار کوچک که نمونه[3] نام دارد استنباط کنیم. آمار استنباطی به‌نوعی گسترش آمار توصیفی است، امّا همیشه این‌گونه نیست. برخی مواقع، دانستن و به دست آوردن میانگین یک مجموعه از نمرات از طریق آمار توصیفی کافی است و دیگر نیازی به کاربست آمار استنباطی نیست. برای مثال، آیا میانگین درس آمار استنباطی دانشجویان رشته روانشناسی از میانگین درس آمار استنباطی دانشجویان رشته مشاوره متفاوت است؟

در آمار استنباطی، پژوهشگر با استفاده مقادیر نمونه آماره ها را محاسبه کرده و سپس با کمک تخمین و یا آزمون فرض آماری، آماره ها را به پارامترهای جامعه تعمیم می دهد. برای تجزیه و تحلیل داده ها و آزمون فرضیه های پژوهش از روش های آمار استنباطی استفاده می شود. چنانچه به‌جای مطالعه کل اعضای جامعه، بخشی از آن را با استفاده از فنون نمونه‌گیری انتخاب و مطالعه کنیم و بخواهیم نتایج حاصل از آن را به‌کل جامعه تعمیم دهیم، از روش‌هایی استفاده می‌شود که موضوع آمار استنباطی است. درواقع، آمار استنباطی به شیوه‌هایی اطلاق می‌شود که از طریق آن‌ها ویژگی‌های گروه‌های بزرگ بر اساس اندازه‌گیری همان ویژگی‌ها در گروه‌های کوچک استنباط می شود. استنباط از نمونه به جامعه با نقش علم احتمال شروع می‌شود. درواقع احتمال، پُل رابط بین آمار توصیفی و استنباطی است. در زمینه اینکه چگونه می‌توان از آمار استنباطی استفاده کرد اجازه دهید مثالی بزنیم. همان‌گونه که می‌دانید یکی از ویژگی‌های علوم انسانی، جبری بودن آن است. به این معنی که وقوع هر رویدادی دلیلی دارد و صرفاً تصادفی نیست.

در همین راستا، می‌خواهیم به علّت زیربنایی مهارت خواندن دانش‌آموزان پی ببریم. در نظر بگیرید که می‌خواهیم تفاوت میانگین‌های دو گروه از دانش آموزان یک مدرسه ابتدایی در یک آزمون مهارت‌های خواندن را ارزیابی کنیم. یک گروه آموزش فوق‌برنامه درسی دریافت کرده است، درحالی‌که گروه دیگر این آموزش را دریافت نکرده است. اکنون دو گروه مورد مقایسه قرار می‌گیرند و نتایج حاصله به جمعیت بزرگ‌تری که دو نمونه از آن انتخاب شده، تعمیم داده می‌شوند. امّا سوال اینجاست که چرا تمامی جمعیت کودکان این مدرسه مورد این ارزیابی قرار نگرفتند؟ در اغلب موارد، جمعیت مورد مطالعه بسیار بزرگ است و در نتیجه هزینه، زمان و نیروی انسانی بسیار بالایی نیاز خواهد بود. دانشمندان در طی سال‌ها مطالعه و تلاش علمی (در طی ۲۰۰ سال گذشته) به روش‌های کارآمد و دقیقی دست‌یافته‌اند که به کمک آن‌ها می‌توان با دقت بسیار بالا، یک نمونه کوچک را به‌گونه‌ای انتخاب کرد که نماینده واقعی و دقیقی از جمعیت بزرگ‌تری باشد که قرار است اطلاعات به آن تعمیم داده شود؛ نکته اینجاست که یک نمونه چقدر

1 . inferential statistics
2 . population
3 . sample

شبیه جامعه‌ای است که از آن گرفته شده است. بنابراین اکنون می‌توان دریافت که میزان مُعرِّف[1] بودن یک نمونه، از کلیدی‌ترین مؤلفه‌های کاربست آمار استنباطی است. مُعرّفِ جمعیت بودنِ یک نمونه، به این موضوع کمک می‌کند که تا چه حد می‌توان نتایج حاصله از تحلیل داده‌های نمونه را به جامعه هدف تعمیم داد. میزان بسیار بالای دقت نتایج یک فرایند استنباطی[2] به این موضوع بستگی دارد که تا چه حد نمونه به‌خوبی از جامعه انتخاب شده است. هر مقدار که یک نمونه بیشتر و دقیق‌تر مُعرّف جمعیت باشد، با اطمینان بیشتری می‌توان شاخص‌ها و ویژگی‌های جمعیت را از نمونه استنباط کرد و با احتمال بسیار بالاتری نیز می‌توان نتایج را از یک نمونه به جمعیت یا جامعه مادر تعمیم داد. درمجموع، آمار توصیفی به ما اجازه می‌دهد که ویژگی‌های یک مجموعه داده را توصیف کنیم، درحالی‌که آمار استنباطی به ما اجازه می‌دهد این مشاهدات آمار توصیفی را به گروه‌های بزرگ‌تر تعمیم دهیم.

برای پرسش‌های بیشتر، پرسش‌های ۶۰، ۷۰ و ۷۷ را مطالعه کنید.

1 . Representativeness
2 . inferential procedure

پرسش ۵: من نمی‌خواهم متخصص آماری شوم، پس چرا باید دوره آمار بگذرانم؟

پاسخ به این پرسش فراتر از یک پاسخ بسیار واضح و ساده نظیر این است که بالاخره شما نیازمند گذراندن یک دوره آماری برای کامل کردن واحدهای درسی و در نهایت اخذ مدرک تحصیلی‌تان هستید و یا می‌بایست برای یک امتحان درسی مطالعه کنید و بخشی از دانش مورد نیاز امتحان شما مبحث آمار باشد. امّا درواقع، حداقل چهار دلیل خوب برای گذراندن یک دوره آمار برای شما وجود دارد:

الف) درس آمار یک موضوع چالش‌برانگیز و از نظر عقلانی جالب است، به‌گونه‌ای که ممکن است در دوره‌های آموزشی قبلی خود آن را تجربه نکرده باشید. درس آمار یک شاخه علمی است که مهارت‌های شما را افزایش می‌دهد. در واقع یک تلاش ارزشمند و مهارت محور است.

ب) در جامعه‌ای که تمام تصمیم‌گیری‌هایش به‌گونه‌ای روزافزون مبتنی برداده و شواهد است، درک و فهم آمار بسیار ارزشمند و حیاتی است. درس آمار شمار را آماده می‌کند تا درک عمیقی از الگو و روند داده‌های در دسترس و همچنین استفاده از آن‌ها در اتخاذ تصمیم آگاهانه داشته باشید. شما نخستین کسی خواهید بود که خواهد پرسید که آیا داده‌هایی در دسترس است که بتوان از طریق آن‌ها فهمید که کدام روش درمان پزشکی اثربخش‌تر است؛ آیا یک روش آموزش ریاضی بهتر از روش دیگر است؛ یا آیا مطالعه کتاب‌های درسی به‌صورت الکترونیک اثربخش‌تر از کتاب‌های چاپی است؟ علم آمار یک مجموعه از روش‌هایی است که به شما نشان می‌دهد که چگونه سؤالات مبتنی بر پژوهش[1] را که ممکن است در ذهنتان ایجاد شود، طرح کنید و پاسخ دهید.

ج) با مطالعه علم آمار شما شهروند آگاه‌تری خواهید بود به‌گونه‌ای که در تعامل با همکلاسی‌ها، همکاران، اساتید و دیگران و در روابط اجتماعی هوشمند خواهید بود و آمادگی بیشتری خواهید داشت. برای نمونه، هنگامی‌که نتایج یک مقاله علمی مورد بحث قرار می‌گیرد و یا هنگامی‌که بحث و نتیجه‌گیری یک مقاله علمی بر اساس ابعاد اجتماعی، پزشکی و یا رفتاری مورد بحث واقع می‌شود، قادر خواهید بود که مباحث مطروحه را عمیقاً درک کنید و به ترتیب و با روش آگاهانه و منطقی به آن‌ها پاسخ دهید.

د) تکالیف دوره درسی آمار، یک مسیر بسیار عالی و ارزشمند برای فرصت‌های حرفه‌ای و تحصیلی‌تان نظیر یک حرفه دانشگاهی فراهم می‌کند و اگر شما اخیراً دانش‌آموخته شده‌اید، برای آینده‌تان می‌تواند یک فرصت شغلی فراهم کند. تمامی این‌ها از طریق آموزش آمار فراهم می‌شود؛ بنابراین درک عمیق شما از علم آمار به‌ویژه آمار توصیفی و استنباطی، پتانسیل زیادی برای موفقیت حرفه‌ای شما فراهم خواهد کرد.

برای پرسش‌های بیشتر، پرسش‌های ۱، ۲ و ۶ را مطالعه کنید.

---

1 . research based question

پرسش ۶: از کدام بسته‌های نرم‌افزاری آماری می‌توانم استفاده کنم؟ SPSS یا اِکسل و یا سایر بسته‌های نرم‌افزاری؟

انواع متنوعی از بسته‌های نرم‌افزاری تحلیلی آمار در دسترس هستند. تعدادی رایگان‌افزار[1]، برخی اشتراک‌افزار[2] و برخی نیز محصولات تجاری هستند که می‌بایست خریداری شوند. رایگان‌افزار و اشتراک‌افزار هر دو به برنامه‌ها یا نرم‌افزارهایی گفته می‌شود که می‌توانید بدون پرداخت هیچ‌گونه هزینه‌ای از اینترنت دانلود کنید و از آن‌ها استفاده کنید؛ اما به‌هرحال این دو باهم تفاوت دارند. شما می‌توانید از رایگان‌افزارها بدون محدودیت زمانی و برای همیشه استفاده کنید درصورتی‌که اشتراک‌افزارها را فقط می‌توانید برای مدت‌زمان محدود معمولاً ۳۰ روزه برای آزمایش و آزمون استفاده کنید. اینکه از کدام‌یک استفاده می‌کنید به عوامل متعددی بستگی دارد.

SPSS که مخفف بسته آماری برای علوم اجتماعی[3] است و برای اولین بار در سال ۱۹۶۸ توسط محققان[4] علوم اجتماعی طراحی شد، یک بسته نرم‌افزاری بسیار معروف است و توسط محققان علوم اجتماعی و رفتاری مورد کاربست زیادی قرار می‌گیرد. این بسته آماری گران‌قیمت است، امّا بسیاری از مؤسسات آموزش عالی[5] آن را بر روی سِرورهای محوطه دانشگاه[6] نصب می‌کنند؛ بنابراین اگر دانشجو هستید می‌توانید از آن بدون هیچ‌گونه هزینه‌ای استفاده کنید. بعلاوه، نسخه‌های دانشجویی[7] ارزان‌قیمت آن نیز موجود است که البته قابلیت‌های محدودتری دارند.

اکسل[8] که معروف‌ترین برنامه صفحه گسترده جهان است، به شما اجازه می‌دهد تا ردیف‌ها و ستون‌ها اعداد و کلمات را تنظیم و طراحی کنید. هرچند اِکسل یک برنامه صرفاً تحلیل آماری نیست، امّا دربرگیرنده بسیاری از کارکردها و روش‌های تحلیل آماری است که تقریباً می‌تواند جایگزین نرم‌افزاری نظیر Spss شود. اِکسل روش‌های ترسیم بسیار قدرتمندی دارد و بخشی از بسته مایکروسافت آفیس است و می‌توان به‌راحتی نتایج تحلیل را به فضای word انتقال داد. همانند Spss، بسیاری از مؤسسات از اِکسل نیز استفاده می‌کنند. در این کتاب از اِکسل برای مثال‌هایمان استفاده کردیم زیرا این نرم‌افزار بسیار سهل‌الوصول است و تمام کاربران از سطح مبتدی تا متوسط می‌توانند به‌راحتی از آن استفاده کنند.

برای کارتان می‌بایست از کدام بسته نرم‌افزاری استفاده کنید؟ هم Spss و هم اِکسل برای تقریباً تمامی مثال‌هایی که در این کتاب موجود است، کاربرد دارند. امّا در مجموع، هنگامی‌که می‌خواهید تصمیم بگیرید که از کدام بسته نرم‌افزاری در کار خود استفاده کنید، پرسش‌های پیش رو را مورد توجه قرار دهید:

---

1 . Freeware
2 . Shareware
3 . Statistical Package for the Social Sciences (SPSS)
4 . Norman H. Nie, Dale H. Bent, and C. Hadlai Hull
5 . postsecondary institutions
6 . campus-wide servers
7 . student versions
8 . Excel

- این برنامه چقدر هزینه‌بر است؟ آیا از طریق سِرورهای محوطه دانشگاه در دسترس است؟
- آیا نسخه دانشجویی آن موجود است و آیا قادر به خرید آن هستید؟
- آیا برنامه، پیمانه‌ای است؟ اگر چنین است به کدام پیمانه نیاز دارید؟

برنامه‌نویسی پودمانی یا پیمانه‌ای[1] یک روش طراحی نرم‌افزار است برای افزایش وسعت نرم‌افزاری که از قسمت‌های مجزایی به نام پودمان (پیمانه) تشکیل شده است. یک ماژول (پودمان یا پیمانه)، تکه برنامه‌ای است که با یک استاندارد و قالب مشخص طراحی‌شده و انجام یکسری از فعالیت‌ها را بر عهده دارد. در عوض مدیریت امکانات مورد نیاز ماژول و همچنین مدیریت بر روی خود ماژول توسط سایت انجام می شود. مثلاً شما با افزودن ماژول تحلیل آماری، تمامی قابلیت‌های یک سایت دارای تحلیل‌های آماری آنلاین را به سایت موجود خود خواهید افزود. در عوض ماژول جدید به دلیل ارث بردن و استفاده از قابلیت‌های پایه‌ای سایت، چنان بر روی سایت شما عمل می‌نماید که گویی از ابتدا درون سایت بوده و بخشی جدا از برنامه نیست. برای خرید یک نرم‌افزار پرسش‌های پیش رو را مورد نظر قرار دهید:

- آیا شرکت سازنده آن، حمایت فنی تلفنی یا اینترنتی را به‌آسانی در دسترس گذاشته است؟
- آیا سیستم شما، از نظر سخت‌افزاری، حافظه، مکان ذخیره و ... برای اجرای آن نرم‌افزار مناسب است؟
- آیا سیستم کامپیوتر شما، از نظر سخت‌افزاری قابلیت اجرای آن نرم‌افزار را دارد؟
- قبل از خرید نرم‌افزار، مطمئن شوید که نیازهای شما را برطرف خواهد کرد. امّا همیشه به خاطر داشته باشید که تقریباً تمامی برنامه‌ها همه نیازمندی‌های آماری بنیادی را برطرف نخواهند کرد.

برای پرسش‌های بیشتر، پرسش‌های ۱۶، ۷۹ و ۸۰ را مطالعه کنید.

1 . Modular programming

بخش ۲: درک مقیاس‌های گرایش مرکزی

پرسش ۷: شاخص‌های گرایش مرکزی کدام‌اند و چرا به‌کاربرده می‌شوند؟

شاخص‌های مرکزی[1] به مقدار یا عددی که ویژگی گروه را به بهترین شکل نشان می‌دهد، اشاره می‌کنند. مترادف شاخص مرکزی، نمره‌ی متوسط یا معدل[2] است؛ بنابراین شاخص‌های مرکزی یا شاخص‌های گرایش به مرکز، شاخص‌هایی هستند که با استفاده از آن‌ها مجموعه‌ای از داده‌ها در یک مقدار یا عدد که نماینده‌ی آن مجموعه است، خلاصه می‌شود.

شاخص‌های گرایش مرکزی عبارت‌اند از: میانه، میانگین و نما. این سه شاخص از نوع اول آمار توصیفی محسوب می‌شوند که به‌عنوان معیارهایی برای توصیف ویژگی‌های یک مجموعه داده به‌کاربرده می‌شوند. نوع دوم آمار توصیفی شاخص‌های تغییرپذیری هستند که در ادامه درباره آن‌ها بحث خواهد شد. شاخصهای گرایش مرکزی که به شاخص‌های "نمره‌ی متوسط یا معدل" نیز شناخته می‌شوند، نشان‌دهنده مرکزی‌ترین و یا مُعرِّف‌ترین نقطه یک مجموعه داده هستند. به‌عبارت‌دیگر، اگر شما می‌بایست فقط یک عدد برای معرفی کردن مجموعه داده‌های خود داشته باشید، باید یکی از شاخص‌های گرایش مرکزی را انتخاب کنید. برای مثال اگر از شما به‌عنوان فروشنده درخواست شود که فروش اتومبیل در شرکت هوندا در ماه اوت را با بهترین شاخص گزارش کنید، می‌بایست از شاخص گرایش مرکزی "میانگین" در آمار توصیفی استفاده کنید. نکته مهم این است که شاخص گرایش مرکزی خاصی که هر محقق برای توصیف داده‌هایش به کار می‌برد، در ابتدا به نوع داده‌ها بستگی دارد.

"میانه" نمره‌ای است که نقطه میانی دامنه‌ای از نمرات را که از کوچک‌ترین تا بزرگ‌ترین مرتب شده‌اند، نشان می‌دهد. میانه نقطه‌ای است که ۵۰ درصد نمرات در بالای آن و ۵۰ درصد نمرات در پایین آن جای گرفته‌اند. در اکثر مواقع، هنگامی‌که در مجموعه نمرات، اعداد کرانه‌ای (نمره‌های بسیار بالا و یا بسیار پایین) وجود دارند، از میانه استفاده می‌شود تا حد وسط نمرات از اعداد کرانه‌ای متاثر نشود. میانه دقیقاً در وسط یک مجموعه از مقدارها و اعداد قرار دارد. برای مثال اگر سه دانش‌آموز در یک آزمون نمره‌های ۸۱، ۸۴ و ۹۳ گرفته باشند، شاخص گرایش مرکزی آن برابر ۸۴ خواهد بود.

"نما" یا مُد که در میان شاخص‌های گرایش مرکزی مقیاس بسیار غیردقیقی است، عددی است که تعداد بیشترین نمره تکرار شده در یک مجموعه نمره را نمایان می‌سازد. در یک تعریف ساده، نما به داده‌ای گفته می‌شود که در یک مجموعه از اعداد بیشتر از همه رخ دهد و بیشترین فراوانی را داشته باشد. برای مثال اگر در یک کلاس هفت نفر ۱۲ ساله، ده نفر ۱۳ ساله و چهار نفر ۱۴ ساله باشند، نما ۱۳ خواهد شد زیرا در این مجموعه اعداد سن، بیش از همه دانش آموز ۱۳ ساله

1 . Measures of Central Tendency
2 . average

حضور دارد. نما در اغلب مواقع زمانی به‌کاربرده می‌شود که ماهیت داده‌ها اسمی[1] یا طبقه‌ای[2] باشند.

"میانگین"، معدل ریاضی یک مجموعه نمره است. بااین‌حال، چند نوع متفاوت از میانگین وجود دارد. میانگین ریاضی نقطه‌ای است که به‌نوعی نقطه ثقل یک مجموعه داده محسوب می‌شود. درواقع نقطه‌ای است که نمرات مختلف دقیقاً در آن نقطه معیار میانی را می‌یابند. میانگین معمولی یا میانگین حسابی (که گاهی اوقات متوسط نیز نامیده می‌شود) برابر با حاصل‌جمع تمام مقدارها تقسیم بر تعداد آنها است. برای مثال اگر سه دانش‌آموز در یک آزمون نمره‌های ۸۱، ۸۴ و ۹۳ گرفته باشند، میانگین برابر است با ۸۶.

علاوه بر این، یکی از پرسش‌های جالب کاربردی هنگامی‌که با نوعی از داده سروکار داریم این است که فرق بین میانگین و معدل[3] چیست؟ در بسیاری از متون، این دو به‌جای یکدیگر به کار می‌روند، اما در متون علمی میانگین بیشتر به کار می‌رود. برخی ریاضی‌دانان تفاوتی بین این دو قائل‌اند، به این صورت که میانگین را یکی از روش‌های بیان معدل می‌دانند. در واقع، میانگین نیز نوعی نمره متوسط یا معدل است. پیشنهاد ما این است که در متون علمی از میانگین که متداول‌تر است استفاده کنید

برای پرسش‌های بیشتر، پرسش‌های ۹، ۱۱ و ۱۳ را مطالعه کنید.

1 . nominal
2 . categorical
3 . average

پرسش ۸: آیا مثالی برای چگونگی کاربست شاخص‌های گرایش مرکزی وجود دارد؟

معدل (نمره متوسط) در اغلب مواقع به‌عنوان یک نقطه مرجع برای ویژگی‌های یک مجموعه از نمرات بکار برده می‌شود؛ نظیر معدل قد زنان و مردان. معدل‌ها در حقیقت حد وسط مجموعه نمرات هستند به این معنا که معدل، بهترین مُعرّفِ تمامی نمرات یک مجموعه نمره هستند. در آمار توصیفی و استنباطی، معدل‌ها (نما، میانه، میانگین) به‌صورت بسیار گسترده به‌عنوان شاخص‌هایی بکار برده می‌شوند که می‌توانند جایگزین یکدیگر شده و تعیین کنند که آیا بین دو یا چند گروه تفاوت وجود دارد یا ندارد؟

در یکی از مطالعات، با استفاده از نظریه مثلث عشق[1] استرنبرگ[2] که عشق را به سه ضلع مثلث یعنی اشتیاق، صمیمیت و تعهد مفروض می‌داند، ادراک از عشق در گستره عمر مورد بررسی قرار گرفت. نتایج این مطالعه حاصل تکمیل مقیاس کوتاه مثلث عشق توسط ۳۰۰۰ فرد ۱۲ تا ۸۸ سال در دو جنس بود. در این مطالعه، معدل‌ها چه کاربردی دارند؟ محققان در پاسخ به سؤالات خود، از آزمون آماری تی که برای مقایسه میانگین‌های نمرات دو گروه از شرکت‌کنندگان بکار برده می‌شود، استفاده کردند. آزمون تی، از میانگین حسابی به‌عنوان معدل استفاده می‌کند و لذا به پژوهشگر کمک می‌کند تا به این نتیجه برسد که آیا تفاوت میانگین‌های مشاهده شده بین دو گروه، ناشی از شانس و یا ناشی از تفاوت موجود در میزان متغیر مورد مطالعه (مقیاس عشق) بوده است.

نتایج نشان داد که نوجوانان سنین ۱۲ تا ۱۷ ساله در مقایسه با جوانان ۱۸ تا ۳۰ ساله میانگین کمتری در مؤلفه‌های عشق را گزارش کردند؛ بزرگسالان بالای ۵۰ سال در مقایسه با جوانان و میان‌سالانِ ۳۰ تا ۵۰ سال، میانگین کمتری از مؤلفه‌های اشتیاق و صمیمیت و سطوح یکسانی از مؤلفه تعهد گزارش کردند. بعلاوه، از نظر ادراک در مورد سه مؤلفه عشق در بین دو جنس مرد و زن در میانگین تفاوت معناداری مشاهده شد، امّا این میزان تفاوت کمتر از آن چیزی بود که محققان انتظارش را داشتند. بنابراین، استفاده از میانگین به‌عنوان شاخصی که نمایانگر تفاوت‌ها باشد در تحلیل‌های آماری ابتدایی بسیار رایج است. بعلاوه، در ابتدای بخش آماری تمامی مطالعات منتشر شده و تمام پژوهش ها، در کنار متغیرهای جمعیت‌شناختی[3] نظیر جنسیت، قومیت، نژاد، سن، سطوح تحصیلات، طبقه اجتماعی و اقتصادی، به ویژگی های توصیفی متغیرهای مورد بررسی در نمونه مورد مطالعه پرداخته می شود که میانگین، میانه و نما از ضروری ترین این شاخص ها به شمار می روند. برای مطالعه بیشتر درباره مثال موجود در این پرسش به مطالعه مرجع زیر بپردازید.

Sumter, S. R., Valkenburg, P. M., & Peter, J. (2013). Perceptions of love across the lifespan: Differences in passion, intimacy, and commitment. International Journal of Behavioral Development, 37(5), 417-427.

برای پرسش‌های بیشتر، پرسش‌های ۷۴، ۸۲ و ۸۳ را مطالعه کنید.

---

1 . triangular theory of love
2 . Sternberg R
3 . demographic variable

پرسش ۹: میانگین چیست و چگونه محاسبه می‌شود؟

میانگین پرکاربردترین، مشهورترین، معتبرترین و باثبات‌ترین شاخص از شاخص‌های گرایش مرکزی است. میانگین از نظر آماری، مجموع نمرات یک گروه تقسیم بر تعداد اعضای همان گروه است. در یک تعریف فنی‌تر می‌توان گفت میانگین نقطه‌ای است که مجموع انحراف نمرات از آن صفر است. نوع میانگینی که اینجا در مورد آن بحث می کنیم میانگین ریاضی[1] نام دارد.

فرمول میانگین برای یک مجموعه داده به این صورت است که $\bar{x}$ نشانگر، $\sum x$ برابر است با مجموع اعداد و N برابر است با تعداد کل داده‌ها:

$$\bar{x} = \frac{\sum x}{N}$$

توجه شود که در آمار، n نشانگر حجم نمونه و N نشانگر حجم جامعه است. برای محاسبه میانگین، گام های زیر را بردارید:

- فهرستی از تمامی داده ها فراهم کنید؛
- داده ها را جمع کنید؛
- عدد جمع بدست آمده را بر تعداد مشاهدات تقسیم کنید؛

برای مثال جمع اعداد ۴، ۶، ۷، ۸، ۵ برابر با ۳۰ و میانگین آنها برابر با ۶ است. میانگین را در اغلب مواقع با $\bar{x}$ و یا M نشان می دهند. یک هشدار در کابرد شاخص گرایش مرکزی میانگین این است که این شاخص نبست به نمرات کرانه ای حساس نیست. برای مثال، میانگین نمرات ۴، ۶، ۷، ۸ و ۲۰ برابر با ۹ است، امّا این نمره، در این مثال شاخص مناسب و یا حداقل بهترین شاخص نمایانگر معدل این مجموعه نمرات نیست، چراکه معدل این مجموعه نمره توسط یک نمره انتهایی یعنی ۲۰ بالا کشیده شده است. یکی از انواع میانگین حسابی، میانگین وزنی[2] است که برای هر مقدار، وزنی به عنوان میزان اهمیت آن در محاسبه میانگین در نظر می‌گیرد. میانگین وزنی، گاهی معدل[3] نیز نامیده می‌شود. اگر وزن یا اهمیت مقدارها یکسان در نظر گرفته شود، میانگین وزنی به میانگین حسابی تبدیل خواهد شد. از میانگین وزنی در آمار بخصوص در مطالعات مربوط به جمعیت استفاده زیادی می‌شود. این نوع از میانگین را با نماد x_w نشان می‌دهند. برای محاسبه میانگین وزنی هر مقدار در وزن آن مقدار ضرب شود و مقادیر مرحله اول با هم جمع شوند. برای مثال میانگین وزنی مقادیر ۱، ۳، ۵، ۷ و ۱۰ با وزن‌های ۰/۲، ۰/۲، ۰/۴، ۰/۱ و ۰/۱ برابر با ۴/۵ خواهد بود.

برای پرسش‌های بیشتر، پرسش‌های ۷، ۱۵ و ۱۶ را مطالعه کنید.

1 . Arithmetic Mean
2 . Weighted Mean
3 . Average

پرسش ۱۰: آیا مثالی برای چگونگی کاربست میانگین وجود دارد؟

بیش از پنجاه سال است که ایالات‌متحده آمریکا به یک جامعه اکثریت-اقلیت تبدیل شده است، به‌گونه‌ای که اقلیت های نژادی مختلف در حال ساختن یک جامعه یکپارچه هستند ولی همچنان جمعیت اکثریت آمریکا را تشکیل نمی‌دهند؛ بنابراین برای دانشمندان علوم رفتاری و اجتماعی بسیار مهم است که اطلاعات بیشتری درباره جمعیت‌های اقلیت نظیر لاتین‌ها، آسیایی‌ها، آسیایی-آفریقایی‌ها و ... در آمریکا کسب کنند. در همین راستا کریستوفر الیسون و همکاران[1] به مطالعه اهمیت رشد سریع جمعیت لاتین‌تبارها در ایالات‌متحده و نقشی که کاتولیزم[2] در شکل‌دهی نگرش شان درباره خانواده، پرداختند. هرچند که حدود یک‌سوم لاتین تبارها کاتولیک نیستند. یک فعالیت متداول در میان محققان تحلیل و بررسی داده‌های در دسترس و موجود است. در این مطالعه، داده‌های مطالعه ملی درباره مذهب و زندگی خانوادگی که در سال ۲۰۰۶ جمع‌آوری شده بود، مورد تحلیل قرار گرفت. در این مطالعه افراد ۱۸ تا ۵۹ ساله شاغل به کار از ۴۸ ایالت آمریکا مورد بررسی قرار گرفتند. این مطالعه به بررسی ارتباط بین ابعاد چندگانه مذهب (ایمان، حضور در کلیسا، دعا و نیایش، باور به کتاب مقدس) و نگرش لاتین تبارها درباره ازدواج، طلاق، زندگی غیررسمی زوجین و رابطه جنسی اتفاقی پرداخت. نتایج نشان داد که در مقایسه با کاتولیک‌ها، پروتستان‌ها تمایل به نگرش‌های محافظه‌کارانه‌تری نسبت به مسائل مربوط به خانواده داشتند. لاتین‌تبارهایی که منظم‌تر و مداوم‌تر در مراسم مذهبی و نیایش شرکت می‌کردند، دیدگاه سنتی‌تری داشتند. تحلیل داده‌ها نشان داد که متغیرهای مذهبی همانند متغیرهای جمعیت شناختی و اجتماعی-اقتصادی، این قابلیت را داشتند که در تبیین نگرش‌های لاتین تبارها نقش معناداری داشته باشند. برای رسیدن به این نتایج، تحلیل داده‌ها از طریق محاسبه و مقایسه میانگین‌ها انجام شد. قسمتی از نتایج در جدول زیر نشان داده شده است. علامت ** و علامت * به ترتیب نشان دهنده معنادار بودن تفاوت با ۹۹ و ۹۵ درصد اطمینان است.

متغیر وابسته	کل جمعیت	کاتولیک	پروتستان	سایر اقلیتها	غیر اقلیت
طلاق خوب است	۳/۲۵ (۷۶۵)	۳/۱۸ (۴۸۳)	۳/۵۹ (۱۴۷)	۳/۴۶ (۳۵)	۲/۹۸ (۸۰)
آزمون تی دو گروه مستقل	-	-	۲/۵۵ **	۰/۹۱	-۰/۹۹
زوجین عاشق جدا نمی شوند	۲/۲۴ (۷۸۴)	۲/۰۳ (۴۸۹)۲	۲/۹۰ (۱۵۴)	۲/۸۱ (۳۶)	۱/۸۰ (۸۴)
آزمون تی دو گروه مستقل	-	-	۶/۴۳ **	۳/۱۴ **	-۱/۴۱ *

در این جدول می‌توانید مؤلفه‌های نگرش (برای مثال طلاق مشکلی ندارد) و میانگین نمرات پاسخ‌دهندگان ازجمله کاتولیک‌ها، پروتستان‌ها، سایر اقلیت‌ها و افرادی که در اقلیت خاصی

1 . Christopher Ellison, Nicholas Wolfinger, Ramos-Wada
2 . Catholicism

نبودند، را ملاحظه کنید. این میانگین‌ها نیازمند تحلیل‌های بیشتری هستند، امّا همین هم می‌تواند مثال خوبی از چگونگی سطوح میانگین گروه‌های مختلف درباره نگرششان نسبت به مسئله طلاق باشد.

در یک مثال برای میانگین وزنی، یک دانشجو در آزمون‌های نهایی، نمرات زیر را در دروس مختلف کسب کرده است. از آنجایی که میزان اهمیت هر درس در محاسبه معدل با تعداد واحد آن درس مطابقت دارد، برای محاسبه معدل دانشجو تعداد واحد به عنوان ضریب یا وزن آن نمره محسوب می شود.

نام درس	تعداد واحد (W)	نمره (X)	حاصلضرب
آمار استنباطی	۴	۱۵	۶۰
روانشناسی رسانه	۱	۲۰	۲۰
روانشناسی شخصیت	۲	۲۰	۱۰
روانشناسی تحولی	۳	۱۹	۵۷
مجموع	۱۰		۱۷۷
معدل (میانگین وزنی)			۱۷۷÷۱۰=۱۷/۷

همانطور که دیده می‌شود، با اینکه اکثر نمرات دانشجو بیشتر از ۱۷ است ولی کمبود نمره در درس آمار استنباطی ۴ واحدی باعث شده معدل وی کاهش یابد. برای توضیح بیشتر می توان افزود که میانگین وزنی یا میانگین حسابی وزنی یا میانگین موزون[1] عبارت است از میانگین حسابی یک مجموعه داده‌های نابرابر و ناموزون. در محاسبه میانگین وزنی یک مجموعه از داده های (عوامل) نابرابر، برای هر یک از عامل‌ها، وزن یا ارزش معینی در نظر گرفته می‌شود و سپس آن عامل در وزن معین ضرب می‌گردد. آن‌گاه جمع این ارقام به دست آمده بر مجموع وزنها تقسیم می‌شود.

$$\bar{x} = \frac{\sum_{i=1}^{n} w_i x_i}{\sum_{i=1}^{n} w_i}$$

برای مطالعه بیشتر درباره مثال موجود در این پرسش به مطالعه مرجع زیر بپردازید.

Ellison, C. G., Wolfinger, N. H., & Ramos-Wada, A. I. (2013). Attitudes toward marriage, divorce, cohabitation, and casual sex among working-age Latinos: Does religion matter? Journal of Family Issues, 34(3), 295ت322.

برای پرسش‌های بیشتر، پرسش‌های ۷، ۹ و ۱۶ را مطالعه کنید.

[1] . Weighted mean

پرسش ۱۱: میانه چیست و چگونه محاسبه می شود؟

میانه[1] یکی دیگر از شاخص‌های گرایش مرکزی و نوع دیگری از نمره‌ی متوسط یا معدل است. میانه، نقطه‌ای از توزیع داده‌ها است که ۵۰ درصد موارد پایین‌تر از آن و ۵۰ درصد موارد بالاتر از آن قرار دارند. میانه نقطه وسط در توزیع نمرات و توزیع فراوانی است. هنگامی‌که نمرات توزیع فراوانی به ترتیب از پایین‌ترین به بالاترین نمره مرتب شود، نصف نمرات بالای میانه و نصف نمرات پایین میانه قرار می‌گیرد؛ بنابراین، میانه نقطه ۵۰ درصدی است؛ یعنی عددی است که توزیع نمره‌ها را به دو قسمت مساوی تقسیم می‌کند. اندازه یا حجم واحدهای اندازه‌گیری در میانه تأثیر ندارد. در محاسبه میانه، مقدار بزرگ‌ترین عدد توزیع که در بالاترین وضعیت قرار دارد، در مقدار میانه تأثیری ندارد. این ویژگی یکی از مزایای میانه است. میانه اغلب مواقع با حرف اختصاری Md و برخی مواقع نیز با حرف M نمایه می‌شود. برای محاسبه میانه مراحل زیر را سپری کنید:

۱. تمامی اعداد موجود در مجموعه داده‌ها را از کم به زیاد فهرست کنید؛

۲. عددی که دقیقاً در وسط مجموعه داده‌ها وجود دارد را انتخاب کنید.

در مجموعه اعداد ۲۳، ۴۵، ۶۵، ۷۲ و ۸۷، میانه ۶۵ است زیرا دو عدد بالاتر و دو عدد پایین‌تر از آن قرار دارد. اگر در یک مجموعه داده، تعداد اعداد زوج باشد، میانه از میانگین دو عددی که در وسط مجموعه داده قرار می‌گیرند، محاسبه می‌شود. برای مثال در مجموعه داده‌های ۲۳، ۴۵، ۶۵، ۶۷، ۷۲ و ۸۷ که شش عدد است، میانه برابر خواهد بود با حاصل جمع ۶۵ و ۶۷ تقسیم‌بر ۲ که برابر با ۶۶ خواهد بود. بنابراین وقتی در توزیع نمرات، اعداد خیلی بزرگ و خیلی کوچک وجود داشته باشد و پراکندگی نمرات زیاد باشد و یا داده‌های پرت وجود داشته باشد و یا نمودار دارای کجی باشد، استفاده از میانه کاربرد و دقت بیشتری دارد. برای مثال مجموعه داده ۲۸۸، ۲۵، ۲۱، ۲۰، ۱۸، ۱۷، ۱۵، ۱۴، ۹ را در نظر بگیرید. چنانچه عدد ۲۸۸ در داده‌ها نمی‌بود، توزیع داده‌ها متقارن بوده و میانگین مناسب‌ترین شاخص مرکزی بود. امّا وقتی اعداد خیلی بزرگ یا خیلی کوچک در بین داده‌ها وجود دارد، میانه مناسب‌تر از میانگین است؛ بنابراین در این توزیع به خاطر وجود عدد ۲۸۸ توزیع نامتقارن شده است؛ بنابراین کاربرد شاخص مرکزی میانه مناسب‌تر است. نکته این که میانه در هر نوع توزیع با کجی‌های راست و چپ، بین میانگین و نما قرار دارد؛ در توزیع دارای کجی منفی، نما بیشتر از میانه و میانه بیشتر از میانگین است؛ در توزیع دارای کجی مثبت، میانگین بیشتر از میانه و میانه بیشتر از نما است؛ در مجموع، میانه نمره‌ای است که به تعداد مشاهدات یا موارد موجود در مجموعه داده‌ها مرتبط است و همیشه صدک[2] پنجاهم است، و در واقع نقطه‌ای است که مجموعه داده‌ها را دقیقاً به دو نیمه برابر تقسیم می‌کند. برای پرسش‌های بیشتر، پرسش‌های ۷، ۱۲ و ۱۶را مطالعه کنید.

1 . Median
2 . percentile

پرسش ۱۲: آیا مثالی برای چگونگی کاربست میانه وجود دارد؟

حتماً از پرسش قبل به یاد دارید که سودمندترین کاربرد میانه به‌عنوان یک شاخص گرایش مرکزی در کار بر روی مجموعه داده‌های دارای نمرات انتهایی یا کرانه‌ای است. در یک جمعیت آماری، میانه را می‌توان با مرتب کردن آن و انتخاب عددی که در وسط قرار می‌گیرد، تعیین کرد. درصورتی‌که جمعیت برابر با عددی زوج باشد آنگاه میانه برابر است با میانگین دو عددی که در وسط قرار می‌گیرند. یکی از پرکاربردترین استفاده‌های میانه در ایالات متحده و اکثر کشورهای توسعه‌یافته محاسبه درآمد سالانه مردم است. دلیل این امر این است که درآمد مردم در این کشورها پراکندگی زیادی دارد و درنتیجه نمرات انتهایی زیادی مشاهده می‌شود و دامنه درآمدها بسیار متنوع و گسترده است.

برای مثال، فهرست حقوق هفت ساله از پایین‌ترین تا بالاترین درآمد به ترتیب عبارت است از ۴۱/۶۷۵، ۴۵/۶۸۷، ۴۷/۶۹۸، ۵۴/۶۷۵، ۵۸/۷۶۸، ۶۰/۱۰۰، ۲۳۵/۴۷۵ دلار. شما می‌دانید که یک شاخص گرایش مرکزی می‌بایست بهترین نماینده و مُعرّف تمام داده‌های یک مجموعه داده باشد. با یک نگاه سرانگشتی می‌توان مشاهده کرد که تمامی این اعداد به‌جز یک عدد، در طیف ۴۱۰۰۰ دلار تا ۶۱۰۰۰ دلار هستند و برای نشان دادن نمره‌ی متوسط یا معدل حدوداً ۵۰۰۰۰ دلار را تصور می‌کنید. امّا میانگین این درآمدها ۷۷/۷۲۸ دلار است که درواقع ۲۲۰۰۰ دلار بالاتر از حساب سرانگشتی شماست. امّا اگر از شاخص گرایش مرکزی میانه استفاده کنید، درآمد ۵۴/۶۷۵ دلار نماینده مجموعه درآمدهای سالیانه قرار می‌گیرد که به تخمین سرانگشتی شما بسیار نزدیک‌تر است. این تفاوت بین میانگین و میانه به دلیل حساسیت میانگین به نمرات انتهایی یا کرانه‌ای است. در این مجموعه درآمدهای سالیانه، درآمد ۲۳۵/۴۷۵ دلار همان نمره انتهایی است که میانه به آن غیر حساس و میانگین به آن حساس است.

چگونه متوجه شویم که از میانه استفاده کنیم یا میانگین ؟ برای این انتخاب، در ابتدا یک ارزیابی دیداری[1] از مجموعه داده‌ها داشته باشید؛ دوم اینکه اگر داده‌ها از نوع داده‌های ترتیبی (رتبه کلاسی، رتبه کنکور، ...) باشند، میانه بسیار کارآمدتر و دقیق‌تر از میانگین می تواند نماینده و مُعرّف نمره‌ی متوسط یا معدل باشد.

برای پرسش‌های بیشتر، پرسش‌های ۷، ۱۱ و ۱۶ را مطالعه کنید.

1 . visual examination

پرسش ۱۳: نما چیست و چگونه محاسبه می‌شود؟

نما یا مُد عددی است که بیشترین فراوانی را در یک مجموعه داده دارد و از طریق شمارش تعداد تکرار یک ارزش در یک مجموعه نمره محاسبه می‌شود. در آمار و ریاضیات، به ارزش یا مقداری که بیشتر از سایر مقادیر در یک مجموعه داده رُخ دهد، نما یا مُد گفته می‌شود. نما ضعیف‌ترین و ساده‌ترین شاخص مرکزی است. نما از طریق مشاهده‌ی توزیع فراوانی و تعیین عددی که دارای بیشترین فراوانی است، تعیین می‌گردد. عددی که دارای بیشترین فراوانی است، غالباً نزدیک مرکز توزیع فراوانی قرار دارد. در چنین شرایطی نمایه شاخص مرکزی است. امّا نما همیشه در مرکز توزیع فراوانی قرار ندارد. به همین دلیل نمی‌توان به‌عنوان یک شاخص مرکزی به آن اطمینان داشت. در واقع نما در میان شاخص‌های گرایش به مرکز، شاخصی بی‌ثبات است.

برای مثال به مجموعه داده‌های زیر توجه کنید. در این مجموعه داده، به هر یک از داده‌ها که در اینجا نوع رنگ است، یک برچسب عددی می‌زنیم. در واقع به هر داده (رنگ) یک عدد اختصاص می‌دهیم.

مجموعه داده: قرمز، آبی، آبی، خاکستری، بنفش، زرد، آبی، آبی، خاکستری، بنفش، زرد، آبی، آبی، خاکستری، بنفش، زرد، آبی، خاکستری، بنفش، بنفش، آبی، آبی، خاکستری، بنفش، بنفش.

اگر تعداد تکرار هر نوع رنگ در مجموعه داده را بشمارید، مشخص می‌شود که بیشترین رنگ تکرار شده آبی با ۹ فراوانی است. بنابراین نما یا مُد در این مجموعه داده رنگ آبی خواهد بود.

یکی از رایج‌ترین اشتباهاتی که در هنگام محاسبه نما رُخ می‌دهد این است که عدد نماینده نما، صرفاً نشانگر فراوانی یا تعداد وقوع یک رویداد است و به‌خودی‌خود ارزش عددی ریاضی ندارد؛ بنابراین در مثال بالا، نما عدد ۹ نیست بلکه رنگ آبی است.

نکته مهم این است که اگر تمامی داده‌های موجود در یک مجموعه داده برابر باشند آن مجموعه داده نما نخواهد داشت. به توزیع فراوانی که فقط یک نما دارد، یک نمایی[1] و به توزیع دارای بیش از دو نما، توزیع چند نمایی و اگر دو نما وجود داشته باشد به آن مجموعه داده، دونمایی[2] می‌گویند.

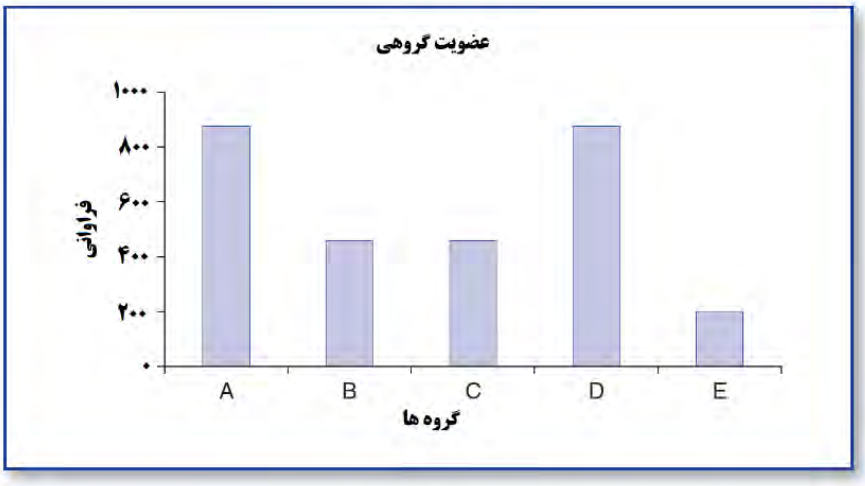
برای پرسش‌های بیشتر، پرسش‌های ۷، ۱۴ و ۱۶ را مطالعه کنید.

1 . uni modal
2 . bimodal

۱۴: آیا مثالی برای چگونگی کاربست نما وجود دارد؟

نما کم‌دقت‌ترین شاخص گرایش مرکزی است زیرا با خودِ عدد سروکاری ندارد و صرفاً با فراوانی[1] میزان رخداد یک داده سروکار دارد. بااین‌حال، اگر کسی بخواهد نمره‌ی متوسط یا معدل یا درواقع مرکزی‌ترین ارزش یک مجموعه از برچسب‌ها را محاسبه کند، مُد یا نما مناسب‌ترین شاخص مرکزی خواهد بود. برای مثال، در ادامه فهرستی از تعداد اعضای گروه‌های پنج‌گانه سیاسی ارائه شده است: گروه A دارای ۵۸۷ عضو، گروه B دارای ۴۵۶ عضو، گروه C دارای ۴۵۴ عضو، گروه D دارای ۸۷۶ عضو و گروه E دارای ۱۹۴ عضو. همانگونه که ملاحظه می کنید، گروه D با ۸۷۶ عضو، بیشترین فراوانی را دارد.

در ادامه فهرستی از تعداد دیگری از اعضای گروه‌های پنج‌گانه سیاسی ارائه شده است: گروه A دارای ۸۷۶ عضو، گروه B دارای ۴۵۶ عضو، گروه C دارای ۴۵۴ عضو، گروه D دارای ۸۷۶ عضو و گروه E دارای ۱۹۴ عضو. همانگونه که ملاحظه می کنید، گروه D و A با ۸۷۶ عضو، بیشترین فراوانی را دارند و چون تعداد اعضای دارای گرایش‌های سیاسی A و D برابر با ۸۷۶ است، به آن مجموعه دونمایی[2] گفته می شود. همانگونه که در تصویر زیر مشاهده می کنید، یک توزیع دو نمایی دو نقطه برافراشته یا سرآمد دارد.



برای پرسش‌های بیشتر، پرسش‌های ۷، ۱۳ و ۱۶ را مطالعه کنید.

1 . frequency
2 . bimodal

پرسش ۱۵: چگونه می‌توانم تصمیم بگیرم که از کدام شاخص گرایش مرکزی استفاده کنم؟

شاخص گرایش مرکزی خاصی که می‌بایست در کارتان از آن استفاده کنید در ابتدا به نوع داده‌هایی بستگی دارد که قصد تحلیل آن را دارید. اکنون به برخی معیارها برای کمک به شما در انتخاب شاخص گرایش مرکزی مناسب اشاره می‌کنیم.

۱- اگر داده‌های شما اسمی یا طبقه‌ای است، از مُد یا نما استفاده کنید. داده‌های اسمی یا طبقه‌ای، اطلاعاتی هستند که در یک‌زمان فقط در یک طبقه قرار می‌گیرند و درواقع داده‌هایی هستند که صرفاً با برچسب‌ها یا اسامی متمایز می‌شوند. این داده‌ها ارزش‌های کیفی ندارند. برای ذکر چند مثال دراین‌باره می‌توان به رنگ مو، گرایش‌های سیاسی، مدل اتومبیل و گروه بیس‌بال موردعلاقه اشاره کرد. برای مثال در یک نشست مباحثه سیاسی، دموکرات‌ها، جمهوری خواهان و مستقل‌ها حضور داشتند. شخصی که ارزیاب این نشست بود، قصد داشت تعداد افراد مربوط به هر یک از گرایش‌های سیاسی را مشخص کند. اگر در این نشست، افرادی که حضور دارند خود را جمهوری‌خواه معرفی کنند، بنابراین جمهوری‌خواه بودن در این نشست سیاسی فراوانی بیشتری داشته و نما محسوب می‌شود و در حقیقت دقیق‌ترین شاخص گرایش مرکزی خواهد بود که نماینده یا مُعرِّف این نمونه است.

۲- اگر داده‌ها نمرات انتهایی داشته باشند، دقیق‌ترین و مناسب‌ترین شاخص گرایش مرکزی میانه خواهد بود. برای مثال، اگر شخصی قصد داشته باشد که معدل قیمت منزل در یک شهر بزرگ را داشته باشد، با علم به این‌که در یک شهر بزرگ، منازل بسیار گران‌قیمت و بسیار ارزان‌قیمت وجود دارد، دقیق‌ترین و مناسب‌ترین شاخص گرایش مرکزی میانه خواهد بود، چراکه به نمرات انتهایی حساس نیست.

۳- اگر داده‌ها طبقه‌ای نباشند و نمرات انتهایی نداشته باشند، دقیق‌ترین و مناسب‌ترین شاخص گرایش مرکزی میانگین خواهد بود. برای مثال، اگر قصد داشته باشید که معدل یا نمره متوسط نمره هجی کردن یک گروه از دانش آموزان کلاس ششم را داشته باشید و نمره هجی کردن آن‌ها بر مبنای تعداد واژگان درست هجی شده از ۲۰ محاسبه شود، مناسب‌ترین شاخص گرایش مرکزی میانگین خواهد بود. اکنون دو مسئله مهم در تمایز قائل شدن بین این سه شاخص گرایش مرکزی وجود دارد.

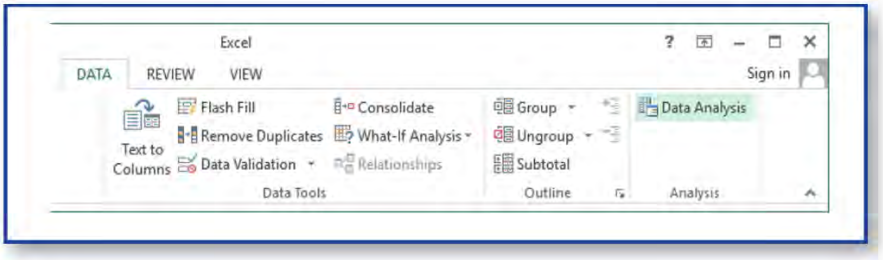
۱- میانگین بسیار دقیق‌تر از میانه و میانه بسیار دقیق‌تر از نما است. پس تا جایی که ممکن است از میانگین استفاده کنید. ۲- اگر یک تحلیل را بتوان بر اساس شاخص گرایش مرکزی انجام داد، پس یقیناً قابلیت تحلیل بر اساس میانه و نما را نیز دارد. در واقع میانگین میانه و نما، و میانه نما را در بردارد.

برای پرسش‌های بیشتر، پرسش‌های ۹، ۱۱ و ۱۳ را مطالعه کنید.

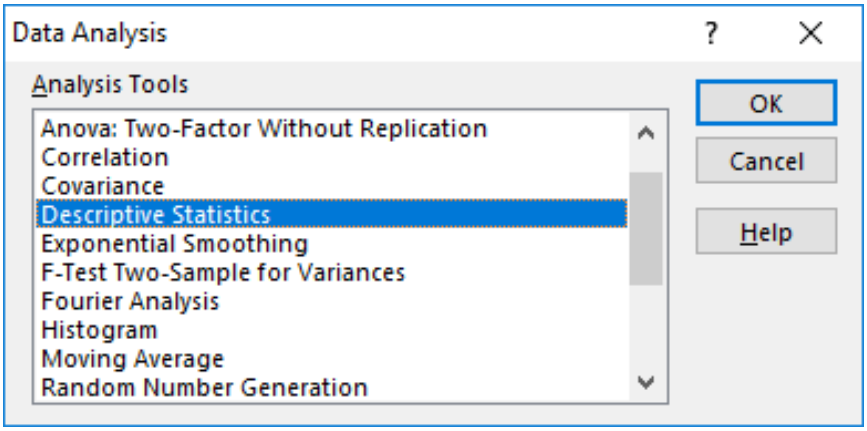
پرسش ۱۶: چگونه نرم‌افزار اکسل برای محاسبه شاخص‌های گرایش مرکزی به‌کاربرده می‌شود؟

نرم‌افزار اِکسل با پیروی از چند گام، شاخص‌های گرایش مرکزی را محاسبه می‌کند. فرض کنید نمرات مربوط به پیشرفت تحصیلی ۵۰ نفر را در اختیار داریم. دامنه این نمرات از ۷۰ تا ۱۰۰ است.

گام ۱- بر روی نوار Data کلیک کرده و سپس روی گزینه Data Analysis کلیک کنید.



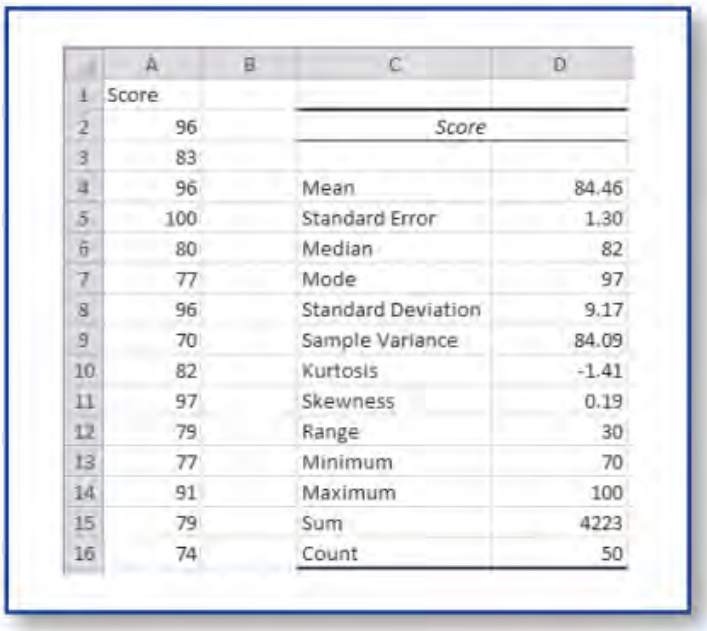
گام ۲- در بر روی گزینه Descriptive Statistics دو بار کلیک کنید.



گام ۳- در پنجره بازشده، دامنه درونداد Input Range و دامنه برونداد Output Range را تعریف کنید. همچنین گزینه‌های Summary statistics و Labels in first row را علامت بزنید. در پنجره بازشده، دامنه درونداد Input Range را می‌توانید با انتخاب خانه‌های مربوط به داده‌هایی که قصد تحلیل آنها را دارید بر روی صفحه اصلی اِکسل مشخص کنید.

دامنه برونداد Output Range را می‌توانید با انتخاب خانه‌ای که قصد دارید تحلیل در آنجا نمایه شود، انتخاب کنید.

گام ۴- بر روی OK کلیک کنید و نتایج را ملاحظه نمایید.



	A	B	C	D
1	Score			
2	96		*Score*	
3	83			
4	96		Mean	84.46
5	100		Standard Error	1.30
6	80		Median	82
7	77		Mode	97
8	96		Standard Deviation	9.17
9	70		Sample Variance	84.09
10	82		Kurtosis	-1.41
11	97		Skewness	0.19
12	79		Range	30
13	77		Minimum	70
14	91		Maximum	100
15	79		Sum	4223
16	74		Count	50

همان‌گونه که ملاحظه می‌کنید، نرم‌افزار اِکسل اطلاعات توصیفی ارزشمندی را فراهم می‌کند. در این جدول میانگین ۸۱/۴۴، میانه ۷۹/۵ و نما (مُد) ۹۶ به‌دست آمده است. سایر شاخص‌های به ترتیب انحراف استاندارد، واریانس، کشیدگی، چولگی (کجی)، دامنه تغییرات، کمینه، بیشینه، جمع و تعداد نیز نمایه شده است.

شما می‌توانید از نرم‌افزار اِکسل برای محاسبه میانگین، میانه و نما از طریق توابع[1] نیز استفاده کنید. تابع، یک فرمول از پیش طراحی‌شده[2] برای محاسبه یک شاخص است. برای محاسبه میانگین می‌توانید از تابع AVERAGE= استفاده نمایید. برای محاسبه میانه می‌توانید از تابع MEDIAN= و برای محاسبه نما می‌توانید از تابع MODE = استفاده نمایید. برای کاربست یک تابع، نام تابع را در یک خانه صفحه اِکسل وارد کنید و سپس دامنه خانه‌هایی که داده‌های موردنظرتان در آن‌ها می‌باشند را در آن وارد نمایید. برای مثال برای محاسبه نما در مجموعه داده‌های مثال صفحه قبل، تابع MODE(A2:A51)= را در یک خانه از صفحه اِکسل وارد کنید. حال اگر روی Enter کلیک کنید، عددی را نشان می‌دهد که همان نما است.

برای پرسش‌های بیشتر، پرسش‌های ۹، ۱۱ و ۱۳ را مطالعه کنید.

1 . functions
2 . predesigned formula

بخش ۳: درک شاخص‌های پراکندگی

پرسش ۱۷: شاخص‌های پراکندگی چه هستند و چرا به‌کاربرده می‌شوند؟

شاخص‌های پراکندگی نظیر دامنه تغییرات[1]، انحراف استاندارد[2] و واریانس[3]، نوع دوم آمار توصیفی محسوب می‌شوند و در توصیف تمام‌عیار ویژگی‌های یک مجموعه داده بسیار مهم و ضروری هستند. شاخص‌های پراکندگی با نام‌های شاخص‌های توزیع یا انتشار نیز شناخته می‌شوند و نشان می‌دهند که داده‌های موجود در یک مجموعه، چگونه و چقدر از یکدیگر و از شاخص‌های گرایش مرکزی فاصله دارند. شاخص‌های پراکندگی نشان می‌دهند که داده‌های یک مجموعه داده چقدر باهم متفاوت‌اند. به بیان دقیق‌تر، پراکندگی عبارت است از فاصله یا مقدار تفاوتی که هر نمره در یک مجموعه نمره، از یک نمره خاص دارد.. در اغلب موارد، آن نمره خاص که معیار قرار می‌گیرد و تفاوت‌های سایر نمرات و داده‌ها از آن سنجیده می‌شود، میانگین است؛ اما ممکن است میانه و نما نیز باشد.

سه شیوه آماری توصیفی برای ارزیابی و نشان دادن پراکندگی نمرات به‌کار برده می‌شود که عبارت‌اند از: دامنه تغییرات، واریانس و انحراف استاندارد. در این میان، دامنه تغییرات، کمترین دقت را دارد. دامنه تغییرات عبارت است از فاصله بین کوچک‌ترین و بزرگ‌ترین نمره‌های موجود در یک مجموعه نمره. این شاخص سریع‌ترین شیوه محاسبه را دارد و یک تخمین بسیار کلی از چگونگی تفاوت نمرات از هم فراهم می‌آورد. به نظر می‌رسد که انحراف استاندارد مقدار میانگین تفاوت تمامی نمرات از یک نمره استاندارد که میانگین همان مجموعه نمرات است می‌باشد. در آمار، انحراف استاندارد یکی از شاخص‌های پراکندگی است که نشان می‌دهد به‌طور میانگین، داده‌ها چه مقدار از مقدار متوسط فاصله دارند. اگر انحراف استاندارد مجموعه‌ای از داده‌ها نزدیک به صفر باشد، نشانه آن است که داده‌ها نزدیک به میانگین هستند و پراکندگی اندکی دارند؛ درحالی‌که انحراف معیار بزرگ بیانگر پراکندگی قابل‌توجه داده‌ها است. انحراف استاندارد بیشترین کاربرد را در میان شاخص‌های پراکندگی دارد. واریانس، مجذور انحراف استاندارد است. همانند انحراف استاندارد، واریانس نیز به عنوان یک شاخص در بسیاری از آماره‌های پیشرفته کاربرد دارد.

هرچند تمامی این شاخص‌های پراکندگی اطلاعات مهمی درباره درک و فهم ویژگی‌های یک مجموعه داده فراهم می‌آورند، امّا پراکندگی به‌عنوان یک مفهوم، در تمام مباحث آمار بسیار مهم تلقی می‌شود چراکه بسیاری از ایده‌های بنیادی نظیر استنباط آماری، از پراکندگی به‌عنوان شاخص خطا[4] استفاده می‌کنند. خطا نباید با اشتباه و یا بی‌دقتی هم معنی در نظر گرفته شود. در واقع، شاخص خطا عبارت است از تفاوت‌های فردی‌ای که بصورت طبیعی رخ می‌دهند[5] که

1 . Range
2 . Standard Deviation
3 . variance
4 . Error
5 . naturally occurring incidence of individual differences

البته در هر پژوهشی اتفاق می‌افتد و این پراکندگی بین شرکت کنندگان می بایست در نظر گرفته شده و فهمیده شود. در واقع این نقش پراکندی در نظریه و کاربرد است.
برای پرسش‌های بیشتر، پرسش‌های ۱۹، ۲۱ و ۲۵ را مطالعه کنید.

پرسش ۱۸: آیا مثالی برای چگونگی کاربست شاخص پراکندگی وجود دارد؟

پراکندگی به‌عنوان یک مفهوم آماری موضوعی شگفت‌انگیز است چراکه بینش ارزشمندی درباره تغییرات درون‌فردی[1] و تغییرات بین‌فردی[2] رفتار انسان‌ها را آشکار می‌کند. این مسئله به تفاوت‌های فردی اشاره دارد. هنگامی‌که شاخص‌های پراکندگی در موضوعاتِ مورد مطالعه نظیر چاقی کودکان به‌کاربرده می‌شوند، می‌توانند علاوه بر نشان دادن میزان پراکندگی نمرات یک مجموعه داده از یکدیگر و یا از یک شاخص گرایش مرکزی نظیر میانگین، معنای فراتری نظیر شاخص تغییر نیز داشته باشند. برای مثال، مطالعات نشان داده‌اند که چاقی کودکان با بدکارکردی سیستم اعصاب[3] آن‌ها مرتبط است. دانشمندان به مطالعه این موضوع پرداختند که چگونه کودکان در حال رسیدن به بلوغ، به دلیل دارا بودن شاخص‌های کنترل متابولیک مشکل‌ساز[4] که با چاقی مرتبط است، مستعد ابتلا به مشکلات سیستم اعصاب خودمختار نظیر کاهش تغییرپذیری ضربان قلب[5] می‌شوند. در این مطالعه به بررسی تاثیر تحولات دوره بلوغ بر سیستم اعصاب خودمختار ۸۴ کودک دارای افزایش وزن و چاق و ۸۷ کودک با وزن بهنجار پرداخته شد. سیستم اعصاب خودمختار کودکان از طریق ارزیابی تغییرپذیری ضربان قلب بررسی شد. یافته‌ها نشان داد که کودکان دارای اضافه وزن و چاق در مقایسه با کودکان دارای وزن بهنجار به‌طور معناداری میزان تغییرپذیری ضربان قلب کمتری داشتند و این مشکل با سطوح فعالیت جسمانی‌شان همبستگی مثبت داشت. این محققان نتیجه گرفتند که چاقی به‌صورت معکوس بر کارکرد سیستم اعصاب خودمختار کودکان، به‌ویژه در دوره بلوغ تأثیر سوء می‌گذارد و لذا این کودکان می‌بایست در دوره بلوغ در برنامه‌های فعالیت‌های جسمانی مشارکت بالایی داشته باشند تا کارکرد سیستم اعصاب خودمختارشان بهبود یابد چراکه سبب کاهش چاقی‌شان خواهد شد. از منظر یک رویکرد کاربردی، این محققان پیشنهاد کردند که گروه پزشکی مدارس می‌بایست از علائم بدکارکردی سیستم اعصاب خودمختار نظیر خستگی، تحریک‌پذیری، رخوت و بی‌حالی آگاهی کامل داشته باشند و همچنین با معلمان تربیت‌بدنی مدارس مشارکت و همفکری داشته باشند تا برنامه فعالیت‌های جسمانی متناسبی تدوین شود که از طریق آن بتوان یک محیط سالم و مناسب برای فعالیت‌های روزمره دانش آموزان فراهم آورد.

برای مطالعه بیشتر درباره مثال موجود در این پرسش به مطالعه مرجع زیر بپردازید.

Chen, S., Chiu, H., Lee, Y., Sheen, T., & Jeng, C. (2012). Impact of pubertal development and physical activity on heart rate variability in overweight and obese children in Taiwan. The Journal of School Nursing, 28(4), 284ـ290.

برای پرسش‌های بیشتر، پرسش‌های ۱۷، ۲۲ و ۲۵ را مطالعه کنید.

---

1 . within-individual changes
2 . between-individual changes
3 . dysfunction of the nervous system
4 . problematic metabolic control
5 . heart rate variability

پرسش ۱۹: دامنه تغییرات چیست و چگونه محاسبه می‌شود؟

دامنه تغییرات که یکی از شاخص‌های کلی پراکندگی است، از تفریق بزرگ‌ترین عدد از کوچک‌ترین عدد موجود در یک مجموعه داده به دست می‌آید. دامنه تغییرات یک شاخص کلی درباره فاصله بین دو نقطه را به دست می‌دهد بدون آنکه سایر نمرات موجود بین این دو نقطه را در نظر بگیرد. به همین دلیل است که دامنه تغییرات بسیار کلی و غیردقیق است. امّا با این‌حال یک شاخص سودمند در فراهم آوردن یک دیدگاه کلی درباره میزان تقریبی پراکندگی نمرات در یک مجموعه نمره است. دو نوع دامنه تغییر وجود دارد:

نوع اول دامنه تغییر انحصاری[1] است که از طریق فرمول R=H-I محاسبه می‌شود. در این فرمول R دامنه تغییرات، H بزرگ‌ترین نمره موجود در مجموعه نمرات و I کوچک‌ترین نمره موجود در مجموعه نمرات است. برای مثال، اگر در مجموعه نمرات مربوط به پیشرفت تحصیلی دانش آموزان، بالاترین عدد ۸۸ و کوچک‌ترین عدد ۴۷ باشد، دامنه تغییر انحصاری برابر با ۴۱ خواهد بود؛ به این معنا که تغییرپذیری نمرات ۴۱ عدد است.

نوع دوم، دامنه تغییر فراگیر[2] است که از طریق فرمول R=H-I+۱ محاسبه می‌شود. در این فرمول R دامنه تغییرات، H بزرگ‌ترین نمره موجود در مجموعه نمرات و I کوچک‌ترین نمره موجود در مجموعه نمرات است. در انتها نیز با عدد ۱ جمع بسته می‌شود. هرچند که در اغلب پژوهش‌های گزارش‌شده از شاخص دامنه تغییرات انحصاری استفاده شده است، اما می‌توان از شاخص دامنه تغییرات فراگیر نیز استفاده نمود. درمجموع، می‌توان از هر دو نوع دامنه تغییرات استفاده کرد. شاخص دامنه تغییرات در فراهم آوردن یک تخمین کلی از اینکه نمرات چقدر باهم فرق دارند سودمند است؛ امّا ازآنجاکه در ارائه این تخمین، تمامی نمرات موجود در مجموعه نمرات را مورد تحلیل قرار نمی‌دهد، کاربردش محدود است.

برای پرسش‌های بیشتر، پرسش‌های ۱۷، ۱۸ و ۲۵ را مطالعه کنید.

1 . exclusive range
2 . inclusive range

پرسش ۲۰: آیا مثالی برای چگونگی کاربست دامنه تغییرات وجود دارد؟

هرچند که محاسبه و فهم شاخص دامنه تغییرات ساده است، امّا اهمیت و معنای آن بر اساس بافتاری که این شاخصِ پراکندگی در آن مورد استفاده قرار می‌گیرد، متفاوت است.

سه محقق به بررسی دامنه تغییرات نمرات سیالی خواندن و درک مطالب ۱۱۴۹ دانش‌آموز در پنج مدرسه ابتدایی مخصوص دانش آموزان سرآمد و با استعداد پرداختند. یافته‌های آن‌ها نشان داد که یک شاخص نظیر دامنه تغییرات در برخی مواقع چقدر می‌تواند سودمند و کارآمد باشد. نتایج نشان داد که دامنه تغییرات در نمرات درک مطلب در کلاس سومی‌ها برابر ۹/۲، در کلاس چهارمی‌ها برابر ۳/۱۱ و در کلاس پنجمی‌ها برابر ۱۱/۶ بود. همچنین این محققان یافتند که دامنه تغییرات زیادی در نمرات سیالی خواندن شفاهی بین تمامی دانش آموزان ابتدایی از زیر صدک دهم تا بالای صدک نودم وجود دارد. اینجاست که می‌توان گفت شاخص دامنه تغییرات می‌تواند اطلاعات بسیار ارزشمندی در اختیارمان قرار دهد. در این مثال، دامنه تغییرات زیاد در نمرات درک مطالب خواندن، توجه ما را این‌گونه هدایت می‌کند که معلمان می‌بایست، محتوا و مطالب خواندن و درک مطالب و برنامه‌ریزی‌های متفاوتی بر اساس پایه تحصیلی و دامنه وسیع توانمندی‌های دانش آموزان ارائه دهند. درواقع، دامنه تغییرات زیاد توانمندی‌های دانش آموزان، نیازمند دامنه تغییرات زیاد مفاهیم و مطالب درسی نیز است تا تمامی سطوح وسیعی که دانش آموزان دارند را پوشش دهد و از این طریق به آن‌ها در رسیدن به موفقیت کمک شود.

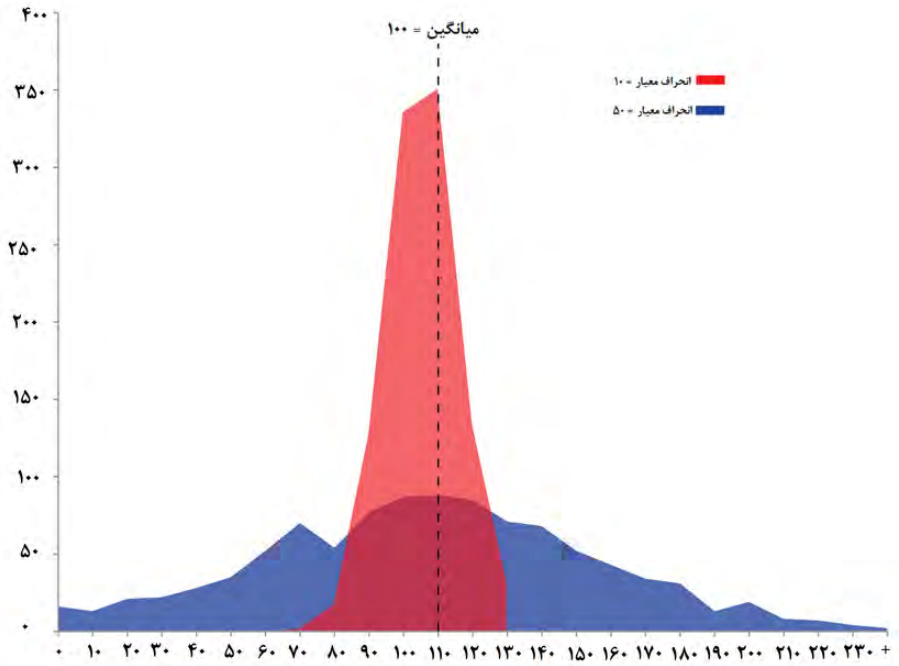
برای مطالعه بیشتر درباره مثال موجود در این پرسش به مطالعه مرجع زیر بپردازید.

Firmender, J. M., Reis, S. M., & Sweeny, S. M. (2013). Reading comprehension and fluency levels ranges across diverse classrooms: The need for differentiated reading instruction and content. Gifted Child Quarterly, 57(1), 3ات4.

برای پرسش‌های بیشتر، پرسش‌های ۱۷، ۱۸ و ۱۹ را مطالعه کنید.

پرسش ۲۱: انحراف استاندارد چیست و چگونه محاسبه می‌شود؟

پراکندگی یکی از مهم‌ترین مفاهیم در آمار است. هدف از اندازه‌گیری معمولاً پیدا کردن تغییرات و توجیه آن‌هاست. هرچه پراکندگی کمتر باشد، پیش‌بینی مقدار یک متغیر تصادفی با کمک مقدار میانگینش دقیق‌تر می‌شود؛ به عبارت دیگر، پراکندگی می‌تواند دقتِ یک پیش‌بینی را نشان دهد. همه شاخص‌های پراکندگی، اعدادی غیر منفی و حقیقی هستند که در صورت یکسان بودن همه داده‌ها صفر می‌شوند و هرچه اختلاف داده‌ها از میانگین بیش‌تر باشد، این شاخص‌ها بزرگ‌تر می‌شوند. انحراف استاندارد یکی از همین شاخص های پراکندگی است. انحراف استاندارد یک معدل و نمره متوسط مجموعه‌ای از نمرات است و نشان می‌دهد که به‌طور میانگین، نمرات یک مجموعه داده، چقدر از شاخص استاندارد که معمولاً میانگین است، فاصله دارند و ارزیابی می‌کند که در یک مجموعه نمرات، چقدر پراکندگی و تغییرپذیری وجود دارد. در حقیقت انحراف استاندارد معدل تفاوت نمرات یک مجموعه نمره از میانگین آن مجموعه نمره است. در شکل زیر انحراف استاندارد دو مجموعه داده که میانگین یکسان دارند، امّا انحراف معیار یک گروه ۱۰ و دیگری ۵۰ است نمایه شده است.



نام انحراف معیار نخستین بار از سوی کارل پیرسون در سال ۱۸۹۴ پیشنهاد شد. پیش از او نام‌های دیگری برای این مفهوم پیشنهاد شده بود برای نمونه، گاوس به آن خطای میانگین می‌گفت. انحراف معیار برای یک مجموعه متناهی، برابر است با جذر میانگین مربعات اختلاف داده‌ها با میانگین‌شان.

فرمول انحراف استاندارد به قرار زیر است:

$$s = \sqrt{\frac{\sum(x - \bar{x})^{۲}}{n - ۱}}$$

در این فرمول s انحراف استاندارد، x نمره، $\bar{x}$ میانگین نمرات، Σ علامت سیگما یا جمع و n تعداد نمونه است. برای مثال، نمرات مربوط به تعداد واژگان صحیح یک آزمون املاء بیست پرسشی عبارت است از: ۱۶، ۱۴، ۱۰، ۱۵، ۱۴، ۱۲، ۱۹، ۱۵، ۸، ۷؛ برای محاسبه انحراف استاندارد از گام‌های زیر پیروی کنید:

۱- تمامی نمرات را فهرست کنید؛

۲- میانگین نمرات را محاسبه کنید؛ در اعداد ذکر شده میانگین برابر ۱۳ است؛

۳- هر یک از اعداد را از میانگین کسر کنید؛

۴- مجذور اعداد کسر شده میانگین را محاسبه کنید؛

۵- تمامی اعداد مجذور شده را باهم جمع کنید؛

۶- عدد جمع به‌دست‌آمده را بر n-۱ تقسیم کنید؛

۷- جذر عدد به‌دست‌آمده را محاسبه کنید که همان انحراف استاندارد است.

نکته

- هنگامی‌که انحراف استاندارد را محاسبه می‌کنیم، تمامی انحراف‌ها را به توان دو می‌رسانیم. با این کار انحراف‌های مثبت و منفی از میانگین همه مثبت می‌شوند و مجموع مجذورات صفر نمی‌شود. چراکه همیشه مجموع انحرافات از میانگین صفر است؛ و اگر تک‌تک انحراف‌ها از میانگین مجذور نشوند، هدف ما از کسب اطلاعات درباره پراکندگی نمرات حول میانگین محقق نمی‌گردد؛
- در آخرین گام محاسبه انحراف استاندارد، اعداد مجذور شده به ریشه خود برمی‌گردند تا اعداد به واحدهای اصلی خود برگردند؛
- هر انحراف استاندارد نمایانگر یک فاصله معین یکی از نمرات از میانگین یا معدل نمرات است؛
- واریانس که یکی از شاخص‌های پراکندگی است، مجذور انحراف استاندارد است؛

برای پرسش‌های بیشتر، پرسش‌های ۱۷، ۱۹ و ۲۳ را مطالعه کنید.

پرسش ۲۲: چرا از برآورد بدون تورش n-1 بجای برآورد تورش دار n در محاسبه واریانس و انحراف استاندارد استفاده می‌کنیم؟

از بحث مربوط به محاسبه انحراف استاندارد به یاد دارید که انحراف استاندارد، معدل فاصله‌های هر یک از نمرات از شاخص مرکزی[1] یک مجموعه نمره است. انحراف استاندارد در حقیقت یک فاصله متوسط[2] است. همچنین به یاد دارید که میانگین، از طریق تقسیم جمع نمرات بر تعداد نمونه محاسبه می‌شود. امّا در محاسبه واریانس و انحراف استاندارد همانند محاسبه میانگین عمل نمی‌کنیم. بلکه در مخرج کسر فرمول انحراف استاندارد بجای استفاده از n که تعداد نمونه است از N-۱ استفاده می‌کنیم. چرا؟

باور عمومی که از کاربرد علم آمار ریشه می‌گیرد این است که آمار، نتایج و ارزش‌هایی را به دست می‌دهد که قابل‌اعتماد هستند، زیرا منطق آمار این است که نتایج تا حد ممکن به‌گونه‌ای به دست آید که نمایانگر هر چه بیشتر دنیای واقعی و جامعه حقیقی باشند. در نظر داشته باشید که ما از واریانس و انحراف استاندارد نمونه برای برآورد واریانس و انحراف استاندارد جامعه استفاده می‌کنیم تا بتوانیم ویژگی‌های جامعه اصلی را تخمین بزنیم.

هنگامی‌که واریانس و انحراف استاندارد نمونه را برآورد می‌کنیم، این شبهه وجود دارد که شاخص‌های توصیفی نمونه، سوگیری داشته باشند به این معنی که ممکن است شاخص‌های توصیفی جامعه را بیش برآورد[3] کنند. برای جبران این نقص در سوگیری، در مخرج کسرِ فرمول محاسبه واریانس و انحراف استاندارد نمونه، ۱ نمره کسر می‌شود تا آنچه به دست می‌آید یک ارزش بدون سوگیری[4] باشد[5]. به همین دلیل، ارزش محاسبه‌شده ممکن است ارزش جامعه را کم برآورد[6] کند، امّا باور بر این است که با این روش ارزش‌های دقیق‌تری به دست آید. مقدار انحراف معیار به‌دست‌آمده در صورتی درست است که از همه جمعیت موجود استفاده شود. اگر نمونه‌ای تصادفی از داده‌ها انتخاب شود و انحراف معیار برای آن نمونه‌ها به دست آید، باید یک واحد از مقدار مخرج کم شود (در این مثال، باید به‌جای ۱۰، مقدار ۹ قرار گیرد)؛ بنابراین هنگامی که میانگین جامعه در متغیر مورد مطالعه مشخص نیست، برای محاسبه واریانس و انحراف استاندارد نمونه از n-۱ در فرمول استفاده می شود تا شاخص های محاسبه شده برآوردکننده تورش دار[7] نباشند.

نکته دیگر اینکه، معمولاً با افزایش تعداد داده‌ها توزیع آن‌ها به منحنی توزیع بهنجار میل پیدا می‌کند. اکنون از فرمول تورش دار محاسبه انحراف استاندارد با حجم نمونه ۱۰ استفاده می‌کنیم.

---

1 . central point
2 . average distance
3 . overestimate
4 . unbiased
۵ . این تغییر را اصلاح بِسِل می‌نامند. فردیریک ویلهم بسل دانشمند ریاضیدان و فیزیکدان آلمانی بود که برای نخستین بار فواصل قابل اعتماد بین خورشید با سایر ستاره‌ها را با روس پارالاکس مورد بررسی قرار داد.
6 . underestimate
7 . biased estimator

در این فرمول خواهیم داشت (۴/۴۷)؛ اکنون بنگرید وقتی‌که میزان پراکندگی نمرات در صورت کسر ثابت بماند و درعین‌حال تعداد نمونه به ۱۰۰ افزایش یابد، چه تفاوتی ایجاد می‌شود (۱/۱۴).

$$s = \sqrt{\frac{۲۰۰}{۱۰}} = ۴/۴۷ \quad ؛ \quad s = \sqrt{\frac{۲۰۰}{۱۰۰}} = ۱/۱۴$$

همان‌گونه که ملاحظه می‌کنید یک تغییر چشمگیر ایجاد می‌شود. چرا؟ زیرا هر چه تعداد نمونه افزایش می‌یابد، در واقع به حجم جامعه نزدیک‌تر می‌شود و این یعنی به تخمین‌های دقیق‌تری نزدیک شده‌ایم و همان‌گونه که ملاحظه می‌شود مقدار انحراف استاندارد کاهش یافته است. در بیشتر مواقع، محققین علاقه‌مند به گزارش تخمین‌های بدون تورش هستند.

برای پرسش‌های بیشتر، پرسش‌های ۱۷، ۲۱ و ۲۳ را مطالعه کنید.

پرسش ۲۳: آیا مثالی برای چگونگی کاربست انحراف استاندارد وجود دارد؟

همانگونه که می دانید، قبل از ارائه یک دارو یا نوع جدید درمان به قشر وسیعی از بیماران، یک آزمایش اعتبارسنجی وسیع درباره اثربخشی آن صورت می‌گیرد. برای اینکه یک شرکت داروسازی، منابعی را برای ایجاد و پخش یک دارو اختصاص دهد، عوامل متعددی نظیر پیچیدگی فرمول و نوع دارو، نیاز به دارو، هزینه‌های تولید و رقابت موجود در بازار را بررسی می‌کند. یکی از مهم‌ترین عواملی که در تولید یک دارو نقش دارد این است که اگر اثرات دارویی در یک منطقه یا کشور آزمون شود تا چه میزان قابلیت تعمیم نتایج آن به سایر کشورها وجود دارد؟ این پرسش بسیار مهم و مورد توجه است. هنگامی‌که کارآزمایی‌های بالینی چندکشوری[1] انجام می‌شوند، بسیار مهم است که ویژگی‌های هر کشور و تغییرپذیری داده‌های گزارش‌شده با شاخص‌هایی نظیر ضریب پراکندگی[2] و انحراف استاندارد ارائه شود. تغییرپذیری داده[3] همانند محاسبه حجم نمونه در یک مطالعه، یکی از مهم‌ترین عوامل در تخمین دقیق اثرات درمان است. محققان در مطالعه‌ای، به مقایسه داده‌های کارآزمایی‌های بالینی ژاپن با سایر کشورها از منظر ۲۹ دارو پرداختند. این محققان یافتند که تغییرپذیری داده‌های مربوط به این کارآزمایی‌های بالینی در ژاپن و سایر کشورها بسیار نزدیک به هم بود و مشابهت بسیاری بین آنها وجود داشت. آن‌ها نتیجه گرفتند که این تشابه نشانگر این است که اثرات دارو در میان ملیت‌ها، کشورها و مناطق مختلف یکسان است.

همان‌گونه که ملاحظه می‌شود، هر چه انحراف استاندارد بیشتر باشد، نمرات از میانگین پراکنده‌تر و دورترند که در هر پژوهش می‌تواند تبیین و تحلیل خاص خود را داشته باشد. مثلا اگر همه ی داده ها با هم برابر باشد انحراف معیار برابر صفر است و بالعکس یعنی اگر انحراف معیار صفر باشد تمام داده ها با هم مساوی هستند.

برای مطالعه بیشتر درباره مثال موجود در این پرسش به مطالعه مرجع زیر بپردازید.

Kanmuri, K., & Narukawa, M. (2013). Investigation of characteristics of Japanese clinical trials in terms of data variability. Therapeutic Innovation & Regulatory Science, 47(4), 430-437.

برای پرسش‌های بیشتر، پرسش‌های ۱۷، ۲۱ و ۲۴ را مطالعه کنید.

1 . multicountry clinical trials
2 . coefficient of variation
3 . Data variability

پرسش ۲۴: برای چگونگی کاربست واریانس چه مثالی وجود دارد؟

خشونت‌های خانگی و خانوادگی چنان ادامه پیدا کرده‌اند که تبدیل به یک معضل روان‌شناختی شده‌اند و بر بازماندگان این پدیده اثرات زیان‌بار پایداری می‌گذارند. این مسئله یک مشکل سلامت عمومی عمده در تمامی طیف‌های اجتماعی اقتصادی است و یک چالش برای قانون‌گذاران و مقامات قضایی به شمار می‌رود.

بر هر حال، پژوهش در این زمینه همیشه بر الگوی مردسالاری تمرکز داشته است و مردها اغلب عامل ظلم و جرم و زن‌ها مصدومان همیشگی این قضیه به شما رفته‌اند. مطالعات تجربی در زمینه آسیب‌های فرزندان در چنین خانواده‌هایی در ادبیات پژوهشی جایگاه چندانی نداشته است. تحلیل واریانس گروه‌ها، متداول‌ترین روشی است که توسط اکثر تحلیل گران در تحلیل های آماری به کار می‌رود. در یک مطالعه، محققان به بررسی این موضوع پرداختند که آیا بر اساس مقیاس تاکتیک‌های تعارض، تفاوتهای جنسیتی معناداری درباره مورد سواستفاده یا قربانی واقع شدن در دوران کودکی وجود دارد یا خیر؟ آزمون تحلیل واریانس چندمتغیری[1] نشان داد که از منظر سوءاستفاده از فرزندان و مصدومیت جسمانی و عاطفی ناشی از خشونت خانوادگی، بین دو گروه دختران و پسران تفاوت معناداری وجود ندارد، امّا دختران از نظر سوءاستفاده‌های جنسی بیشتر از پسران مورد آسیب واقع می‌شدند.

برای مطالعه بیشتر درباره مثال موجود در این پرسش به مطالعه مرجع زیر بپردازید.

Morrill, M., & Bachman, C. (2013). Confronting the gender myth: An exploration of variance in male versus female experience with sibling abuse. Journal of Interpersonal Violence, 28(8), 1693 تا 1708.

برای پرسش‌های بیشتر، پرسش‌های ۱۷، ۲۲ و ۲۵ را مطالعه کنید.

1 . multivariate analyses of variance

پرسش ۲۵: چگونه نرم‌افزار اِکسل برای محاسبه شاخص‌های پراکندگی به‌کاربرده می‌شود؟

هم Spss و هم اِکسل می‌توانند به‌آسانی برای محاسبه دامنه تغییرات، انحراف استاندارد و واریانس به کار روند. برای کاربست اِکسل در محاسبه شاخص‌های تغییرپذیری می‌توانید از توابع موجود در نرم‌افزار یا نوارابزارهای تحلیل داده استفاده کنید. ما فرض می‌کنیم که هم‌اکنون چند داده را وارد نرم‌افزار اِکسل کرده‌اید. برای کاربست توابع از گام‌های زیر پیروی کنید.

۱. برای محاسبه انحراف استاندارد، یکی از توابع زیر را در هر خانه‌ای که مایلید وارد کنید. البته اگر دقیقاً در زیر داده‌ها وارد شود، بسیار بهتر خواهد بود.

برای محاسبه انحراف استاندارد جامعه تابع STDEV.P= را وارد خانه مورد نظر نمایید؛

برای محاسبه انحراف استاندارد نمونه تابع STDEV.S= را وارد خانه موردن ظر نمایید؛

۲. بعد از وارد کردن این فرمول‌ها در خانه مورد نظر، دامنه داده‌هایی را که مدن ظرتان است با کلیک بر روی آن‌ها مشخص نمایید و سپس با کلید Enter، محاسبه بلافاصله انجام خواهد شد؛

۳. برای محاسبه واریانس جامعه تابع VAR.P= را وارد خانه مورد نظر نمایید؛

برای محاسبه واریانس نمونه تابع VAR.S= را وارد خانه مورد نظر نمایید؛

۴. بعد از وارد کردن این فرمول‌ها در خانه مورد نظر، دامنه داده‌هایی را که مد نظرتان است با کلیک بر روی آن‌ها مشخص نمایید و سپس با کلید Enter، محاسبه بلافاصله انجام خواهد شد؛

البته تمام این توابعی که در اِکسل وجود دارد را می‌توانید از طریق گزینه تحلیل توصیفی در نوارابزار Analysis در اِکسل انجام دهید. برنامه افزودنی ToolPak صرفاً برای نسخه ویندوز[1] موجود است.

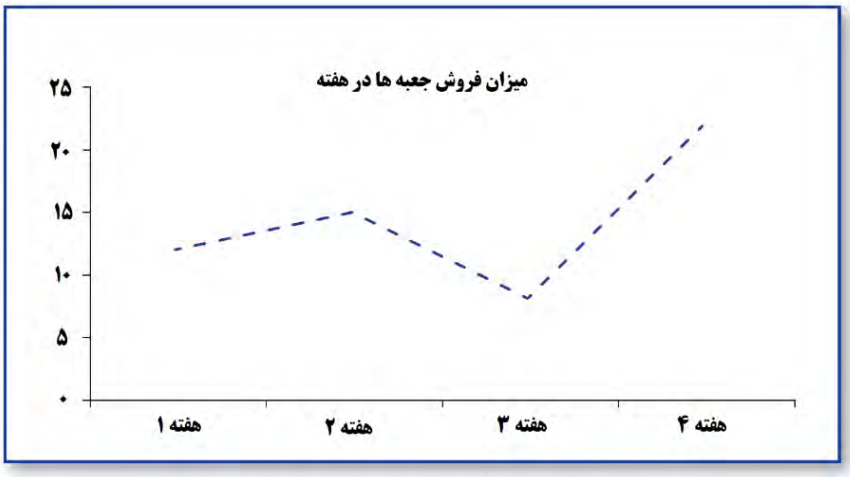
برای پرسش‌های بیشتر، پرسش‌های ۱۷، ۷۹ و ۸۰ را مطالعه کنید.

1 . Windows version of Excel

بخش ۴: نمایش داده

پرسش ۲۶: آیا واقعاً یک تصویر یا نمودار ارزش هزاران واژه را دارد؟ چرا داده‌ها را نمایه می‌کنند؟

قبل از پاسخ به این پرسش، اجازه دهید به یک مجموعه داده نگاهی بیندازیم. در این مجموعه داده، میزان فروش در هفته های مختلف ارائه شده است. در هفته اول ۱۲ جعبه، در هفته دوم ۱۵ جعبه، در هفته سوم ۸ جعبه و در هفته چهارم ۲۲ جعبه فروخته شده است. اکنون به کمک یک نمودار خطی ساده[1] به نمایه سازی داده ها پرداخته شده است.



ما از نمودار[2] خطی به این دلیل استفاده کردیم که این نمودار بهترین قالب برای ثبت و تصویرسازی[3] تغییرات در طول زمان است. شما خواهید آموخت که چگونه بتوانید یک نمودار خطی ساده شبیه نمودار بالا با استفاده از چند کلیک ساده موس ترسیم کنید. در واقع به نمایش بصری مفاهیم، ایده‌ها، ساختارها، روابط، داده‌های آماری و غیره به گونه‌ای ساده و شکل‌یافته نمودار می‌گویند. از نمودارها برای ساده کردن و ملموس کردن موضوعات گوناگون استفاده می‌شود. نمودار، نمایش نمادین دو بعدی هندسی اطلاعات بر پایه برخی فنون تجسمی است. سوال مهمی که هنگام مقایسه این دو روش فوق مطرح می‌شود این است که کدام‌یک از این دو روش یعنی نمایش داده‌ها با جدول و یا نمایش داده‌ها با نمودار آسان‌تر است؟ یکی از قوی ترین و سودمندترین پاسخ‌های ما این است که ما بر اساس ماهیت انسان بودن خود، از حس دیداری اطلاعات بسیار سریع و دقیقی به دست می‌آوریم. در حقیقت، آنچه به‌صورت فشرده، مستقیم و سریع از مشاهده نمودار و تغییر داده‌ها در طول زمان از طریق حس بینایی درک می‌کنیم، هرگز به این دقت و سرعت از ملاحظه داده‌های چیده شده در جدول به دست نمی‌آوریم. ما می‌توانیم با یک نیم نگاه سریع به بالا و پایین نمودار، سریع‌ترین و دقیق‌ترین درک از تغییرات داده‌ها را

1 . simple line chart
2 . Diagram
3 . visualizing

کسب کنیم. برای مثال، تفاوت بین هفته سوم و چهارم که ۱۴ جعبه بیشتر فروخته شده است به‌روشنی و دقت با یک نگاه اجمالی به نمودار خطی بین این دو هفته مشخص می شود، به‌گونه‌ای که به‌راحتی می‌توان تفاوت هفته های سوم به بعد با هفته‌های اول و دوم را مقایسه کرد. نمایه دیداری[1] داده‌ها به ما اجازه می‌دهد که تغییرات داده‌ها در طی زمان را به‌راحتی و به‌سرعت ملاحظه کنیم. همیشه سعی کنید خلاصه داده‌هایتان را در قالب جدول و نمودار ارائه کنید.

برای پرسش‌های بیشتر، پرسش‌های ۲۷، ۳۴ و ۳۹ را مطالعه کنید.

1 . visual representation

پرسش ۲۷: برای نمایه‌سازی کارآمد داده‌ها چه دستورالعمل‌هایی وجود دارد؟

صدها مسیر برای نمایه کردن داده‌ها به‌صورت نمودار وجود دارد تا کمک کنند که داده‌ها بسیار جالب و غیر خسته‌کننده به نظر آیند. نمودارها، نشان‌دهنده روابط عددی و نسبی ای هستند که خواننده را قادر به درک معنای خاص یک مجموعه از اطلاعات پیچیده می‌کنند. هنگامی‌که قصد دارید با نمودارهایی نظیر نمودار خطی، دایره‌ای و میله‌ای[1] برای نمایش داده‌هایتان استفاده کنید، پنج عامل وجود دارد که می‌بایست جدای از نوع نمودار مورد کاربست و نوع داده‌های مورد تحلیل، موردنظر قرار بگیرند. این پنج عامل عبارت‌اند از:

۱. تا می‌توانید به طراحی نمودار بپردازید. چند کاغذ نقشه‌کشی[2] تهیه کنید و به ترسیم هر نوع نموداری که مایلید بپردازید. همچنین سعی کنید نوع نمودار، اندازه، عنوان افقی و عمودی و همچنین سایر مؤلفه‌های مهم را در ترسیم وارد کنید.

۲. برای هر ایده، یک نمودار ترسیم کنید. اگر در یک نمودار چند ایده را ترسیم کردید، سعی کنید فقط بر یک نمودار تمرکز داشته باشید. در واقع هدف نمودار می‌بایست روشن و واضح باشد به‌گونه‌ای که خواننده متن یا نمودار در ادراک نمودار درگیر اشتباه یا تفسیر نادرست نشود.

۳. اطمینان حاصل کنید که مقیاس محورهای افقی و عمودی باهم متناسب باشند. محورهای افقی و عمودی می‌بایست با نسبت‌های درست ترسیم شده باشند تا تصنعی به نظر نرسند.

۴. سادگی بهترین است. هدف شما می‌بایست نمایش دادن داده‌ها باشد و برای این کار از نموداری استفاده کنید که درک را افزایش و ابهام و سوءتعبیر را کاهش می‌دهد. نمودار باید ساده باشد تا یک ایده را به‌خوبی نشان دهد. اگر به ارائه اطلاعات بیشتری نیاز دارید که ممکن است در قالب نمودار بیایند و این سبب پیچیدگی و ابهام نمودار شوند، می‌توانید از پاورقی یا جمله همراه نمودار[3] استفاده کنید.

۵. بیهودگی ممنوع. هر زمانی که یک تحلیل انجام می‌دهید به همراه نمودار اصلی، نمودارهای بیهوده‌ای نیز ایجاد می‌شوند که نرم‌افزار آن‌ها را پیشنهاد می‌دهد نظیر طرح‌های متفاوت، الگوها، شکل‌ها و غیره. کاربست تمامی این‌ها بسیار جالب خواهد بود امّا آن‌ها در دادن اطلاعات مناسب و جدید نکته قابل‌توجهی ندارند و اطلاعاتی به ما نمی‌دهند. در واقع محتاط باشید: تا آنجا که ممکن است از کمترین نمودارها بیشترین مطالب را نمایه کنید.

برای پرسش‌های بیشتر، پرسش‌های ۲۶، ۳۸ و ۳۹ را مطالعه کنید.

1 . line charts, pie charts, and bar charts
2 . graph paper
3 . footnote or accompanying text

پرسش ۲۸: جداول توزیع فراوانی و توزیع فراوانی تراکمی چه هستند و چگونه محاسبه می‌شوند؟

یک جدول توزیع فراوانی عبارت است از خلاصه‌ای از داده‌ها که نشان می‌دهد وضعیت فراوانی و رخداد هر داده در مجموعه داده‌ها چگونه است. این روش یک شیوه مناسب برای سازمان‌دهی مجموعه‌ای از داده‌ها در قالب طبقات مختلف یا همان طبقه بندی[۱] است.

طبقه بندی داده‌ها مستلزم محاسبه مرحله به مرحله دامنه تغییرات، تعداد طبقات، فاصله طبقات و انواع فراوانی‌های (مطلق، نسبی، تراکمی، نسبی درصدی، تراکمی درصدی، ...) با استفاده از فرمولهای مشخص است. طبقه بندی داده‌ها در یک جدول به نام جدول توزیع فراوانی[۲] گردآوری می‌شود و این جدول باید مبنایی برای محاسبه شاخص‌های مرکزی[۳]، شاخص‌های پراکندگی[۴] و مقایسه گروهی از داده‌ها با گروههای دیگر جهت استنباط آماری است.

برای مثال به توزیع ۲۵ نمره توجه کنید: ۲۰، ۱، ۱، ۴، ۱، ۲۴، ۱۵، ۴، ۱۳، ۳، ۱۱، ۲۳، ۲۰، ۲۳، ۵، ۱۰، ۱۳، ۱۳، ۱۷، ۴، ۲، ۱۱، ۱، ۱۴، ۳؛

برای ایجاد یک جدول فراوانی گام های زیر را بردارید:

۱. یک فاصله طبقاتی[۵] یا یک دامنه نمره برای هر طبقه در نظر بگیرید. سعی کنید برای هر طبقه بین ۵ تا ۱۰ نمره فاصله طبقاتی در نظر بگیرید. محاسبه طبقات از تقسیم دامنه تغییرات بر تعداد طبقات از طریق فرمول ($i=R/K$) به دست می‌آید که در آن k تعداد طبقات، R دامنه تغییرات و i فاصله طبقات است؛ در این مثال، ما فاصله طبقاتی را ۵ نمره در نظر می‌گیریم و در واقع فواصل طبقاتی در این مثال برابر ۰-۴، ۵-۹، ۱۰-۱۴، ۱۵-۱۹، ۲۰-۲۴ خواهند بود.

۲. اکنون، از یک جدول دو ستونی[۶] استفاده می‌کنیم و تعداد تکرار هر نمره‌ای که در هر فاصله طبقاتی جدول توزیع فراوانی قرار دارند را می‌شماریم و در خانه مربوط به خود در جدول وارد می‌کنیم. معمولا نوشتن طبقات را از پایین و با عددی شروع می‌کنند که فاصله طبقات مضربی از آن باشد.

توزیع فراوانی تراکمی[۷]، یک فراوانی با فراوانی ردیف قبلی خود جمع می‌شود و فراوانی تراکمی را به دست می‌دهد. فراوانی تراکمی این اجازه را به ما می‌دهد تا بدانیم که هر طبقه چند نفر را در برگرفته است. به علاوه درصد فراوانی و درصد فراوانی تراکمی نیز همین قابلیت‌های ذکر شده را در قالب درصد نشان می‌دهند. فراوانی تراکمی را می‌توان از فرمول مقابل محاسبه نمود.

---

1 . classification
2 . frequeny table
3 . central index
4 . dispersion index
5 . class interval
6 . two-column table
7 . cumulative frequency distribution

به این معنی که فراوانی تراکمی یا تجمعی هر طبقه عبارت است از: فراوانی آن طبقه به‌اضافه‌ی فراوانی طبقات ماقبل آن. برای نمونه فراوانی تراکم عدد ۷ در مجموعه اعداد ۳، ۴، ۵، ۷، ۷، ۷، ۸، ۱۳ برابر ۶ است. جدول توزیع فراوانی و فراوانی تراکمی در ادامه نمایه شده است.

فاصله طبقاتی	فراوانی	فراوانی تراکمی
۲۰–۲۴	۵	۲۵
۱۵–۱۹	۲	۲۰
۱۰–۱۴	۷	۱۸
۵–۹	۱	۱۱
۰–۴	۱۰	۱۰

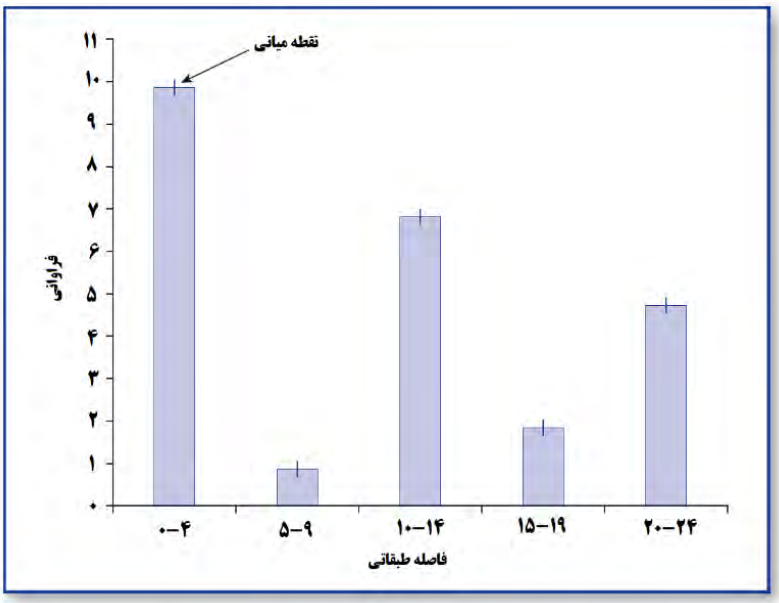
توزیع‌های فراوانی یک نوع نمودار نیستند، چراکه می‌دانیم نمودارها ماهیتاً دیداری هستند، امّا اولین گام در ترسیم نمودارها فهرست داده‌ها و گام دوم ترسیم جداول توزیع فراوانی آن‌هاست. گام سوم نیز ترسیم نمودار بر اساس جدول توزیع فراوانی است.

برای پرسش‌های بیشتر، پرسش‌های ۲۶، ۲۹ و ۳۱ را مطالعه کنید.

پرسش ۲۹: نمودار هیستوگرام چیست و چگونه می‌توان آن را ترسیم کرد؟

نمودار هیستوگرام[1] (مستطیلی) یک نمایه دیداری از یک توزیع فراوانی است. نمودار هیستوگرام یک روش ساده در نمایش مؤثر تعدادی از نمرات است که در طبقات مختلف یک جدول توزیع فراوانی قرار دارند. بهترین نمودار برای داده‌های کمّی، استفاده از نمودار هیستوگرام است. اگر متغیر ها در سطح سنجش فاصله‌ای یا نسبی باشند نمودار مناسب برای توصیف ویژگی های نمونه، نمودار بافت نگار یا هیستوگرام است. نمودار هیستوگرام اعداد را به دسته‌های مختلف تقسیم‌بندی می‌کند و از روی آن می‌توان تصمیم گرفت که کدام دسته انتخاب شود؛ در واقع بُردار افقی پیوسته است. حدود دسته هر یک از میله‌ها همچنین به نام فاصله طبقاتی[2] شناخته می‌شوند. نمودارهای مستطیلی روش خوبی برای نمایش داده‌های پیوسته نظیر قد، وزن و مدت زمان و ... هستند. این نمودار، فراوانی یا درصد فراوانی هر کدام از طبقات را به صورت میله های بهم پیوسته در یک بازه یا طیف نشان می‌دهد. اندازه هر میله، فراوانی آن میله را در بین بازه‌ها نشان می دهد. نمودار هیستوگرام شبیه نمودار ستونی است با این تفاوت که هیستوگرام گروهی از داده‌ها را به دامنه ای از داده ها (بازه) محدود می‌کند و این ما هستیم که انتخاب می‌کنیم کدام بازه از داده‌ها مورد استفاده قرار گیرند.

هرچند در پرسش ۳۰ به شما نشان خواهیم داد که چگونه با استفاده از اِکسل یک نمودار هیستوگرام رسم کنید، امّا ترسیم دستی آن نیز اهمیت دارد. برای رسم یک نمودار هیستوگرام، از داده‌های پرسش ۲۸ استفاده کنید.



1 . Histogram chart
2 . Class Interval

برای ترسیم نمودار هیستوگرام، گام‌های زیر را به ترتیب انجام دهید:

۱. از یک کاغذ نقشه‌کشی میلی‌متری استفاده کنید (این کاغذها را می‌توانید به‌دلخواه خود و بر اساس هدفتان از یک سایت[1] اینترنتی تهیه کنید و سپس پرینت بگیرید). محور عمودی را با عنوان "فراوانی" و محور افقی را با عنوان "فاصله طبقه" نام‌گذاری کنید. در این مثال، محور عمومی می‌تواند از ۰ تا ۱۲ برچسب زده شود و محور افقی می‌تواند برچسب‌های ۰-۴، ۵-۹، ۱۰-۱۴، ۱۹-۱، ۲۴-۲۰ را داشته باشد.

۲. یک ستون برای هر نقطه میانی هر فاصله طبقه ترسیم کنید. نقطه میانی، مرکز فاصله طبقه است. ارتفاع هر ستون به میزان فراوانی نمرات موجود در آن فاصله طبقه مرتبط است. برای مثال ارتفاع ستون در فاصله طبقه ۱۰-۱۴ که نقطه میانی‌اش ۱۲ است، برابر ۷ است. همچنین می‌توانید از روش ترسیم چوب‌خط برای ارزیابی فراوانی یا رخداد هر نمره در مجموعه نمره استفاده کنید، امّا ارزش نمودار هیستوگرام در این است که به‌راحتی می‌توانید میزان فراوانی هر فاصله طبقاتی، بیشترین فراوانی، کمترین فراوانی، دامنه تغییرات تقریبی نمرات و تخمین تقریبی میانگین یا میانه نمرات یک مجموعه داده را به دست آورید. همان‌گونه که در تصویر بالا مشاهده می‌کنید، هر چه فاصله طبقاتی کمتر شود، نمودار ستونی به منحنی بهنجار شبیه‌تر می‌شود.

برای پرسش‌های بیشتر، پرسش‌های ۲۶، ۳۰ و ۳۹ را مطالعه کنید.

[1] . www.printfreegraphpaper.com

پرسش ۳۰: چگونه می‌توان از طریق نرم‌افزار اِکسل یک نمودار هیستوگرام ترسیم کرد؟

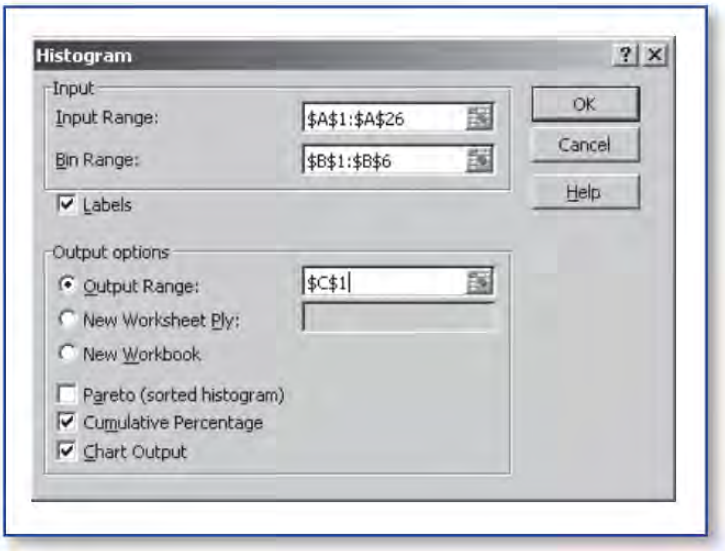
به‌منظور ایجاد یک نمودار هیستوگرام از طریق نرم‌افزار اِکسل داده‌های ۲۸ را در نظر بگیرید و مراحل زیر را به ترتیب رعایت کنید:

۱. فواصل طبقات را از ۰-۴، ۵-۹، ۱۰-۱۴، ۱۵-۱۹، ۲۰-۲۴ به ۱، ۲، ۳، ۴، ۵ تغییر دهید. همانند تصویر زیر، در صفحه اِکسل برای هر طبقه یک عدد که نشانگر میزان فراوانی نمرات موجود در آن طبقه است را در نظر بگیرید.

	A	B
1	Score	Bin
2	1	1
3	1	2
4	1	3
5	1	4
6	1	5
7	2	
8	2	

۲. یک ستون با نام "bins" ایجاد کنید که نشانگر فاصله طبقه است. در واقع در این ستون، ۱ همان فاصله طبقه ۰-۴؛ ۲ همان فاصله طبقه ۵-۹؛ ۳ همان فاصله طبقه ۱۰-۱۴؛ ۴ همان فاصله طبقه ۱۵-۱۹؛ ۵ همان فاصله طبقه ۲۰-۲۴ است.

۳. بر روی Data در نوارابزار صفحه اِکسل کلیک، Data Analysis را انتخاب کرده و با باز شدن پنجره زیر، Histogram را انتخاب کنید.

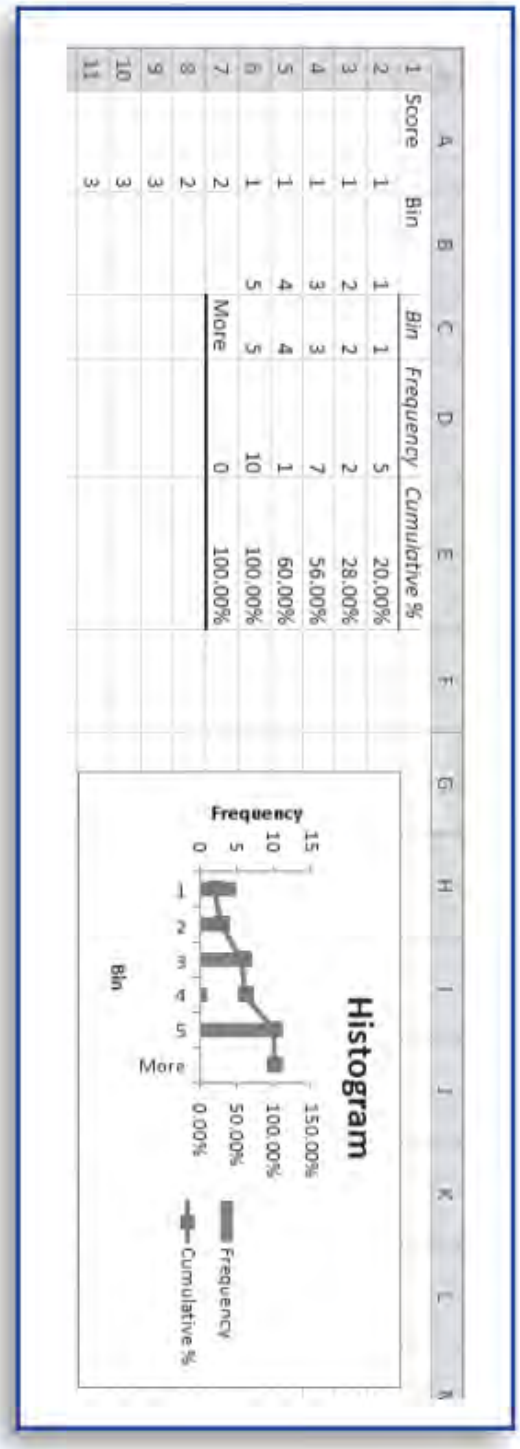


۴. با باز شدن پنجره Histogram، دامنه ورودی، دامنه خروجی و محل نمایش جدول توزیع فراوانی و نمودار را مشخص کنید.

۵. گزینه Labels box را علامت بزنید.

۶. بر روی درصد فراوانی تراکمی و جعبه‌های برونداد نمودار علامت بزنید.

۷. اکنون با کلیک بر روی OK شکل زیر نمایه می‌شود.



همان‌گونه که قبلاً گفتیم، ارزش نمودار هیستوگرام در این است که به‌راحتی می‌توانید میزان فراوانی هر فاصله طبقاتی، بیشترین فراوانی، کمترین فراوانی، دامنه تغییرات تقریبی نمرات و تخمین تقریبی میانگین یا میانه نمرات یک مجموعه داده را به دست آورید.

برای پرسش‌های بیشتر، پرسش‌های ۲۹، ۳۲ و ۳۹ را مطالعه کنید.

پرسش ۳۱: نمودار اجایو یا تراکمی چیست و چگونه می‌توان آن را ترسیم کرد؟

یک نمودار هیستوگرام یک نمایه دیداری[1] از یک توزیع فراوانی است. یک نمودار اجایو[2] یا تراکمی، یک نمایه دیداری از توزیع فراوانی تراکمی[3] است. این نمودار وقتی مفید است که پژوهشگر علاقه مند باشد وضعیت یک نمره یا یک فرد را نسبت به بقیه نمره ها یا افراد مشخص کند.ر به عنوان مثال پژوهشگری می خواهد تعیین کند که یک نمره از چند درصد نمره های دیگر بیشتر یا کمتر است.

فاصله طبقاتی	فراوانی	نقطه میانی	فراوانی تراکمی
۲۰–۲۴	۵	۲۲	۲۵
۱۵–۱۹	۲	۱۷	۲۰
۱۰–۱۴	۷	۱۲	۱۸
۵–۹	۱	۷	۱۱
۰–۴	۱۰	۲	۱۰

این نمودار بسیار سودمند است زیرا تغییر تعداد نمرات در درون هر فاصله طبقه را نشان می‌دهد.

به‌منظور ایجاد یک نمودار اجایو از طریق اِکسل از گام‌های زیر تبعیت کنید:

۱. در یک کاربرگ[4] اِکسل ، همانند آنچه ما در پرسش ۳۰ انجام دادیم، در یک ستون فاصله طبقه و در ستون دیگر تعداد فراوانی یک نمره در آن فاصله طبقه را وارد کنید. برای ستون اول برچسب فاصله طبقه و برای ستون دوم برچسب فراوانی را در نظر بگیرید؛

۲. اکنون دو ستون دیگر ایجاد کنید. در ستون اول نقطه میانی[5] آن فاصله طبقه را وارد کنید و در ستون چهارم فراوانی تراکمی یعنی مجموع تمام فراوانی‌های طبقات قبل تا همین طبقه را وارد نمایید.

۳. ستون دوم (نقطه میانی) و ستون چهارم (فراوانی تراکمی) را هم‌زمان در اِکسل با موس هایلایت (پُرنورتر یا درخشان) کنید؛

۴. بر روی نوارابزار، بر روی Insert و سپس بر Scatter کلیک کنید؛

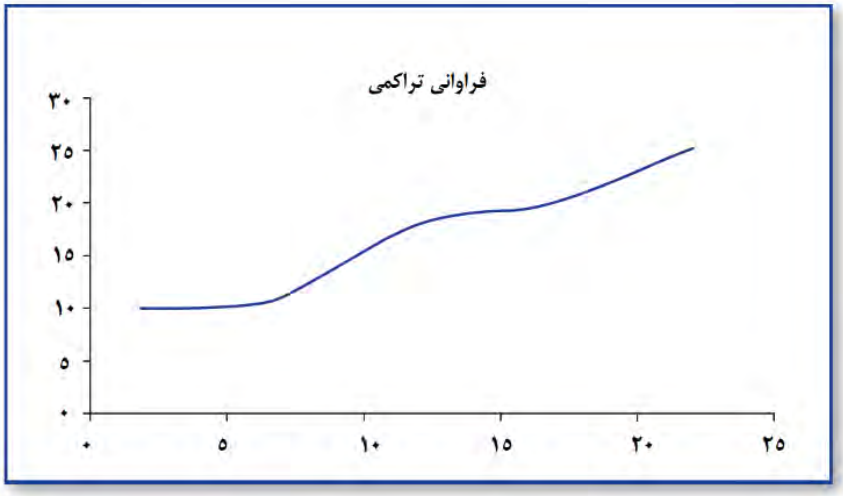
۵. در آیکُن Scatter بر روی گزینه Scatter with Smooth Lines کلیک کنید؛

اگر بر روی گزینه more نیز کلیک کنید، انواع نمودار Scstter نمایه می‌شود که به‌راحتی می‌توانید انتخاب نمایید.

---

1 . visual representation
2 . ogive (/ˈoʊdʒaɪv/ OH-jyv)
3 . cumulative frequency distribution
4 . worksheet
5 . Midpoint

با OK کردن این پنجره، نمودار فراوانی تراکمی مجموعه داده‌ها نمایه می‌شود؛

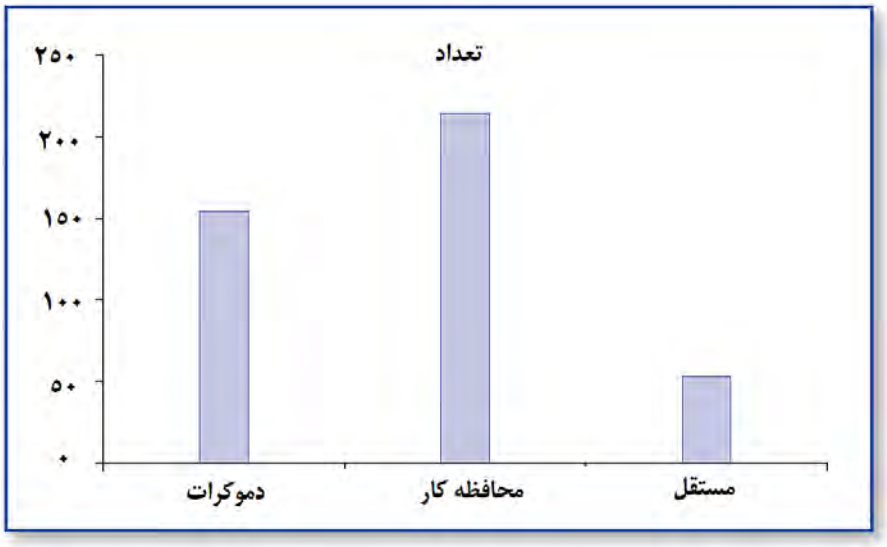


برای پرسش‌های بیشتر، پرسش‌های ۲۸، ۳۰ و ۳۲ را مطالعه کنید.

پرسش ۳۲: نمودار ستونی چیست و چگونه می‌توان آن را ترسیم کرد؟

نمودار ستونی[1] یک نمایه دیداری از داده‌های طبقه‌ای است که به‌صورت ستون‌های عمودی، فراوانی‌ها یا ارزش‌های محور افقی را نشان می‌دهد. یک نمودار ستونی همانند نمودار هیستوگرام است، با این تفاوت که در نمودار ستونی، ستون‌ها مجزا از یکدیگر هستند چون متغیرهای مورد استفاده در آن از نوع گسسته هستند. امّا در نمودار هیستوگرام چون از متغیرهای پیوسته استفاده می‌شود، نمودار از ستون‌های به هم چسبیده تشکیل شده است؛ بنابراین جایگزین مناسب برای نمودار هیستوگرام، نمودار ستونی است. نمودارهای کیفی همچون ستونی (میله‌ای) و دایره‌ای برای داده‌های کیفی که مقیاس اسمی یا ترتیبی دارند به کار می‌رود. تهیه این نوع نمودار نیز بسیار آسان است، به‌این‌ترتیب که هر دسته از اطلاعات را در یک ستون قرار می‌دهیم. از نمودار ستونی بیشتر برای مقایسه استفاده می‌شود که ستون‌ها را با گذاشتن فاصله‌هایی میان آن‌ها از هم جدا می‌کنند. برای داده های ۱۵۴ دموکرات، ۲۱۳ جمهوری خواه و ۵۴ مستقل، نمودار زیر را خواهیم داشت. به‌منظور ایجاد یک نمودار ستونی در اِکسل از مراحل زیر پیروی کنید:

۱. در یک کاربرگ[2] اِکسل ، در یک ستون طبقه و در ستون دیگر تعداد فراوانی یک نمره در آن طبقه را وارد کنید. برای ستون اول برچسب فاصله طبقه و برای ستون دوم برچسب فراوانی را در نظر بگیرید؛ ۲. تمامی داده‌ها را هایلایت کنید؛ ۳. بر روی نوارابزار، بر روی Insert و سپس بر آیکن Column کلیک کنید؛ ۴. در آیکن Column بر روی گزینه 2-D Column icon کلیک کنید؛ که در شکل نشان داده شده است.



برای پرسش‌های بیشتر، پرسش‌های ۲۹، ۳۰ و ۳۹ را مطالعه کنید.

1 . column chart
2 . worksheet

پرسش ۳۳: نمودار میله‌ای چیست و چگونه می‌توان آن را ترسیم کرد؟

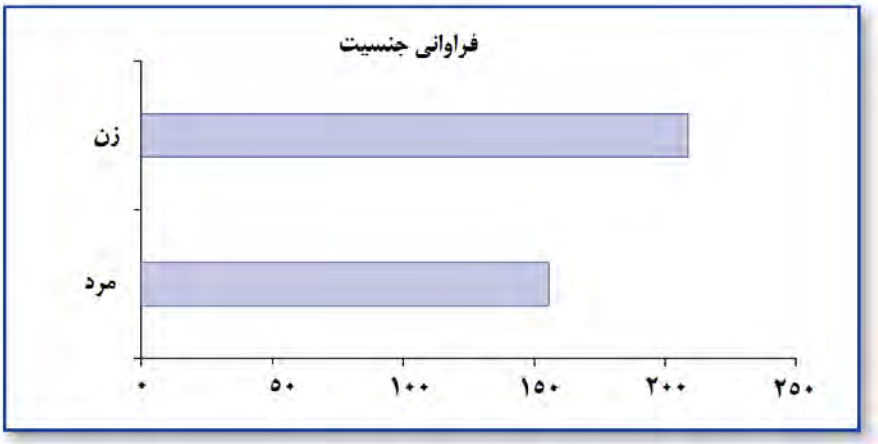
یک نمودار میله‌ای[1] یک نمایه دیداری از داده‌های طبقه‌ای است که به‌صورت ستون‌های افقی یا عمودی، فراوانی‌ها یا ارزش‌های محور افقی را نشان می‌دهد. به‌منظور ایجاد یک نمودار ستونی در اِکسل برای ۱۵۶ مرد و ۲۱۰ زن، از مراحل زیر پیروی کنید:

۱. در یک کاربرگ اِکسل ، در یک ستون طبقه و در ستون دیگر تعداد فراوانی یک نمره در آن طبقه را وارد کنید؛

۲. تمامی داده‌ها را هایلایت کنید؛

۳. بر روی نوارابزار، بر روی Insert و سپس بر آیکن Bar کلیک کنید؛

۴. در آیکن Bar بر روی گزینه 2-D Bar کلیک کنید که در شکل زیر نشان داده شده است.



همان‌طور که در شکل بالا نمایش داده شده است، نمودار ستونی و میله‌ای در اکسل خود اشکال متفاوتی دارند، مکعبی دوبعدی و سه‌بعدی، سیلندری یا استوانه‌ای[2]، مخروطی[3] و هرمی[4]، کاربران می‌توانند نمودار دلخواه را انتخاب نمایند، امکان تغییر نوع نمودار و فرمت آن بعد از رسم نمودار در اکسل هم ممکن است، پس می‌توانید بعد از اتمام رسم نمودار در مورد شکل آن تصمیم بگیرید. نمودار ستونی و میله‌ای در اکسل مشابه یکدیگر هستند و کاربرد یکسانی دارند. تفاوت عمده‌ی نمودار میله‌ای و ستونی در اکسل، عمودی و افقی بودن نمودارها است، که همین امر سبب ایجاد تفاوت در ماهیت داده‌های کاربردی برای این دو نمودار می‌گردد. نمودارهای میله‌ای[5] در اکسل از میله‌های افقی برای نشان دادن کمیت‌ها استفاده می‌نماید و نمودار ستونی[6] در اکسل ستون‌های عمودی را برای مقایسه‌ی کمیت‌ها به کار می‌برد.

برای پرسش‌های بیشتر، پرسش‌های ۲۶، ۳۲ و ۳۹ را مطالعه کنید.

---

1 . bar chart
2 . Cylinder
3 . Cone
4 . Pyramid
5 . Bar Charts
6 . Column Chart

پرسش ۳۴: نمودار خطی چیست و چگونه می‌توان آن را ترسیم کرد؟

یک نمودار خطی[1] (چندضلعی) یک نمایه دیداری از داده‌های غیرطبقه‌ای[2] است که با استفاده از خط، فراوانی‌ها یا ارزش‌های محور افقی را نشان می‌دهد. اگر وسط مستطیل‌های نمودار هیستوگرام را با پاره‌خط به یکدیگر متصل کنیم، یک نمودار خطی به دست می‌آید.

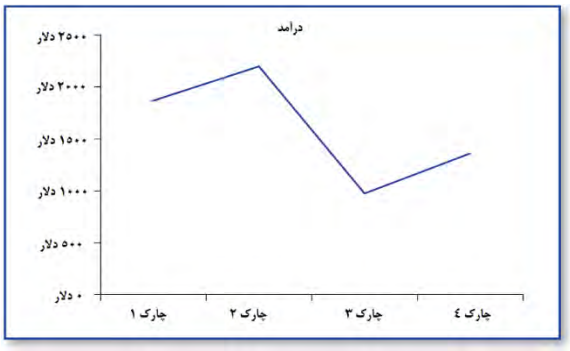
نمودار خطی که ساده‌ترین نوع نمودار است، هنگامی به کار می رود که داده‌ها به‌صورت پیوسته باشد و همچنین در نشان دادن ارتباط میان دو سری اطلاعات بسیار مفید است. هنگامی از این نوع استفاده می‌شود که اطلاعات بسیار زیادی در دست باشد. در این نمودار تنظیم اعداد بر روی محور افقی از چپ به راست و بر روی محور عمودی از پایین به بالا صورت می‌گیرد و همواره باید نقطه صفر مشخص شود و مقیاس فواصل مساوی باشد. برای رسم نمودارهای خطی می‌توان از کاغذهای شطرنجی استفاده کرد. محور افقی در این نمودار معمولاً اندازه‌های متغیر مستقل و محور عمودی صفت اندازه‌گیری شده را نشان می‌دهد. کاربرد مهم نمودار خطی یا چندضلعی رسم دو سری از داده‌ها بر روی یکی صفحه مختصات و مقایسه آن‌ها با یکدیگر است. به‌منظور ایجاد یک نمودار ستونی در اِکسل، چارک درآمد شرکت کنندگان در واحد "هزاردلار" ذکر شده است. برای چارک ۱: ۱/۸۶۷ دلار؛ چارک ۲: ۲/۱۹۳ دلار؛ چارک ۳: ۹۸۹ دلار و چارک ۴: ۱/۳۵۸ دلار. برای رسم نمودار خطی این مثال، از مراحل زیر پیروی کنید:

۱. در یک کاربرگ اِکسل، در یک ستون طبقه و در ستون دیگر تعداد فراوانی یک نمره در آن طبقه را وارد کنید. در جدول زیر یک نمونه درباره چارک درآمد به هزار دلار ارائه شده است.

۲. تمامی داده‌ها را هایلایت کنید.

۳. بر روی نوارابزار، بر روی Insert و سپس بر آیکن Line کلیک کنید.

۴. در آیکن Line بر روی گزینه 2-D Line کلیک کنید که در شکل نشان داده شده است.



در شکل زیر نمودار خطی نمایه شده است. نکته مهمی که می‌توان به آن اشاره کرد این است که اگر بخواهیم وضعیت یک نمونه را با نمونه‌ی دیگری مقایسه کنیم، بهترین روش، استفاده از نمودار خطی یا چندضلعی است.

برای پرسش‌های بیشتر، پرسش‌های ۳۷، ۳۸ و ۳۹ را مطالعه کنید.

1 . line chart
2 . noncategorical data

پرسش ۳۵: نمودار دایره‌ای چیست و چگونه می‌توان آن را ترسیم کرد؟

یک نمودار دایره‌ای[1] یک نمایه دیداری از داده‌های طبقه‌ای و کیفی است که با استفاده از بخش-های یک دایره یا "بُرِش"، فراوانی‌های سطوح مختلف یک متغیر را نشان می‌دهد. در این نوع نمودار شعاع‌های دایره به‌طور عمودی کشیده می‌شوند و بخش‌های تشکیل‌دهنده نمودار در جهت حرکت عقربه ساعت از بزرگ به کوچک تنظیم می‌شود. در یک دایره ۳۶۰ درجه‌ای نسبت عددها، برحسب تعداد درجه‌های هریک از بخش‌های نمودار مشخص می‌شود.نمودار دایره‌ای ساده‌ترین و درعین‌حال مناسب‌ترین روش برای مقایسه و نمایش داده‌های جمع‌آوری‌شده از متغیرهای گسسته است.

درجه (زاویه یا کمان) مربوط به هر داده بر روی نمودار دایره‌ای از فرمول $d = \frac{Fi}{n} \times ۳۶۰$ به دست می‌آید. برای مثال، گروه خونی ۴۵۰ نفر به‌صورت نمودار دایره‌ای رسم شده است. زاویه‌ای که برای نمایش نسبت فراوانی گروه خونی AB به‌دست‌آمده برابر ۷۲ درجه خواهد بود. تعداد افراد دارای گروه خونی AB برابر خواهد بود با:

$$۷۲= \frac{Fi}{۴۵۰} \times ۳۶۰ : Fi = \frac{۴۵۰ \times ۷۲}{۳۶۰} = ۹۰$$

به‌منظور ایجاد یک نمودار ستونی در اِکسل از مراحل زیر پیروی کنید:

۱. در یک کاربرگ اِکسل، در یک ستون طبقه و در ستون دیگر تعداد فراوانی آن طبقه را وارد کنید. در ادامه یک نمونه از انواع مدل‌های اتومبیل فروخته شده ارائه شده است: مدل ولوو[2] ۵۶۴ عدد؛ مدل شولت[3] ۳۴۳۴ عدد؛ مدل هوندا[4] ۴۳۳۱ عدد و مدل مرسدس[5] ۳۱۲ عدد.

۲. تمامی داده‌ها را هایلایت کنید.

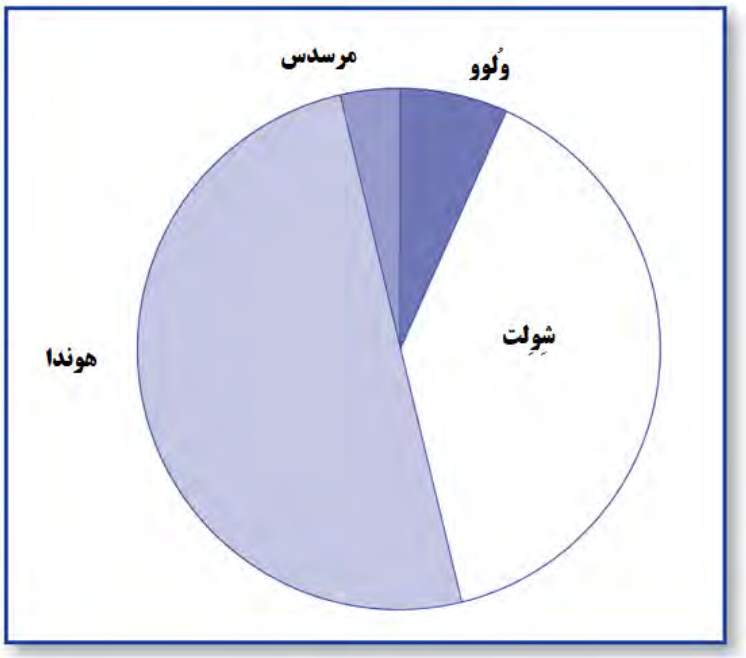
۳. بر روی نوارابزار، بر روی Insert و سپس بر آیکن Pie کلیک کنید.

۴. در آیکن Pie بر روی گزینه 2-D Pie کلیک کنید.

نکته مهمی که می‌توان به آن اشاره کرد این است که اگر بخواهیم داده‌های کیفی نظیر اطلاعات جمعیتی را نمایه کنیم، بهترین روش استفاده از نمودار دایره‌ای است.

---

1 . pie /paɪ/ chart
2 . Volvo
3 . Chevrolet
4 . Honda
5 . Mercedes



برای پرسش‌های بیشتر، پرسش‌های ۲۶، ۳۷ و ۳۸ را مطالعه کنید.

پرسش ۳۶: نمودار پراکندگی چیست و چگونه می‌توان آن را ترسیم کرد؟

یک نمودار پراکندگی (اِسکَتر)[1] که به نمودار نقطه‌ای نیز شناخته می‌شود، یک نمایه دیداری از دودسته نقاط داده‌ای در یک مجموعه داده است. این نمودار نقطه‌هایی روی خود دارد که ارتباط بین دو نوع از داده‌ها را نمایش می‌دهند و مقادیر متغیر ورودی را در محور افقی و مقادیر متناظر پاسخ را در محور عمودی ترسیم می‌کند. به‌منظور ایجاد یک نمودار ستونی در اِکسل از مراحل زیر پیروی کنید:

۱. در یک کاربرگ اِکسل، در یک ستون طبقه و در ستون دیگر تعداد فراوانی آن طبقه را وارد کنید. در جدول زیر یک نمونه از قد و وزن ۱۰ شرکت‌کننده در برنامه کاهش وزن ارائه شده است.

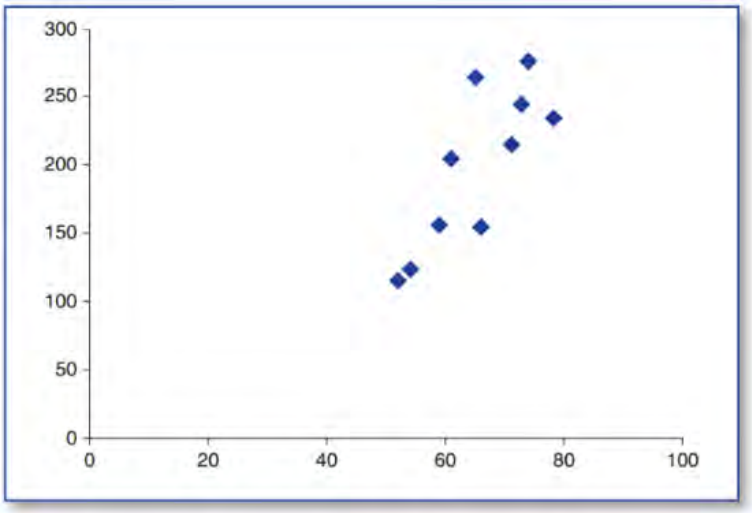
شرکت کننده	قد (اینچ)	وزن (پوند)
۱	۶۶	۱۵۴
۲	۵۴	۱۲۳
۳	۷۱	۲۱۴
۴	۶۵	۲۶۵
۵	۷۴	۲۷۶
۶	۶۱	۲۰۵
۷	۵۲	۱۱۶
۸	۷۳	۲۴۵
۹	۷۸	۲۳۵
۱۰	۵۹	۱۵۶

۲. تمامی داده‌ها را هایلایت کنید.

۳. بر روی نوارابزار، بر روی Insert و سپس بر آیکن Scatter کلیک کنید.

۴. در آیکن Scatter بر روی گزینه Scatter کلیک کنید که در شکل زیر نشان داده شده است. همان‌گونه که در نمودار نقطه‌ای یا پراکندگی بالا مشاهده می‌کنید،

1 . pie chart

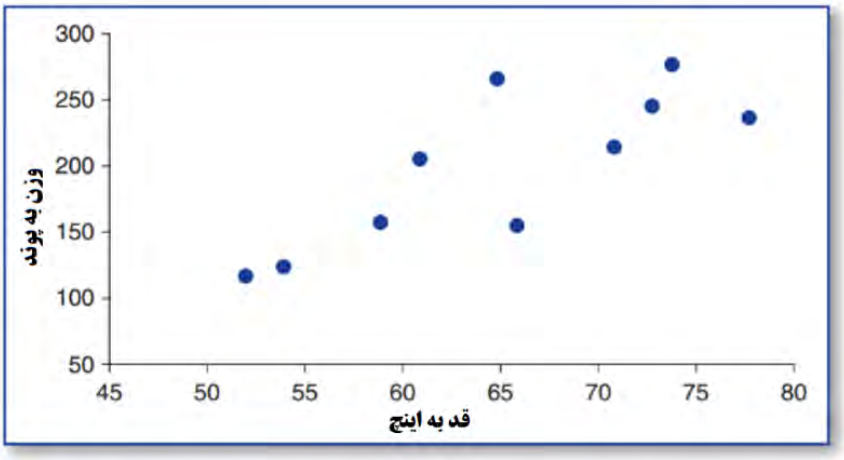


هر نقطه نماینده دو عدد است: یک عدد در محور افقی و یک عدد در محور عمودی. در حقیقت نقطه، حاصل تقاطع ارزش‌های محورهای عمودی و افقی است.

برای پرسش‌های بیشتر، پرسش‌های ۳۸، ۴۲ و ۴۳ را مطالعه کنید.

پرسش ۳۷: چگونه می‌توانم نمودارهایی را که در اِکسل رسم کرده‌ام ویرایش کنم؟

در حقیقت هزاران مسیر برای ویرایش یک نمودار ترسیم شده وجود دارد. در نمودار زیر می‌توانید ملاحظه کنید که در عناوین افقی و عمودی نسبت به نمودار پرسش قبل ویرایش صورت گرفته است.



برای ویرایش نمودار، از گام‌های زیر پیروی کنید:

۱. موس را رو نمودار قرار داده و راست کلیک کنید و در پنجره بازشده Chart Area را انتخاب کنید.

۲. در نوارابزار اِکسل بر روی گزینه‌های طرح‌بندی یا Chart Layout کلیک کنید.

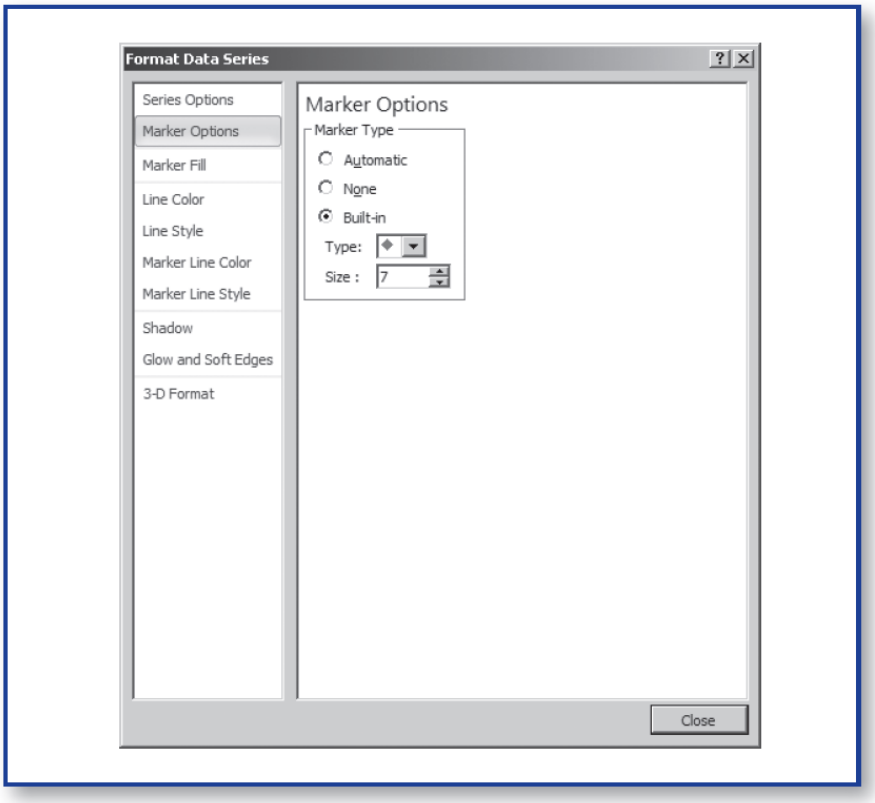
۳. بر Chart Title که بالای نمودار نوشته شده است، دو بار کلیک کنید و سپس عبارت مورد نظر را در آن مکان تایپ کنید.

۴. بر محور عمودی دو بار کلیک کنید و سپس عبارت مورد نظر را در آن مکان تایپ کنید.

۵. بر محور افقی دو بار کلیک کنید و سپس عبارت مورد نظر را در آن مکان تایپ کنید.

۶. بر روی برچسب Series1 کلیک کنید و سپس با کلید Delete آن را حذف نمایید.

۷. بر روی گزینه diamond-shaped marker در نمودار کلیک کنید تا پنجره مربوط به گزینه Format Data Series باز شود.



۸. بر روی گزینه Marker Options و سپس Built-in کلیک کنید. اکنون در پنجره ظاهرشده، گزینه solid circle را انتخاب کنید.

۹. اکنون بر روی Close کلیک کنید.

نکته مهم این است که در ویرایش یک نمودار می‌توانید با دو بار کلیک بر هر قسمتی از نمودار، امکان ویرایش آن وجود خواهد داشت.

برای پرسش‌های بیشتر، پرسش‌های ۲۶، ۳۶ و ۳۹ را مطالعه کنید.

پرسش ۳۸: چگونه می‌توانم یک نمودار را در سایر اسنادی که دارم ترکیب کنم؟

هنگامی‌که با مجموعه‌ای از داده‌ها کار می‌کنید، بعد از ایجاد یک برونداد که ممکن است جدول یا نمودار باشد، می‌بایست آن را به یک فضای دیجیتال دیگر نظیر Word یا داده‌پردازهای دیگر انتقال دهید. بدین منظور، از گام‌های زیر پیروی کنید:

۱. بر روی نمودار کلیک کنید؛

۲. اکنون بر روی نمودار کلیک راست کنید و سپس نمودار را Copy کنید؛

۳. اکنون فضایی نظیر Word، Power point[۱]، اِکسل و یا هر فضای دیگری که قصد انتقال نمودار یا جدول را به آن دارید باز کنید؛

۴. اکنون با کلیک راست و paste، انتقال را انجام دهید؛

اکنون انتقال انجام شده است و نمودار یا جدول یا هر برونداد دیگری که موردنظر شماست در سند جدید نمایه می‌شود.

نکته: برخی مواقع، برخی پژوهشگران از نموداری که در صفحه اِکسل ایجاد شده است، از طریق دستورالعمل Ctrl+PrtSc در کیبورد عکس می‌گیرند و بعد از انتقال به Paint و ویرایش آن، نمودار نهایی را به سند جدید نظیر Word انتقال می‌دهد. امّا روش عالی‌تری وجود دارد و آن اینکه، اگر نمودار را به‌صورت مستقیم Copy کنید و در سند Word که موردنظر دارید paste کنید، با تغییر داده‌ها در اِکسل نمودار انتقال داده شده به Word نیز تغییر خواهد کرد. این قابلیت معمولاً در نرم‌افزارهایی که در یک شرکت ساخته شده باشند وجود دارد. برای مثال نرم‌افزارهای اِکسل و Word هر دو توسط شرکت Microsoft ساخته شده‌اند.

برای پرسش‌های بیشتر، پرسش‌های ۲۶، ۳۷ و ۳۹ را مطالعه کنید.

---

۱. پرده نگار

پرسش ۳۹: چه هنگامی می‌بایست از یک نمودار استفاده کنم؟ نوع انتخاب نمودار به چه عواملی بستگی دارد؟

پاسخ به این پرسش آسان است. هر موقع که تشخیص می‌دهید نمایه دیداری داده‌هایی که مورد تحلیل واقع شده‌اند به درک خوانندگان کمک می‌کند، می‌توانید از نمودار استفاده کنید. هرچند در اغلب موارد نمودارها بسیار کمک‌کننده هستند، امّا کاربرد نمودار نیاز به یک استدلال دارد. درواقع هر گوشه‌ای از داده‌ها نیازمند نمایه دیداری نیست. در ادامه به خلاصه‌ای از معیارهای استفاده از نمودار در داده‌های مختلف و مثالی از کاربست داده‌ها پرداخته خواهد شد. توجه داشته باشید که برای یک مجموعه داده، از نمودارهای بسیاری می‌توان استفاده کرد امّا برخی نمودارها در نمایه‌سازی برخی داده‌ها بسیار کارآمدتر و سودمندترند. جدول زیر در ابعاد وسیعتر در صفحه بعد ارائه شده است.

نوع نمودار	کاربرد	مثال
نمودار ستونی	مقایسه مولفه‌ها با یکدیگر، یا بررسی تغییرات در طول زمان. در این نمودار ستون‌ها عمودی هستند.	تعداد موافقان و مخالفان یک نامزد انتخاباتی
نمودار میله‌ای	مقایسه مولفه‌ها با یکدیگر، یا بررسی تغییرات در طول زمان. در این نمودار میله‌ها افقی هستند.	میزان درآمد سه شرکت اقتصادی در سال ۲۰۱۷
نمودار خطی	نشان دادن داده هایی که در طول زمان مداومت دارند. این نمودار بهترین کاربست را در نمایش روند تغییر داده ها با فواصل یکسان دارد.	تغییرات ماهانه ایجاد شده در فروش محصولات یک شرکت در یک سال
نمودار دایره ای	برای نشان دادن اندازه یک مولفه در میان چند مولفه بصورت بُرش و بخش ها مجزا بصورت دایره	فروش انواع مختلف شکلات
نمودار مساحت[1]	برای تاکید بر میزان تغییرات در طی زمان	رشد در جمعیت یک کشور در مقایسه با نه کشور در حال توسعه
نمودار پراکندگی یا نقطه ای	برای نمایش ارتباط بین دو مجموعه داده بصورت نقطه	میانگین تحصیلی دانش آموزان دبیرستانی
نمودار دایرای دونات[1]	شبیه نمودار دایره ای به منظور نشان دادن میزان بخش بخش هر مولفه در میان چندین مولفه به کار می رود. امّا اجازه می دهد بیش از یک سری از داده ها در یک زمان نشان داده شوند.	فروش چهارماهه در طول سه سال

برای پرسش‌های بیشتر، پرسش‌های ۲۷، ۳۷ و ۳۸ را مطالعه کنید.

نوع نمودار	کاربرد	مثال
نمودار ستونی	مقایسه مولفه‌ها با یکدیگر، یا بررسی تغییرات در طول زمان. در این نمودار ستون‌ها عمودی هستند.	تعداد موافقان و مخالفان یک نامزد انتخاباتی
نمودار میله‌ای	مقایسه مولفه‌ها با یکدیگر، یا بررسی تغییرات در طول زمان. در این نمودار میله‌ها افقی هستند.	میزان درآمد سه شرکت اقتصادی در سال ۲۰۱۷
نمودار خطی	نشان دادن داده هایی که در طول زمان مداومت دارند. این نمودار بهترین کاربست را در نمایش روند تغییر داده ها با فواصل یکسان دارد.	تغییرات ماهانه ایجاد شده در فروش محصولات یک شرکت در یک سال
نمودار دایره ای	برای نشان دادن اندازه یک مولفه در میان چند مولفه بصورت بُرش و بخش ها مجزا بصورت دایره	فروش انواع مختلف شکلات
نمودار مساحت[1]	برای تاکید بر میزان تغییرات در طی زمان	رشد در جمعیت یک کشور در مقایسه با نه کشور در حال توسعه
نمودار پراکندگی یا نقطه ای	برای نمایش ارتباط بین دو مجموعه داده بصورت نقطه	میانگین تحصیلی دانش آموزان دبیرستانی
نمودار دایرای دونات[1]	شبیه نمودار دایره ای به منظور نشان دادن میزان بخش بخش هر مولفه در میان چندین مولفه به کار می رود. امّا اجازه می دهد بیش از یک سری از داده ها در یک زمان نشان داده شوند.	فروش چهارماهه در طول سه سال

بخش ۵: درک روابط

پرسش ۴۰: ضریب همبستگی چیست و چگونه به کار می‌رود؟

حتماً تاکنون درباره انواع شاخص‌های گرایش مرکزی نظیر میانگین، میانه و نما و شاخص‌های پراکندگی و تغییرپذیری نظیر انحراف استاندارد و واریانس مطالب زیادی می‌دانید. امّا برخی مواقع بسیار مهم است که بتوانید ارتباط بین دو یا چند متغیر را توصیف کنید. ضریب همبستگی ابزاری آماری برای تعیین نوع و درجه رابطه یک متغیر کمی با متغیر کمی دیگر است. ضریب همبستگی، یکی از معیارهای مورداستفاده در تعیین همبستگی دو متغیر است. ضریب همبستگی شدت رابطه و همچنین نوع رابطه (مستقیم یا معکوس) را نشان می‌دهد. این ضریب بین ۱ تا ۱- است و در عدم وجود رابطه بین دو متغیر، برابر صفر است. در مباحث آماری، ضریب همبستگی پیرسون یا ضریب همبستگی گشتاوری پیرسون[1] میزان همبستگی خطی بین دو متغیر تصادفی را می‌سنجد. این ضریب که کاربرد فراوانی در آمار دارد، توسط کارل پیرسون بر اساس ایده اولیه فرانسیس گالتون تدوین شد. نماد ضریب همبستگی حرف r است و مقدار این ضریب بین ۱- تا ۱ تغییر می‌کند که «۱» به معنای همبستگی مثبت کامل، «۰» به معنی نبود همبستگی و «۱- » به معنی همبستگی منفی کامل است. همبستگی بین دو متغیر را اصطلاحاً همبستگی دومتغیری می‌نامند. همبستگی بین چند متغیر به‌راحتی قابل محاسبه است. برای نمونه اگر متغیر سن، قد و وزن باهم ارتباط داشته باشند، این همبستگی‌ها را می‌توان به‌صورت زیر نمایه کرد.

همبستگی پیرسون برای محاسبه همبستگی بین متغیرهایی به کار می‌رود که ماهیتاً پیوسته باشند. متغیر پیوسته، متغیری است که بین دو واحد آن، هر نقطه یا ارزشی را می‌توان انتخاب کرد. در این متغیر درجات مختلف اندازه‌گیری وجود دارد و دقت وسیلهٔ اندازه‌گیری، تعداد این درجات را تعیین می‌کند. به‌عنوان‌مثال، وزن یک متغیر پیوسته است و می‌تواند بین صفر تا بی‌نهایت باشد. قد، زمان، طول یا ارتفاع پرش، درصد چاقی بدن و سطح هموگلوبین خون متغیرهای پیوسته هستند. ناگفته نماند که در عمل تشخیص بین متغیر پیوسته و گسسته به‌صورت نظری امکان‌پذیر نیست. دلیل این امر فقدان وسایل اندازه‌گیری دقیق و مناسب است. در خیلی از متغیرهای پیوسته ما ناگزیریم اعداد را به‌صورت کلی برای اندازه‌گیری به کاربریم. بهرهٔ هوشی ازنظر تئوری یک متغیر پیوسته است؛ اما در عمل، آزمونی که برای اندازه‌گیری هوش به‌کاربرده می‌شود، به‌گونه‌ای است که نمره‌ها را به‌صورت کلی یا نمره‌های گسسته نشان می‌دهد. برای مثال، ممکن است یک محقق به بررسی رابطه همبستگی پیرسون بین سال‌های تحصیل و نمرات آزمون استعداد تحصیلی بپردازد امّا اگر متغیرها پیوسته نباشد بلکه نظیر جنسیت یا نژاد یا نظرات سیاسی طبقه‌ای باشند، از سایر شاخص‌های همبستگی استفاده می‌شود.

برای پرسش‌های بیشتر، پرسش‌های ۴۱، ۴۲، ۵۰ را مطالعه کنید.

1 . Pearson product - moment correlation coefficient

پرسش ۴۱: آیا مثالی برای چگونگی کاربست ضریب همبستگی وجود دارد؟

همبستگی‌ها همیشه برای قضاوت درباره شدت رابطه[1] بین متغیرها به‌کاربرده می‌شوند. برخی مواقع صرفاً به‌صورت توصیفی و برای نشان دادن اینکه آیا متغیرها باهم ارتباط دارند به‌کاربرده می‌شوند. بعلاوه شاخص همبستگی برای استنباط بکار برده می‌شود به این معنا که تا چه حد نتایج یک نمونه به یک جامعه قابل‌تعمیم است. در همین راستا، یک مطالعه به بررسی ارتباط بین فرهنگ مشارکت در مدرسه و پیشرفت تحصیلی دانش آموزان در مدرسه پرداخت. داده‌های مربوط به فرهنگ مدارس که شامل شش عامل بود از معلمین ۸۱ مدرسه در ایندیانا[2] جمع‌آوری شد. یافته‌ها نشان داد که همه این شش عامل با پیشرفت تحصیلی دانش آموزان ارتباط داشت که نشان داد خرده مؤلفه‌های فرهنگ مشارکتی بین معلمین نظیر حمایت از یکدیگر، با نمرات دانش آموزان در آزمون‌های درسی همبستگی داشت. تمامی شش عامل سازنده فرهنگ مشارکت که عبارت بودند از رهبری مشارکتی[3]: مشارکت تسهیل‌کننده بین معلمین؛ مشارکت معلمین[4]: رفتارهای بیانگر مشارکت و همکاری معلمین با یکدیگر؛ پیشرفت حرفه‌ای[5]: نگرش معلمین به ایده‌های جدید؛ هدف مشترک[6]: چگونگی تأثیر قوانین مدیریتی مدارس بر شیوه تدریس؛ حمایت آموزشگاهی[7]: مشارکت آموزشگاهی بین معلمین؛ همکاری در یادگیری[8]: کیفیت ارتباطات بین معلمین و والدین؛ با پیشرفت تحصیلی دانش آموزان همبستگی داشت. این یافته‌ها پیشنهاد می‌کنند که ازآنجاکه بین این مؤلفه‌ها با پیشرفت تحصیلی ارتباط مشاهده‌شده است، بنابراین با ارائه آموزش رسمی به معلمین در جهت ارتقاء مهارت‌های مشارکتی می‌توان پیشرفت تحصیلی دانش آموزان را افزایش داد.

برای مطالعه بیشتر درباره مثال موجود در این پرسش به مطالعه مرجع زیر بپردازید.

Gruenert, S. (2005). Correlations of collaborative school cultures with student achievement. NASSP Bulletin, 89, 43-55.

برای پرسش‌های بیشتر، پرسش‌های ۴۰، ۴۴ و ۴۹ را مطالعه کنید.

1 . the strength of the relationship
2 . Indiana
3 . Collaborative Leadership
4 . Teacher Collaboration
5 . Professional Development
6 . Unity of Purpose
7 . Collegial Support
8 . Learning Partnership

پرسش ۴۲: انواع شدت ضرایب همبستگی کدام‌ها هستند؟

میزان یا اندازه (۱- تا ۱+) و جهت (منفی یا مثبت) یک ضریب همبستگی، اطلاعات زیادی درباره ارتباط بین متغیرها و چگونگی تغییراتشان در مقابل یکدیگر در اختیارمان قرار می‌دهد. ضریب همبستگی شدت رابطه و همچنین نوع رابطه (مستقیم یا معکوس) را نشان می‌دهد. این ضریب بین ۱ تا ۱- است و در عدم وجود رابطه بین دو متغیر، برابر صفر است. اگر متغیرها در یک جهت تغییر کنند، ضریب همبستگی مستقیم یا مثبت[1] خواهد بود. برای مثال، هر چه کودکان بلندقدتر باشند، وزنشان بیشتر خواهد بود. این جمله به این معناست که قد و وزن با یکدیگر به‌صورت مستقیم یا مثبت همبستگی دارند. همبستگی مثبت از ۰ تا ۱+ ارزش می‌گیرد.

اگر متغیرها با جهت مخالف یکدیگر تغییر یابند، همبستگی بین آنها غیرمستقیم یا منفی[2] خواهد بود. برای مثال، هر چه یک فرد با سرعت بیشتری یک آزمون را تکمیل کند، احتمال پاسخ‌های خطای او زیادتر خواهد بود، به این معنا که زمان پاسخ آزمون و میزان خطا، همبستگی غیرمستقیم یا منفی دارند. همبستگی منفی از ۰ تا ۱- ارزش می‌گیرد.

در جدول زیر یک خلاصه از نوع تغییر متغیرها در برابر یکدیگر، نوع همبستگی، میزان همبستگی و یک نمونه از هر کدام ارائه شده است.

متغیر X	متغیر Y	نوع همبستگی	میزان همبستگی	مثال: رابطه بین ...
افزایش	افزایش	مستقیم یا مثبت	از ۰ تا ۱	ساعات بیشتر مطالعه و نمرات بیشتر امتحان
کاهش	کاهش	مستقیم یا مثبت	از ۰ تا ۱	ساعات کمتر مطالعه و نمرات کمتر امتحان
افزایش	کاهش	غیرمستقیم یا منفی	از ۰ تا ۱-	ورزش بیشتر و وزن کمتر
کاهش	افزایش	غیرمستقیم یا منفی	از ۰ تا ۱-	ورزش کمتر و وزن بیشتر

تحلیل گران و متخصصان آماری، دامنه‌های متفاوتی را برای تفسیر ضریب همبستگی در نظر گرفته‌اند. در اینجا به مهم‌ترین این دسته‌بندی‌ها و نحوه تفسیر آن‌ها که کاربرد بیشتری دارد اشاره می‌کنیم:

دامنه همبستگی ۰/۲ تا ۰/۳۵: همبستگی بین این دامنه، هرچند ازنظر آماری ممکن است معنی‌دار باشد، تنها رابطه نسبتاً ضعیفی را بین متغیرها نشان می‌دهد. همبستگی ۰/۲ نشان می‌دهد که تنها چهار درصد واریانس بین دو متغیر مشترک است.

دامنه همبستگی ۰/۳۵ تا ۰/۶۵: در این دامنه، همبستگی‌ها ازنظر آماری بالاتر از سطح یک در صد معنی‌دار هستند. وقتی همبستگی حدود ۰/۴۰ باشد، پیش‌بینی ساده گروهی امکان‌پذیر است.

1 . direct or positive
2 . indirectly or negatively

دامنه ۰/۶۵ تا ۰/۸۵: با استفاده از همبستگی‌های این دامنه، می‌توان پیش‌بینی‌های گروهی انجام داد که به‌اندازه کافی دقیق باشند. از ضرایب همبستگی بالاتر از ۰/۷۵، می‌توان برای پیش‌بینی‌های فردی تا حد قابل قبولی استفاده کرد.

همبستگی‌های بالاتر از ۰/۸۵: همبستگی‌های بالاتر از ۰/۸۵ نشان‌دهنده رابطه نزدیک بین دو متغیر است. همبستگی ۰/۸۵ حاکی از آن است که اندازه مورداستفاده برای پیش‌بینی حدود ۷۲ درصد با عملکرد مورد پیش‌بینی، واریانس مشترک دارد. مطالعات پیش‌بینی در علوم اجتماعی و تربیتی به‌ندرت تا این حد بالاست. بااین‌حال، اگر همبستگی در این حد به دست آمد، برای پیش‌بینی گروهی و انفرادی بسیار مفید است.

یکی از شیوه‌های بسیار مؤثر در نمایش چگونگی همبستگی بین متغیرها استفاده از اشکال هندسی است.

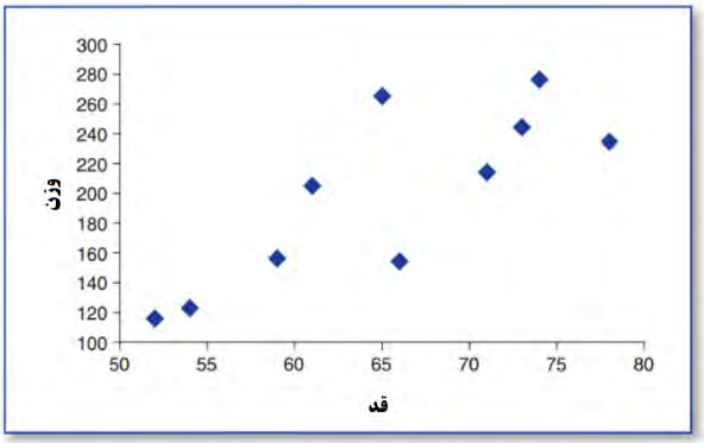
برای پرسش‌های بیشتر، پرسش‌های ۴۰، ۴۱، ۴۶ را مطالعه کنید.

پرسش ۴۳: نمودارهای پراکندگی چگونه به فهم ضرایب همبستگی کمک می‌کنند؟

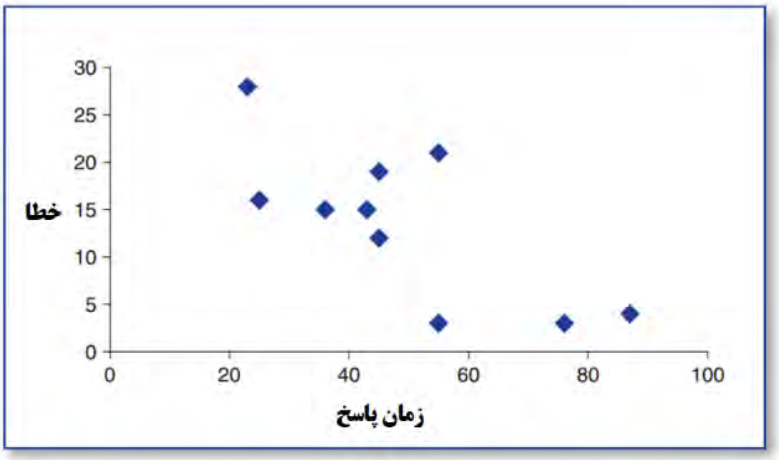
همان‌گونه که در پرسش‌های قبل آموختید، یک نمودار پراکندگی، یک نمایه دیداری از نقاطی است که بیانگر ارتباط دو مجموعه داده است. ایجاد یک نمودار پراکندگی، همان‌گونه که در جدول زیر ملاحظه می‌کنید، یک روش ساده برای نشان دادن نقاط تقاطع دو ارزش یا عدد بر روی نمودار است.

شرکت کننده	قد (اینچ)	وزن (پوند)
۱	۶۶	۱۵۴
۲	۵۴	۱۲۳
۳	۷۱	۲۱۴
۴	۶۵	۲۶۵
۵	۷۴	۲۷۶
۶	۶۱	۲۰۵
۷	۵۲	۱۱۶
۸	۷۳	۲۴۵
۹	۷۸	۲۳۵
۱۰	۵۹	۱۵۶

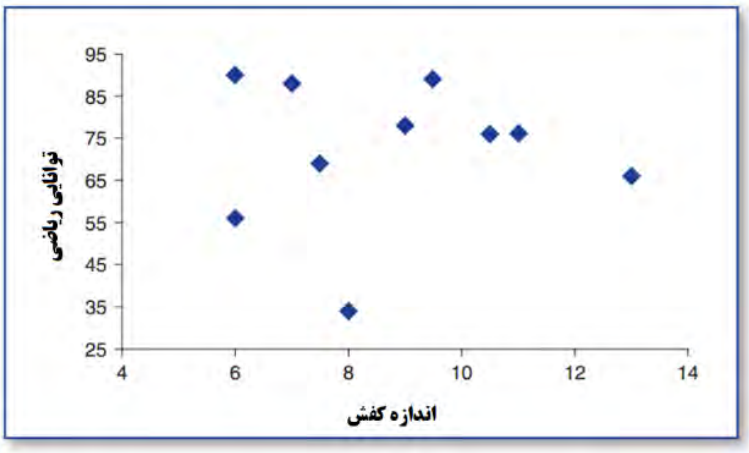
هنگامی‌که به نمودار پراکندگی زیر نگاه می‌کنید، می‌توانید ببینید نقاط روی نمودار که بیانگر تقاطع نقاط موجود در محورهای افقی (X) و عمودی (Y) هستند، از گوشه پایین سمت چپ نمودار تا گوشه بالای سمت راست نمودار سازمان یافته‌اند. این وضعیت و چینش نشان‌دهنده این است که این دو متغیر باهم همبستگی مثبت دارند. بنابراین نمودار زیر نشانگر یک همبستگی مثبت است. به این معنی که با افزایش قد، وزن افزایش می‌یابد و با کاهش قد، وزن کاهش می‌یابد.



در نمودار زیر، یک نمایه دیداری از همبستگی منفی ارائه شده است. در این نمودار، نقاط روی نمودار که بیانگر تقاطع نقاط موجود در محورهای افقی (X) و عمودی (Y) هستند، از گوشه بالای سمت چپ نمودار تا گوشه پایین سمت راست نمودار سازمان یافته‌اند. این وضعیت و چینش نشان‌دهنده این است که این دو متغیر باهم همبستگی منفی دارند. بنابراین نمودار زیر نشانگر یک همبستگی منفی است. به این معنی که با افزایش زمان پاسخگویی به آزمون، میزان خطا کاهش می‌یابد و با کاهش زمان پاسخگویی به آزمون، میزان خطا افزایش می‌یابد.



سرانجام، وضعیت‌هایی وجود دارد که بین دو یا چند متغیر هیچ‌گونه ارتباطی وجود ندارد و هیچ‌گونه اشتراک و تعاملی بین آنها وجود ندارد. در نمودار زیر چنین وضعیتی که در آن همبستگی ۰ است، نمایه شده است.



نمودارهای پراکندگی بسیار سودمند هستند زیرا نشانه‌های دیداری ارزشمندی از ارتباط بین متغیرها به دست می‌دهند. این نشانه‌ها بیانگر و روشن‌کننده ماهیت ارتباط بین متغیرهاست.

برای پرسش‌های بیشتر، پرسش‌های ۳۶، ۴۲ و ۴۴ را مطالعه کنید.

پرسش ۴۴: چگونه می‌توانم ضریب همبستگی پیرسون را محاسبه کنم؟

ضریب همبستگی از طریق فرمول ضریب همبستگی زیر محاسبه می‌شود:

$$r_{xy} = \frac{n\Sigma XY - \Sigma X\Sigma Y}{\sqrt{[n\Sigma X^{۲} - (\Sigma X)^{۲}][n\Sigma Y^{۲} - (\Sigma Y)^{۲}]}}$$

در این فرمول، r_{xy} شاخص ضریب همبستگی، n شاخص حجم نمونه؛ X نمرات افراد در متغیر X، Y نمرات افراد در متغیر Y، XY حاصل‌ضرب هر X در Y متناظر خودش؛ $X^{۲}$ مجذور هر X و Y^2مجذور هر Y است. برای مثال، ضریب همبستگی داده‌های نمایه شده در جدول زیر را محاسبه کنید.

XY	Y^2	X^2	Y	X	
۳	۹	۱	۳	۱	
۲۸	۴۹	۱۶	۷	۴	
۴۰	۶۴	۲۵	۸	۵	
۲۷	۸۱	۹	۹	۳	
۳۰	۳۶	۲۵	۶	۵	
۵۴	۸۱	۳۶	۹	۶	
۵۶	۶۴	۴۹	۸	۷	
۴۰	۶۴	۲۵	۸	۵	
۵۴	۸۱	۳۶	۹	۶	
۲۸	۴۹	۱۶	۷	۴	
۳۶۰	۵۷۸	۲۳۸	۷۴	۴۶	**جمع**

با جایگزینی اعداد در فرمول ضریب همبستگی خواهیم داشت:

$$r = \frac{۱۰(۳۶۰) - (۴۶)(۷۴)}{\sqrt{[۱۰(۲۳۸) - (۴۶)^{۲}][۱۰(۵۷۸) - (۷۴)^{۲}]}} = ۰/۶۹$$

همان‌گونه که ملاحظه می‌شود r_{xy} در این فرمول برابر ۰/۶۹ به‌دست‌آمده است که مثبت و مستقیم است.

برای پرسش‌های بیشتر، پرسش‌های ۴۰، ۴۵ و ۴۶ را مطالعه کنید.

پرسش ۴۵: چگونه می‌توانم از طریق اِکسل ضریب همبستگی را محاسبه کنم؟

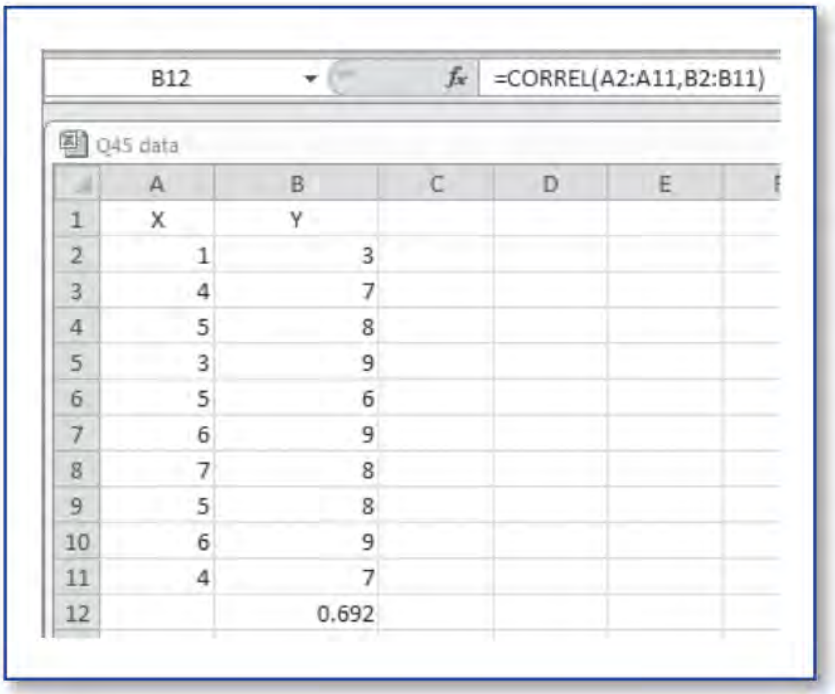
برای محاسبه ضریب همبستگی با کمک اِکسل می‌توانید از تابع CORREL= استفاده کنید. می‌توانید از داده‌های پرسش ۴۴ استفاده کنید. برای این منظور از گام‌های زیر پیروی کنید:

۱. تابع CORREL= را به خانه‌ای که قصد دارید ضریب همبستگی در آن نشان داده شود وارد کنید؛

۲. سپس با فعال شدن تابع، دو ستون داده‌ای که قصد محاسبه ضریب همبستگی‌شان را دارید هایلایت (برجسته) کنید؛

برای مثال فرمول زیر برای داده‌های جدول زیر به کار ببرده می‌شود. این فرمول‌ها به‌صورت پیش‌فرض در اِکسل وجود دارند و تنها با واردکردن "CORREL=" در هر خانه‌ای از اِکسل به‌راحتی مشهود است که چگونه می‌توان بقیه مراحل را انجام داد.

= CORREL (A2:A11; B2:B11)



	A	B	C	D	E
1	X	Y			
2	1	3			
3	4	7			
4	5	8			
5	3	9			
6	5	6			
7	6	9			
8	7	8			
9	5	8			
10	6	9			
11	4	7			
12		0.692			

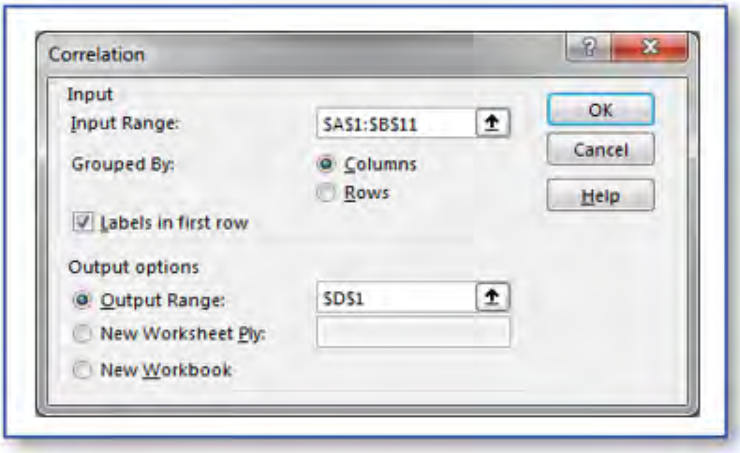
۳. اکنون با فشردن کلید Enter محاسبه انجام شده است و ضریب همبستگی که برابر ۰/۶۹ است در خانه‌ای که فرمول در آن وارد شده نمایه می‌شود.

همچنین می‌توانید از روش دیگری برای محاسبه ضریب همبستگی استفاده کنید. برای این کار می‌توانید از Data Analysis استفاده نمایید. بنابراین از گام‌های زیر پیروی کنید:

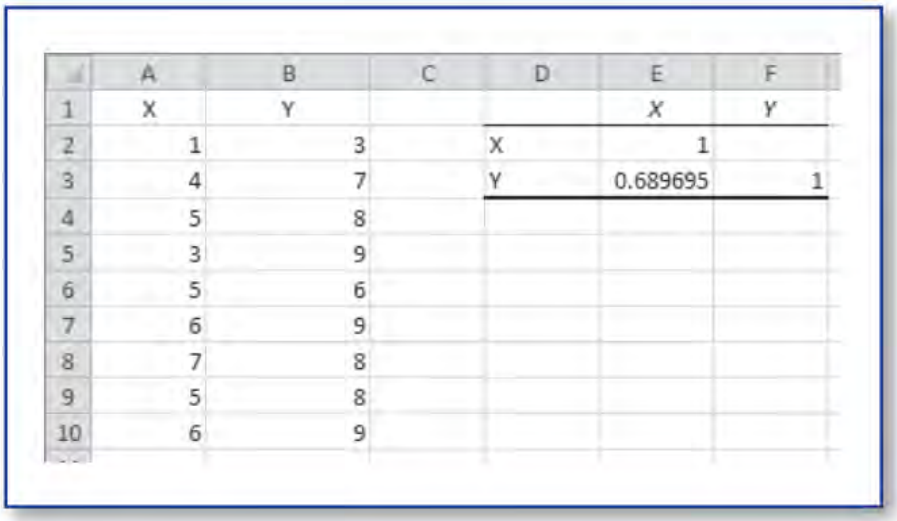
۱. بر Data و سپس Data Analysis کلیک کنید تا پنجره Data Analysis باز شود؛

۲. بر گزینه Correlation دو بار کلیک کنید؛

۳. پنجره Correlation را تکمیل کنید. برای این کار دامنه درونداد و دامنه برونداد و خانه‌ای که مایلید جدول همبستگی در آن نمایه شود را مشخص نمایید؛



۴. اکنون با فشردن کلید Enter محاسبه انجام شده و جدول همبستگی نمایه می‌شود.



	A	B	C	D	E	F
1	X	Y			X	Y
2	1	3		X	1	
3	4	7		Y	0.689695	1
4	5	8				
5	3	9				
6	5	6				
7	6	9				
8	7	8				
9	5	8				
10	6	9				

برای پرسش‌های بیشتر، پرسش‌های ۴۰، ۴۴ و ۴۶ را مطالعه کنید.

پرسش ۴۶: مستقیم‌ترین مسیر برای تفسیر ارزش‌های یک ضریب همبستگی چیست؟

ضریب همبستگی یک ارزش عددی است که نشان‌دهنده درجه ارتباط بین دو متغیر است. ارزش هر ضریب همبستگی از ۱- تا ۱+ است. اما این ضرایب چگونه تفسیر می‌شوند؟ و منظور این ضرایب و علامت‌های مثبت و منفی چیست؟.

بهترین و ساده‌ترین مسیر برای تفسیر ضرایب همبستگی کاربرد روش تحلیل دیداری است. بر این اساس، میزان شدت۰ تا ۰/۲ نشان‌دهنده عدم همبستگی یا همبستگی بسیار ضعیف؛ ۰/۲ تا ۰/۴ نشان‌دهنده همبستگی ضعیف؛ ۰/۴ تا ۰/۶ نشان‌دهنده همبستگی متوسط؛ ۰/۶ تا ۰/۸ نشان‌دهنده همبستگی قوی؛ ۰/۸ تا ۱ نشان‌دهنده همبستگی بسیار قوی است. هرچند که کاربرد روش تحلیل دیداری بالا دقیق‌ترین راه برای تفسیر ضرایب همبستگی نیست، امّا یقیناً یک منطق مناسب برای تفسیر شدت و درجه همبستگی بین چند متغیر را فراهم می‌کند. برای روش‌های دقیق‌تر می‌توانید به پرسش ۴۸ که به بررسی ضریب تعیین[1] پرداخته است مراجعه کنید.

برای پرسش‌های بیشتر، پرسش‌های ۴۰، ۴۲ و ۴۷ را مطالعه کنید.

1 . coefficient of determination

پرسش ۴۷: ضریب تعیین چیست و چگونه محاسبه می‌شود؟

هرچند محاسبه مقدار ضریب همبستگی می‌تواند یک ارزیابی کلی از شدّت رابطه بین دو متغیر فراهم کند، امّا تا ارزیابی دقیق آن راه بسیار است. ضریب تعیین[1] که با نماد r_{xy}^2 مشخص می‌شود، مجذور ضریب همبستگی است و حاصل میزان پراکندگی در یک متغیر بر اساس تغییرات ایجادشده در یک متغیر دیگر است. ضریب همبستگی، فقط اندازه همبستگی بین دو متغیر را نشان می‌دهد و یا نشان می‌دهد که بین دو متغیر همبستگی بالا یا پایین وجود دارد. امّا درباره درصد تبیین متغیر y بر اساس متغیر x اطلاعاتی به ما نمی‌دهد. ضریب تعیین، واریانس مشترک بین دو متغیر است و با فرمول زیر محاسبه می‌شود:

$$V= (r_{xy})^2 \times ۱۰۰$$

اجازه دهید به مثال همبستگی بین قد و وزن در دانش آموزان کلاس ششم که ۰/۸۵ به دست آمد نگاهی بیندازیم. بنابراین همبستگی گشتاوری پیرسون یا r_{xy} برابر ۰/۸۵ شده است. ضریب تعیین یا r_{xy}^2 برابر ۰/۷۲۲۵ خواهد بود که نشان می‌دهد ۷۲/۲۵ درصد از پراکندگی یا واریانس قد را می‌توان بر اساس پراکندگی یا واریانس وزن دانش‌آموزان کلاس ششم تعیین کرد. پراکندگی یا واریانس نیز به این معناست که قد یا وزن دانش آموزان کلاس ششم چقدر از یکدیگر متفاوت است. یک موضوع مهم که می‌بایست درباره استفاده از ضریب تعیین در ذهن داشته باشیم این است که هر چه ضریب همبستگی قوی‌تر باشد، واریانس بیشتری بر اساس متغیر دیگر محاسبه می‌شود. برای مثال، اگر ضریب همبستگی بین دو متغیر ۰/۴ باشد، ۱۶ درصد واریانس متغیر اول بر اساس متغیر دوم تعیین می‌گردد. اگر ضریب همبستگی بین دو متغیر ۰/۶ باشد، ۳۶ درصد واریانس متغیر اول بر اساس متغیر دوم تعیین می‌گردد. اگر همبستگی بین دو متغیر صفر باشد، دو متغیر چیزی برای به اشتراک هم‌زمان گذاشتن ندارند. هر چه همبستگی بین آنها بیشتر باشد، به این معنی است که پراکندگی مشترک بین آنها بطور هم‌زمان افزایش یافته و تغییرپذیری هم‌زمان بیشتری خواهند داشت.

برای پرسش‌های بیشتر، پرسش‌های ۴۴، ۴۵ و ۴۶ را مطالعه کنید.

1 . coefficient of determination

پرسش ۴۸: مؤلفه‌های مهمی که درباره فهم و کاربست ضرایب همبستگی باید به یاد داشته باشم، کدام‌اند؟

چند نکته مهم کلیدی در فهم و کاربست ضرایب همبستگی در مطالعه آمار وجود دارد.

۱. ضرایب همبستگی میزان تغییرات یا پراکندگی‌هایی است که بین دو متغیر وجود دارد. هر چه میزان تغییرپذیری و مشترک بودن تغییرات بین دو متغیر بیشتر باشد، رابطه بین دو متغیر قوی‌تر خواهد بود. اگر یک متغیر تغییرپذیری نداشته باشد و بین افراد یک گروه هیچ تفاوتی ازنظر متغیر مورد نظر وجود نداشته باشد، همبستگی صفر خواهد بود. در واقع در چنین وضعیتی، دو متغیر هیچ اشتراکی نخواهند داشت.

۲. ضریب همبستگی گشتاوری پیرسون با حرف کوچک r نشان داده می‌شود و با اندیس حرف اول دومتغیری که همبستگی‌شان را نشان می‌دهد نظیر r_{xy} برای متغیرهای X و Y نمایه می‌شود.

۳. دو نوع ضریب همبستگی وجود دارد. ضریب همبستگی مستقیم یا مثبت که نشان‌دهنده رابطه هم‌جهت دو متغیر است. ضریب همبستگی غیرمستقیم یا منفی که نشان‌دهنده رابطه متضاد دو متغیر است.

۴. میزان ضریب همبستگی از ۱- تا ۱+ متغیر است. ۱- همبستگی منفی کامل و ۱+ همبستگی مثبت کامل است.

۵. مقدار ضریب تعیین که درواقع مجذور ضریب همبستگی است از ۰ تا ۱۰۰ درصد متغیر است.

۶. مقدار قدر مطلق[1] ضریب همبستگی نشان‌دهنده میزان ارتباط بین دو متغیر است. درواقع، ضریب همبستگی ۰/۷۰- قوی‌تر از ضریب همبستگی ۰/۵۰+ است.

۷. همبستگی‌های منفی یا غیرمستقیم که علامت منفی دارند از همبستگی‌های مثبت یا مستقیم که علامت مثبت دارند، نه بهترند و نه بدتر، بلکه فقط متفاوت‌اند.

۸. همبستگی همیشه بیانگر وضعیتی است که حداقل دو سری داده یا دو مجموعه داده وجود داشته باشد.

۹. یک ضریب هیچ‌وقت چیزی درباره رابطه علّی[2] به ما نمی‌گوید، بلکه بیانگر شدت ارتباط بین حداقل دو متغیر است.

۱۰. نمودارهای پراکندگی[3] بهترین گزینه برای نشان دادن ارتباط دیداری بین نقاطی هستند که با یکدیگر ارتباط دارند.

برای پرسش‌های بیشتر، پرسش‌های ۴۰، ۴۶ و ۴۷ را مطالعه کنید.

---

1 . absolute value
2 . causal relationship
3 . Scatter charts

پرسش ۴۹: چگونه می‌توانم از یک ماتریس برای نمایش چند ضریب همبستگی استفاده کنم؟

یک ضریب همبستگی همیشه بیانگر یک عدد است که میزان شدت ارتباط بین دو متغیر را نشان می‌دهد. اگر بیش از دو متغیر باهم همبسته باشند، استفاده از ماتریس همبستگی[1] می‌تواند برای فهم ساده‌تر ارتباط بین متغیرها بسیار کمک‌کننده باشد. برای مثال، به داده‌های جدول زیر توجه کنید. در این جدول چهار متغیر وجود دارد که می‌خواهیم ضرایب همبستگی بین آنها را محاسبه کنیم. پرانتزهای مقابل هر متغیر نیز دامنه نمرات را نشان می‌دهد.

پایه تحصیلی (۱–۴)	نمره کلاسی (۰–۴)	ساعات مطالعه (۰–۴۰)	نمره خستگی (۱–۱۰)
۱	۳/۳	۲۳	۱۶
۲	۳/۲	۲۴	۲۶
۲	۲/۸	۱۲	۲۸
۳	۳	۱۲	۱۱
۱	۱/۸	۱۸	۵۷
۲	۱/۸	۷	۶۹
۳	۲/۸	۲۲	۵۸
۴	۲/۳	۲۵	۴۴
۳	۴	۳۰	۴
۴	۲/۷	۹	۲۹

جدول زیر ماتریس همبستگی را نشان می‌دهد که در آن همبستگی بین چهار متغیر را به‌صورت زوجی نشان می‌دهد. این چهار متغیر در مقابل هم به‌صورت زوجی ده ضریب همبستگی را نشان می‌دهند که به‌آسانی قابل‌فهم هستند.

	پایه تحصیلی	نمره کلاسی	ساعات مطالعه	نمره خستگی
پایه تحصیلی	۱	۰/۱۳	۰	−۰/۳۳
نمره کلاسی		۱	۰/۵۶	−۰/۸۶
ساعات مطالعه			۱	−۰/۳۳
نمره خستگی				۱

1 . correlation matrix

همان‌گونه که ملاحظه می‌شود هر خانه از جدول بالا همبستگی بین دو متغیر را نشان می‌دهد. برای مثال، همبستگی بین معدل تحصیلی و تعداد ساعات مطالعه در هفته برابر ۰/۵۶ است که نشان می‌دهد هر چه دانش آموزان ساعات بیشتری را در هفته به مطالعه بپردازند، معدل تحصیلی‌شان بیشتر خواهد بود. هر یک از این ضرایب همبستگی بیانگر شدت ارتباط بین دو متغیر است. ضریب همبستگی در خانه‌های قطری[1] ۱ شده است زیرا همبستگی بین یک متغیر با خودش همیشه ۱ یا کامل است.

ازآنجاکه ضریب همبستگی دلیلی بر علّیت نیست و می‌تواند به‌صورت r_{xy} یا r_{yx} نشان داده شود، خانه‌های زیر قطرِ جدول خالی است. در واقع، ضریب همبستگی بین پایه تحصیلی و خستگی تحصیلی، دقیقاً برابر با ضریب همبستگی بین خستگی تحصیلی و پایه تحصیلی است. بنابراین خانه‌های خالی به دلیل جلوگیری از تکرار ضرایب همبستگی خالی است.

برای پرسش‌های بیشتر، پرسش‌های ۴۴، ۴۶ و ۴۷ را مطالعه کنید.

1 . diagonal

پرسش ۵۰: سایر شاخص‌های همبستگی کدام ها هستند و چگونه به کار می‌روند؟

ضریب گشتاوری پیرسون، یکی از روش‌های محاسبه همبستگی بین دو متغیر با ماهیت پیوسته است. در پرسش‌های قبلی گفته شد که متغیر پیوسته به متغیری گفته می‌شود که هر ارزشی عددی را بتوان به آن نسبت داد. مثلاً نمره افراد در آزمون املاء می‌تواند یک مورد از متغیر پیوسته باشد. امّا هنگامی‌که متغیرها پیوسته نباشند و هر ارزشی را نتوان به آن‌ها نسبت داد، باید از ضرایب همبستگی دیگری استفاده کرد. برخی متغیرها نظیر جنسیت (زن و مرد) و گرایش‌های سیاسی (جمهوری‌خواه، دموکرات) متغیرهای طبقه‌ای هستند. نمودار زیر انواع ضرایب همبستگی را در انواع متغیرهای مختلف نشان می‌دهد. در این جدول، متغیرهای طبقه‌ای به اسمی[1]، متغیرهای رتبه‌بندی شده به رتبه‌ای[2]، و متغیرهای پیوسته به فاصله‌ای[3] نام گرفته‌اند.

متغیر X	متغیر Y	نوع همبستگی	مثال: رابطه بین ...
متغیر اسمی (ترجیح سیاسی)	متغیر اسمی (جنسیت)	ضریب فای[4]	رابطه بین جنسیت و ترجیح سیاسی (جمهوری خواه و دموکرات)
متغیر اسمی (طبقه اجتماعی)	متغیر رتبه ای (رتبه کلاسی)	ضریب دو رشته ای رتبه ای	رابطه بین طبقه اجتماعی و رتبه کلاسی
متغیر اسمی (جنسیت)	متغیر فاصله ای (معدل)	ضریب دو رشته ای نقطه ای	رابطه بین جنسیت و معدل تحصیلی
متغیر رتبه ای (رتبه قد)	متغیر رتبه ای (رتبه وزن)	ضریب رتبه ای اسپیرمن	رابطه بین قد و وزن
متغیر فاصله ای (قد)	متغیر فاصله ای (وزن)	ضریب پیرسون	رابطه بین قد و وزن

محاسبه انواع مختلف ضرایب همبستگی فراتر از حیطه این کتاب است، امّا چندین کتاب بسیار سودمند و چندین سایت اینترنتی و بسته‌های نرم‌افزاری رایانه‌ای برای یادگیری بهتر و محاسبات و تفسیرهای بهتر وجود دارد که می‌توانید به آن‌ها رجوع کنید.

برای پرسش‌های بیشتر، پرسش‌های ۴۱، ۴۶ و ۴۸ را مطالعه کنید.

1 . nominal
2 . ordinal
3 . interval

بخش ۶: درک اندازه‌گیری و اهمیت آن

پرسش ۵۱: چرا اندازه‌گیری، یک موضوع مهم در میان متخصصین آمار است؟

در ابتدای بحث درباره اندازه گیری، مجموعه مقدماتی از ابزارهایی در اختیار داریم که می توانند در ایجاد فهم بهتری درباره جمع آوری، طرح ریزی و انجام بخشی از پژوهش به کار برده شوند. اندازه گیری در حقیقت نسبت دادن ارزشها یا برچسب ها به پیامدها یا نتایج است. برای نمونه می‌توان به تعداد پاسخ‌های درست در یک آزمون پیشرفت تحصیلی (نظیر نمره ۷۲)، تعداد و نوع اتومبیل‌های فروخته‌شده در ماه گذشته (نظیر ۸۲ اتومبیل هیوندا)، یا رنگ مورد علاقه در یک نقاشی (نظیر کد ۱ برای آبی پررنگ و کد ۲ برای رنگ آبی کم‌رنگ) اشاره کرد. تمامی این اندازه‌گیری‌ها بیانگر وقوع یک رویدادی هستند که در بافتی که پژوهش و فرضیه در آن شکل‌گرفته است، به وقوع پیوسته‌اند. در سه مثال ذکرشده، ۷۲ به معنی پیشرفت تحصیلی بالا در بین دانش آموزان مورد مطالعه،۸۲ به معنی فروش بالای اتومبیل هیوندا در ماه گذشته است. ۱ و ۲ نشان‌دهنده تمایل نقاش به رنگ خاص است. چندین دلیل وجود دارد که مطالعه اندازه‌گیری در چارچوب علم آمار اهمیت دارد.

اول اینکه همه آماردان‌ها با اندازه‌گیری نتایج مطالعات خود سروکار دارند. بنابراین فهم فرایند اندازه‌گیری و کاربردهای متعدد و متفاوت آن از اصول مقدماتی علم آمار است. دوم اینکه برای اندازه‌گیری یک پدیده، در ابتدا می‌بایست ماهیت پدیده مورد اندازه‌گیری شناسایی و درک شود. چندین روش برای ارزیابی و سنجش یک نتیجه وجود دارد. مثلاً می‌توان پرسید که کدام‌یک از دانش آموزان کلاس بلندقدتر از بقیه است؟ یا می‌توان پرسید که بلندقدترین دانش‌آموز کلاس چند سانتی‌متر قد دارد؟ یا می‌توان پرسید آیا همانند همکلاسی‌هایش بلندقد است؟ یا می‌توان پرسید او چقدر بلندقد است؟. سرانجام اینکه، بدون اندازه‌گیری دقیق، فرضیه‌ها به‌صورت دقیق آزمون نمی‌شوند. در حقیقت، اگر ابزار اندازه‌گیری، ناقص باشد، بسیار دشوار خواهد بود که بگوییم فرضیه مورد آزمون، تبیین منطقی و مستدلی برای آنچه مشاهده می‌کنیم، ارائه می‌دهد. ممکن است روشی که با آن نتایج پژوهش اندازه‌گیری شده‌اند ناقص و غیردقیق باشد که این مسئله باعث می‌شود نتایج حاصل از آزمون فرضیه به‌صورت دقیق درک نشود و صرفاً پاسخ‌های کلی و مبهمی به سؤالات مطرح‌شده، داده می‌شود. درواقع، اندازه‌گیری دقیق شامل کاربست ابزارهای موثق و مورد اعتماد است که نتایج باثبات و پایداری به دست می‌دهند. هرچند کاربست ابزارهای دقیق اندازه‌گیری، به‌تنهایی پاسخی به سؤالات و فرضیات ما نیست، امّا کاربست این ابزارهای دقیق اندازه‌گیری، اولین گام در مسیر پاسخ به سؤالات و فرضیات محقق است.

برای پرسش‌های بیشتر، پرسش‌های ۵۲، ۵۳ و ۵۶ را مطالعه کنید.

پرسش ۵۲: سطوح اندازه‌گیری کدام‌اند و چرا اهمیت دارند؟

سطوح اندازه‌گیری[۱] یا مقیاس‌های اندازه‌گیری[۲]، سطوح ویژه‌ای هستند که بر اساس آنها نتیجه یک یافته آماری ارزیابی می‌شود. چهار مقیاس یا سطوح اندازه‌گیری بانام‌های اسمی، رتبه‌ای، فاصله‌ای و نسبی یا نسبتی وجود دارد که هرکدام ویژگی‌های خاص خود را دارند. این مقیاس‌ها یا سطوح اندازه‌گیری[۳] توسط استیونس[۴] که از روانشناسان معروف آمریکایی و استاد دانشگاه هاروارد[۵] بود معرفی شدند.

ابتدایی‌ترین مقیاس اندازه‌گیری، مقیاس اسمی است. در این مقیاس افراد یا اشیاء بر اساس یک ملاک معین در طبقه‌ها که کیفی هستند و نه کمّی، جایگزین می‌شوند. در این مقیاس، اندازه-گیرنده باید بتواند طبقه‌ها را از یکدیگر تشخیص دهد و ملاکی را که بر اساس آن افراد یا اشیاء را در طبقه‌های مختلف جایگزین می‌کند، بشناسد. ویژگی اصلی سطح اندازه‌گیری اسمی این است که داده‌های مورد بررسی در این سطح می‌بایست طبقه‌ای[۶] باشند. نظیر گرایش‌های سیاسی (گروه اول، گروه دوم و گروه سوم)؛ رنگ مو (قهوه‌ای، سفید، سیاه)؛ جنسیت (زن و مرد). تمامی این متغیرهای طبقه‌ای یا اسمی هستند. این سطح اندازه‌گیری، کم‌دقت‌ترین نوع سطوح اندازه‌گیری به شمار می‌رود.

دومین مقیاس اندازه‌گیری، ترتیبی است که کلیه ویژگی‌های مقیاس اسمی را دارد. ویژگی اصلی سطح اندازه‌گیری رتبه‌ای این است که داده‌های مورد بررسی در این سطح، می‌بایست ترتیبی[۷] باشند. نظیر رتبه کلاسی، رتبه ورزشکاران در مسابقه دوومیدانی، رتبه بهترین تا بدترین تکلیف کلاسی. تمامی این مثال‌ها از نوع متغیرهای ماهیتاً رتبه‌ای هستند.

مقیاس فاصله‌ای دارای کلیهٔ ویژگی‌های مقیاس‌های اسمی و ترتیبی است و علاوه بر آن‌ها، در این مقیاس فاصله هر صفت تا مبدأ آن نیز مشخص است. در این مقیاس نه‌تنها ترتیب اشیاء یا صفت‌های مورد اندازه‌گیری مشخص است، ویژگی اصلی سطح اندازه‌گیری فاصله‌ای این است که داده‌های مورد بررسی در این سطح، می‌بایست داده‌های پیوسته‌ای باشند که فاصله بین ارزش‌هایشان برابر باشد. نظیر نمره آزمون املاء که از ۰ تا ۲۰ نمره می‌گیرد و نمرات بین ۱ و ۲ یک واحد است و این یک واحد در کل دامنه ۰ تا ۲۰ یکسان است. امّا چنین وضعیتی در مقیاس رتبه‌ای رعایت نشده است زیرا رتبه اول تا رتبه دوم، دقیقاً برابر رتبه دوم تا رتبه سوم نیست. در مقیاس فاصله‌ای تمام اعمال ریاضی کاربرد دارد به این معنا که عدد ۲۰ دقیقاً دو برابر عدد ۱۰ است و عدد ۸ دقیقاً نصف عدد ۱۶ است، درحالی‌که این اعمال ریاضی در مقیاس رتبه‌ای کاربرد ندارند.

---

1 . Levels of measurement
2 . scales of measurement
3 . levels of measurement
4 . Stanley Smith Stevens (November 4, 1906ت January 18, 1973)
5 . The Harvard University
6 . categorical
7 . rank

مقیاس نسبی دارای کلیهٔ ویژگی‌های مقیاس‌های فاصله‌ای، ترتیبی و اسمی است. این مقیاس بالاترین سطح اندازه‌گیری است و در آن صفر واقعی وجود دارد. در این مقیاس، برای مقایسه دو ارزش یا دو واحد می‌توان از نسبت استفاده کرد. متر که برای اندازه‌گیری طول به‌کاربرده می‌شود و دارای مبدأ صفر است، یک مقیاس نسبی است. مقیاس نسبتی، سطحی از اندازه‌گیری است که در آن صفر مطلق وجود دارد به این معنا که صفر در یک پدیده یعنی آن پدیده وجود ندارد. هرچند در علوم زیست‌شناسی و فیزیک وضعیت‌هایی وجود دارد که صفر مطلق معنا پیدا می‌کند نظیر نبود نور یا دمای صفر درجه یا سن نوزاد تازه متولدشده، امّا در علوم رفتاری و اجتماعی نظیر علم روانشناسی چنین موقعیت‌های وجود خارجی نخواهند داشت. برای مثال بسیار خیال‌پردازانه است که در یک شخص هیچ‌گونه استرس، اضطراب، افسردگی، خشم، هوش، عزت‌نفس و ... وجود نداشته باشد.

در ادامه چند نکته مهم درباره سطوح اندازه‌گیری که می‌بایست در یاد داشته باشید:

۱. سطوح اندازه‌گیری به ما کمک می‌کنند که بدانیم چگونه نتایج را تحلیل کنیم؛

۲. سطوح اندازه‌گیری از ترتیب اسمی، رتبه‌ای، فاصله‌ای و نسبتی برخوردارند؛

۳. سطح اندازه‌گیری نسبتی از همه سطوح دیگر دقیق‌تر است؛

۴. سطح اندازه‌گیری نسبتی، تمامی سه سطح اندازه‌گیری اسمی، رتبه‌ای، فاصله‌ای را در برگرفته است. برای مثال، شما می‌توانید قد افراد یک گروه را به چند صورت بررسی کنید. اگر برخی افراد بر اساس قد به گروه A و برخی به گروه B اسم‌گذاری شوند، از سطح اندازه‌گیری اسمی استفاده کرده‌اید. اگر به بررسی گروه قدبلند و گروه قدبلندتر هستید، از سطح اندازه‌گیری رتبه‌ای استفاده کرده‌اید. اگر میانگین دو گروه A و B را ازنظر قد محاسبه می‌کنید، از سطح اندازه‌گیری نسبتی استفاده کرده‌اید. درواقع هنگامی‌که از سطح اندازه‌گیری نسبتی استفاده می‌کنید، می‌دانید به‌راحتی دریابید که میانگین کدام گروه کمتر و یا بیشتر است، کدام گروه قدبلند و قدبلندتر است و اینکه هر یک از افراد به کدام گروه تعلق می‌گیرند.

برای پرسش‌های بیشتر، پرسش‌های ۵۱، ۵۳ و ۵۶ را مطالعه کنید.

پرسش ۵۳: پایایی چیست و چه کاربردی دارد؟

پایایی[1]، کیفیت یک ابزار سنجش در پایا بودن نتایجی است که به دست می‌دهد. پایایی از چند روش مختلف محاسبه می‌شود. در کنار اعتبار (روایی[2])، پایایی یکی از دو کیفیت‌های بسیار مهم یک آزمون است. اگر آزمونی قابلیت اعتماد یا پایایی نداشته باشد، قادر نخواهد بود که یک ویژگی، خصیصه یا سطح عملکرد را به‌صورت پایدار و پیوسته ارزیابی کند و به همین دلیل ارزشش به‌عنوان یک آزمون زیر سؤال می‌رود.

درک پایایی، داشتن اطلاعات درباره سه مؤلفه درباره یک آزمون را در برمی‌گیرد. این سه مؤلفه می‌توانند را می‌توان با ملاحظه معادله زیر درک نمود:

$$\text{Observed score} = \text{True score} + \text{Error score}$$

نمره خطا + نمره درست = نمره مشاهده‌شده

نمره مشاهده شده عبارت است نمره‌ای که فرد در آزمون یا هر ابزار سنجش کسب می‌کند. نظیر نمره ۸۹ در آزمون هفتگی درس علوم، یا اخذ نمره ۴ از یک پرسشنامه پنج سؤالی درباره شادکامی. این نمره کسب‌شده، نمره واقعی[3] است. یعنی همان نمره‌ای که توسط استاد بر اساس تعداد پاسخ‌های صحیح شما در بالای برگه امتحانی شما نوشته می‌شود.

نمره درست[4]، نشان‌دهنده سطح واقعی عملکرد[5] است که به‌صورت مستقیم قابل‌مشاهده نیست. درواقع نمره واقعی نشان‌دهنده سطح واقعی عملکرد بدون بررسی سایر عوامل اثرگذار است.

نمره خطا نشانگر تمامی رویدادهایی است که به‌عنوان تفاوت بین نمره مشاهده‌شده و نمره واقعی در نظر گرفته می‌شوند. نمره خطا می‌تواند شامل خطاهایی باشند که منبع آن‌ها خود افراد مورد بررسی نظیر خستگی بود و مطالعه نکردن، و یا موقعیت‌های اخذ آزمون نظیر نور کم یا اتاق گرم باشد.

اگر یک آزمون کاملاً پایایی داشته باشد، نمره خطا وجود در کار نخواهد بود و نمره مشاهده‌شده دقیقاً برابر نمره واقعی خواهد بود. امّا واقع‌بینانه این است که عوامل متعددی وجود دارند که بر عملکرد افراد در یک آزمون تأثیر می‌گذارند که همه آن‌ها از مصادیق نمره خطا محسوب می‌شوند.

در نهایت معادله پایایی به‌صورت زیر خواهد بود:

$$\text{Reliability (پایایی)} = \frac{\text{True score (درست نمره)}}{\text{True score (درست نمره)} + \text{Error score(خطا نمره)}}$$

همان‌گونه که در معادله صفحه قبل ملاحظه می‌کنید، مقدار کمتر نمره خطا، پایایی آزمون را افزایش می‌دهد و اگر نمره خطا صفر باشد، که بندرت رخ می‌دهد، پایایی کامل یا ۱۰۰ درصد خواهد بود. در بیشتر موارد، پایایی با استفاده از ضریب همبستگی محاسبه می‌شود به این نحو

1 . relliability
2 . validity
3 . real score
4 . true score
5 . actual level of performance

که ضریب همبستگی بین دو مجموعه نمره از دو آزمون متفاوت که یک چیز را می‌سنجند معیار پایایی واقع می‌شود. محققان دانشگاه آریزونا علاقه‌مند به بررسی پایایی ابزار سنجش دوره کودکی اولیه دورکس[1] بودند. این محققان پایایی این آزمون را از طریق بررسی ضریب همبستگی بین فرم والد و فرم معلّم این آزمون بررسی کردند. آن‌ها دریافتند که ضرایب همبستگی پایایی دارند، امّا اذعان نمودند که این آزمون اعتبار (روایی) مناسبی ندارد و لذا پیشنهاد کردند که مطالعات بیشتری در رابطه با این آزمون و ویژگی‌های روان‌سنجی آن صورت پذیرد.

برای مطالعه بیشتر درباره مثال موجود در این پرسش به مطالعه مرجع زیر بپردازید.

Otilia, C. B., Levine-Donnerstein, D., Marx, R. W., & Yaden, D. B., Jr. (2013). Reliability and validity of the Devereux Early Childhood Assessment (DECA) as a function of parent and teacher ratings. Journal of Psychoeducational Assessment, 31(5), 469-481.

برای پرسش‌های بیشتر، پرسش‌های ۵۴، ۵۵ و ۵۹ را مطالعه کنید.

1 . the Devereux Early Childhood Assessment (DECA)

پرسش ۵۴: انواع مختلف پایایی چیست و کاربرد آن‌ها کجاست ؟

چهار نوع پایایی وجود دارد و هر یک برای خود هدف ویژه‌ای دارند و در هر یک از آن‌ها از انواع ضرایب همبستگی استفاده می‌شود تا یک معیار عددی از شاخص قابلیت اعتماد یا پایایی فراهم شود.

پایایی بازآزمایی[1] به بررسی ثبات و پایداری نتایج یک آزمون در طی یک دوره زمانی می‌پردازد. ثبات به‌دقت اندازه‌گیری حاصل با گذشت زمان اشاره دارد. این روش برای وضعیت‌هایی مناسب است که از یک آزمون در دو یا چند زمان مختلف در یک نمونه استفاده می‌شود.

پایایی فرم های موازی[2] به بررسی ثبات و پایداری نتایج یک آزمون، در مواقعی که از دو آزمون شبیه به هم استفاده می‌شود، یا زمانی که برای یک آزمون، دو فرم مشابه به شرکت‌کنندگان داده می‌شود مناسب است. یکی از روش‌های متداول ارزیابی پایایی در اندازه‌گیری‌های روانی- تربیتی استفاده از آزمون‌های همتا است که شباهت زیادی با یکدیگر دارند ولی کاملاً یکسان نیستند.

پایایی همسانی درونی[3] به بررسی این موضوع می‌پردازد که آیا همه سؤالات یک آزمون، به‌صورت پایدار یک بعد یا مؤلفه را می‌سنجند یا نه. درواقع هنگامی‌که برایمان مهم است که بدانیم که آیا یک آزمون یک مؤلفه یا مفهوم را می‌سنجد یا نه؟ استفاده از این روش مناسب است. این روش برای محاسبه هماهنگی درونی ابزار اندازه‌گیری ازجمله پرسشنامه‌ها یا آزمون‌هایی که خصیصه‌های مختلف را اندازه‌گیری می‌کند بکار می‌رود. در این‌گونه ابزارها، پاسخ هر سؤال می‌تواند مقادیر عددی مختلف را اختیار کند. برای محاسبه پایایی همسانی درونی از فرمول آلفای کرونباخ است. مقدار صفر این ضریب نشان‌دهنده عدم قابلیت اعتماد و ۱+ نشان‌دهنده قابلیت اعتماد کامل است.

پایایی بین ارزیاب‌ها[4] به بررسی این موضوع می‌پردازد که آیا تمامی کسانی که به آزمون پاسخ می‌دهند پاسخ‌های همسان ارائه می‌دهند یا خیر؟ هنگامی‌که دو یا چند کاربر یک آزمون را تکمیل می‌کنند، محاسبه درجه توافق آن‌ها در پاسخگویی، دلیلی بر پایایی بین ارزیاب‌ها است.

برای پرسش‌های بیشتر، پرسش‌های ۵۴، ۵۵ و ۵۹ را مطالعه کنید.

1 . Test-retest reliability
2 . Parallel forms reliability
3 . Internal consistency reliability
4 . Interrater reliability

پرسش ۵۵: چگونه می‌توان اعتماد (پایایی) یک آزمون را افزایش می‌یابد ؟

اگر با یاد داشته باشید، فرمول پایایی عبارت بود از:

$$\text{پایایی} = \frac{\text{نمره واقعی}}{\text{نمره خطا} + \text{نمره واقعی}}$$

که در آن هر چه میزان نمره خطا کمتر باشد، ضریب پایایی به نمره واقعی نزدیک می‌شود و درنتیجه ارزیابی کامل‌تر خواهد بود. می‌توانید در معادله زیر ملاحظه کنید که نمره خطا، از دو مؤلفه خطای ویژگی و خطای روش تشکیل شده است.

نمره خطا [۱] = خطای ویژگی + خطای روش

خطای ویژگی[۲] به تفاوت‌های فردی بین تکمیل‌کنندگان یک آزمون گفته می‌شود، نظیر ساعات مطالعه آن‌ها، میزان سلامتی آن‌ها، وضعیت روان‌شناختی عمومی آن‌ها و انگیزش و هیجان آن‌ها. خطای روش[۳] به تفاوت‌هایی گفته می‌شود که به ویژگی‌های خود افراد برنمی‌گردد بلکه به محیط پیرامون اشاره دارد، نظیر ویژگی‌های فیزیکی خود آزمون، مکان برگزاری آزمون و شرایط برگزاری آزمون. بهترین روش برای افزایش پایایی یک آزمون این است که نمره خطا آن آزمون به حداقل برسد و این به این معناست که خطای ویژگی و خطای روش را به حداقل برسانیم. از منظر کاربردی، کاهش واریانس خطای روش بسیار آسان‌تر از کاهش واریانس خطای ویژگی است. برای مثال، تلاش برای یکسان‌سازی شیوه آزمون گیری، خود آزمون و شرایط آزمون بسیار آسان‌تر از تلاش برای یکسان‌سازی اضطراب در آزمون دهندگان است.

چند منبع احتمالی خطای ویژگی عبارت‌اند از: سلامتی شرکت‌کنندگان در آزمون، آمادگی کم، انگیزش پایین، عدم علاقه به آزمون. چند منبع احتمالی خطای رو عبارت‌اند از: دستورالعمل غیرواضح و مبهم، سؤالات ناخوانا و مبهم، سؤالات نادرست، ناآشنایی با فرمت سؤالات، وابسته بودن سؤالات به یکدیگر، پاسخ‌های جایگزین غیرمحتمل.

درمجموع پایایی یک آزمون تحت تأثیر عوامل مختلفی قرار دارد که عبارت است از: ۱. آزمون‌های طولانی‌تر از پایایی بالاتری در مقایسه با آزمون‌های کوتاه برخوردارند. ۲. سؤال‌های مشابه از نظر محتوا و با سطح دشواری متوسط، پایایی را افزایش می‌دهد. ۳. ماهیت متغیر اندازه‌گیری شده نیز ضریب پایایی را تحت تأثیر قرار می‌دهد. لازم به ذکر است که پایایی در تفسیر نتایج مهم بوده و شرط اساسی در انتخاب یک ابزار اندازه‌گیری است.

در ادامه می‌توان افزود، ۱- با افزودن تعداد سؤالات بیشتر به یک آزمون، پایایی آن افزایش می‌یابد؛ با متجانس‌تر و همگون‌تر کردن سؤالات یک آزمون، پایایی آن بیشتر می‌شود؛ ۳- با افزودن تعدادی سؤال با ضریب تمیز زیاد به یک آزمون، پایایی آن بالا می‌رود؛ ۴- با افزودن

---

1 . Error score = Trait error + Method error
2 . Trait error
3 . Method error

تعدادی سؤال با ضریب دشواری متوسط به یک آزمون، پایایی آن افزایش می‌یابد؛ ۵- اگر یک آزمون را در مورد گروهی از افراد که ازنظر توانایی موردسنجش تجانس کمتری با یکدیگر دارند، اجرا کنیم، آزمون پایایی بیشتری نشان خواهد داد؛ ۶- معمولاً آزمون‌های سرعت ضریب پایایی بیشتری را نسبت به آزمون‌های قدرت نشان می‌دهند.

از عوامل غیر فنی مؤثر بر پایایی نیز می‌توان به چند مورد اشاره کرد: ۱ – کیفیت مطلوب تکثیر و ترکیب آزمون: باید مطمئن شد که سؤالات خوانا و روشن نوشته شده و صفحه شلوغ و در هم نیست؛ ۲- دستورالعمل واضح و صریح برای آزمون‌شوندگان؛ ۳-دستورالعمل واضح و صریح برای تصحیح کنندگان؛ ۴- موقعیت کنترل‌شده و مطلوب برای اجرای آزمون.

برای پرسش‌های بیشتر، پرسش‌های ۵۳، ۵۴ و ۵۹ را مطالعه کنید.

پرسش ۵۶: روایی چیست و چه کاربردی دارد؟

روایی، دومین ویژگی روان‌سنجی بسیار مهم آزمون، مقیاس یا ابزاری است که رفتار را اندازه‌گیری می‌کند. روایی ویژگی‌ای است که نشان می‌دهد آیا یک آزمون آنچه را که برای آن ساخته شده است ارزیابی و اندازه‌گیری می‌کند یا خیر. در حقیقت هنگامی‌که نتایج یک ابزار، مقیاس یا پرسشنامه روایی داشته باشد، نتایج آن معنادار و ارزشمند خواهد بود و مقیاس و ابزاری که روایی نداشته باشد، نتایج بی‌معنی، بی‌ارزش و غیرقابل استفاده‌ای به دست می‌دهد. اساساً مفهوم روایی به این سؤال پاسخ می‌دهد که ابزار اندازه‌گیری تا چه حد خصیصه مورد نظر را می‌سنجد. بدون آگاهی از روایی ابزار اندازه‌گیری نمی‌توان به‌دقتِ داده‌های حاصل از آن اطمینان داشت. ابزار اندازه‌گیری ممکن است برای اندازه‌گیری یک ویژگی روایی داشته باشد، درحالی‌که برای سنجش همان خصیصه بر روی جامعه دیگر از هیچ‌گونه روایی برخوردار نباشد.

برای مثال یک آزمون ریاضی ممکن است برای سنجش توانایی ریاضی دانش آموزان پایه پنجم ابتدایی از روایی لازم برخوردار باشد اما برای سنجش توانایی ریاضی دانش آموزان پایه سوم راهنمایی فاقد اعتبار باشد. حضور برخی عوامل خارجی بسیار مهم و حیاتی است. برای مثال اگر یک شاهد بیرونی با دیدن چند نمونه سؤال از یک مقیاس، تشخیص دهد که متغیر خاصی توسط آن مقیاس مورد سنجش واقع می‌شود، با نوع خاصی از روایی روبرو خواهیم شد. یا هنگامی‌که نمره یک فرد در یک آزمون جدید با نمره‌اش در یک آزمون از قبل ساخته شده مشابه باشد، با نوع خاصی از روایی روبرو خواهیم شد. به‌هرحال، ازآنجایی‌که نمی‌توانیم به مفهوم روایی، نمره عددی دقیقی نسبت دهیم، لذا در گزارش این مفهوم از درجه‌بندی ضعیف تا قوی استفاده می‌کنیم. برای مثال، برای ساخت مقیاسی درباره نگرش افراد درباره کاربرد بودجه‌های عمومی در ساخت مجتمع‌های حرفه‌ای ورزشی، محققان ممکن است از نمونه‌ای که شامل معماران، ورزشکاران، طرفداران گروه‌های ورزشی، متخصصان علوم ورزشی و روانشناسان و ... استفاده کنند تا به‌دقت و درستی مؤلفه‌هایی که مقیاس نگرش سنج خود بکار بسته‌اند پی ببرند و متوجه شوند که آیا مقیاس آن‌ها به‌درستی و با دقت تمامی ابعاد این مسئله را به‌صورت دقیق و کامل توصیف کرده است یا خیر.

در همین راستا در یک مطالعه گروهی از محققان به بررسی روایی و پایایی پرسشنامه بهینه‌سازی درمان میگرن از دیدگاه بیماران[1] پرداختند. بدین منظور این محققان از ۳۰۰ بیمار میگرنی خواستند که پرسشنامه‌های مرتبط با بیماری میگرن و پرسشنامه تحت ارزیابی در این مطالعه را تکمیل کنند. این بیماران پرسشنامه‌های از قبل ساخته‌شده و دارای روایی و پایایی مناسب نظیر پرسشنامه تأثیر سردرد[2] را که به بررسی تأثیر سردرد بر کیفیت زندگی می‌پردازد، تکمیل کردند.

1 . the patient-rated Migraine Treatment Optimization Questionnaire (M-TOQ)
2 . the Headache Impact Test

هدف این بود که بررسی شود که چگونه با کمک پرسشنامه‌های موجود قبلی می‌توان به بررسی روایی و پایایی ابزار در حال ساخت پرداخت.

برای مطالعه بیشتر درباره مثال موجود در این پرسش به مطالعه مرجع زیر بپردازید.

Lipton, R. B., Kolodner, K., Bigal, M. E., Valade, D., Lغinez, M. J. A., Pascual, J. . . . Parsons, B. (2009). Validity and reliability of the Migraine-Treatment Optimization Questionnaire, Cephelalgia, 29(7), 751ت759.

برای پرسش‌های بیشتر، پرسش‌های ۵۱، ۵۷ و ۵۸ را مطالعه کنید.

پرسش ۵۷: انواع مختلف اعتبار (روایی) چیست و چگونه محاسبه می‌شوند ؟

سه نوع روایی وجود دارد و همه آن‌ها به صحت این موضوع می‌پردازند که آیا ابزار یا مقیاس یا پرسشنامه طراحی‌شده آن چیزی را می‌سنجد که برای آن طراحی شده است یا خیر.

روایی محتوایی[1] برای بررسی و ارزیابی میزان پوشش موضوع مورد بررسی توسط ابزار سنجش به کار می رود. برای مثال در یک مقیاس پیشرفت، روایی محتوایی تعیین می کند که تا چه میزان تمامی حیطه های مورد بررسی، در آزمون تعبیه شده اند.

روایی ملاکی[2] عبارت است از کارآمدی یک ابزار اندازه‌گیری در پیش‌بینی رفتار یک فرد در موقعیت‌های خاص. برای این منظور عملکرد هر فرد در آزمون با یک ملاک مقایسه می‌شود. به‌عبارت‌دیگر، روایی ملاکی یک ابزار اندازه‌گیری عبارت است از همبستگی بین نمرات آزمون و نمره ملاک. روایی ملاکی بر دو نوع است: ۱. روایی پیش‌بین[3]، ۲. روایی هم‌زمان[4].

الف) روایی پیش‌بین: در اندازه‌گیری‌های روانی-تربیتی، روایی پیش‌بین عبارت است از بررسی رابطه نمره‌های آزمونی که برخی ویژگی‌ها را می‌سنجد و آنچه ادعای پیش‌بینی آن را دارد.

ب) روایی هم‌زمان: روایی هم‌زمان در مواردی به کار می‌رود که داده‌های حاصله از دو اندازه‌گیری در یک زمان در دسترس باشد. در این‌گونه موارد عملکرد در یک آزمون به عملکرد در آزمون دیگر مرتبط می‌گردد.

۳-روایی سازه[5] یک ابزار اندازه‌گیری بیانگر آن است که ابزار اندازه‌گیری تا چه اندازه یک سازه یا خصیصه‌ای را که مبنای نظری دارد، می‌سنجد. در بررسی روایی سازه باید به تدوین فرضیه-هایی درباره مفاهیم اندازه‌گیری شده، آزمودن این فرضیه‌ها و محاسبه همبستگی نتایج با اندازه‌گیری اولیه پرداخت.

برای پرسش‌های بیشتر، پرسش‌های ۵۶، ۵۸ و ۵۹ را مطالعه کنید.

1 . Content validity
2 . Criterion validity
3 . predictive validity
4 . Concurrent validity
5 . Construct validity

پرسش ۵۸: چگونه می‌توان اعتبار (روایی) یک آزمون را افزایش می‌یابد؟

افزایش پایایی یک آزمون، مستقیماً به کاهش خطاهای ویژگی، روش و افزایش تعداد سؤالات آزمون بستگی دارد. افزایش روایی یک آزمون به‌راحتی افزایش پایایی آن نیست. برای افزایش روایی یک آزمون می‌توانید روش‌های را بکار ببرید:

- الف) اگرچه افزایش پایایی الزاماً افزایش روایی را در پی ندارد، امّا پایایی یکی از پیش‌شرط‌های مهم روایی به شمار می‌رود. بنابراین از پایا بودن مقیاس اطمینان داشته باشید. ضریب پایایی کم سبب محدود شدن ضریب روایی بالا می‌شود، به این معنا که هر چه پایایی افزایش یابد، روایی نیز افزایش می‌یابد.
- ب) اطمینان حاصل کنید که آنچه مقیاس، مورد سنجش قرار می‌دهد، آن چیزی باشد که شما قصد سنجش آن را دارید. برای مثال در حیطه آزمون‌های پیشرفت تحصیلی برخی متخصصان به دو بُعد پیشرفت تحصیلی توجه می‌کنند و در ساخت آزمون پیشرفت تحصیلی دو بُعد دانش نظری و مهارت عملی را مدنظر قرار می‌دهند. لذا وقتی یک دانش‌آموز در یک آزمون موفق تلقی می‌شود که هم دانش نظری و هم مهارت عملی داشته باشد. درواقع، آزمون پیشرفت تحصیلی دو بُعدی در نظر گرفته و ساخته می‌شود. چراکه قصد متخصصین ارزیابی دو بُعد از پیشرفت تحصیلی بوده است.
- پ) اگر یک آزمون آن‌گونه که شما در نظر داشتید کارکرد نداشت، می‌توانید با بازبینی سؤالات اطمینان حاصل کنید که آیا سؤالات بر مبنای محتوا، سازه و مهارت‌های مورد نظر شما بوده‌اند یا نه.
- ج) روایی یابی یک آزمون بر اساس نمونه‌های بزرگ‌تر و متنوع‌تر، می‌تواند اطلاعات دقیق‌تر و صحیح‌تری از اینکه آیا سؤالات مقیاس، تمام محتوا و سازه‌ای که قصد دارید ارزیابی‌شان کنید را پوشش داده‌اند یا نه.
- د) با یک پیش‌آزمون دقیق، اطمینان حاصل نمایید که آزمون‌ها، مقیاس‌ها و پرسشنامه‌ها و سایر ابزارهایی که ساخته‌اید، از تمامی کیفیت‌های یک آزمون خوب برخوردارند. این کیفیت‌ها عبارت‌اند از شفاف بودن دستورالعمل آزمون، کاربرد آسان، دسترسی‌پذیری به طیف وسیعی از شرکت‌کنندگان و قابل‌فهم بودن آزمون. تمامی این مؤلفه‌ها می‌توانند در افزایش روایی و پایایی یک آزمون نقش مؤثر داشته باشند.

برای پرسش‌های بیشتر، پرسش‌های ۵۶، ۵۷ و ۵۹ را مطالعه کنید.

پرسش ۵۹: ارتباط بین پایایی و روایی چیست؟

همان‌گونه که تاکنون متوجه شده‌اید، پایایی عبارت است از کیفیتی که نشان‌دهنده پایداری و ثبات نتایج یک آزمون است. اگر یک آزمون پایا باشد، اجرای آن در زمان‌ها و مکان‌های مختلف نتایج یکسانی را به دست خواهد داد. همچنین، روایی عبارت است از کیفیتی که درستی، صحت و اعتبار نتایج یک آزمون است. اگر یک آزمون صحت، درستی و اعتبار داشته باشد، چیزی را می‌سنجد که برای آن ایجادشده است. امّا آنچه نشان می‌دهد یک ابزار تازه ساخته‌شده، کارکرد و عملکرد مناسبی خواهد داشت یا نه، به ارتباط بین روایی و پایایی برمی‌گردد.

بسیار محتمل است که یک مقیاس، پایایی داشته باشد، ولی روایی نداشته باشد. به‌هرحال، شما نمی‌توانید یک مقیاس معتبر داشته باشید مگر اینکه در ابتدا می‌بایست پایا باشد. برای مثال، یک آزمون ممکن است بارها و بارها نتایج پایا و همسانی به دست دهد، امّا هرگز آن چیزی را که برای آن ساخته‌شده باشد ارزیابی نکند. اما اگر یک آزمون به‌صورت دقیق آن چیزی را بسنجد که برای آن ساخته شده است (روایی)، یقیناً نتایج مشابه و یکسانی در طی سنجش‌های متعدد به دست می‌دهد. چرا که بسیار بعید است که متغیرهای روان‌شناختی هرلحظه در حال تغییر و دگرگونی شدید باشند، مگر در اختلالات روان‌شناختی!!!.

برای مثال، سؤال چندگزینه‌ای زیر را در نظر بگیرید:

در طی تابستان ۱۷۷۶، نشست کنگره قاره در کجا برقرار شد؟

۱. واشنگتن دی سی ۲. فیلادلفیا ۳. نیویورک ۴. بوستون

اگر یک آزمون شامل ۵۰ سؤال شبیه سؤال بالا باشد، یقیناً پایایی بازآزمایی این آزمون بسیار مناسب به دست خواهد آمد. امّا اگر این آزمون به‌عنوان آزمون درس روانشناسی مقدماتی و یا درس مطالعات بین‌المللی از دانش آموزان گرفته شود، یقیناً روایی آن صفر خواهد بود زیرا به‌هیچ‌عنوان برای ارزیابی روانشناسی مقدماتی یا درس مطالعات بین‌المللی تدوین نشده است. به‌بیان‌دیگر، اگر این آزمون برای سنجش دانش آموزان در درس تاریخ آمریکا گرفته شود، روایی و البته پایایی مناسبی خواهد داشت.

یک نکته بسیار مهم این است که هم روایی و هم پایایی، کیفیت‌های ضروری ابزارهای اندازه‌گیری هستند. چراکه حیاتی‌ترین مؤلفه یک پژوهش، فرضیه[1] است که بر اساس ابزارهای دارای روایی و پایایی مورد تائید یا رد قرار می‌گیرند. اگر یک آزمون پایا نباشد، در نتیجه روا نیز نخواهد بود، و بر اساس نتایجی که به دست می‌دهد نمی‌توان به سؤالات و فرضیات یک پژوهش به‌صورت دقیق، منصفانه و بر اساس واقعیت پاسخ داد. در نهایت، به‌عنوان یک نکته فنی، سطح بیشینه روایی با مجذور ریشه ضریب پایایی برابر است. برای مثال، اگر ضریب پایایی یک آزمون استعداد مکانیکی ۰/۸۷ باشد، ضریب روایی آن بزرگ‌تر از ۰/۹۳ نخواهد بود، زیرا ۰/۸۷ حاصل مجذور ۰/۹۳ است. بنابراین ضریب پایایی برابر است با مجذور ریشه ضریب روایی. این نکته به این

1 . hypothesis

معناست که روایی یک آزمون به این بستگی دارد که میزان پایایی آن در چه وضعیتی قرار دارد. بنابراین قبل از هر چیزی می‌بایست به این موضوع مهم توجه کنیم که نتایجی که یک آزمون به دست می‌دهد، پایا و قابل‌اعتماد باشد.

برای پرسش‌های بیشتر، پرسش‌های ۵۱، ۵۳ و ۵۶ را مطالعه کنید.

بخش ۷: درک نقش فرضیه در علم آمار

پرسش ۶۰: فرضیه چیست و چرا در پژوهش علمی اهمیت دارد؟

فرضیه یک حدس علمی است. عمده‌ترین تفاوت بین روش علمی و سایر روش‌های کسب دانش تدوین فرضیه است. در حقیقت در روش علمی، پژوهشگر در ابتدا چنین استدلال می‌کند که اگر این فرضیه درست باشد چه به دست خواهد آمد و سپس به‌طور منظم برای آنکه بتواند فرضیه را رد یا قبول کند، مشاهده را آغاز می‌کند. چندین نوع فرضیه وجود دارد که در پرسش‌های بعدی به آن‌ها خواهیم پرداخت. نقش بسیار مهم فرضیه مطرح کردن یک سؤال پژوهشی نظیر "آیا پسران و دختران دبیرستانی ازنظر پیشرفت کلام و ریاضی باهم متفاوت‌اند" است و به ما اجازه می‌دهد سؤالات و حدس‌های علمی را به‌صورت جملات خبری قابل آزمون مطرح کنیم، نظیر "پسران و دختران دبیرستانی ازنظر پیشرفت کلام و ریاضی باهم متفاوت‌اند". تفاوت بین سؤال و فرضیه بدیهی به نظر می‌رسد، ولی درواقع این‌گونه نیست. فرضیه این امکان را فراهم می‌کند که یک سؤال به‌صورت یک ایده قابل‌اجرا درآید و قابلیت پژوهش پذیری بیابد.

فرضیه عبارت است از حدس یا گمان اندیشمندانه درباره ماهیت، چگونگی و روابط بین پدیده‌ها، اشیاء و متغیرها، که محقق را در تشخیص نزدیک‌ترین و محتمل‌ترین راه برای کشف مجهول کمک می‌نماید؛ بنابراین، فرضیه گمانی است موقتی که درست بودن یا نبودنش باید مورد آزمایش قرار گیرد. فرضیه بر اساس معلومات کلی و شناخت‌های قبلی یا تجارب محقق پدید می‌آید. این شناخت‌ها ممکن است بر اساس تجارب یا مطالعات قبلی باشد، از منابع شفاهی به‌دست‌آمده باشد، یا در جریان مطالعه ادبیات تحقیق حاصل شده باشد.

بهترین روش برای درک اهمیت یک فرضیه این است که گام‌های یک فرایند علمی فهرست شود و به‌صورت مختصر درباره آن‌ها بحث شود.

- پرسیدن یک سؤال پژوهشی؛
- شناسایی عوامل مهم در سؤال پژوهشی؛
- فرمول‌بندی یا بیان یک فرضیه؛
- جمع‌آوری داده‌هایی که به فرضیه ما مرتبط‌اند؛
- آزمون فرضیه؛
- بازبینی یا بررسی فرضیه؛
- بازبینی یا بررسی نظریه؛
- پرسیدن سؤالات جدید؛

هرچند تمامی این هشت مرحله، به‌صورت مستقیم با فرمول‌بندی و آزمون فرضیه سروکار ندارند، امّا ماهیت فرضیه به‌واسطه هر یک از این مراحل تحت تأثیر قرار می‌گیرد.

مراحل۳، ۵ و ۶ در درک فرضیه بسیار مهم‌اند. مرحله ۳، که به فرمول‌بندی یا بیان یک فرضیه اشاره دارد؛ به محقق اجازه می‌دهد که یک عبارت آزمون‌پذیر منطقی را بر اساس پرسش پژوهشی ارائه دهد.

کنجکاوی‌ای که دانشمندان درباره پروژه‌های تحقیقاتی‌شان دارند منجر به هیجان‌زدگی و انتظارهای بسیار ارزشمند درباره اهمیت یافته‌های پژوهشی‌شان دارد. امّا این عوامل و یافته‌های نوین می‌بایست در چارچوب یک فرایند نظام‌مند باشد تا به سؤالاتی که از سؤال اصلی مستخرج شده‌اند پاسخ دقیق و مناسب داده شود.

مرحله ۵، به آزمون فرضیه پرداخته می‌شود. این سؤالات به‌صورت جملات خبری آزمون‌پذیر مطرح می‌شوند، به‌گونه‌ای که متغیرهای مورد بررسی و همچنین انتظاراتی که پژوهشگران از نتایج مطالعه خود دارند، به‌صورت دقیق مشخص شوند. برای مثال، اگر مهارت‌های ریاضی و کلامی در یک گروه از کودکانی که در گروه‌های جنسیتی مختلط حضور دارند مورد بررسی قرار گیرد، متغیرهای جنسیت، پیشرفت کلامی و پیشرفت ریاضی و شاید ارتباط بین آن‌ها مورد بررسی قرار گیرد و برای بررسی آن‌ها، فرضیه‌هایی ارائه شود.

در مرحله ۶، که به بازبینی یا بررسی فرضیه‌ها اختصاص دارد، محقق بر اساس یافته‌های به‌دست‌آمده، به بررسی و بازبینی نظریه‌ای که فرضیه‌ها از آن مستخرج شده است و همچنین به بازبینی خود فرضیه‌ها می‌پردازد. در این چرخه بازنگری و بازبینی، با بازخوردهایی که به فرضیه و نظریه داده می‌شود، یک فرایند خوداصلاحی[1] شکل می‌گیرد که درنهایت منجر به طرح سؤالات، فرضیه‌ها و نظریه‌های دقیق‌تری خواهد شد.

برای پرسش‌های بیشتر، پرسش‌های ۶۱، ۶۴ و ۶۵ را مطالعه کنید.

1 . self-repair

پرسش ۶۱: ویژگی‌های یک فرضیه خوب چیست؟

تفاوت اصلی بین تلاش‌های پژوهشی موفق و ناموفق در یک‌چیز است، و آن فرضیه‌ای است که بسیار مناسب و دقیق نوشته شده باشد. هنگامی فرضیه بسیار دقیق و مناسب نگاشته می‌شود که پژوهشگر مفهوم پروژه در دسترس را به‌خوبی درک، پیشینه پژوهشی را بسیار دقیق بررسی و بر اساس منطق، بین متغیرها رابطه فرض کرده باشد. در ادامه به چند مؤلفه مهم یک فرضیه خوب اشاره کرده‌ایم:

اول اینکه، یک فرضیه خوب به‌صورت یک جمله خبری مطرح می‌شود، نه به‌صورت یک سؤال. برای مثال، در یک پژوهش ممکن است این سؤال مطرح شود که "آیا میزان اقامت دانشجویان سال اول در دانشگاه‌های ایالتی در تعطیلات بین دو ترم به دلیل فقر مالی کم است؟". امّا بامطالعه دقیق پیشینه پژوهشی می‌توان این سؤال را در قالب دقیق‌تر و صحیح‌تر فرضیه مطرح کرد که "میزان اقامت دانشجویان سال اول در دانشگاه‌های ایالتی نسبت به میانگین، به دلیل کمبود منابع مالی ناشی از هزینه بالای تعطیلات بین دو ترم کمتر است". این فرضیه بسیار دقیق‌تر و واضح‌تر از سؤال مطرح‌شده است.

دوم اینکه، یک فرضیه خوب، رابطه بین متغیرها را هدف‌گیری می‌کند. در مثال قبل، متغیرهای زیادی ممکن است در اقامت یا عدم اقامت دانشجویان در تعطیلات بین دو ترم نقش داشته باشند. امّا صرفاً به منابع مالی اشاره شد.

سوم اینکه، یک فرضیه خوب، مبتنی بر بررسی پیشینه پژوهشی و مطالعات انجام‌شده قبلی ارائه می‌شود. این موضوع بسیار مهم است که پیشینه پژوهشی و مطالعات انجام‌شده قبلی بسیار دقیق بررسی شده باشند، زیرا این کار کمک بسیار ارزشمندی به کشف روابط بین متغیرها و اهمیت موضوع مورد بررسی می‌کند.

چهارم اینکه، یک فرضیه خوب، مختصر و مبتنی بر هدف است. یک فرضیه خوب یک مرور ادبیات تحقیقی و یا یک عبارت فلسفی یا منطقی نیست. ازآنجایی‌که یک فرضیه خوب، روابط بین متغیرها را بسیار دقیق و واضح نشان می‌دهد، هر شخص دیگری قادر خواهد بود با خواندن فرضیه، هدف اصلی پژوهش را کاملاً درک کند.

سرانجام اینکه، یک فرضیه خوب آزمون‌پذیر است. هم روابط بین متغیرها و هم خود متغیرها به‌خوبی فهمیده می‌شوند. در مثالی که مطرح شد، هم استطاعت مالی، میزان اقامت در دانشگاه، تعطیلات بین دو ترم قابل‌درک هستند. در حقیقت نکته مهم این است که شیوه ارائه یک فرضیه سبب می‌شود، فرضیه روشن، دقیق، مبتنی بر هدف، مختصر و آزمون‌پذیر باشد و که اگر این‌گونه باشد سبب خواهد شد اطلاعات جدیدی در حوزه مورد بررسی فرضیه کشف شود که همه زمینه‌ساز فرضیات بعدی باشد و در فرضیات بعدی قابل آزمون و کاربرد باشد.

برای پرسش‌های بیشتر، پرسش‌های ۶۰، ۶۳ و ۶۴ را مطالعه کنید.

پرسش ۶۲: یک نمونه و یک جامعه چه تفاوت‌هایی باهم دارند؟

یکی از کارکردهای عمده آمار استنباطی این است که به پژوهشگران اجازه می‌دهد که نتایج حاصل از آزمون فرضیه‌های نمونه مورد مطالعه را به جامعه هدف تعمیم دهند. به این معنا که نتایج و یافته‌های حاصل از بررسی نمونه را به جامعه‌ای که نمونه از آن مستخرج شده تعمیم دهند. یک نمونه به زبان ساده، یک زیرمجموعه از جامعه است. استفاده از نمونه به‌جای جامعه چند مزیت دارد. ازآنجایی‌که نمونه کوچک‌تر است و منابع مالی، زمانی، تجهیزاتی و نیروی انسانی کمتری را صرف می‌کند، سبب شده است که محققان از نمونه‌های مستخرج از جامعه استفاده کنند. برای مثال، فرض کنید که ما علاقه‌مند به بررسی میزان قد دانش آموزان کلاس ششم در مدارس یک ایالت باشیم. بررسی نشان می‌دهد که این ایالت دارای ده هزار دانش‌آموز کلاس ششم است. اگر یک نمونه مناسب و مُعرّف انتخاب کنیم، می‌توانیم با بررسی قد ۱۰۰ تا ۲۰۰ نفر از دانش آموزان کلاس ششم، به‌صورت دقیق میانگین قد همه ده هزار نفر از آن‌ها را محاسبه کنیم. امّا سؤال اینجاست که چقدر اطمینان داریم که شاخص‌های آماری نمونه (آماره) که تعداد اندکی از افراد یک جامعه است، به‌صورت دقیق می‌تواند نشان‌دهنده شاخص‌های آماری جامعه (پارامتر) باشد؟.

اینکه شاخص‌های نمونه (آماره) تا چه حد می‌توانند نشان‌دهنده دقیق شاخص‌های جامعه باشند، خطای نمونه‌گیری[1] نام دارد. خطای نمونه‌گیری به تفاوت بین ارزش‌های نمونه (آماره نمونه) و ارزش‌های جامعه (پارامتر جامعه) گفته می‌شود. بین ویژگی‌های یک نمونه و ویژگی‌های جامعه‌ای که نمونه از آن انتخاب می‌شود، تفاوت وجود دارد. این تفاوت برای نمونه تصادفی قابل برآورد است و به آن خطای نمونه‌گیری گفته می‌شود. خطای نمونه‌گیری تابع اندازه حجم نمونه است هر چه اندازه نمونه کوچک‌تر باشد خطای نمونه‌گیری زیاد است. هر چه میزان خطای نمونه‌گیری بیشتر باشد، سبب می‌شود پژوهشگر با دقت کمتر و با دشواری بیشتری بتواند شاخص‌های نمونه را به جامعه تعمیم دهد و یا شاخص‌های جامعه را از شاخص‌های نمونه تخمین بزند. درواقع، وظیفه هر پژوهشگر این است که تا حدی که امکان دارد خطای نمونه‌گیری را به حداقل برساند تا با دقت بیشتری بتواند پارامترهای جامعه را از آماره‌های نمونه تخمین بزند.

برای پرسش‌های بیشتر، پرسش‌های ۶۰، ۶۱ و ۶۵ را مطالعه کنید.

1 . sampling error

پرسش ۶۳: فرضیه صفر چیست و چگونه به کار می‌رود؟

نخستین گام در استنباط آماری[1] تدوین فرضیه صفر است. در آمار و احتمالات، برای بررسی یک فرضیه، در آغاز فرض صفر[2] (فرض پوچ) را می‌پذیریم تا با رد آن فرض یک (فرض جایگزین[3]) را اثبات نماییم. فرضیه صفر بیانگر تشابه و برابری است. در واقع برای آزمون فرضیه آماری[4] در آغاز از نوعی فرایند استدلالی معکوس استفاده می‌کنیم؛ یعنی در ابتدا می‌گوییم که هیچ ارتباط یا تفاوتی وجود ندارد. فرضیه صفر که در حقیقت نقطه آغاز یک تلاش و پژوهش علمی است، یک الگوی مفهومی است که بیانگر عدم وجود ارتباط بین متغیرهاست. فرضیه صفر به پژوهشگر نشان می‌دهد که اطلاعاتی درباره ارتباط بین متغیرها وجود ندارد. فرضیه صفر که به فرضیه آماری یا فرضیه پوچ نیز موسوم است وجود رابطه، اثر یا تفاوت بین متغیرها را رد و انکار می‌کند. فرض صفر را می‌توان به‌صورت $H_0: \mu_1 = \mu_2$ نوشت. در این معادله H_0 نماد فرضیه صفر، μ_1 پارامتر جامعه برای نمونه اول و μ_2 پارامتر جامعه برای گروه دوم است. توجه داشته باشید که فرضیه صفر را می‌توان برای رابطه متغیرها، اثر متغیرها بر یکدیگر و تفاوت بین گروه‌ها در متغیرهای مختلف ارائه کرد که در معادله ذکرشده به تفاوت بین دو گروه از منظر یک متغیر اشاره شده است. برای مثال، این فرضیه صفر را در نظر بگیرید: بین سطح رضایت شغلی کارگران جدیدالورودی که برنامه کارآموزی تابستانی را سپری کردند در مقایسه با کارگران جدیدالورودی که برنامه کارآموزی تابستانی را سپری نکردند، تفاوت معناداری وجود ندارد. این فرضیه این‌گونه فرض می‌کند که بین شرکت در برنامه کارآموزی تابستانی و رضایت شغلی ارتباط معناداری وجود ندارد. همان‌گونه که توجه نمودید، در فرضیه مطرح‌شده هیچ‌گونه صحبتی از ارتباط بین متغیرها به چشم نمی‌خورد و فرضیه بر عدم تفاوت دو گروه از شرکت‌کنندگان و عدم شرکت کندگان در یک برنامه کارآموزی تأکید کرده است و دو گروه را ازلحاظ رضایت شغلی برابر و یکسان فرض کرده است. ارائه فرضیه صفر به سه دلیل اهمیت دارد:

اول اینکه، فرضیه صفر یک نقطه آغاز است زیرا مهم‌ترین و ارزشمندترین اطلاعاتی که به ما می‌دهد این است که بسیار نامحتمل و بعید است که بین دو یا چند متغیر ارتباط وجود داشته باشد. درواقع فرض صفر بیان می‌کند که در حال حاضر هیچ‌گونه اطلاعاتی درباره ارتباط بین متغیرها، یا تفاوت بین گروه‌ها وجود ندارد. به همین دلیل سبب می‌شود پژوهشگران دست به پژوهش بزنند تا بتوانند فرضیه صفر را رد کنند و این امر سبب می‌شود که هیچ‌گونه سوگیری در ذهن محقق در ابتدای امر وجود نداشته باشد. به تعبیر کارل پوپر[5]، اگر ما منتقد نباشیم، همواره چیزهایی را می‌یابیم که می‌خواهیم؛ جستجو می‌کنیم و آنچه را که ما را تائید می‌کند می‌یابیم، نگاه محدود داریم و آنچه را که تئوری‌های دوست‌داشتنی ما را به خطر می‌اندازد، نمی‌بینیم.

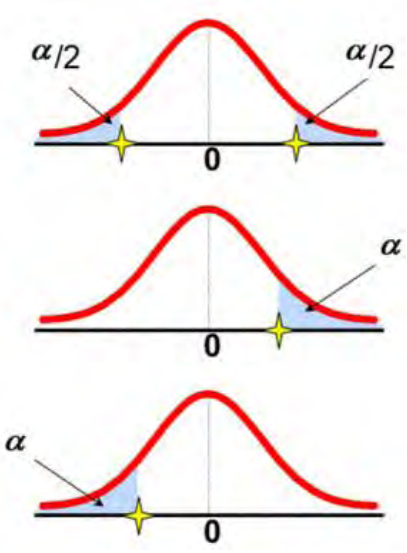
---

1 . Statistical Inference
2 . Null hypothesis
3 . Alternative hypothesis
4 . Statistical hypothesis testing
5 . Sir Karl Raimund Popper

دوم اینکه، فرضیه صفر معیاری است که می‌توانیم بر اساس آن دست به جمع‌آوری داده‌ها بزنیم و از پراکندگی و بی‌هدفی در پژوهش نجات پیدا کنیم. هنگامی‌که نقطه آغاز ما مشخص باشد، درواقع هدف و مسیر و مراحل جستجوی گری ما نیز مشخص خواهد بود.

سوم اینکه، فرض یه صفر در واقع یک نقطه آغازین مفهومی است و بندرت در مقالات علمی و گزارش‌های پژوهشی مشاهده می‌شود. درواقع به‌نوعی در پشت‌صحنه پژوهش‌های علمی خودنمایی می‌کند، و این در حالی است که از ابتدا تا انتهای هر مطالعه علمی مبتنی بر فرضیه دقیق و مناسب است و هر جایی که یک پرسش علمی مطرح شود، پای فرضیه صفر در میان است!. در ادامه به چند نکته مهم اشاره شده است:

- فرض صفر، فرضی است که محقق مایل به رد کردن آن است؛
- در فرض صفر، این‌گونه قضاوت می‌شود که روابط یا تفاوت‌های مشاهده‌شده یا ناشی از تفاوت‌ها و روابط ظاهری حقیقی‌اند و یا صرفاً ناشی از خطای نمونه‌گیری هستند؛
- فرض یه صفر همیشه با نمادهای = یا ≤ یا ≥ مطرح می‌شود؛
- فرض خلاف، انتظار پژوهشگر را درباره نتایج آتی پژوهش نشان می‌دهد؛
- فرض یه یک همیشه با نمادهای = یا > یا < مطرح می‌شود؛
- چنانچه فرض صفر رد شود، فرض خلاف قابل‌قبول باقی می‌ماند؛
- فرض صفر به‌طور غیرمستقیم رد می‌شود؛
- آزمون فرضیه با آزمون‌های جهت‌دار پارامتریک قوی‌تر از آزمون‌های دودامنه یا بدون جهت است؛ برای مثال در حجم نمونه بینهایت در جدول آزمون t می‌توان مشاهده کرد که مقدار بحرانی برای رد فرضیه صفر با ۹۵٪ اطمینان برای فرضیه جهت‌دار (یک‌دامنه) ۱/۶۴ و برای فرضیه بی‌جهت (دودامنه) ۱/۹۶ است؛ مقدار بحرانی برای رد فرضیه صفر با ۹۹٪ اطمینان برای فرضیه جهت‌دار (یک‌دامنه) ۲/۳۳ و برای فرضیه بی‌جهت (دودامنه) ۲/۵۸ است؛ باید افزود که هرگاه فرضیه یک‌دامنه باشد، به نقاط بحرانی مربوط به آلفای ۰/۰۵ و ۰/۰۱ در قسمت مقادیر یک‌دامنه رجوع می‌کنیم که به ترتیب ۱/۹۶ و ۲/۵۸ خواهد بود و هرگاه فرضیه دودامنه باشد، به نقاط بحرانی مربوط به آلفای ۰/۰۵ و ۰/۰۱ در قسمت مقادیر دودامنه رجوع می‌کنیم که به ترتیب ۱/۶۴ و ۲/۳۳ خواهد بود و این در حالی خواهد بود که آلفا به ترتیب ۰/۰۲۵ و ۰/۰۰۵ خواهد بود.



- خطای برآورد یا اشتباه برآورد عبارت است از احتمال تائید H_0 درست؛
- قاعده رد[1]، در موقعیت آزمون فرضیه، می‌بایست قبل از آغاز آزمایش یا مطالعه تعیین شود. این قاعده مشخص می‌کند که چه نتایجی موجب رد فرضیه صفر به سود فرضیه خلاف می‌شود.
- ضریب اطمینان $(1-\alpha)$ عبارت است از احتمال عدم رد H_0 درست؛
- فاصله اطمینان یک فرضیه عبارت است از $(1-\alpha)$۱ ×۱۰۰ ٪؛
- توان یک آزمون آماری $(1-\beta)$ عبارت است از احتمال رد H_0 نادرست؛

برای پرسش‌های بیشتر، پرسش‌های ۶۰، ۶۲ و ۶۴ را مطالعه کنید.

1 . reject rule

پرسش ۶۴: فرضیه تحقیق چیست و چگونه به کار می‌رود؟

همان‌گونه که در سؤال قبل مطرح شد، فرضیه صفر بر برابری[1] تأکید دارد و به پژوهشگر نشان می‌دهد که ارتباطی بین متغیرها وجود ندارد و اگر هم ارتباطی بین متغیرها و یا تفاوتی در یک متغیر بین دو یا چند گروه مشاهده شود، ناشی از تصادف است. H_1 یا فرضیه یک[2]، که به فرضیه تحقیق، فرضیه محقق، فرضیه جایگزین، فرضیه خلاف، فرضیه مقابل یا فرضیه بدیل نیز معروف است، برعکس فرضیه صفر است. فرض یک را می‌توان به‌صورت $H_1: \mu_1 \neq \mu_2$ نوشت. در این معادله H_1 نماد فرضیه یک، μ_1 پارامتر جامعه برای نمونه اول و μ_2 پارامتر جامعه برای گروه دوم است. توجه داشته باشید که فرضیه یک را می‌توان برای رابطه متغیرها، اثر متغیرها بر یکدیگر و تفاوت بین گروه‌ها در متغیرهای مختلف ارائه کرد، که در معادله ذکرشده به تفاوت بین دو گروه در یک متغیر اشاره شده است. برای مثال، فرضیه یک ممکن است این‌گونه مطرح شود که بین دانشجویان روانشناسی و دانشجویان ریاضی ازنظر هوش هیجانی تفاوت معناداری وجود دارد؛ بین هوش هیجانی و بیماری قلبی ارتباط معناداری وجود دارد؛ بین مشاهده فیلم‌های خشن و فشارخون ارتباط معناداری وجود دارد. درواقع، فرضیه تحقیق از احتمال وجود رابطه یا اثر و یا تفاوت بین متغیرها خبر می‌دهد. در فرضیه تحقیق بر اساس مرور منابع و ادبیات پژوهشی و مجموع اطلاعاتی که محقق دارد، سبب می‌شود که پژوهشگر یک رابطه یا فرضیه را مفروض بداند.

دو نوع فرضیه تحقیق وجود دارد: جهت‌دار[3] و بی‌جهت[4]. در فرضیه تحقیق بی‌جهت، بر تفاوت گروه‌ها و یا رابطه بین متغیرها تأکید می‌شود، امّا جهت این تفاوت یا رابطه مشخص نمی‌شود. برای مثال، این فرضیه تحقیق را در نظر بگیرید: "بین میزان پرخاشگری افرادی که فیلم‌های خشن می‌بینند در مقایسه با افرادی که فیلم‌های خشن نمی‌بینند تفاوت معناداری وجود دارد". در این فرضیه همان‌گونه که متوجه شدید، جهت تفاوت مشخص نشده است به این معنی که هیچ اطلاعاتی درباره اینکه کدام گروه میزان پرخاشگری بیشتری دارند به دست نمی‌دهد و صرفاً از وجود تفاوت در میزان خشونت بحث می‌کند.

در فرضیه تحقیق جهت‌دار، بر تفاوت گروه‌ها و یا رابطه بین متغیرها تأکید می‌شود، امّا جهت این تفاوت یا رابطه مشخص می‌شود. برای مثال، این فرضیه تحقیق را در نظر بگیرید: "میزان پرخاشگری افرادی که فیلم‌های خشن می‌بینند در مقایسه با افرادی که فیلم‌های خشن نمی‌بینند به‌طور معناداری بیشتر است". در این فرضیه همان‌گونه که متوجه شدید، جهت تفاوت مشخص‌شده است به این معنی که اطلاعاتی درباره اینکه کدام گروه میزان پرخاشگری بیشتری

1 . equality
2 . research hypothesis
3 . directional
4 . nondirectional

دارند به دست می‌دهد. اکنون که فرضیه جهت‌دار است فرضیه تحقیق به‌صورت $H_1: \mu_1 \geq \mu_2$ نگاشته می‌شود.

در این معادله H_1 نماد فرضیه یک، μ_1 آماره (پرخاشگری) گروه اول و μ_2 آماره (پرخاشگری) گروه دوم است. علامت $\geq$ نیز نشان‌دهنده بزرگ‌تر یا مساوی بودن آماره‌های دو گروه است.

برای پرسش‌های بیشتر، پرسش‌های ۶۰، ۶۱ و ۶۳ را مطالعه کنید.

پرسش ۶۵: فرضیه‌های صفر و تحقیق چه تفاوت‌هایی با یکدیگر دارند؟

آزمون فرض آماری[1]، قانون یا دستوری است که بر طبق آن، بر اساس مشاهدات انجام‌شده تصمیم به رد یا قبول فرض موردنظر گرفته می‌شود. وقتی‌که فرض یا ادعایی در مورد یک پارامتر جامعه بیان می‌شود، ممکن است درست یا نادرست باشد، بنابراین دو فرض مطرح می‌شود یکی آنکه ادعا درست است و دیگری آنکه ادعا درست نیست. لذا یک آزمون فرض آماری شامل دو فرض آماری است که در مقابل یکدیگر قرار می‌گیرند، یکی را فرض صفر[2] و دیگر را فرض مقابل[3] (فرض یک) می‌گویند.

فرضیه صفر و فرضیه یک از جهات مختلفی باهم فرق دارند:

۱. اول اینکه فرضیه صفر از برابری صحبت می‌کند و فرضیه یک بیانگر نابرابری است؛

۲. عموماً فرضیه صفر در مقالات مجلات ذکر نمی‌شود، امّا فرضیه یک همیشه در مقالات ذکر می‌شود؛

۳. فرضیه صفر، جامعه را مد نظر قرار می‌دهد امّا فرضیه پژوهش (یک) به نمونه توجه دارد. درواقع نمونه‌های انتخاب‌شده از جامعه مورد بررسی و تحلیل واقع می‌شوند (فرضیه پژوهش) و نتایج به جامعه تعمیم داده می‌شود (فرضیه صفر).

۴. در بیان فرضیه صفر همیشه پارامتر جامعه بیان می‌شود نظیر $H_0: \mu_1 = \mu_2$ که بیانگر برابر بودن میانگین نمونه‌های دو جامعه با یکدیگر است، درحالی‌که در بیان فرضیه پژوهش (یک) همیشه آماره نمونه[4] بیان می‌شود نظیر $H1: X_1 \neq X_2$ که بیانگر برابر نبودن آماره‌های دو نمونه است. نماد میانگین مطرح‌شده در فرضیه صفر μ است و نماد میانگین مطرح‌شده در فرضیه یک X است.

۵. ازآنجاکه تمامی اعضای یک جامعه را نمی‌توان مورد ارزیابی قرار داد، بنابراین ناممکن است که با قطعیت بگوییم که آیا فرضیه صفر درست است یا نادرست. در واقع تعمیم نتایج بررسی نمونه به جامعه با قطعیت کامل همراه نیست. می‌توان گفت که فرضیه صفر به‌طور غیرمستقیم و فرضیه پژوهش به طور مستقیم آزمون می‌گردند.

برای پرسش‌های بیشتر، پرسش‌های ۶۰، ۶۱ و ۶۲ را مطالعه کنید.

---

1 . Hypothesis testing
2 . Null hypothesis
3 . Alternative hypothesis
4 . sample statistics

بخش ۸: درک منحنی بهنجار و احتمال

پرسش ۶۶: چرا در مطالعه آمار، مفهوم احتمال مهم است؟

بیشتر مطالبی که در مطالعه آمار می‌آموزید، به نوعی با منحنی بهنجار یا زنگوله‌ای[1] در ارتباط است. درک و فهم منحنی بهنجار ما را قادر می‌سازد دریابیم که احتمال و درصد احتمال رخداد یک پدیده تا چه حد با یک نتیجه معین در ارتباط است.

برای مثال، چقدر احتمال دارد که یک دانش‌آموز در یک آزمون نمره ۸۷ بگیرد، درحالی‌که میانگین نمره آن کلاس ۹۳ است؛ یا فروش یک مدل از اتومبیل شرکتت هیوندا در یک ایالت آمریکا، تا چه میزان نماینده فروش آن محصول در کل ایالات‌متحده است. درواقع بررسی احتمال وقوع یک پدیده یا رویداد به ما اجازه می‌دهد تا به سؤالات مشابه مطرح‌شده پاسخ دهیم. ما می‌توانیم مشخص کنیم که احتمال وقوع یک نتیجه بالا یا پایین است و سپس بر اساس آن دست به تصمیم بزنیم و تعیین کنیم که آیا احتمال به‌دست‌آمده قابل‌پذیرش است یا خیر؟ پدیده‌ها را می‌توان به دو صورت مورد بررسی قرار داد: به‌صورت قطعی و بی‌چون‌وچرا و به‌صورت احتمالی. نوع اول مربوط به پدیده‌های است که ساختار واقعی آن‌ها شناخته‌شده است و حرکت بعدی آن‌ها کاملاً روشن است. نو ع دوم پدیده‌هایی هستند که به‌طور تصادفی پیش می‌آیند. در زندگی عادی هم اغلب به ارزیابی‌هایی از این نوع می‌پردازیم.

از دیرباز تنها رهیافت تکامل‌یافته ریاضی برای حل مسائل در شرایط عدم قطعیت[2]، نظریه احتمال[3] بوده است. احتمال، نظریه‌ای ریاضی است که اولین بار در قرن هفدهم در فرانسه در رابطه با بازی‌های شانسی به وجود آمد و در ۱۹۳۳ در کتابی با عنوان مبانی نظریهٔ احتمال توسط ریاضیدان روسی کولموگروف به‌صورت کامل ارائه شد. بر اساس باور عمومی، در بسیاری از محیط‌های تصمیم، داده‌های موجود جنبه آماری دارند و بنابراین با روش‌های نظریه احتمال می‌توان بر عدم قطعیت ناشی از جنبه‌های تصادفی فائق آمد. مطالعه و کاربست قوانین احتمال به ما اجازه می‌دهد که مشخص کنیم که تا چه درجه‌ای از اطمینان می‌توانیم وقوع یک نتیجه خاص را درست پیش‌بینی کنیم. برای مثال، اگر مشاهده کنیم که مردها در مقایسه با زن‌ها، میزان پرخاشگری متفاوتی دارند، با چه درجه از اطمینان می‌توانیم بیان کنیم که این ادعا درست است؟. ممکن است این ادعا، صرفاً نتیجه یک مطالعه آزمایشی ضعیف باشد، یا ممکن است به دلیل خطای نمونه‌گیری باشد، زیرا برخی موارد نمونه انتخاب‌شده به‌خوبی معرف جامعه مادر نیست. درنهایت می‌توان گفت، مفهوم کلی قانون احتمال ارتباط تنگاتنگی با نقش فرضیه‌های صفر و تحقیق دارد. فرضیه تحقیق با استفاده از نمونه‌ای از جامعه مادر مورد آزمون قرار می‌گیرد. درک نقش قانون احتمال به ما اجازه می‌دهد که یافته‌های مستخرج از نمونه را به جمعیت مادر تعمیم دهیم، البته با احتیاط در تعمیم نتایج حاصل از نمونه به جامعه، به این دلیل که پژوهشگران، جامعه را به‌صورت مستقیم بررسی نمی‌کنند، بلکه بر اساس یک حدس عالمانه بر

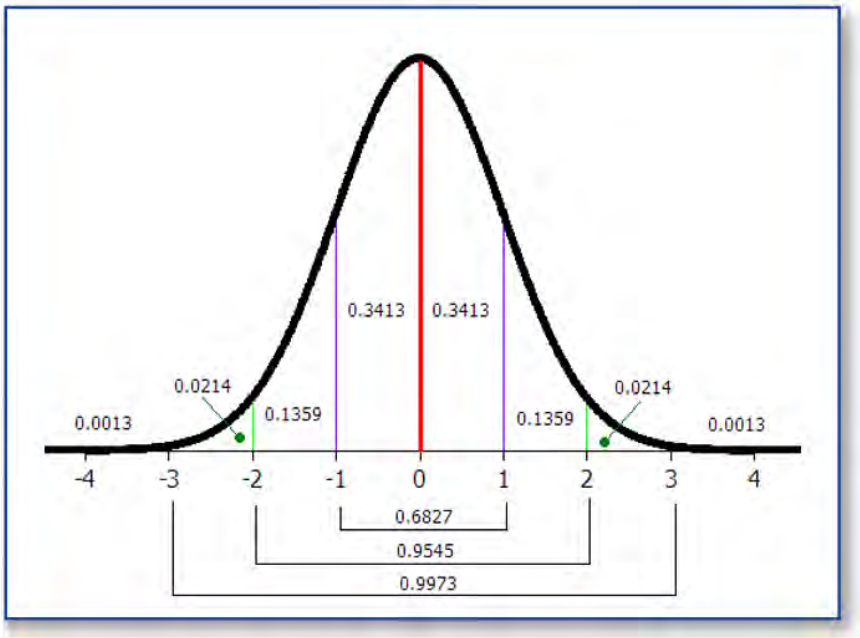
1 . bell-shaped curve
2 . Uncertainty
3 . Probability Theory

اساس احتمال، دست به تعمیم نتایج می‌زنند. حدس عالمانه بر اساس احتمال در اینکه یک نمونه چقدر مشابه جامعه است و یک نمونه چقدر دقیق انتخاب شده است.

برای پرسش‌های بیشتر، پرسش‌های ۶۷، ۶۹ و ۷۲ را مطالعه کنید.

پرسش ۶۷: منحنی بهنجار یا زنگوله‌ای چیست؟

همان‌گونه که در تصویر زیر ملاحظه می‌کنید، منحنی بهنجار یا زنگوله‌ای نمایه دیداری توزیع یک مجموعه داده است. این منحنی، مبنای بسیاری از آزمون‌های آماری به شمار می‌رود. این منحنی سه ویژگی مهم دارد. این توزیع گاهی به دلیل استفاده کارل فردریک گاوس از آن در کارهای خود با نام توزیع یا تابع گوسی[1] (گاوسی) نامیده می‌شود؛ همچنین به دلیل شکل تابع احتمال این توزیع، با نام منحنی زنگوله‌ای[2] (زنگدیس) نیز معروف است. تابع احتمال این توزیع دارای دو پارامتر است که یکی تعیین‌کننده مکان (μ) و دیگری تعیین‌کننده مقیاس (σ) توزیع هستند.



ویژگی اول این است که در منحنی بهنجار ، میانگین، میانه و نما برابر هستند. درواقع، نقطه میانی این نمودار نشان‌دهنده میانگین، میانه و نما است؛ ویژگی دوم این است که منحنی بهنجار قرینه است، یعنی نیمه چپ کاملاً مشابه نیمه راست است؛ ویژگی سوم این است که منحنی بهنجار ، مجانب[3] است. مجانب بودن به این معناست که دامنه‌های سمت چپ و راست منحنی بهنجار ، هرچند که به محور افقی نزدیک و نزدیک‌تر می‌شوند، ولی هرگز محور افقی (X) را قطع نمی‌کنند.

مهم اینجاست که هر عددی که روی محور افقی (X) قرار می‌گیرد، با توجه به جایگاه آن عدد، می‌توان احتمال رخداد آن را مشاهده کرد. چه نیمه چپ منحنی بهنجار باشد و چه در نیمه

1 . Gaussian Function
2 . Bell Shaped
3 . asymptotic

راست. برای مثال، تقریباً ۶۸٪ از کل اعدادی که از یک توزیع طبیعی گرفته شوند، فاصله‌ای برابر یا کمتر از یک برابر انحراف معیار توزیع نسبت به میانگین توزیع دارند. تقریباً ۹۵٪ از کل اعدادی که از یک توزیع طبیعی گرفته شوند، فاصله‌ای برابر یا کمتر از دو برابر انحراف معیار توزیع نسبت به میانگین توزیع دارند.

نکته

- آزمون کولموگروف-اسمیرنوف[1] یک آزمون ناپارامتریک است که توسط آندری کالموگروف[2] و نیکولای اسمیرنوف[3] که هر دو از ریاضیدانان روسیه بودند ابداع شد و از طریق آن می‌توان بهنجار بودن توزیع را ارزیابی کرد؛ برخی مواقع برای سنجش شباهت توزیع نمرات یک نمونه با توزیع منحنی بهنجار به کار می‌رود، که از طریق آزمون کولموگروف-اسمیرنوف تک نمونه‌ای[4] استفاده می‌شود؛ برخی مواقع برای سنجش شباهت توزیع نمرات یک نمونه با توزیع نمرات یک نمونه دیگر به کار می‌رود، که از طریق آزمون کولموگروف-اسمیرنوف دو نمونه‌ای[5] استفاده می‌شود؛
- آزمون شاپیرو-ویلک[6] نیز یک آزمون ناپارامتریک است که توسط پروفسور ساموئل سنفورد شاپیرو[7] و پروفسور مارتین ویلک[8] که هر دو از ریاضیدانان آمریکایی و کانادایی بودند در سال ۱۹۶۵ ابداع شد و از طریق آن می‌توان بهنجار بودن توزیع را ارزیابی کرد؛ اگر حجم نمونه بیشتر از ۲۰۰۰ باشد، آزمون کولموگروف-اسمیرنوف مناسب‌تر از شاپیرو-ویلک است و اگر حجم نمونه کم باشد، هر دو آزمون نتایج یکسانی دارند؛
- هنگام بررسی بهنجار بودن داده‌ها ما فرض صفر مبتنی بر اینکه توزیع داده‌ها بهنجار است را در سطح خطای ۵٪ آزمون می‌کنیم. بنابراین اگر سطح معناداری[9] (sig) آزمون بزرگ‌تر یا مساوی ۰/۰۵ به دست آید، در این صورت دلیلی برای رد فرض صفر مبتنی بر اینکه داده بهنجار است، وجود نخواهد داشت. به عبارت دیگر، توزیع داده‌ها بهنجار خواهد بود. برای آزمون توزیع بهنجار فرض‌های آماری به‌صورت زیر تنظیم می‌شود:

  H_0: توزیع داده‌های مربوط به هر یک از متغیرها بهنجار است

  H_1: توزیع داده‌های مربوط به هر یک از متغیرها بهنجار نیست
- پروفسور گرولامو کاردانو[10] که از ریاضیدانان ایتالیایی بود شکل اولیه قانون اعداد بزرگ[11] را که احتمالاً معروف‌ترین نتیجه در نظریه احتمالات است معرفی کرد. به‌طور خلاصه، کاربرد

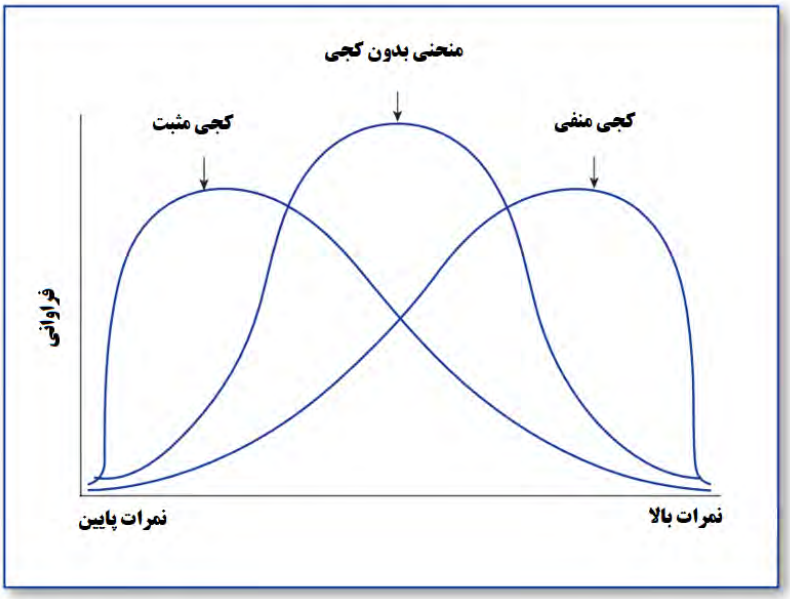
1 . the Kolmogorov–Smirnov test (K–S test or KS test)
2 . Andrey Nikolaevich Kolmogorov (25 April 1903 – 20 October 1987)
3 . Nikolai Vasilyevich Smirnov (4 October 1900 – 2 June 1966)
4 . one-sample Kolmogorov–Smirnov test
5 . two-sample Kolmogorov–Smirnov test
6 . The Shapiro–Wilk test
7 . Samuel Sanford Shapiro (born July 13, 1930)
8 . Martin Bradbury Wilk, OC (18 December 1922 – 19 February 2013)
9 . significance level
10 . Gerolamo Cardano (24 September 1501 – 21 September 1576)
11 . law of large numbers

این قانون در زمینه استنباط آماری است. قانون اعداد بزرگ بیان می‌دارد که هرگاه حجم نمونه تصادفی انتخابی به‌اندازه کافی «بزرگ» باشد، این نمونه می‌تواند با دقت خوبی نماینده مشخصات اساسی جامعه‌ای باشد که از آن استخراج شده است،

برای پرسش‌های بیشتر، پرسش‌های ۶۶، ۶۸ و ۷۳ را مطالعه کنید.

پرسش ۶۸: کجی و کشیدگی چیست و توزیع‌های نمرات چگونه ازنظر این ویژگی متفاوت‌اند؟

توزیع همه مجموعه داده‌ها بهنجار نیست. در حقیقت، هرچند توزیع پراکندگی برخی مجموعه داده‌ها بهنجار یا شبیه بهنجار است، امّا توزیع‌های پراکندگی بسیاری نیز وجود دارند که بهنجار نیستند. برخی اوقات یک توزیع پراکندگی دارای کجی یا چولگی[1] است که به معنی این است که منحنی یک توزیع پراکندگی متقارن نیست، یعنی نیمه چپ کاملاً مشابه نیمه راست نیست. در این مواقع، یک نیمه از منحنی، طولانی‌تر از نیمه دیگر است. همان‌گونه که در شکل زیر مشاهده می‌کنید، منحنی‌های A و C دارای کجی یا چولگی هستند و صرفاً منحنی A بهنجار است.

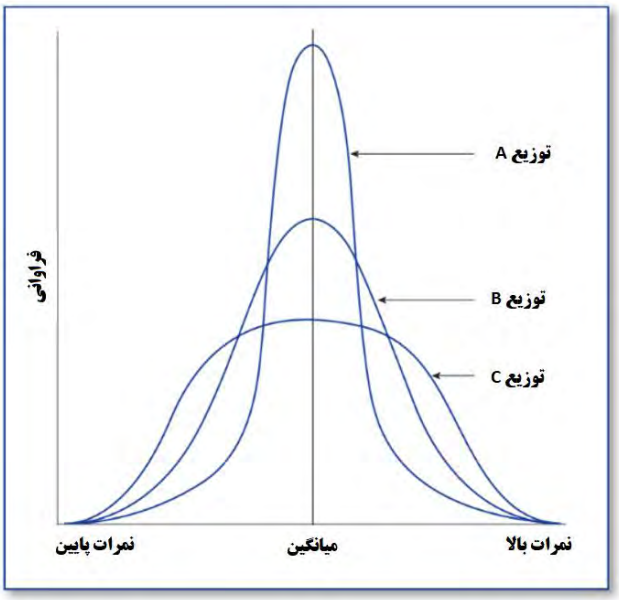


در آمار و نظریه احتمالات برجستگی یا کشیدگی، توصیف‌کننده‌ی میزان قله‌ای بودن یک توزیع احتمالی است. هرچقدر شکل منحنی توزیع داده‌ها یا تابع چگالی احتمال[2] بلندتر یا کشیده‌تر باشد، حتماً قله‌ای‌تر[3] و میزان شاخص کشیدگی برای آن بیشتر است.

اگر دامنه نیمه چپ منحنی طولانی‌تر از دامنه نیمه راست منحنی باشد، به معنی است که تعداد نمرات بالا بسیار بیشتر از تعداد نمرات پایین است، منحنی کجی منفی دارد. برای مثال، هنگامی‌که به بررسی وضعیت قد ۱۰۰ بازیکن گروه بسکتبال بپردازیم، متوجه می‌شویم که ۷۵ ورزشکار قدبلند و ۱۵ ورزشکار دارای قدکوتاه یا متوسط هستند. بنابراین منحنی توزیع قد ورزشکاران، کجی منفی دارد.

1 . Skewness
2 . probability density function
3 . heavy-tailed

اگر دامنه نیمه راست منحنی طولانی‌تر از دامنه نیمه چپ باشد، به معنی است که تعداد نمرات پایین بسیار بیشتر از تعداد نمرات بالا است، منحنی کجی مثبت دارد. برای مثال، هنگامی‌که به بررسی وضعیت نمرات دانشجویان درس روانشناسی در درس آمار بپردازیم، متوجه می‌شویم که از میان ۳۰ دانشجو، ۲۵ نفر نمرات بسیار پایین و ۵ دانشجو نمره مناسب کسب کرده‌اند. این مسئله نشان می‌دهد یا اغلب دانشجویان به مطالعه مناسب نپرداخته‌اند و یا سطح سؤالات دشوار بوده است. بنابراین منحنی توزیع نمره دانشجویان، کجی مثبت دارد. کجی یا چولگی در حقیقت معیاری از وجود یا عدم تقارن تابع توزیع است. برای یک توزیع کاملاً متقارن کجی صفر و برای یک توزیع نامتقارن با کشیدگی به سمت مقادیر بالاتر کجی مثبت و برای توزیع نامتقارن با کشیدگی به سمت مقادیر کوچک‌تر مقدار کجی منفی است. برای محققین علوم رفتاری توجه به چولگی دارای اهمیت است زمانی که در داده‌ها چنین موردی رخ می‌دهد. چولگی بزرگ، محقق را برای کنکاش داده‌های پرت رهنمود می‌کند. توزیع بهنجار دارای چولگی صفر است. اگرچه، یک توزیع می‌توان کاملاً متقارن باشد، امّا از شکل توزیع بهنجار بسیار متفاوت باشد. اولین بار کارل پیرسون[1] (۱۸۹۵) اندازه‌گیری چولگی را به‌وسیله‌ی استاندارد نمودن تفاوت بین میانگین و مد یعنی "تفاضل میانگین از مد تقسیم‌بر انحراف استاندارد" را پیشنهاد نمود. دومین مسیر تفاوت منحنی‌های توزیع پراکندگی نمرات با یکدیگر، به شاخص کشیدگی[2] برمی‌گردد.



پیرسون در سال ۱۹۰۵ کشیدگی را به‌عنوان اندازه‌ای از میزان همواری بالای یک توزیع متقارن ارائه نمود تا آن را با توزیع بهنجار با واریانس مشابه مقایسه نماید. همان‌گونه که در تصویر زیر

1 . Karl Pearson FRS (27 March 1857ت 27 April 1936)
2 . kurtosis

ملاحظه می‌کنید، منحنی A بیشترین کشیدگی، منحنی B کشیدگی متوسط و منحنی C در مقایسه با دو منحنی دیگر، کمترین کشیدگی را داراست. به‌عبارت دیگر، کشیدگی معیاری از تیزی منحنی در نقطه ماکزیمم است. مقدار کشیدگی برای توزیع بهنجار برابر ۳ است. در بین سطوح متفاوت کشیدگی منحنی‌ها، به یک منحنی نظیر منحنی A که کشیدگی بیشتری نسبت به منحنی بهنجار دارد، منحنی دارای کشیدگی مثبت (لپتوکُرتیک)[1] می‌گویند. در منحنی لپتوکُرتیک، بیشتر نمرات تمایل به تجمیع در مرکز توزیع داده‌ها را دارند و نمرات کمی در فاصله‌های دور از مرکز تجمیع می‌شوند. این وضعیت هنگامی رخ می‌دهد که پراکندگی و تغییرپذیری داده‌ها بسیار کم و اغلب نمرات نزدیک به میانگین باشند. برای مثال، هنگامی‌که اکثر دانشجویان روانشناسی در یک کلاس که دارای میانگین نمره درس آمار ۱۸ است، نمرات خیلی نزدیک به ۱۸ کسب می‌کنند.

به یک منحنی نظیر منحنی C که کشیدگی کمتری نسبت به منحنی بهنجار دارد، منحنی دارای کشیدگی مثبت (پِلَتی‌کُرتیک)[2] می‌گویند. در منحنی لپتوکُرتیک، بیشتر نمرات تمایل به توزیع در تمام فضای منحنی توزیع داده‌ها را دارند. این وضعیت هنگامی رخ می‌دهد که پراکندگی و تغییرپذیری داده‌ها پراکنده و یکنواخت در تمام طول منحنی باشند. برای مثال، هنگامی‌که اکثر دانشجویان روانشناسی در یک کلاس نمرات شبیه به یکدیگر کسب می‌کنند.

در یک جمع‌بندی می‌توان گفت، کجی یا چولگی در آمار، مقداری از تقارن توزیع یک متغیر در اطراف میانگین است. در واقع کجی، انحراف منحنی بهنجار از حالت تقارن است. مقدار کجی می‌تواند مثبت یا منفی باشد. در حالت کجی مثبت، میانگین بزرگ‌تر از میانه و میانه بزرگ‌تر از مد است و در حالت کجی منفی، مد بزرگ‌تر از میانه و میانه بزرگ‌تر از میانگین است.کشیدگی بیان‌کننده نحوه انباشته شدن نمره‌ها در مرکز توزیع یک متغیر است. درواقع کشیدگی، برآمدگی یا فرورفتگی منحنی توزیع بهنجار است. در حالت کشیدگی تقارن توزیع حفظ می‌شود و دو نیمه منحنی متقارن هستند، اما نقطه اوج منحنی بهنجار دچار تغییر می‌شود. مقدار کشیدگی می‌تواند مثبت یا منفی باشد. در حالت کشیدگی منفی، منحنی توزیع متغیر فرورفته تر از حالت بهنجار می‌شود و نقطه اوج توزیع متغیر، پایین‌تر از توزیع بهنجار است. در حالت کشیدگی مثبت، منحنی توزیع متغیر برآمده‌تر از منحنی بهنجار می‌شود و نقطه اوج توزیع متغیر، بالاتر از توزیع بهنجار است. در حقیقت کجی یا چولگی، میزان تقارن منحنی یک توزیع را نشان می‌دهد؛

برای پرسش‌های بیشتر، پرسش‌های ۶۶، ۶۷ و ۷۳ را مطالعه کنید.

1 . leptokurtic
2 . platykurtic

پرسش ۶۹: قضیه حد مرکزی چیست و چرا مهم است؟

در پرسش ۶۷ نمای کلی منحنی بهنجار را ملاحظه کردید و متوجه شدید که بسیاری از مفاهیم پایه‌ای آمار توصیفی و استنباطی مبتنی بر شکل و ویژگی‌های منحنی بهنجار است. امّا اگر توزیع پراکندگی یک مجموعه داده بهنجار نباشد و دارای کجی و کشیدگی باشد چه باید کرد؟. آیا بازهم می‌توان از مؤلفه‌های آمار توصیفی و استنباطی استفاده کرد؟. پاسخ مثبت است. قضیه حد مرکزی[1]، کمک می‌کند که در تمام انواع توزیع‌های بهنجار و غیربهنجار از قواعد استنباط آماری استفاده شود. قضیه حد مرکزی در ابتدا توسط پیر سیمون معروف به مارکوس لاپلاس[2] ریاضیدان فرانسوی بیان و اثبات شد. او از مشاهده خطای اندازه‌گیری‌های خود که دارای توزیع بهنجار بودند، بدین قضیه دست‌یافت. وی در ضمن یک ستاره‌شناس (معروف به نیوتن فرانسه) و یکی از بزرگ‌ترین نویسندگان پیشین در مباحث آمار و احتمال بود. همچنین عامل اشاعه فرهنگ به‌کارگیری احتمال در زندگی روزمره بود. کاربرد قضیه حد مرکزی در نشان دادن اینکه مقادیر خطاهای اندازه‌گیری دارای توزیع تقریباً بهنجاری هستند، کمک شایانی به علوم قلمداد می‌شود. ازاین‌رو در قرون ۱۷ و ۱۸ از قضیه حد مرکزی به نام (قانون فراوانی خطاها) یاد می‌شده است و به‌عنوان یک پیشرفت عمده در علوم تلقی می‌شود. قضیه حد مرکزی پیشنهاد می‌کند که حتی در مجموعه داده‌های غیربهنجار، نمونه‌گیری‌های مکرر میانگینی به دست خواهد داد که برابر با میانگین نمونه‌گیری‌های مکرر یک مجموعه داده با توزیع بهنجار است.

با توجه به اینکه محاسبه تمام نمونه‌های ممکن از یک جامعه نامتناهی غیرممکن است سؤالی که در اینجا مطرح می‌شود این است که آیا می‌توان توزیع نمونه‌گیری را بدون به دست آوردن تمام نمونه‌های ممکن محاسبه نمود؟. در پاسخ می‌توان گفت، ۱. اگر جامعه دارای توزیع بهنجار باشد، برای نمونه‌هایی با هر حجم مشخص، توزیع پراکندگی نمرات نمونه بهنجار است. ۲. اگر جامعه دارای توزیع بهنجار نباشد، بر اساس قضیه حد مرکزی، در نمونه‌گیری تصادفی ساده از یک جامعه با میانگین μ و انحراف معیار σ، وقتی‌که حجم نمونه بزرگ باشد، توزیع پراکندگی نمرات نمونه بسیار نزدیک به بهنجار خواهد بود. درواقع، قضیه حد مرکزی می‌گوید که خواه توزیع جامعه پیوسته، گسسته، متقارن یا چوله باشد، مادام که واریانس جامعه متناهی است، توزیع میانگین نمونه، اگر حجم نمونه بزرگ باشد ($n>۳۰$) تقریباً بهنجار است.

اجازه دهید بیشتر توضیح دهیم.

اگر از یک جامعه ۸۰ میلیون نفری نمونه‌های ۱۰۰۰ نفری بگیریم و میانگین فشارخون افراد هر نمونه را تعیین کنیم، این میانگین‌ها لزوماً باهم برابر نخواهند بود. این میانگین‌ها لزوماً با میانگین کل جامعه ۸۰ میلیون نفری نیز برابر نیستند. اگر این نمونه‌برداری را بارها تکرار کنیم، می‌توان برای این میانگین‌ها هم یک نمودار توزیع رسم کرد. به چنین نموداری اصطلاحاً «نمودار توزیع

1 . central limit theorem

2 . Pierre-Simon, marquis de Laplace (23 March 1749ت 5 March 1827)

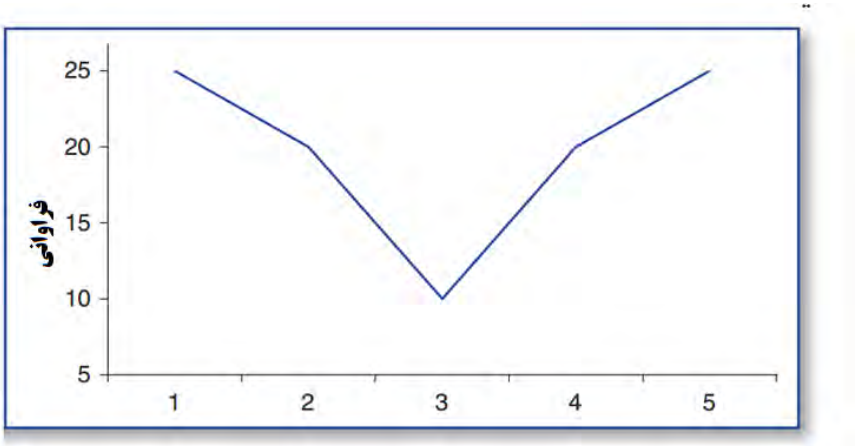
میانگین‌ها» یا «نمودار توزیع نمونه‌برداری[1]» گفته می‌شود. توزیع نمونه‌برداری، توزیع احتمالی یکی از آماره‌های نمونه نظیر میانگین، میانه یا نما است که با تکرار نمونه‌گیری از جامعه حاصل می‌شود.

- میانگین نمونه‌ها برابر است با میانگین جامعه ($\mu_x=\mu$)؛
- انحراف معیار میانگین نمونه‌ها برابر است با انحراف معیار جامعه تقسیم‌بر جذر حجم نمونه و خطای معیار میانگین نامیده می‌شود:

$$\delta_{\bar{x}} = \frac{\delta}{\sqrt{n}}$$

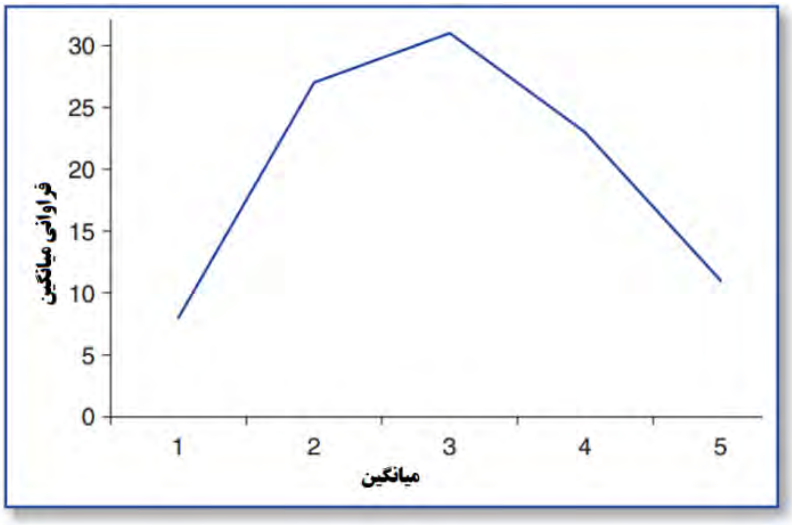
درمجموع اگر جامعه بهنجار باشد، برای هر تعداد از حجم نمونه توزیع میانگین نمونه‌ها که نمودار توزیع نمونه‌برداری یا نمودار توزیع میانگین‌ها نام دارد، بهنجار خواهد بود. بر اساس قضیه حد مرکزی، اگر یک نمونه تصادفی با حجم ۳۰ نفر یا بیشتر از یک جامعه دلخواه (دارای کجی منفی یا دارای کجی مثبت) با میانگین μ و انحراف معیار σ انتخاب شود، توزیع میانگین نمونه‌ها (نمودار توزیع نمونه‌برداری یا نمودار توزیع میانگین‌ها) به‌صورت بهنجار خواهد بود و مقدار میانگین میانگین‌ها برابر مقدار میانگین جامعه مادر μ خواهد بود.

برای مثال، فرض کنید که ۱۰۰ نمره داریم که دامنه آن‌ها از ۱ تا ۵ است که فراوانی آن‌ها در نمودار زیر نمایه شده است.



اکنون اجازه دهید یک نمونه تصادفی ۵ نفرِ از این مجموعه داده ۱۰۰ نفرِ گرفته و میانگینش را محاسبه کنیم. مجدداً اجازه دهید یک نمونه تصادفی ۵ نفرِ از این مجموعه داده ۱۰۰ نفرِ گرفته و میانگینش را محاسبه کنیم. می‌خواهیم این کار را صدها بار انجام دهیم. بعد از انجام این صدها بار نمونه‌گیری و محاسبه نمونه‌ها، این بار قصد داریم فراوانی میانگین‌های محاسبه‌شده را در نمودار زیر ترسیم کنیم.

1 . Sampling Distribution of the Mean



همان‌گونه که ملاحظه کردید، این توزیع جدید فراوانی میانگین‌ها، بسیار شبیه به منحنی بهنجار است و تقارن و نمای ظاهری آن را دارد و دامنه‌های نمرات محورهای افقی را قطع نکرده‌اند. حالا اگر این نمونه‌گیری، محاسبه میانگین و ترسیم نمودار را بی‌نهایت باز انجام دهیم، دقیقاً منحنی بهنجار به دست می‌آید. درسی که اینجا می‌توان آموخت این است که فارغ از شکل توزیع پراکندگی مجموعه‌ای که نمونه‌ها از آن انتخاب می‌شوند، فراوانی توزیع میانگین‌های تعداد زیاد و یا بی‌نهایت نمونه، شبیه منحنی بهنجار خواهد بود.

برای پرسش‌های بیشتر، پرسش‌های ۶۷، ۷۲ و ۷۳ را مطالعه کنید.

پرسش ۷۰: نمره استاندارد Z چیست و چگونه محاسبه می‌شود؟

مطالعات متعددی وجود دارد اگر بخواهیم در میان آن‌ها به مقایسه متغیرها در توزیع‌های پراکندگی بپردازیم، قادر نخواهیم بود، زیرا معیار واحدی برای سنجش آن‌ها در مطالعات مختلف به کار نرفته است. در واقع، نکته اینجاست که در تمام مطالعات مختلف، میانگین و انحراف استاندارد متغیر مورد نظر یکسان به دست نیامده است. بنابراین متفاوت بودن شاخص‌های مرکزی و پراکندگی در مطالعات مختلف فرصت مقایسه یک متغیر را به ما نمی‌دهد. نمرات استاندارد به ما اجازه می‌دهند قادر به مقایسه متغیرها در مطالعات مختلف شویم. برای مثال، اگر نمره شما در درس ریاضی ۱۶ باشد وضعیت درسی خود را چگونه ارزیابی می‌کنید؟ اگر در درس ریاضی ۱۴ و در درس فیزیک ۱۶ گرفته باشید در کدام درس قوی‌تر هستید؟ برای جواب دادن به این‌گونه سؤالات نمی‌توان تنها نمره کسب شده را لحاظ کرد بلکه باید وضعیت جامعه‌ای (کلاسی) که در آن بوده‌اید را نیز در نظر گرفت. به‌عنوان‌مثال اگر نمره ۱۶ را در کلاسی با میانگین ۱۸ و انحراف معیار ۲ کسب کرده باشیم مشخص است که نمره ما یک انحراف معیار در زیر میانگین قرار گرفته است ولی اگر میانگین کلاس ۱۴ و انحراف معیار ۱ باشد دو انحراف معیار بالای میانگین کلاس قرارگرفته‌ایم و جزء دانشجویان فعال به‌حساب می‌آیم. یا اگر معدل و انحراف معیار کلاس ریاضی به ترتیب ۱۳ و ۲ و معدل و انحراف معیار کلاس فیزیک به ترتیب ۱۶ و ۱ باشد آنگاه در درس ریاضی قوی‌تر از فیزیک عمل کرده‌ایم. درواقع، نمرات استاندارد با یکدیگر قائل قیاس هستند زیرا همه‌شان بر اساس یک انحراف استاندارد مقایسه می‌شوند. در علوم اجتماعی و رفتاری، همانند سایر علوم، پرکاربردترین نمره استاندارد، نمره Z است. نمره Z را با نام‌های ارزش Z، نمره بهنجار و متغیر استانداردشده[1] می‌شناسند. نوع دیگر نمرات استاندارد نمرات استاندارد T هستند. نمره استاندارد z از طریق فرمول زیر محاسبه می‌شود:

$$z = \frac{x - \mu_s}{\sigma_s}$$

در این فرمول z نمره استاندارد، x نمره‌ای که قصد دارید به نمره استاندارد تبدیلش کنید، $\bar{x}$ میانگین نمونه و s انحراف استاندارد نمونه است. همچنین می‌توان در برخی مواقع از فرمول $Z \times \sigma_s + \mu_s$ برای محاسبه نمره خام استفاده کرد.

برای مثال، میانگین یک نمونه ۷۸ و انحراف استانداردش ۳ است. برای عدد ۸۱ در این مجموعه نمره، نمره استاندارد Z برابر خواهد بود با ۱+ خواهد بود.

$$z = \frac{81-78}{3} = +1$$

حال اگر نمره استاندارد z مربوط به عدد ۷۷ را بخواهیم محاسبه کنیم، خواهیم داشت.

$$z = \frac{77-78}{3} = -0/33$$

1 . z value, normal score, standardized variable

اگر نمره استاندارد پایین‌تر از میانگین قرار بگیرد، علامت منفی خواهد گرفت و اگر بالاتر از میانگین قرار بگیرد، علامت مثبت خواهد گرفت.

البته نمرات استاندارد دیگری نیز وجود دارند که در ادامه به تعدادی از آن‌ها اشاره شده است که حفظ کردن آن‌ها خالی از لطف نیست:

نمره T از طریق فرمول مقابل محاسبه می‌شود. در این نمره، میانگین و انحراف استاندارد به ترتیب ۵۰ و ۱۰ است:

$$T= ۱۰ Z + ۵۰$$

نمره نه‌گانه[1] از طریق فرمول مقابل محاسبه می‌شود. در این نمره، میانگین و انحراف استاندارد به ترتیب ۵ و ۲ است:

$$T= ۲ Z + ۵$$

نمره آزمون هوشی وکسلر[2] از طریق فرمول مقابل محاسبه می‌شود. در این نمره، میانگین و انحراف استاندارد به ترتیب ۱۰۰ و ۱۵ است:

$$T= ۱۵ Z + ۱۰۰$$

نمره تراز شده آزمون‌های ورودی دانشگاه‌ها[3] از طریق فرمول مقابل محاسبه می‌شود. در این نمره، میانگین و انحراف استاندارد به ترتیب ۵۰۰ و ۱۰۰ است:

$$T= ۱۰۰ Z + ۵۰۰$$

در ادامه به یک مجموعه داده خواهیم پرداخت که نمرات استاندارد Z آن‌ها نیز نمایه شده است.

نمره خام	نمره Z
۵	-۰/۲۹
۷	۰/۵۸
۶	۰/۱۹
۴	-/۷۷
۵	-۰/۲۹
۶	۰/۱۹
۱	-۲/۲۳
۸	۱/۱۶
۶	۰/۱۹
۸	۱/۱۶

1 .stanine score
2 .stanine score
3 .College Entrance Examination board (CEEB)

در ادامه به اهمیت نمرات استاندارد Z در آمار توصیفی و استنباطی خواهیم پرداخت، امّا در اینجا به‌صورت مختصر به جدول بالا اشاره می‌کنیم:

۱. نمراتی که بالای میانگین قرار می‌گیرند دارای نمره استاندارد مثبت Z و نمراتی که پایین‌تر از میانگین قرار می‌گیرند دارای نمره استاندارد منفی Z هستند؛

۲. نمره استاندارد مثبت Z همیشه در طرف راست میانگین یعنی طرف راست منحنی بهنجار و نمره استاندارد منفی Z همیشه در طرف چپ میانگین یعنی طرف چپ منحنی بهنجار قرار می‌گیرد؛

۳. اینکه بگوییم نمره استاندارد یک نمره ۱+ شده است، دقیقاً مانند این است که بگوییم نمره استاندارد آن عدد ۱ انحراف استاندارد بالاتر از میانگین واقع شده است؛

۴. نمرات استاندارد توزیع‌های پراکندگی متفاوت، با یکدیگر قابل قیاس هستند. برای مثال، نمره استاندارد Z برای عدد ۱/۲۳۷ در توزیعی که دارای میانگین ۱۰۰ و انحراف استاندارد ۲۲ باشد، دقیقاً برابر نمره استاندارد Z برای عدد ۱/۲۳۷ در توزیعی که دارای میانگین ۵۵/۴ و انحراف استاندارد ۴/۳ است. هر دو عدد دارای ۱/۲۳۷ واحد استاندارد یا انحراف استاندارد از میانگین دارند؛

۵. نمرات استاندارد z در مقیاس فاصله‌ای جای دارند و میانگین ۰ و انحراف استاندارد ۱ دارد؛

۶. منظور از تبدیل نمرات خام به نمرات استاندارد یا تراز شده این است که آن‌ها را قابل مقایسه کنیم؛

برای پرسش‌های بیشتر، پرسش‌های ۷۱، ۷۲ و ۷۳ را مطالعه کنید.

پرسش ۷۱: چگونه می‌توانم از طریق اِکسل نمرات استاندارد Z را محاسبه کنم؟

برای محاسبه نمره استاندارد Z از طریق اِکسل از مراحل زیر پیروی کنید. بعد از اینکه مرحله چهارم را انجام دادید، نمره استاندارد Z همانند جدول زیر در کاربرگ اِکسل نشان داده خواهد شد.

	نمره خام	نمره Z
	۸۹	۰/۶۸
	۷۸	−۰/۱۶
	۴۹	−۲/۳۷
	۸۵	۰/۳۷
	۹۳	۰/۹۸
	۶۸	−۰/۹۲
	۷۹	−۰/۰۸
	۹۰	۰/۷۵
	۸۲	۰/۱۴
	۸۸	۰/۶۰
میانگین	۸۰/۱	
انحراف استاندارد	۳۱/۱	

۱. تمامی داده‌ها را در یک ستون فهرست کنید. برای این ستون داده‌ها عنوان X را در نظر بگیرید؛

۲. از تابع AVERAGE= استفاده کرده و مکان برونداد آن را دقیقاً مانند شکل بالا، در زیر ستون نمرات قرار دهید؛

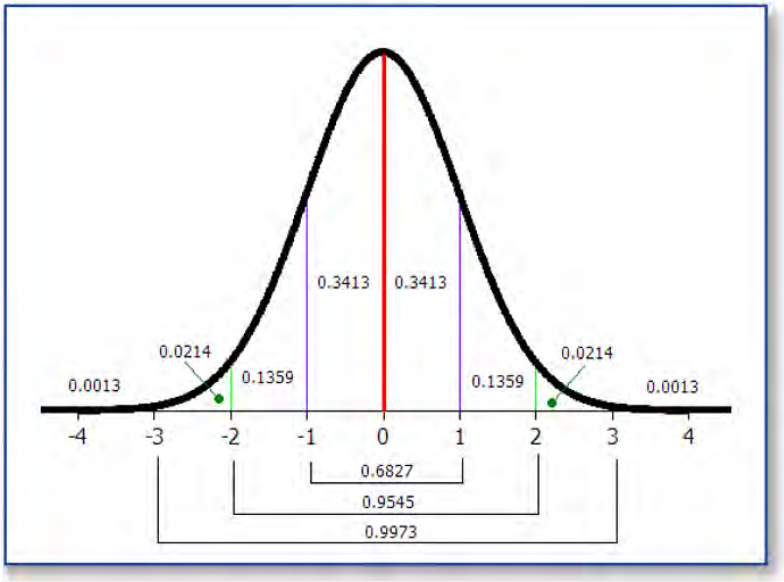
۳. از تابع STDEV.S= استفاده کرده و مکان برونداد آن را دقیقاً مانند شکل بالا، در زیر ستون میانگین قرار دهید؛

۴. در خانه C2 فرمول (B2-B12)/B13= را قرار دهید و برای هر یک از خانه‌های بعدی، به ترتیب B3، B4، B5، ... ، B11 را وارد کنید؛ در این فرمول B12 برابر است با میانگین، B13 برابر است با انحراف استاندارد و خود فرمول هم که فرمول نمره استاندارد Z است. به‌این‌ترتیب نمرات استاندارد Z برای این نمرات محاسبه شده است.

برای پرسش‌های بیشتر، پرسش‌های ۷۰، ۷۲ و ۷۳ را مطالعه کنید.

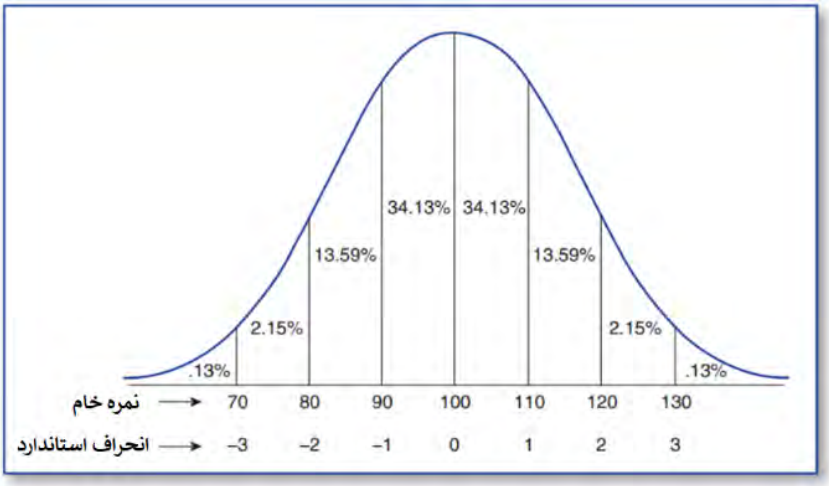
پرسش ۷۲: بین نمرات استاندارد Z و منحنی بهنجار چه ارتباطی وجود دارد؟

همان‌گونه که ذکر شد ویژگی اول منحنی بهنجار، این است که میانگین، میانه و نما برابر هستند؛ ویژگی دوم این است که منحنی بهنجار قرینه است، یعنی نیمه چپ کاملاً مشابه نیمه راست است؛ ویژگی سوم این است که منحنی بهنجار، مجانب[1] است، به این معنا که دامنه‌های سمت چپ و راست منحنی بهنجار هرگز محور افقی (X) را قطع نمی‌کنند. امّا منحنی بهنجار ویژگی‌های دیگری نیز دارد که در کاربست آمار استنباطی بسیار مهم و کاربردی هستند، مخصوصاً زمانی که می‌خواهیم درباره نمرات استاندارد Z به بحث بپردازیم. همان‌گونه که در تصویر زیر مشاهده می‌نمایید، منحنی بهنجار می‌تواند به بخش‌های[2] کوچکی تقسیم شود. این بخش‌های کوچک نشانگر درصد احتمال وقوع یک رویداد در هر یک از آن مناطق است.



همان‌گونه که ملاحظه می‌شود، ۳۴/۱۴ درصد نمرات بین میانگین و ۱ انحراف استاندارد بالاتر یا پایین‌تر از میانگین قرار دارند و ازآنجایی‌که منحنی بهنجار قرینه است، ۶۸/۲۷ درصد نمرات بین میانگین و ۱± انحراف استاندارد از میانگین قرار دارند؛ ۹۵/۴۵ درصد نمرات بین میانگین و ۲± انحراف استاندارد از میانگین قرار دارند؛ ۹۹/۷۳ درصد نمرات بین میانگین و ۳± انحراف استاندارد از میانگین قرار دارند؛ ۱۰۰ درصد نمرات بین میانگین و ۴± انحراف استاندارد از میانگین قرار دارند. پرواضح است که ۵۰ درصد نمرات در نیمه راست و ۵۰ درصد نمرات در نیمه چپ منحنی بهنجار قرار دارند. برای مثال همان‌گونه که در شکل زیر مشاهده می‌کنید، اگر میانگین یک مجموعه داده ۱۰۰ و انحراف استانداردش ۱۰ باشد، ۳۴/۱۴ درصد نمرات بین ۱۰۰ و ۱۱۰ قرار دارند.

1 . asymptotic
2 . segments



با توجه به منحنی بهنجار در شکل بالا، ۳۴/۱۳ درصد نمرات بین میانگین و ۱ انحراف استاندارد بالاتر یا پایین‌تر از میانگین قرار دارند؛ ۱۳/۵۹ درصد نمرات بین ۱ و ۲ انحراف استاندارد بالاتر یا پایین‌تر از میانگین قرار دارند؛ ۲/۱۵ درصد نمرات بین ۲ و ۳ انحراف استاندارد بالاتر یا پایین‌تر از میانگین قرار دارند؛ ۰/۱۳ درصد نمرات بین ۳ و ۴ انحراف استاندارد بالاتر یا پایین‌تر از میانگین قرار دارند. برای مثال اگر در یک مجموعه داده میانگین برابر ۱۰۰ و انحراف استاندارد ۱۰ باشد، ۲/۱۵ درصد احتمال دارد که نمرات بین ۱۲۰ تا ۱۳۰ قرار بگیرند. شخصی که نمره استاندارد ۲ گرفته باشد، در منحنی بهنجار بالاتر از ۹۷/۷۲ درصد (۱۳/۵۹+ ۳۴/۱۳+ ۳۴/۱۳ + ۱۳/۵۹ + ۲/۱۵ + ۰/۱۳) و پایین‌تر از ۲/۲۸ درصد (۰/۱۳+ ۲/۱۵) افراد قرار گرفته است. مفهوم ارزیابی احتمال وقوع یک نتیجه اهمیت بسیاری در اغلب آزمون‌های آمار استنباطی دارد.

برای پرسش‌های بیشتر، پرسش‌های ۶۶، ۶۷ و ۶۹ را مطالعه کنید.

پرسش ۷۳: نمرات استاندارد Z چگونه با آزمون فرضیه ارتباط دارد؟

مهم‌ترین دلیل قدرتمند بودن آزمون‌های آمار استنباطی این است که ما را در راستای تخمین زدن شاخص‌های جامعه از روی شاخص‌های نمونه توانمند می‌سازند. در واقع، این کمک آزمون‌های آماری استنباطی است که می‌توانیم بر اساس تحلیل داده‌های حاصل از نمونه، شاخص‌های جامعه را تخمین بزنیم. همه این تخمین زده‌ها بر قانون احتمال مبتنی است. در حقیقت، ما با کاربست یک مجموعه از قواعد آماری بر اساس اصول احتمال می‌توانیم تصمیم بگیریم که نتایج به‌دست‌آمده را واقعاً به ارتباط متغیرها نسبت دهیم، و یا به‌تصادف و خطای نمونه‌گیری و خطای ابزار. نمرات استاندارد، نماینده انواع متفاوتی از اعداد هستند که به‌احتمال وقوع یک رویداد مرتبط‌اند. برای مثال، فرض کنید که علاقه‌مندید بدانید سکه‌ای که در دست شماست، سالم است یا نه. برای این کار یک سکه را ۱۰ بار به هوا پرتاب می‌کنید. جامعه این پرتاب ۲۰ خواهد بود (پ ر، پ ر، پ ر، پ ر، پ ر، پ ر، پ ر، پ ر، پ ر، پ ر). انتظار دارید که در ۵۰ درصد موارد رو بیاید و در ۵۰ درصد موارد پشت. درواقع احتمال اینکه هر یک از طرفین بیاید پنج از ده، تا به تعبیر صحیح‌تر پنج‌دهم یا پنجاه صدم است. امّا، بعد از ده بار پرتاب پی‌درپی، چقدر اطمینان دارید که ۷ یا ۸ یا ۹ بار رو بیاید؟. در حقیقت، احتمال اینکه از ۱۰ پرتاب، ۸ بار رو بیاید ۰/۰۴ یا ۴ درصد است؛ احتمال اینکه از ۱۰ پرتاب، ۹ بار رو بیاید ۰/۰۱ یا ۱ درصد است؛ احتمال اینکه از ۱۰ پرتاب، ۱۰ بار رو بیاید ۰/۰۰۱ یا ۰/۱ درصد است. حال اگر نتایج تفاوت بسیار زیادی از این داشته باشد، احتمالاً خواهید گفت که سکه سالم نیست. و این تصمیم شما خواهد بود!.

در بیشتر موارد، معیار ۰/۰۵ یا ۵ درصد بسیار مناسب خواهد بود. به‌عبارت دیگر، اگر احتمال وقوع یک نتیجه کمتر از ۵ درصد باشد، می‌توان گفت که آن نتیجه بسیار نامحتمل است و یا ممکن است وقوع آن رویداد به دلیل تصادف و در مثال سکه، ناشی از حقه‌بازی یا سالم نبودن سکه باشد. اکنون اگر در ده پرتاب پی‌درپی، همه رو بیایند، بسیار احتمال دارد که سکه‌هایتان تقلبی باشد یا شاید شما حقه زده‌اید!. نمرات استاندارد Z اولین مقدمه برای تبیین و توضیح این مسئله هستند که وقوع یک رویداد بر اساس یک نظم و قانون بوده است یا مبتنی بر تصادف. ما به این درصد احتمالات در پرسش‌های بعدی خواهیم پرداخت.

برای پرسش‌های بیشتر، پرسش‌های ۶۶، ۶۹ و ۷۲ را مطالعه کنید.

بخش ۹: درک مفهوم معناداری

پرسش ۷۴: استنباط آماری چه کاربردی دارد؟

استنباط فرایندی است که در طی آن، آماردان‌ها بر مبنای نتایج حاصل از تحلیل داده‌های یک نمونه، ویژگی‌ها و شاخص‌های آماری جامعه مادر را برآورد می‌کنند. موفقیت در استنباط به چند عامل برمی‌گردد. نظیر اینکه یک نمونه تا چه میزان مُعرّف جامعه است و اینکه ابزارهای به‌کاربرده شده تا چه حد پایایی و روایی مناسب دارند. البته عوامل دیگری نیز نقش دارند. هنگامی‌که این عوامل به‌خوبی رعایت شده باشند، روش‌های استنباطی بسیار پرقدرت و کارآمد خواهند بود. برای بررسی مراحل انجام استنباط آماری، اجازه دهید با یک مثال آغاز کنیم.

پژوهشگران در مطالعه‌ای به بررسی رابطه فرهنگ و نوع رسانه با میزان یادآوری اخبار شنیده یا مشاهده‌شده پرداختند. شرکت‌کنندگان در این پژوهش از دو فرهنگ آمریکایی و اسپانیایی بودند و از طریق روزنامه، رایانه، رادیو و تلویزیون در معرض اخبار قرار گرفتند. پژوهشگران با تحلیل داده‌ها یافتند که بین فرهنگ‌های مختلف و همچنین شیوه‌های چهارگانه دریافت اخبار ازنظر یادآوری اخبار تفاوت معناداری وجود دارد. در اینجا چهار مرحله برای استنباط آماری وجود دارد که در مثال زیر مطرح خواهیم کرد:

۱. پژوهشگران ۷۲۰ دانشجوی رشته ارتباطات را انتخاب کردند. نظر پژوهشگران این بود که این ۷۲۰ دانشجو به‌خوبی می‌توانند معرف جامعه بزرگ‌تری آمریکایی-اسپانایی که مشغول به تحصیل در رشته ارتباطات هستند؛ ۲. هر شرکت‌کننده در معرض اخبار مختلف از شیوه‌های چهارگانه روزنامه، رایانه، رادیو و تلویزیون قرار می‌گرفت. سپس از طریق پرسشنامه‌های دارای پایایی و روایی، میزان یادآوری جزئیات اخبار تحت شرایط چهارگانه مورد بررسی قرار گرفت. سپس میانگین‌های مربوط به هر گروه در شرایط چهارگانه مختلف محاسبه و مقایسه شد؛ ۳. یکی از نتایج مطالعه این بود که تفاوت بین میانگین‌ها بر اساس شیوه‌های چهارگانه روزنامه، رایانه، رادیو و تلویزیون و فرهنگ‌های آمریکایی-اسپانیایی بوده است، نه بر اساس شانس، خطای نمونه‌گیری و خطای اندازه‌گیری؛ ۴. نتیجه دیگر مطالعه این بود که بین فرهنگ و شرایط با میزان یادآوری رابطه معناداری وجود دارد؛ در مجموع، به‌عنوان یک استنباط آماری، می‌توان یافته‌های حاصل از نمونه ۷۲۰ نفری را به همه دانشجویان رشته ارتباطات تعمیم داد.

برای مطالعه بیشتر درباره مثال موجود در این پرسش به مطالعه مرجع زیر بپردازید.

Facorro, L. B., & Defleur, M. L. (1993). A cross-cultural experiment on how well audiences remember news stories from newspaper, computer, television, and radio sources. Journalism & Mass Communication Quarterly, 70(3), 585-601.

برای پرسش‌های بیشتر، پرسش‌های ۷۵، ۷۷ و ۷۸ را مطالعه کنید.

پرسش ۷۵: مفهوم معناداری چیست و چرا مهم است؟

درک مفهوم معناداری[1] می‌تواند تا حدی چالش‌برانگیز باشد، امّا اساساً معناداری به این موضوع اشاره دارد که روابط بین متغیرها و یا تفاوت‌های بین گروه‌ها ناشی از شانس است یا ناشی از تفاوت‌های واقعی و تأثیرات که پژوهشگر اِعمال کرده است. همان‌گونه که از پرسش ۶۳ به یاد دارید، گفته بودیم که فرضیه صفر به این دلیل بسیار مهم است که نقطه آغاز در مواقعی است که فرض کرده‌ایم بین متغیرها ارتباطی وجود ندارد و یا بین گروه‌ها تفاوتی وجود ندارد.

به‌عبارت‌دیگر، زمانی که به مشاهده روابط و تفاوت‌های بین متغیرها می‌پردازیم، برای عدم مداخله سوگیری و پیش‌زمینه ذهنی در امر قضاوت، مبنا را بر فرضیه صفر قرار داده و مدعی عدم ارتباط و تفاوت بین متغیرها می‌شویم. مگر اینکه شواهد بسیار دقیق و قوی سبب رد شدن فرضیه صفر شود و اینجاست که مفهوم معناداری خود را نشان می‌دهد. آیا روابط یا تفاوت مشاهده‌شده ازنظر آماری معنادار است؟ یا روابط یا تفاوت مشاهده‌شده ناشی از شانس است؟.

برای مثال، اجازه دهید به دو گروه از زنان و مردان میان‌سالی که در برنامه کاهش وزن شرکت کردند، نگاهی بیندازیم. گروه اول به عنوان گروه ازمایش، برنامه کاهش وزن را دریافت کرد، امّا گروه دوم به عنوان گروه کنترل، هیچ‌گونه برنامه‌ای دریافت نکرد. در ابتدا و انتهای برنامه کاهش وزن، وزن دو گروه مقایسه شد. فرضیه صفر این بود که بین گروهی که در مداخله شرکت کردند در مقایسه با گروه کنترل گروه کنترل که هیچ مداخله‌ای دریافت نکرد، تفاوت معناداری از نظر وزن وجود ندارد. فرضیه تحقیق این بود که بین گروه مداخله در مقایسه با گروه کنترل، تفاوت معناداری از نظر وزن وجود دارد. نتایج نشان داد که در پس‌آزمون نسبت به پیش‌آزمون، به‌طور متوسط از گروه آزمایش ۲/۵ کیلو و از گروه کنترل ۰/۵ کیلو وزن کم شد.

امّا سؤالی که وجود دارد این است که کاهش وزن مشاهده شده در پس‌آزمون ناشی از برنامه کاهش وزن بوده است و یا ناشی از شانس، خطای نمونه‌گیری و خطای اندازه‌گیری؟. اینجاست که ما از اصول آمار استنباطی و استنباط آماری استفاده می‌کنیم. اگر تفاوت وزن بین گروه‌ها در پس‌آزمون بیش از ۳ پوند باشد، با درجه‌ای از اطمینان می‌گوییم که این تفاوت ناشی از برنامه کاهش وزن بوده است، نه شانس و خطا. اگر تفاوت وزن بین گروه‌ها در پس‌آزمون کمتر از ۳ پوند باشد، می‌بایست اعتراف کنیم که تفاوت به‌اندازه‌ای نیست که بتوانیم ازنظر آماری معنادار بدانیم. در این مطالعه، متوسط کاهش وزن ۴ پوند بوده است، که از میزان خطا و شانس بیشتر است و می‌توانیم مدعی شویم که این تفاوت ناشی از برنامه کاهش وزن بوده است.

ازنظر آماری، معناداری به‌احتمال آزمون (رد یا تائید کردن) فرضیه صفر گفته می‌شود. درواقع معناداری به ما کمک می‌کند که فرضیه صفر درست را رد نکنیم و فرضیه صفر نادرست را تائید نکنیم. سطح معناداری را برای پژوهش‌های علوم انسانی بازه ۰/۰۱ تا ۰/۰۵ در نظر گرفته‌اند. در حقیقت در پژوهش‌های علوم انسانی سعی می‌شود تا با احتمال ۰/۰۱ تا ۰/۰۵ خطا، یعنی با

1 . Significance

احتمال ۹۹ تا ۹۵ درصد اطمینان، فرضیه درست، تائید شود و فرضیه نادرست، رد شود. اینجاست که می‌توان تصمیم گرفت که آیا یافته‌ها بیانگر رابطه یا تفاوت معنادار است یا نه و اینکه نتایج به دست آمده ناشی از شانس بوده است یا نه.

برای پرسش‌های بیشتر، پرسش‌های ۷۴، ۷۶ و ۷۷ را مطالعه کنید.

پرسش ۷۶: خطای نوع ۱ و ۲ چیست؟

خطای نوع اول[1] یا آلفا (α) احتمال رد کردن فرضیه صفر درست است. درعین‌حال تمامی پژوهشگران قصد کاهش خطای نوع اول و دوم را دارند. حتماً به یاد دارید که فرضیه صفر بر اساس جامعه مادر مطرح می‌شود و ازآنجایی‌که هرگز به‌صورت مستقیم بررسی نمی‌شود، دقیقاً نمی‌توانیم بدانیم که فرضیه صفر درست است یا نادرست. برای مثال، ما یقیناً می‌پذیریم که بین زنان و مردان ازنظر نمره استعداد کلامی تفاوت وجود دارد. فرضیه صفر مطرح می‌کند که بین زنان و مردان ازنظر نمره استعداد کلامی تفاوت وجود ندارد. پژوهشگران با مطرح کردن فرضیه تحقیق و بررسی یافته‌ها دریافتند که بین زنان و مردان در نمره استعداد کلامی در سطح ۰/۰۱ درصد اطمینان تفاوت وجود دارد. درواقع تفاوت مشاهده‌شده ازنظر آماری معنادار بود.

میزان خطای نوع اول (α) به دست پژوهشگر است و می‌تواند از ۰/۰۱ تا ۰/۰۵ متغیر باشد. فرضیه تحقیق در این سطوح خطای نوع اول (α) آزمون می‌شود. اگر میزان خطای نوع اول یا سطح معناداری، کمتر از ۰/۰۵ به دست آید، فرضیه صفر رد می‌شود و فرضیه یک تائید می‌شود و با نماد $P<۰/۰۵$ نشان داده می‌شود که نشانگر وجود رابطه یا تفاوت معنادار است. اگر میزان خطای نوع اول یا سطح معناداری، بیشتر یا مساوی از ۰/۰۵ به دست آید، فرضیه صفر تائید می‌شود و فرضیه یک رد می‌شود و با نماد $P>۰/۰۵$ نشان داده می‌شود که نشانگر عدم وجود رابطه یا تفاوت معنادار است. احتمال خطای نوع اول ازآنجایی‌که توسط محقق انتخاب می‌شود، قبل از انجام پژوهش مشخص است (۰/۰۱ تا ۰/۰۵). برای مثال اگر پژوهشگری احتمال خطای نوع اول یا همان سطح معناداری را ۰/۰۵ در نظر بگیرد، در واقع این‌گونه متصوّر شده است که ۵ درصد یا ۰/۰۵ احتمال وجود دارد که فرض صفر درست به اشتباه رد شود. نکته اینجاست که پرتوان‌ترین آزمون آن است که با احتمال بیشتری بتواند H_0 نادرست را رد کند. به تعبیر صحیح‌تر توان آزمون برابر است با احتمال قبول H_1 درصورتی‌که H_1 درست باشد.

خطای نوع دوم[2] یا بتا (β) احتمال تائید کردن فرضیه صفر نادرست است. برای مثال، درحالی‌که در دنیای واقعی، بین زنان و مردان در استعداد کلامی تفاوت وجود نداشته باشد و فرضیه صفر بر عدم وجود تفاوت بین زنان و مردان تأکید کند، در این صورت با پذیرش فرضیه صفر نادرست، دچار خطای نوع دوم یا بتا خواهیم شد.

1. Type I error
2 . Type II error

جدول زیر می‌تواند در درک دو نوع خطای آلفا و بتا کمک شایانی کند:

جدول انواع خطا		فرض صفر (H_0)	
		درست	نادرست
قضاوت درباره فرض صفر (H_0)	رد	خطای نوع اول (مثبت دروغین)	استنباط درست (مثبت حقیقی)
	عدم رد	استنباط درست (مثبت حقیقی)	خطای نوع دوم (منفی دروغین)

به‌صورت تفصیلی‌تر می‌توان از جدول زیر کمک گرفت:

ماهیت فرضیه صفر		عملکرد شما	
		پذیرش فرضیه صفر	رد فرضیه صفر
فرضیه صفر درست		1 ☺ Bingo! You پذیرش فرضیه صفر درست؛ واقعاً تفاوتی بین گروه ها وجود ندارد؛	2 ☹ رد فرضیه صفر درست؛ واقعاً تفاوتی بین گروه ها وجود ندارد (آلفا)؛
فرضیه صفر نادرست		3 ☹ پذیرش فرضیه صفر نادرست؛ واقعاً تفاوتی بین گروه ها وجود دارد (بتا)؛	4 ☺ رد فرضیه صفر نادرست؛ واقعاً تفاوتی بین گروه ها وجود دارد؛

در یک جمع‌بندی می‌توان گفت، خطای نوع اول یا آلفا (α) درجه‌ای از خطر است. زمانی رخ می‌دهد که فرضیه صفر (عدم وجود رابطه) درست است اما محقق در آزمون فرضیه به‌اشتباه آن را رد کند (ردّ فرض صفر درست). یعنی رابطه وجود ندارد (فرضیه صفر) اما محقق اشتباهاً بگوید رابطه هست. هر چه خطای نوع اول کوچک‌تر باشد، درجه کمتری از خطرِ در اشتباه بودن در کمین است. هیچ آزمون آماریِ کاملاً آزاد از خطا وجود ندارد، زیرا هرگز نمی‌توانید کاملاً دقیق به روابط بین متغیرها پی ببرید. درواقع، پس از انجام آزمون‌های آماری پژوهشگر در مورد رد یا عدم رد فرضیه تصمیم می‌گیرد. اگر نتایج آزمون به‌گونه‌ای باشد که نتوان فرض صفر را رد کرد، جایی برای اثبات یا تائید فرضیه یک باقی نمی‌ماند؛ اما اگر فرضیه صفر رد شود به‌طور غیرمستقیم فرضیه پژوهشی تائید می‌شود.

حال اگر فرضیه صفر در حقیقت صحیح باشد ولی پژوهشگر تصمیم به رد آن بگیرد، خطای نوع اول رخ داده است. به‌بیان‌دیگر، این نوع خطا هنگامی انجام می‌شود که فرض صفر درست است اما پژوهشگر به‌اشتباه، فرض صفر را رد کرده و فرض پژوهش را می‌پذیرد.
خطای نوع دوم هنگامی انجام می‌شود که فرض پژوهش درست است اما پژوهشگر به‌اشتباه فرض پژوهش را رد کرده و فرض صفر را می‌پذیرد. این خطا زمانی رخ می‌دهد که فرضیه صفر (عدم وجود رابطه) غلط است اما محقق به‌اشتباه آن را بپذیرد (پذیرش فرض صفر نادرست). یعنی رابطه وجود دارد (فرضیه خنثی) اما محقق اشتباهاً بگوید رابطه وجود ندارد. خطای نوع دوم را با علامت بتا نشان می‌دهند.

خطاهای نوع اول و دوم دسته‌بندی انواع خطاها در آزمون فرض آماری هستند. دو نوع خطا ممکن است در طی آزمون فرض صفر در مقابل متمم آن رخ دهد: خطای نوع اول (مثبت کاذب[1]): فرض صفر درست باشد و آزمون فرض آن را رد کند؛ خطای نوع دوم (منفی کاذب[2]): فرض صفر درست نباشد و آزمون فرض آن را رد نکند؛

برای مثال، درباره سن دانشجویان، خطای نوع اول زمانی اتفاق می‌افتد که میانگین سن دانشجویان ۲۳ سال باشد و نتیجه بگیریم که ۲۳ سال نیست و خطای نوع دوم زمانی اتفاق می‌افتد که میانگین سن دانشجویان ۲۳ سال نباشد و نتیجه بگیریم ۲۳ سال است.

در یک آزمون آماری:

- میزان خطای نوع دوم را قدرت آزمون می‌نامند. توان آزمون درواقع همان $1-\beta$ است؛
- توان یک آزمون آماری عبارت است از: احتمال اینکه شما از خطای نوع دو دوری کنید؛
- هر چه توان یک آزمون بیشتر باشد احتمال وقوع خطای نوع دوم کمتر خواهد بود؛
- در مطالعات با توان آماری بالا، خیلی کم پیش می‌آید که در تشخیص اثرات تمرین یا نقش متغیر مستقل یا وجود ارتباط واقعی بین متغیرها اشتباه کنند؛
- با افزایش تعداد نمونه، به‌طور هم‌زمان آلفا و بتا کاهش می‌یابد و توان آمون افزایش می‌یابد؛ هر چه تأثیر متغیر مستقل بر وابسته افزایش یابد، خطای نوع دوم کاهش یافته و توان آزمون افزایش می‌یابد؛
- آلفا و بتا با یکدیگر رابطه عکس دارند، یعنی با افزایش آلفا، بتا کاهش می‌یابد و برعکس؛
- در آزمون فرضیه، قبل از دیدن مشاهدات، می‌بایست درباره دودامنه بودن آزمون تصمیم گرفت.

برای پرسش‌های بیشتر، پرسش‌های ۷۴، ۷۵ و ۷۷ را مطالعه کنید.

1 . False positive
2 . false negative

پرسش ۷۷: مراحل کاربرد یک آزمون آماری برای آزمون یک فرضیه پژوهشی چیست؟

هر فرضیه تحقیق، تحلیل آماری متناسب با خود را می‌طلبد. برای مثال، اگر پژوهشگری قصد مقایسه میانگین دو گروه زنان و مردان در بهداشت روان را داشته باشد، می‌بایست حتماً از آزمون آماری t مستقل استفاده کند. به علاوه، هر یک از آزمون‌های آماری توزیع پراکندگی خاص خود را دارند. درواقع، مقایسه شاخص‌ها و ویژگی‌های دو نمونه با یکدیگر از طریق یک آزمون آماری به ما اجازه می‌دهد به این نتیجه برسیم که آیا تفاوت یا تشابه شاخص‌ها و ویژگی‌های دو نمونه ناشی از شانس است یا خیر. نکته ای که اکنون می توان افزود این است که فرضیه ای که برای پارامتر جمعیت تدورین می شود، فرض صفر در نظر گرفته می‌شود و فرضیه مقابل آن فرضی است که به نظر می‌رسد طبق نمونه گرفته شده باید پارامتر جمعیت یا جامعه باشد. همچنین فرض‌هایی که به صورت یک رابطه تساوی بیان می‌شوند ($H_0:\mu=۵$) فرض‌های ساده نامیده می‌شوند ولی ممکن است این فرض‌ها به صورت نامساوی ($H_0:\mu\neq۵$) یا فرض‌های مرکب نیز باشند. تصمیم گیری در هر آزمون به مقدار آماره آن آزمون و سطح معناداری[1] محاسبه شده بر می گردد. در ادامه به مراحل کاربست یک آزمون آماری در یک فرضیه پژوهشی اشاره کرده‌ایم:

۱. یک فرضیه پژوهشی تدوین کنید؛

۲. خطای نوع اول یا آلفا (α) را مشخص کنید؛

۳. آزمون آماری مناسب را به‌کاربرده و ارزش به‌دست‌آمده[2] را بدست آورید؛

۴. ارزش یا مقدار بحرانی[3] مورد نظر را بر اساس جداول نقاط بحرانی معین کنید؛

۵. ارزش به‌دست‌آمده را با ارزش بحرانی مقایسه کنید. این مرحله بسیار مهم است زیرا در این مرحله، ارزش‌های محاسبه‌شده توسط آزمون آماری را با ارزش‌های بحرانی مقایسه می‌کنید؛

۶. اگر ارزش یا مقدار بدست‌آمده برابر یا بیشتر از ارزش بحرانی بود، فرضیه صفر (H_0) رد می شود و فرض یک خودبه‌خود تائید می شود که نشان می دهد روابط یا تفاوت مشاهده‌شده نه به شانس، بلکه به فرضیه تحقیق مربوط است. اگر ارزش به‌دست‌آمده کمتر از ارزش بحرانی بود، فرضیه صفر (H_0) را به‌راحتی تائید کنید!.

برای پرسش‌های بیشتر، پرسش‌های ۷۴، ۷۵ و ۷۶ را مطالعه کنید.

1 . p-value
2 . obtained value
3 . critical value

پرسش ۷۸: تفاوت بین معناداری آماری و معناداری منطقی چیست؟

هر گاه احتمال وقوع تصادفی یک رویداد بسیار اندک باشد و یا به احتمال زیاد دارای دلیلی غیر از شانس بوده است، آن را به لحاظ آماری معنادار گویند. «تفاوت معنادار آماری» به معنای وجود اختلاف زیاد یا مهم بین دو نمونه آماری نیست. عبارت «از نظر آماری معنادار[1]» معیاری است که به ما کمک می‌کند دریابیم که آیا همبستگی بین دو متغیر واقعاً قابل اعتماد است یا صرفاً به علت تصادف بوده است. اگر شما سکه‌ای را سه بار به هوا بیاندازید و هر سه بار شیر بیاید این به احتمال زیاد یک تصادف بوده است. اما اگر شما سکه‌ای را صد بار به هوا بیاندازید و هر صد بار شیر بیاید، شما می‌توانید تا حدی مطمئن باشید که دو روی سکه شیر است. در این شرایط می‌گوییم رویداد اول از نظر آماری «معنادار» نبوده است، اما رویداد دوم از نظر آماری «معنادار» است. در حقیقت این احتمالات نشان می‌دهد که رویداد مورد نظر واقعی است، یا صرفاً ناشی از تصادف بوده است. زمانی یک رابطه از نظر آماری «معنادار» تلقی می‌شود که به احتمال کمتر از ۵ درصد ناشی از تصادف بوده باشد، یعنی اگر پژوهش تکرار شود، به احتمال ۹۵ درصد به همان نتیجه قبلی خواهد انجامید. تعیین عدد ۹۵ دلبخواهی است و استانداردی است که ما انتخاب کرده‌ایم. یک نقطه قراردادی دیگر که اهمیت دارد ۹۹ است. وقتی رابطه بین دو متغیر با ۹۹ درصد اطمینان معنادار باشد، گفته می‌شود که نتیجه از نظر آماری بسیار معنادار است. یک اشتباه رایج در مورد معناداری آماری این است که گاه حتی در نوشته‌های علمی وجود تفاوت معنی‌دار آماری به معنای وجود تفاوتی بزرگ یا مهم تلقی می‌شود. در حالی که تفاوتی ناچیز (و احتمالاً کاملاً بدون اهمیت در عمل) برای دو جمعیت آماری بزرگ می‌تواند به معنای تفاوت معنادار آماری باشد.

امّا در کنار این موضوع، معناداری منطقی[2] نیز بسیار مهم است. برای مثال، در یک مطالعه بین میزان مصرف بستنی و تعداد رفتار خشونت‌آمیز همبستگی مثبت معناداری مشاهده شد. درواقع هر چه افراد بستنی بیشتری مصرف می‌کردند، گویی تعداد رفتار خشونت‌آمیز آن‌ها نیز بیشتر می‌شد. امّا یکی از همکاران پژوهش یادآور شد که متغیرهای دیگری نیز ممکن است وجود داشته باشند که عامل هر دو رفتار مصرف بستنی و تعداد رفتار خشونت‌آمیز باشند، نظیر فصل تابستان. درواقع در فصل تابستان که هوا گرم است، هم خشونت افراد افزایش می‌یابد و هم مصرف بستنی. بنابراین بسیار مهم است که معناداری‌های آماری مبتنی بر منطق باشند. با توجه به مثال ذکرشده، چند نکته مهم درباره تناسب معناداری آماری و منطقی وجود دارد:

۱. معناداری آماری، به‌خودی‌خود، معنادار نیست مگر اینکه رابطه یا تفاوت مفروض بین متغیرها بر اساس منطق شکل گرفته باشد؛

1 . statistically significant
2 . meaningfulness

۲. معناداری آماری، به‌خودی‌خود نمی‌تواند تفسیر شود مگر اینکه در بافتی که داده‌ها جمع‌آوری و تحلیل شده‌اند تفسیر شود. برای مثال، اگر شما سرپرست یک مدرسه باشید، آیا مایلید که دانش آموزان پایه اول را به این دلیل که برنامه تقویت حافظه کودکان پایه اول، توانسته است به‌صورت معناداری قدرت حافظه آن‌ها را افزایش دهد، مجدد در پایه اول ابقا کنید؟

۳. هرچند معناداری آماری، یک مفهوم بسیار مهم است، امّا تمام ماهیت و هویت یک پژوهش علمی نیست. اگر یک مطالعه به‌درستی طراحی شود، تائید فرضیه صفر نیز نتایج بسیار خوبی در بر خواهد داشت. برای مثال اگر یک درمان جدید، اثربخشی مناسبی نداشت، به این معنی است که این درمان ممکن است در آن نمونه یا زمینه اثربخشی نداشته باشد.

برای پرسش‌های بیشتر، پرسش‌های ۷۴، ۷۵ و ۷۶ را مطالعه کنید.

پرسش ۷۹: جعبه‌ابزار اِکسل چیست و چگونه می‌توانم از آن برای آزمون‌های آمار استنباطی استفاده کنم؟

اکسل دارای خصوصیات اصلی تمام نرم‌افزارهای صفحه‌گسترده است. با استفاده از جدولی متشکل از ردیف‌ها و ستون‌ها می‌توان، داده‌ها و اطلاعات را سازمان‌دهی و همچنین با استفاده از همین داده‌ها، به انجام محاسبات پرداخت. می‌توان با استفاده از رسم توابع و نمودارها، به تجزیه‌وتحلیل آماری اطلاعات پرداخت. اکسل به‌عنوان یک ابزار محاسباتی، قادر است به کمک توابعی که در آن قرار دارد، کار نوشتن پیچیده‌ترین فرمول‌ها را ساده کند. آگاهی از این قدرت مخفی اکسل، در وقت و هزینه بسیار صرفه‌جویی خواهد کرد. برای مثال می‌خواهیم به بررسی تفاوت میانگین در میزان پرش قهرمانان در ماه‌های بهار و پاییز بپردازیم. برای انجام تحلیل این مطالعه از آزمون تی مستقل استفاده خواهیم کرد.

برای انجام آزمون تی مستقل در محیط اِکسل مراحل زیر را انجام دهید:

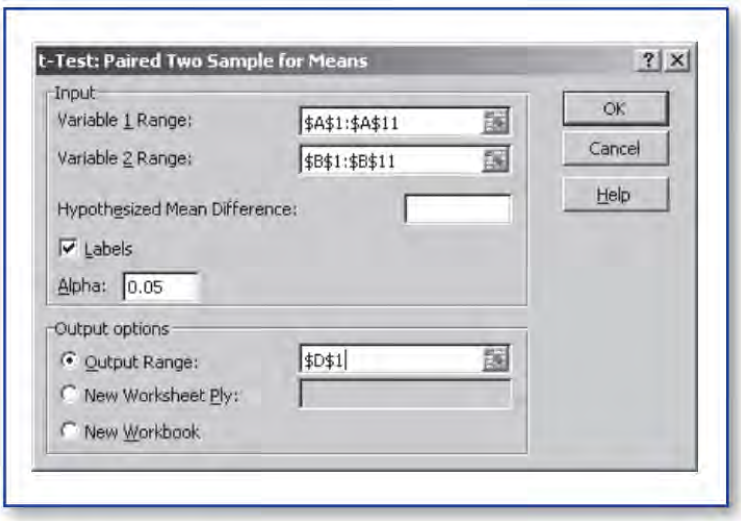
۱. در محیط اِکسل بر روی Data و سپس Data Analysis کلیک کنید؛

۲. در پنجره بازشده، گزینه t-Test: Paired Two Sample for Means را انتخاب کنید؛

۳. گروه‌های اول و دوم و سایر اطلاعات خواسته‌شده را تکمیل کنید؛

۴. گزینه Labels را تیک دار کنید؛

۵. دامنه Output را مشخص کنید. یعنی خانه‌ای که قصد دارید برونداد در آن نمایه شود را تعیین و سپس بر روی OK کلیک کنید؛



	A	B	C	D	E	F
1	Fall	Spring		t-Test: Paired Two Sample for Means		
2	18	19				
3	21	19			*Fall*	*Spring*
4	18	15		Mean	16.40	16.60
5	13	16		Variance	26.27	24.04
6	11	9		Observations	10	10
7	15	16		Pearson Correlation	0.93	
8	21	22		Hypothesized Mean Difference	0.00	
9	22	22		df	9	
10	19	20		t Stat	-0.33	
11	6	8		P(T<=t) one-tail	0.38	
12				t Critical one-tail	1.83	
13				P(T<=t) two-tail	0.75	
14				t Critical two-tail	2.26	

همان‌گونه که ملاحظه می‌کنید، آماره t محاسبه‌شده برابر ۰/۳۳- است؛ آماره t بحرانی در حالت دودامنه یا بی‌جهت برابر ۲/۲۶ است؛ سطح معناداری به‌دست‌آمده یا احتمال خطای نوع اول (آلفا) نیز ۰/۷۵ به‌دست‌آمده است. یعنی ممکن است از هر ۱۰۰ مورد مشاهده و نمونه‌گیری، ۷۵ بار به‌اشتباه فرض صفر درست را رد کنیم!. در حقیقت، سطح خطای به‌دست‌آمده بسیار بیشتر از مقدار سطح خطای قراردادی ۰/۰۱ یا ۰/۰۵ است، یعنی تفاوت مشاهده‌شده در میزان پرش قهرمانان در ماه‌های پاییز و بهار باهم تفاوت ندارد و اگر تفاوتی مشاهده شود ناشی از شانس یا خطای اندازه‌گیری و خطای نمونه‌گیری است. بنابراین فرض صفر تائید می‌شود.

برای پرسش‌های بیشتر، پرسش‌های ۷۴، ۷۷ و ۸۰ را مطالعه کنید.

پرسش ۸۰: یک تابع اِکسل چیست و چگونه می‌توانم از آن برای آزمون‌های آمار استنباطی استفاده کنم؟

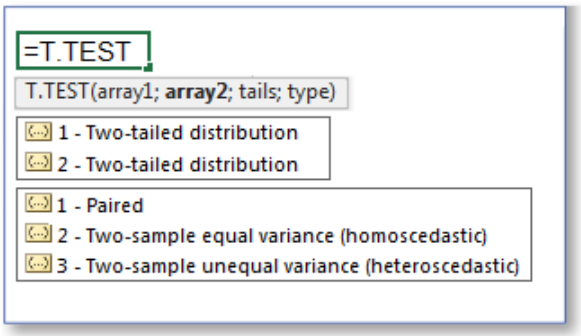
یکی از متداول‌ترین روش‌های موجود در اِکسل برای انجام محاسبات آماری، کاربرد توابع از پیش‌ساخته شده در آن است. برای مثال تابع AVERAGE= برای محاسبه میانگین یک مجموعه داده به‌کاربرده می‌شود. اگر در خانه‌ای تابع AVERAGE(A3:A5)= وارد شود، به این معنی است که میانگین اعداد واردشده در خانه‌های A3, A4, A5 در آن خانه محاسبه خواهد شد. هر یک از توابع موجود در اِکسل مزیت‌های خاص خود را دارد. برای انجام محاسبات آماری از طریق توابع موجود در اِکسل از مراحل زیر پیروی کنید.

۱. خانه‌ای که مایلید برونداد یک تابع یا فرمول در آن نمایه شود، را مشخص کنید؛

۲. تابع را در آن وارد کنید. برای این کار می‌توانید در آن خانه علامت = و حروف اول انگلیسی تابع موردنظر را وارد نمایید. برای نمونه می‌توان تابع AVERAGE= وارد کرد. با این کار مجموعه‌ای از توابع نمایه می‌شوند که می‌توانید تابع موردنظر خود را از میان آن‌ها انتخاب کنید؛

۳. با کلیک بر روی Enter محاسبه بلافاصله انجام خواهد گرفت. بعلاوه، بعد از انجام محاسبه با تغییر در هر یک از اعداد، برونداد به‌صورت خودکار تغییر خواهد کرد.

برای مثال، دو سری نمره وجود دارد که پژوهشگر قصد دارد به مقایسه بین میانگین اضطراب آن‌ها بپردازد و بررسی کند که آیا بین دو گروه ۱ و ۲ از نظر میزان اضطراب تفاوت معناداری وجود دارد یا خیر. نمرات اضطراب ۲ گروه عبارتند از: گروه ۱: ۵، ۶، ۳، ۴، ۶، ۵، ۴، ۶، ۵، ۷ و گروه ۲: ۷، ۶، ۵، ۶، ۷، ۸، ۸، ۹، ۸، ۶. همان‌گونه که از قبل می‌دانید، برای آزمون این فرضیه می‌بایست از آزمون تی مستقل استفاده نمود. به همین منظور در خانه برونداد تابع T.TEST= را وارد می‌کنیم. این تابع چهار نوع اطلاعات را از ما اخذ می‌کند: array1 که به مجموعه داده اول اشاره دارد، array2 که به مجموعه داده اول اشاره دارد، tails که به یک‌دامنه یا دودامنه بودن فرضیه پژوهش اشاره دارد و type که به زوج بودن گروه‌ها یا برابر بودن یا نبودن واریانس‌های دو گروه اشاره دارد. این مشخصات بر اساس اطلاعات پژوهشگر می‌بایست تکمیل شود.



پس از ورود نمرات دو مجموعه به اِکسل و مشخص کردن خانه برونداد، تمامی اطلاعات مذکور می‌بایست به‌درستی وارد شود. پس از انجام تمامی مراحل، ارزش تابع T.TEST= به مقدار ۰/۰۰۳ به دست آمد که به دلیل کمتر بودنش از میزان قراردادی ۰/۰۵ که خطای نوع اول (رد

فرضیه صفر درست) است، بیانگر تفاوت بین میانگین اضطراب گروه‌هاست. بنابراین میزان احتمال خطای به‌دست‌آمده برای سنجش تفاوت دو گروه به مقدار ۰/۰۰۳ است که بسیار کمتر از میزان خطای قراردادی نوع اول ۰/۰۵ است که نشان می‌دهد با این شاخص‌های دو گروه، احتمال ۳ در ۱۰۰۰ وجود دارد که تفاوت مشاهده‌شده میانگین‌ها ناشی از خطای نمونه‌گیری، خطای اندازه‌گیری و یا شانس و تصادف باشد. این در حالی است که متخصصان علم آمار احتمال خطای ۵ در ۱۰۰ را ازنظر علم آمار معقول و استاندارد می‌دانند. بنابراین فرضیه صفر رد می‌شود و فرضیه یک که بیانگر وجود تفاوت بین دو گروه است، تائید می‌گردد.

برای پرسش‌های بیشتر، پرسش‌های ۷۴، ۷۶ و ۷۹ را مطالعه کنید.

بخش ۱۰: درک تفاوت‌های بین گروهی

پرسش ۸۱: چگونه بدانم که از کدام آزمون آماری استفاده کنم؟

همان‌گونه که می‌دانید، یک پژوهشگر برای انجام پژوهش خود به ترتیب چهار مرحله را دنبال می کند: ۱. انتخاب نمونه؛ ۲. انتخاب روش؛ ۳. انتخاب ابزار یا روش جمع‌آوری اطلاعات؛ ۴. انتخاب شیوه تلخیص و جمع‌بندی اطلاعات. در مرحله چهارم که شیوه تلخیص و جمع‌بندی اطلاعات است می‌بایست هوشمندی خود را بکار بگیرد. این در حالی است که بیش از صد نوع مختلف از آزمون‌های آماری وجود دارد که هر یک به‌منظور خاص و برای تحلیل نوع خاصی از معادلات به‌کاربرده می‌شوند. برای مثال اگر بخواهید میانگین دو گروه مستقل را با یکدیگر مقایسه کنید، می‌بایست از آزمون t برای گروه‌های مستقل استفاده کنید؛ اگر بخواهید میانگین دو گروه وابسته را با یکدیگر مقایسه کنید، می‌بایست از آزمون t برای گروه‌های وابسته استفاده کنید؛ و اگر بخواهید میانگین یک گروه را با یک معیار خاص از قبل تعیین‌شده مقایسه کنید، می‌بایست از آزمون t تک نمونه‌ای استفاده کنید؛ و اگر بخواهید به بررسی رابطه بین دو متغیر بپردازید، می‌بایست از آزمون ضریب همبستگی استفاده کنید. بحث درباره تمامی آزمون‌های آماری بسیار گسترده است، امّا در این کتاب سعی کردیم تا حدی به مهم‌ترین آزمون‌هایی که ممکن است برایتان کاربرد داشته باشند اشاره کنیم. در شکل زیر به نمایه آزمون‌های آماری در موقعیت‌های متفاوت اشاره شده است.

از این نمودار می‌توانید برای درک بهتر مواقع کاربست آزمون‌های آماری پی ببرید. این نمودار برای وضوح بیشتردر صفحه بعد نمایه شده است.

برای پرسش‌های بیشتر، پرسش‌های ۸۲، ۸۴ و ۸۶ را مطالعه کنید.

آزمون های آماری
پارامتریک :
• هنگامی که توزیع جامعه نرمال است.
• برای متغیرهایی از نوع فاصله ای و نسبی.
• حداقل شاخص آماری: میانگین و واریانس
ناپارامتریک :
• هنگامی که توزیع جامعه نرمال نیست.
• برای متغیرهایی از نوع اسمی و ترتیبی
• حداقل شاخص آماری: میانه و مد.
T تک نمونه ای
T وابسته
T دو نمونه مستقل
T ولچ
T هتلینگ
تحلیل واریانس (F/ANOVA)
تحلیل واریانس چند عاملی
تحلیل کواریانس چندعاملی
ضریب همبستگی گشتاوری پیرسون
علامت تک نمونه ای
علامت زوجی
ویلکاکسون (علامت)
مان – ویتنی (آزمون U)
کروسکال – والیس (آزمون H)
فریدمن
نیکوئی برازش (خی دو)
کولموگروف - اسمیرنف
تقارن توزیع
میانه
مک نمار
آزمون کوکران
ضریب همبستگی اسپیرمن
برای آزمون فرض پیرامون میانگین یک جامعه استفاده می شود.
برای آزمون فرض پیرامون دو میانگین از یک جامعه استفاده می شود.
جهت مقایسه میانگین دو جامعه استفاده می شود. با فرض برابر بودن واریانس دو جامعه.
جهت مقایسه میانگین دو جامعه استفاده می شود. با فرض عدم برابری واریانس دو جامعه.
برای مقایسه چند میانگین از دو جامعه استفاده می شود. (مقایسه دو جامعه براساس میانگین چندین صفت)
از این آزمون به منظور بررسی اختلاف میانگین چند جامعه آماری استفاده می شود.
از این آزمون (MANOVA) به منظور بررسی اختلاف چند میانگین از چند جامعه آماری استفاده می شود.
چنانچه در MANOVA بخواهیم اثر یک یا چند متغیر کمکی را حذف کنیم، استفاده می شود.
برای محاسبه همبستگی (شدت و جهت ارتباط دو متغیر) دو مجموعه داده استفاده می شود.
برای آزمون فرض پیرامون میانگین یک جامعه استفاده می شود.
برای آزمون فرض پیرامون دو میانگین از یک جامعه استفاده می شود.
همان آزمون علامت زوجی است که در آن اختلاف نسبی تفاوت از میانگین لحاظ می شود.
جهت مقایسه میانگین دو جامعه استفاده می شود.
از این آزمون به منظور بررسی اختلاف میانگین چند جامعه آماری استفاده می شود.
جهت آزمون متغیرهایی با مقیاس رتبه ای در k گروه وابسته و بررسی تفاوت آن در گروه ها.
برای ارزیابی همقوارگی متغیرهای اسمی به کار می‌رود. (مقایسه یک توزیع نظری با توزیع مشاهده شده)
مقایسه یک توزیع نظری با توزیع مشاهده شده و نشان دادن نرمال بودن توزیع داده‌ها (نوعی آزمون نیکوئی برازش)
برای بررسی مشاهدات زوجی درباره متغیرهای دو ارزشی استفاده می شود.
جهت مقایسه میانه دو جامعه استفاده می شود و برای kجامعه نیز قابل تعمیم است.
برای بررسی مشاهدات زوجی درباره متغیرهای دو ارزشی استفاده می شود.
تعمیم آزمون مک نمار در kنمونه وابسته است.
برای محاسبه همبستگی دو مجموعه داده که به صورت ترتیبی قرار دارند استفاده می شود.

پرسش ۸۲: آزمون t میانگین‌های مستقل چیست و یک مثال برای چگونگی کاربست آن کدام است؟

ویلیام سیلی گوسِت[1] (۱۹۳۷-۱۸۷۶) آماردانی مشهور که با نام مستعار دانشجو[2] کارهایش را در زمینه توزیع تی (t) منتشر کرد. گویا وی نمی‌توانست به نام اصلی‌اش مقالاتش را امضا کند و به‌ناچار از نام مستعار استیودنت (دانشجو) استفاده می‌کرد. آزمون تی استیودنت[3] برای ارزیابی میزان هم‌قوارگی یا یکسان بودن و نبودن میانگین نمونه‌ای با میانگین جامعه در حالتی به کار می‌رود که انحراف معیار جامعه مجهول باشد. چون توزیع t در مورد نمونه‌های کوچک با استفاده از درجات آزادی تعدیل می‌شود، می‌توان از این آزمون برای نمونه‌های بسیار کوچک استفاده نمود. همچنین این آزمون مواقعی که خطای استاندارد جامعه نامعلوم و خطای استاندارد نمونه معلوم باشد، کاربرد دارد. گوسِت، توزیعی یا در حقیقت خانواده‌ای از توزیع‌ها است که با استفاده از آن‌ها می‌توان در شرایطی که محقق اطلاعات جامعه را در دست ندارد، با نمونه‌گیری از جامعه، فرضیه‌هایی را آزمون کند. این توزیع‌ها به توزیع t یا منحنی‌های t و یا توزیع t استیودنت معروف هستند. برای به کار بردن این آزمون، متغیر مورد مطالعه باید در مقیاس فاصله‌ای و شکل توزیع آن بهنجار باشد. آزمون تی استیودنت در حالت‌های زیر کاربرد دارد: آزمون t تک نمونه‌ای؛ آزمون tمستقل؛ آزمون t جفت شده که گاهی نوع مکرر، وابسته یا همبسته نیز خوانده می‌شود.

آزمون t تک نمونه‌ای که ساده‌ترین نوع آزمون‌های t است جهت تعیین این‌که آیا میانگین مشاهده‌شده در نمونه که به‌صورت تصادفی از جامعه انتخاب شده است، مقداری برابر با میانگین مفروض جامعه دارد یا خیر، به کار می‌رود. برای مثال میانگین افسردگی دانشجویان دانشگاه با توجه به مقیاس گلدبرگ برابر ۷ محاسبه شده است. حال یکی از روان شناسان دانشگاه معتقد است دانشجویانی که ورزش می‌کنند افسردگی کمتری دارند. بر همین اساس ۱۰۰ نفر از دانشجویان ورزشکار را به‌صورت تصادفی انتخاب کرده و میزان افسردگی آنان را اندازه‌گیری می‌کند، میانگین افسردگی این دانشجویان برابر ۶ به دست می‌آید. آیا این تفاوت ۱ نمره‌ای بین میانگین افسردگی دانشجویان ورزشکار و میانگین افسردگی جامعه دانشجویی یک تفاوت واقعی است یا اینکه بر اساس ادعای فرض صفر، حاصل خطای نمونه‌گیری است و تفاوت واقعی بین نمونه و جامعه وجود ندارد. ازآنجایی‌که این پژوهشگر انحراف استاندارد جامعه را نمی‌داند و می‌خواهد جهت بررسی معنی‌داری این تفاوت از یک آزمون آماری استفاده کند، آزمون t تک نمونه‌ای را به کار می‌برد و انحراف استاندارد نمونه را به‌عنوان برآوردی از انحراف استاندارد جامعه در نظر می‌گیرد. امّا گاهی دو گروه مستقل برای مقایسه وجود دارند. از آزمون t مستقل برای مقایسه میانگین‌های دو نمونه مستقل استفاده می‌گردد. مانند زمانی که می‌خواهیم معدل تحصیلی دو نمونه مستقل (مانند زن و مرد) را باهم مقایسه کنیم. فرضیه صفر مبنی بر این است

1 . William Sealy Gosset
2 . student
3 . T student

که تفاوت معناداری بین میانگین دو نمونه وجود ندارد درحالی‌که فرض یک بر تفاوت معنادار میانگین دو نمونه دلالت دارد. در این آزمون اگر فرض برابری واریانس‌های دو نمونه مطرح باشد، این آزمون از واریانس ادغام‌شده دو گروه استفاده می‌کند ولی اگر واریانس‌های دو نمونه متفاوت باشند، استفاده از آزمون t با واریانس‌های جداگانه پیشنهاد می‌گردد.

برای پرسش‌های بیشتر، پرسش‌های ۸۱، ۸۳ و ۸۴ را مطالعه کنید.

پرسش ۸۳: چگونه می‌توانم از طریق اِکسل به آزمون تفاوت بین میانگین‌های مستقل بپردازم؟

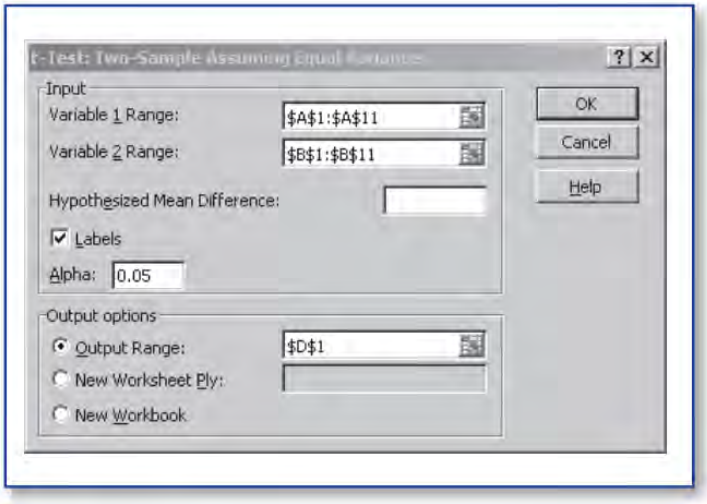
راه های مختلفی برای مقایسه میانگین های دو گروه مستقل از طریق کاربرد نرم افزار اِکسل وجود دارد. دو روش عبارتند از توابع T.DIST= و T.TEST= که در محیط نرم افزاری اِکسل وجود دارند. البته روش بسیار ساده تر و مستقیم تر به کاربردن ابزار تجزیه و تحلیل داده ها[1] است. اکنون داده های مثال زیر را در نظر بگیرید. می خواهیم میانگین این دو گروه مستقل و غیرمرتبط به هم را مقایسه کنیم. گروه اول یک تمرین ۲ ساله بهمراه دریافت ورزشهای تکمیلی را دریافت کرده است، درحالی‌که گروه دوم این مداخله را دریافت نکرده است. متغیر وابسته میزان رضایت مشتریان بود که در یک دامنه ۱ تا ۱۰ نمره گذاری می شد. فرضیه محقق این بود که میانگین رضایت مشتریان در گروه مداخله بصورت معناداری بیشتر از گروه کنترل خواهد بود. فرضیه صفر این بود که بین میانگین نمره رضایت مشتریان در دو گروه مداخله و کنترل، تفاوت معناداری وجود ندارد. نمرات ۲ گروه عبارتند از:

گروه مداخله: ۶، ۸، ۷، ۵، ۹، ۹، ۸، ۷، ۸، ۹

گروه غیرمداخله: ۶، ۴، ۷، ۸، ۶، ۹، ۸، ۹، ۸، ۷

برای انجام آزمون تی مستقل در محیط اِکسل مراحل زیر را انجام دهید:

۱. در محیط اِکسل بر روی Data و سپس Data Analysis کلیک کنید؛



۲. در پنجره بازشده، گزینه Two-Sample t-Test: Assuming Equal Variances را انتخاب کنید؛

۳. دامنه اعداد و سایر اطلاعات خواسته‌شده را تکمیل کنید؛

۴. بر روی OK کلیک کنید؛

1 . Data Analysis ToolPak

	A	B	C	D	E	F
1	Without Training	With Training		t-Test: Two-Sample Assuming Equal Variances		
2	6	6				
3	8	4			*Without Training*	*With Training*
4	7	7		Mean	7.6	7.2
5	5	8		Variance	1.8	2.4
6	9	6		Observations	10	10
7	9	9		Pooled Variance	2.1	
8	8	8		Hypothesized Mean Difference	0	
9	7	9		df	18	
10	8	8		t Stat	0.6	
11	9	7		P(T<=t) one-tail	0.27	
12				t Critical one-tail	1.73	
13				P(T<=t) two-tail	0.55	
14				t Critical two-tail	2.10	

همان‌گونه که ملاحظه می‌کنید، ارزش t محاسبه‌شده برابر ۰/۶ است و سطح معناداری مربوط به آن در حالت یک‌دامنه، ۰/۲۷ است که از مقدار خطای نوع اول ۰/۰۵ بسیار بیشتر است. این نتایج نشان می دهد که بین میانگین نمره رضایت مشتریان در دو گروه مداخله و کنترل، تفاوت معناداری وجود ندارد.

برای پرسش‌های بیشتر، پرسش‌های ۸۱، ۸۲ و ۸۴ را مطالعه کنید.

پرسش ۸۴: آزمون t بین میانگین‌های وابسته چیست و یک مثال برای چگونگی کاربست آن کدام است؟

برخی از مسائل پژوهشی به شرایطی مربوط می‌شوند که در آن‌ها از دو نمونه که مستقل نیستند، استفاده می‌شود. چنین نمونه‌هایی را وابسته[1] می‌نامند. نمونه وابسته با طرح پژوهشیِ اندازه‌های مکرر به‌کاربرده می‌شوند. در این طرح چون هر یک از آزمودنی‌ها در دو شرایط مختلف پژوهشی مورد اندازه‌گیری قرار می‌گیرند، این آزمون گاهی اوقات برای گروه‌های همتا، یعنی گروه‌هایی که بر اساس یک یا چند متغیر به‌جز متغیر مورد تحلیل، یکسان شده‌اند، به‌کاربرده می‌شود. هدف آزمون t برای گروه‌های وابسته همانند آزمون نمونه‌های مستقل عبارت است از پاسخگویی به این سؤال که: آیا تفاوت میانگین‌ها ناشی از عوامل شانس است یا حاصل تفاوت واقعی بین میانگین‌های جامعه‌ای که نمونه از آن‌ها به‌صورت تصادفی انتخاب‌شده است.

از این آزمون زمانی استفاده می‌گردد که هر فرد یا نمونه‌ی یکسانی در دو وضعیت متفاوت مورد آزمون قرار گیرد. یعنی برای یک فرد دو اندازه‌گیری داشته باشیم. مثلاً فشارخون پاسخگویان قبل و بعد از کنکور را ثبت کنیم و سپس آزمون کنیم که آیا میانگین مقادیر در دو نوبت تفاوت معناداری با یکدیگر دارد یا خیر. مزیتی که آزمون تی با نمونه‌های جفت دارد این است که، اگر تفاوتی وجود داشته باشد به‌آسانی کشف می‌شود. مثلاً هنگامی‌که فشارخون تعدادی آزمودنی قبل و بعد از یک امتحان ثبت شود با اطمینان زیادی می‌توان اختلاف مشاهده‌شده بین فشارخون قبل و بعد از امتحان را به امتحان نسبت داد.

خواندن ممناسب یکی از مولفه هایی که دانش آموزان موفق و ناموفق را از یکدیگر متمایز می کند. پژوهشگران این حوزه بر این نکته توجه دارند که چگونه یادگیری خواندن اتفاق م یافتد و موثرترین روش مداخله ای برای یادگیری خواندن اثربخش چیست. در همین راستا، در یک مطالعه ۸۰ کودک پیش دبستانی را وارد دو مداخله آموزشی ۱۸ هفتگی با هدف یادگیری واژگان کردند. در یک روش به آموزش واژگان به شیوه گسترده تر امّا بدون مرور نظام مند واژگان و در روش دیگر به آموزش واژگان به شیوه گسترده تر بهمراه مرور نظام مند واژگان پرداختند. سووال اولیه این بود که ایا بین میانگین تتعداد واژگان آموخته شده توسط کودکان بر اساس شیوه آموزش تفاوت وجود دارد یا خیر. تمامی ۸۰ کودک هر دو شرایط آموزشی را دریافت کردند. همانگوننه که متوجه شده اید نمرات کودکان به هم وابسته است و در چنین شرایطی بهترین شیوه آماری برای محاسبه تفاوت میانگین ها، آزمون تی وابسته است. نتایج نشان داد که میانگین تعداد واژگان یاد گرفته شده در روش مرور نظام مند تقریباً دو برابر روش آموزش واژگان بود.

استفاده از آزمون تی با نمونه‌های جفت منوط به شرایطی است که عبارت‌اند از: توزیع جامعه بهنجار باشد؛ نمونه‌ها به‌صورت اتفاقی انتخاب‌شده باشند؛ متغیر مستقل در سطح سنجش اسمی و متغیر وابسته در سطح سنجش فاصله‌ای (یا نسبی) باشد؛ برای هر فرد یا نمونه آماری دو

1 . dependent

اندازه‌گیری وجود داشته باشد؛ آزمون‌های مک‌نمار، علامت و ویلکاکسون معادل‌های ناپارامتری آزمون تی با نمونه‌های جفت هستند. از این آزمون‌ها هنگامی استفاده می‌گردد که شرایط استفاده از آزمون تی با نمونه‌های جفت فراهم نباشد.

برای مطالعه بیشتر درباره مثال موجود در این پرسش به مطالعه مرجع زیر بپردازید.

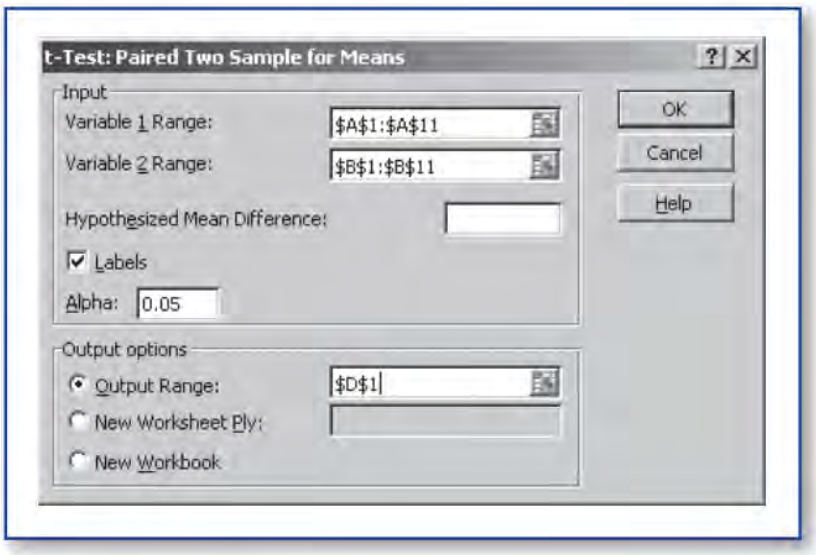
Zipoli, R. P., Jr., Coyne, M. D., & McCoach, B. D. (2011). Enhancing vocabulary intervention for kindergarten students: Strategic integration of semantically related and embedded word review. Remedial and Special Education, 32(2), 131ات43.

برای پرسش‌های بیشتر، پرسش‌های ۸۱، ۸۲ و ۸۵ را مطالعه کنید.

پرسش ۸۵: چگونه می‌توانم از طریق اِکسل به آزمون تفاوت بین میانگین‌های وابسته بپردازم؟

با کمک یک مثال قصد داریم تفاوت بین میانگین های یک گروه که در دو زمان متفاوت اندازه گیری شده اند را بررسی کنیم. فرضیه پژوهش این است که افزایش وزن سبب تغییر در تراکم استخوان می شود. از ۱۵ نفر دانشجویان دختر که همگی دارای ۲۵ سال سن بودند، خواسته شد تا در یک برنامه کاهش وزن یک‌ساله شرکت کنند. تراکم استخوان آن‌ها در پاییز و بهار مورد ارزیابی واقع شد که اعداد مرتبط با تراکم استخوان آنها عبارت شد از: بهار (۵، ۳، ۴، ۴، ۳، ۲، ۳، ۲، ۲، ۳) و پاییز (۳، ۲، ۲، ۳، ۴، ۳، ۳، ۲، ۱ ، ۲). از آزمون t یک دامنه در سطح معناداری ۰/۰۵ استفاده شد و انتظار بر این بود که در زمان دوم اندازه گیری بر تراکم استخوان افزوده شده باشد. متغیر هدف یا وابسته عبارت بود از تراکم استخوان که در یک مقیاس ۱ تا ۵ اندازه گیری می شد. تمامی شرکت کنندگان در برنامه مدیریت وزن شرکت کردند و دو ارزیابی در زمانهای قبل و بعد از مداخله از آنها به عمل آمد.

برای انجام آزمون تی وابسته در محیط اِکسل مراحل زیر را انجام دهید: ۱. در محیط اِکسل بر روی Data و سپس Data Analysis کلیک کنید؛ ۲. در پنجره بازشده، گزینه t-Test: Paired Two Sample for Means را انتخاب کنید؛ ۳. دامنه اعداد و سایر اطلاعات خواسته شده را تکمیل کنید؛ ۴. بر روی OK کلیک کنید؛



	A	B	C	D	E	F
1	Fall Assessment	Spring Assessment		t-Test: Paired Two Sample for Means		
2	3	5				
3	2	3			*Fall Assessment*	*Spring Assessment*
4	2	4		Mean	2.5	3.1
5	3	4		Variance	0.72	0.99
6	4	3		Observations	10	10
7	3	2		Pearson Correlation	0.33	
8	3	3		Hypothesized Mean Difference	0	
9	2	2		df	9	
10	1	2		t Stat	-1.77	
11	2	3		P(T<=t) one-tail	0.06	
12				t Critical one-tail	1.83	
13				P(T<=t) two-tail	0.11	
14				t Critical two-tail	2.26	
15						

همان‌گونه که ملاحظه می‌کنید، ارزش t محاسبه‌شده برابر ۱/۷۷- است و سطح معناداری مربوط به آن در حالت یک‌دامنه، ۰/۰۶ است که از مقدار خطای نوع اول ۰/۰۵ بیشتر است. این نتیجه نشان می‌دهد که برنامه مداخله ای مدیریت وزن بر تراکم استخوان تاثیری نداشته است.

برای پرسش‌های بیشتر، پرسش‌های ۸۰، ۸۱ و ۸۴ را مطالعه کنید.

پرسش ۸۶: تحلیل واریانس ساده چیست و یک مثال برای چگونگی کاربست آن کدام است؟

یک تحلیل واریانس ساده[1] (آنوا) به‌منظور ارزیابی معنادار بودن تفاوت بین دو یا چند میانگین از یک یا چند گروه می‌پردازد. به این تفاوت‌ها، اثرات اصلی[2] گفته می‌شود. واژه "ساده" به این دلیل به‌کاربرده می‌شود که اثر یک متغیر یا عامل مورد ارزیابی قرار می‌گیرد. میزان F به‌دست‌آمده در آزمون تحلیل واریانس ساده حاصل یک کسر است؛ کسری که در صورت آن میزان تفاوت بین گروه‌ها و در خرج آن میزان تغییرپذیری درون هر گروه قرار دارد. هر چه میزان تفاوت‌های بین گروهی افزایش یابد و درعین‌حال تفاوت‌های درون‌گروهی کاهش یابد، F به‌دست‌آمده بزرگ‌تر خواهد بود و درنتیجه نشان‌دهنده آن است که تفاوت‌های مشاهده‌شده واقعی است نه به دلایل خطا و تصادف. در ساده‌ترین شکل آن، ANOVA آزمون آماری را فراهم می‌کند که برابری میانگین‌های گروه‌های متفاوت را می‌آزماید، و درنتیجه آزمون تی استیودنت (t-test) را به بیش از دو گروه تعمیم می‌دهد.

در تحلیل واریانس فرضیه صفر این است که اختلافی بین میانگین جمعیت‌ها وجود ندارد و در مقابل فرضیه یک آن است که حداقل بین میانگین دو گروه از این جمعیت‌ها اختلاف معنی‌داری وجود دارد. به‌عنوان‌مثال، اغلب معلمین و والدین مطرح می‌کنند که یادگیری خواندن در سنین خیلی زود بسیار اهمیت دارد. لذا از سنین کودکی سعی می‌کنند با کتاب‌های تصویری، مهارت‌های خواندن کودکان خود را افزایش دهند. در یک پژوهش، به مطالعه اثربخشی کاربرد کتاب‌های خواندن معمولی کودکان، کتاب‌های دوبعدی کودکان و کتاب‌های تصویری برجسته کودکان بر مهارت خواندن حروف الفبا در کودکان ۳۶ ماهه پرداخته شد. برای تحلیل داده‌ها از آزمون تحلیل واریانس ساده استفاده شد. این پژوهشگران یافتند که شیوه آموزش (کاربرد کتاب‌های خواندن معمولی کودکان، کتاب‌های دوبعدی کودکان و کتاب‌های تصویری برجسته کودکان) تأثیر معناداری بر مهارت خواندن حروف الفبا در کودکان ۳۶ ماهه داشته است. نتایج نشان داد کودکانی که از شیوه کاربرد کتاب‌های خواندن معمولی کودکان استفاده می‌کردند نسبت به کودکان تمرین کرده با دو شیوه دیگر، کلمات بیشتری یاد گرفته بودند. همچنین نتایج نشان داد کتاب‌هایی که تصاویر جذاب‌تری داشتند، الزاماً عملکردی یادگیری بهتری در بر نداشتند. تحلیل واریانس ممکن است یکه راهه، دوراهه، سه‌راهه و عاملی اجرا شود (منظور از راه تعداد متغیرهای مستقل در تحلیل واریانس است)؛ از آزمون لوین[3] برای ارزیابی برابری واریانس-های گروه‌ها استفاده می‌شود؛ به برابری واریانس‌ها، واریانس‌همسانی[4] نیز می‌گویند.

برای مطالعه بیشتر درباره مثال موجود در این پرسش به مطالعه مرجع زیر بپردازید.

---

1 . simple analysis of variance
2 . main effects
3 . levene's test for equality of variances
4 . homoscedasticity

Chiong, C., & DeLoache, J. S. (2013). Learning the ABCs: What kinds of picture books facilitate young children's learning? Journal of Early Childhood Literacy, 13(2), 225–241.

برای پرسش‌های بیشتر، پرسش‌های ۸۱، ۸۷ و ۸۸ را مطالعه کنید.

پرسش ۸۷: چگونه می‌توانم از طریق اِکسل تحلیل واریانس ساده را محاسبه کنم؟

تحلیل واریانس زمانی به‌کاربرده می‌شود که بیش از دو گروه را بخواهیم ازنظر میانگین در یک متغیر مقایسه کنیم. در ادامه با سه مجموعه داده از نمرات نیروهای پلیس شرکت‌کننده در یک دوره آموزشی مواجه هستیم. گروه اول از نیروهای پلیس در گروه آموزشی ۱۰ ساعته دقت هدف‌گیری با سلاح قرار گرفتند؛ گروه دوم از نیروهای پلیس در گروه آموزشی ۲۰ ساعته دقت هدف‌گیری با سلاح قرار گرفتند؛ گروه سوم از نیروهای پلیس در هیچ‌یک از گروهای آموزشی دقت هدف‌گیری با سلاح قرار نگرفتند. متغیر مستقل در این مطالعه، آموزش دقت هدف‌گیری با سلاح بود و متغیر وابسته دقت هدف‌گیری با سلاح نیروهای پلیس.

فرضیه صفر این بود که بین گروه‌های سه‌گانه در دقت هدف‌گیری با سلاح تفاوت معناداری وجود ندارد و فرضیه یک این بود که بین گروه‌های سه‌گانه ازنظر دقت هدف‌گیری با سلاح تفاوت معناداری وجود دارد. برای تحلیل این فرضیه، می‌توان از آزمون تحلیل واریانس بهره برد.

گروه بدون تمرین: ۵۶، ۴۸، ۶۳، ۷۱، ۸۶، ۷۲، ۴۸، ۷۸، ۷۴، ۵۹

گروه با ۱۰ ساعت تمرین: ۵۶، ۷۹، ۷۱، ۸۶، ۶۹، ۸۸، ۷۵، ۵۷، ۸۹، ۷۷

گروه با ۲۰ ساعت تمرین: ۸۷، ۸۹، ۹۹، ۹۲، ۷۸، ۶۱، ۸۷، ۸۰، ۷۹، ۷۶

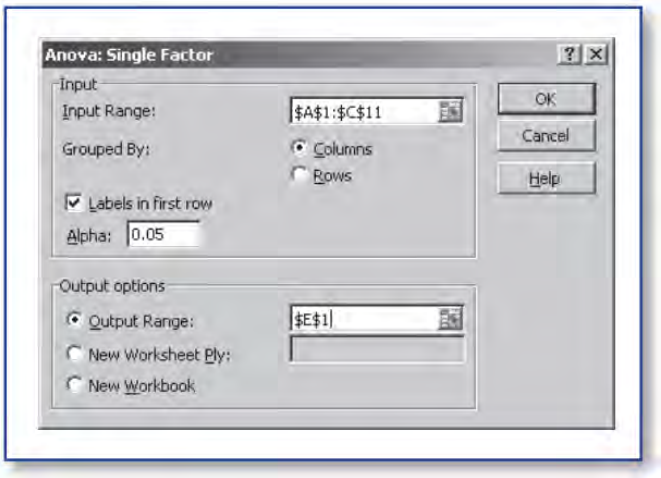
برای انجام آزمون تحلیل واریانس در محیط اِکسل مراحل زیر را انجام دهید:

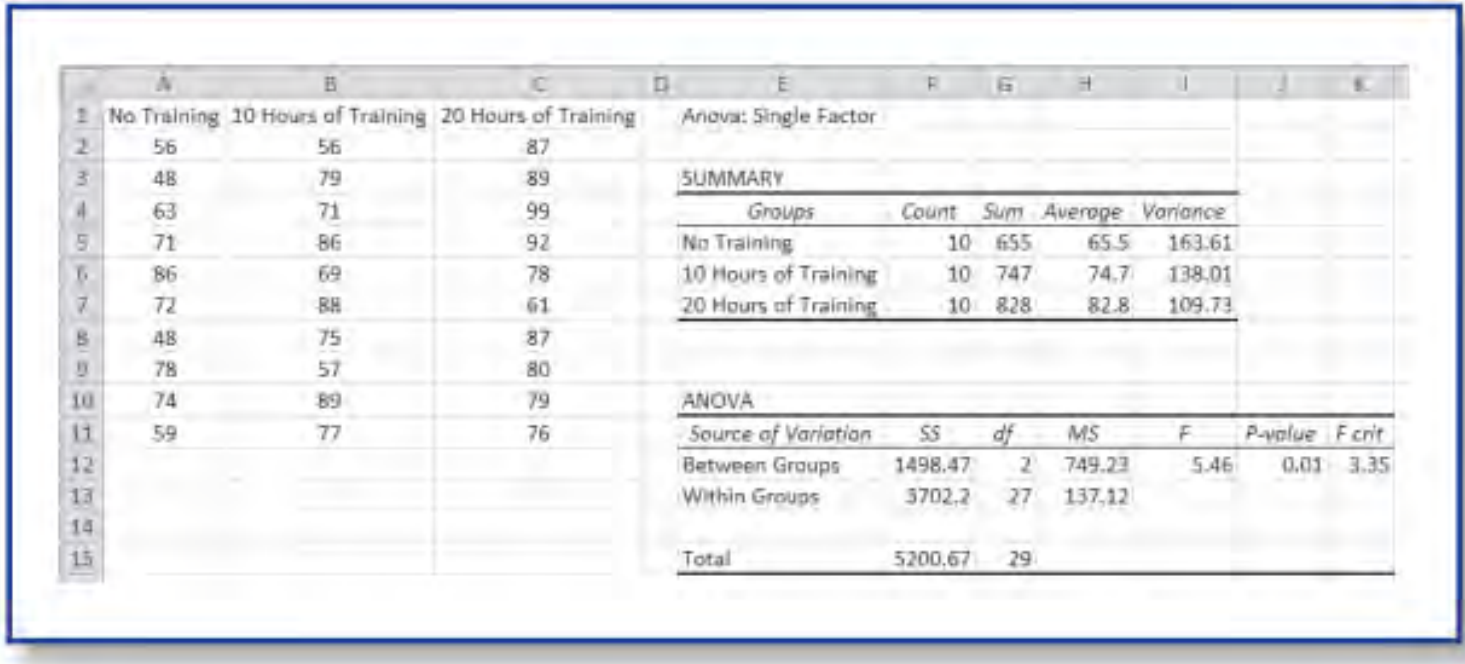
۱. در محیط اِکسل بر روی Data و سپس Data Analysis کلیک کنید؛

۲. در پنجره بازشده، گزینه Anova: Single Factor را انتخاب کنید؛

۳. دامنه اعداد و سایر اطلاعات خواسته‌شده را تکمیل کنید؛

۴. بر روی OK کلیک کنید؛





	A	B	C
1	No Training	10 Hours of Training	20 Hours of Training
2	56	56	87
3	48	79	89
4	63	71	99
5	71	86	92
6	86	69	78
7	72	88	61
8	48	75	87
9	78	57	80
10	74	89	79
11	59	77	76
12			
13			
14			
15			

Anova: Single Factor

SUMMARY

Groups	*Count*	*Sum*	*Average*	*Variance*
No Training	10	655	65.5	163.61
10 Hours of Training	10	747	74.7	138.01
20 Hours of Training	10	828	82.8	109.73

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	1498.47	2	749.23	5.46	0.01	3.35
Within Groups	3702.2	27	137.12			
Total	5200.67	29				

همان‌گونه که در برونداد ملاحظه می‌شود، ارزش F به‌دست‌آمده برابر ۵/۴۶ است و مربوط به اثر گروه بوده و سطح معناداری آن ۰/۰۱ است که نشان می‌دهد بین سه گروه در دقت هدف‌گیری با سلاح تفاوت معناداری وجود دارد. بر اساس میزان خطای محاسبه‌شده که ۰/۰۱ است، با ۹۹ درصد اطمینان می‌توان گفت از تفاوت در جامعه وجود دارد. در واقع، اثر تمرین با سطح خطای ۰/۰۱ بر میزان مهار دقت هدف‌گیری با سلاح در نیروی پلیس معنادار است، و میزان خطای نوع اول به‌دست‌آمده کمتر از ۰/۰۵ است. درواقع ۱ مورد از ۱۰۰ مورد تفاوت‌های مشاهده‌شده در سه گروه ممکن است از روی شانس یا خطا باشد، نه از روی واقعیت موجود در جامعه. بنابراین فرض صفر رد و فرض پژوهش تائید می‌گردد.

برای پرسش‌های بیشتر، پرسش‌های ۸۶، ۸۸ و ۸۹ را مطالعه کنید.

پرسش ۸۸: تحلیل واریانس عاملی چیست و یک مثال برای چگونگی کاربست آن کدام است؟

همان‌گونه که تحلیل واریانس ساده[1] به بررسی تفاوت‌های بین گروه‌ها در یک عامل می‌پردازد، تحلیل واریانس عاملی[2] به بررسی تفاوت‌های بین گروه‌ها در بیش از یک عامل در یک زمان واحد می‌پردازد. در هر تحلیل واریانس عاملی، اثر اصلی[3] هر یک از عامل‌ها و اثر تعاملی[4] آن‌ها آزمون می‌شود. برای مثال، تحلیل واریانس ساده به بررسی تفاوت میانگین مهارت‌های زبانی[5] در سه گروه دانش آموزان پایه‌های اول، دوم و سوم می‌پردازد. امّا تحلیل واریانس عاملی می‌تواند در یک زمان واحد، به بررسی تفاوت میانگین مهارت‌های زبانی در سه گروه دانش آموزان پایه‌های اول، دوم و سوم با در نظر گرفتن جنسیت آن‌ها بپردازد. به‌بیان‌دیگر، از طریق تحلیل واریانس عاملی می‌توان اثر اصلی پایه‌های تحصیلی، اثر اصلی جنسیت و نیز اثر تعاملی پایه‌های تحصیلی× جنسیت را محاسبه کرد. به این وضعیت طرح عاملی می‌گویند و در این مثال یک طرح عاملی ۲×۳ وجود دارد و تعامل سه سطح پایه تحصیلی و دو سطح جنسیت نیز مورد بررسی واقع می‌شود. محققان می‌توانند با در نظر گرفتن سطوح هر یک از متغیرهای مستقل، به‌دقت بیان کنند که چه تحلیلی باید صورت گیرد. همان‌طور که قبلاً گفته شد، سطوح به زیرگروه‌ها یا طبقات هر متغیر مستقل اشاره دارد. در مثال یادشده، دوره آموزش روخوانی سه سطح و جنسیت دو سطح دارد. سطوح می‌توانند با اعدادی پیش از کلمه آنوای نشان داده شوند. در مثال بالا، آن را می‌توان به شکل آنوای ۲×۳ نشان داد. تأثیر متغیرهای مستقل روی‌هم دیگر را تعامل می‌گویند؛ یعنی تأثیر یک عامل در مقایسه با سطح عامل دیگر روی متغیر وابسته متفاوت است. به‌عبارت‌دیگر، تعامل به معنی تأثیر دوطرفه متغیرهای مستقل روی متغیر وابسته است. تعامل زمانی وجود دارد که تفاوت بین سطوح یک متغیر مستقل با سطوح متغیر مستقل دیگر هم‌خوانی نداشته باشد. به‌عبارت دیگر، اگر اثر یک متغیر بر سطوح مختلف متغیر دوم متفاوت باشد، یک تعامل وجود دارد. اکنون روشن است که چگونه یک تحلیل واریانس عاملی در مقایسه با تحلیل‌های واریانس یک‌طرفه می‌تواند اطلاعات بیشتری در اختیار محقق قرار دهد.

برای مطالعه بیشتر درباره مثال موجود در این پرسش به مطالعه مرجع زیر بپردازید.

Sikes, P. L. (2013). The effects of specific practice strategy use on university string players performance. Journal of Research in Music Education, 61(3), 318-333.

برای پرسش‌های بیشتر، پرسش‌های ۸۶، ۸۷ و ۸۹ را مطالعه کنید.

---

1 . simple analysis of variance
2 . factorial analysis of variance
3 . main effect
4 . interaction effect
5 . language skills

پرسش ۸۹: چگونه می‌توانم از طریق اِکسل تحلیل واریانس عاملی را محاسبه کنم؟

اجازه دهید این پرسش را با این مثال آغاز کنیم که پژوهشگری قصد دارد بررسی کند که آیا بین مهارت‌های زبانی دختران و پسرانی که در سه برنامه مهارت زبانی شرکت کردند، تفاوت معناداری وجود دارد یا خیر. همان‌گونه که درست حدس زدید، دو متغیر مستقل (جنسیت و برنامه مهارت زبانی) و یک متغیر وابسته (مهارت زبانی) وجود دارد. عامل اول این پژوهش متغیر جنسیت با دو سطح دختر و پسر است و عامل دوم متغیر برنامه زمانی با سه سطح اول، دوم و سوم است. عامل اول در ستون اول و عامل دوم در ردیف اول قرار دارند. بعد از انجام مداخله به همه شرکت‌کنندگان پرسشنامه مهارت‌های زبانی داده شد. اطلاعات جمع‌آوری‌شده وارد اِکسل شد و در جدول زیر نمایه گردید. فرض صفر بر این مبناست که بین دختران و پسران و همچنین بین گروه‌های سه‌گانه آموزشی ازنظر مهارت‌های زبانی تفاوت معناداری وجود ندارد و اگر تفاوتی نیز مشاهده شود ناشی از شانس، خطای اندازه‌گیری، خطای نمونه‌گیری و تفاوت‌های درون فردی بوده است. فرض یک بر این مبناست که بین دختران و پسران و همچنین بین گروه‌های سه‌گانه آموزشی ازنظر مهارت‌های زبانی تفاوت معناداری وجود دارد. در ادامه به تحلیل داده‌ها می‌پردازیم.

مردان			زنان		
برنامه ۱	برنامه ۲	برنامه ۳	برنامه ۱	برنامه ۲	برنامه ۳
۶	۵	۸	۷	۸	۷
۵	۶	۷	۵	۷	۱۰
۶	۵	۶	۶	۶	۸
۵	۴	۶	۷	۷	۶
۴	۳	۵	۶	۷	۷
۵	۴	۶	۴	۹	۶
۶	۵	۷			

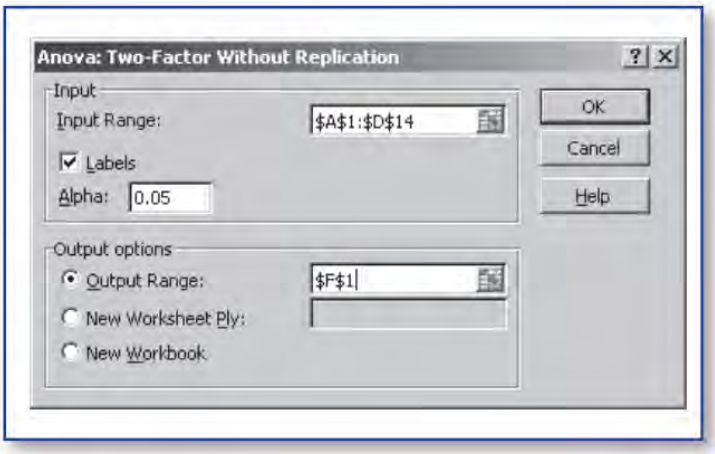
برای انجام تحلیل واریانس عاملی در محیط اِکسل مراحل زیر را انجام دهید؟

۱. در محیط اِکسل بر روی Data و سپس Data Analysis کلیک کنید؛

۲. در پنجره بازشده، گزینه Anova: Two Factor Without Replication option را انتخاب کنید؛

۳. دامنه اعداد و سایر اطلاعات خواسته‌شده را تکمیل کنید؛

۴. بر روی OK کلیک کنید؛



همان‌گونه که در برونداد صفحه بعد ملاحظه می‌شود، ۲ ارزش F محاسبه شده است. اولین ارزش F که ۲/۰۲ به‌دست‌آمده، مربوط به اثر جنسیت بوده و سطح معناداری آن ۰/۰۷ است و دومین ارزش F که ۴/۴۰ به‌دست‌آمده، مربوط به اثر نوع برنامه بوده و سطح معناداری آن ۰/۰۲ است. اثر جنسیت با سطح معناداری ۰/۰۷ بر میزان مهارت‌های زبانی معنادار نیست، زیرا میزان خطای نوع اول به‌دست‌آمده بیشتر از ۰/۰۵ است. اثر برنامه با سطح معناداری ۰/۰۲ بر میزان مهارت‌های زبانی معنادار است، زیرا میزان خطای نوع اول به‌دست‌آمده کمتر از ۰/۰۵ است. نرم‌افزار اِکسل قادر به سنجش اثر تعاملی جنسیت (زن و مرد) در برنامه (اول، دوم، سوم) نیست.

	A	B	C	D	E	F	G	H	I	J	K
1		Program 1	Program 2	Program 3		ANOVA					
2	Male	6	5	8		*Source of Variation*	*SS*	*df*	*MS*	*F*	*P-value*
3		5	6	7		Rows	33.44	12	2.79	2.02	0.07
4		6	5	6		Columns	12.15	2	6.08	4.40	0.02
5		5	4	6		Error	33.18	24	1.38		
6		4	3	5							
7		5	4	6		Total	78.77	38			
8		6	5	7							
9	Female	7	8	7							
10		5	7	10							
11		6	6	8							
12		7	7	6							
13		6	7	7							
14		4	9	6							

برای پرسش‌های بیشتر، پرسش‌های ۸۶، ۸۸ و ۹۱ را مطالعه کنید.

پرسش ۹۰: چگونه می‌توانم از آزمون‌های ناپارامتریک برای آزمون معناداری استفاده کنم؟

تاریخ آمار ناپارامتری[1] مانند آمار پارامتری[2] به اوایل قرن هیجدهم میلادی برمی‌گردد. در سال ۱۹۷۰ میلادی مقاله‌ای منتشر شد و طی آن بر اساس آمار نوزادان شهر لندن در فاصله‌های ۱۶۲۹ و ۱۷۱۰ میلادی، ادعا شد که مشیت الهی بر این است که تعداد نوزادان پسر بیشتر از تعداد نوزادان دختر باشد. در حقیقت این ادعا یک آزمون ناپارامتری معروف است که امروز به نام آزمون علامت[3] شهرت دارد. بااین‌حال، آمار ناپارامتری بیش از دو قرن ناشناخته بود، تا اینکه پیشرفت آن با انتشار دو مقاله، یکی توسط آماردانی به نام فرانک ویلکاکسون[4] آغاز شد. قبل از انتشار این دو مقاله، اصطلاح ناپارامتری برای اولین بار در رسالهٔ دکترای یک آمار دان به نام ولفویتز[5] در سال ۱۹۴۲ به کار رفته است.

برای سنجش فرضیه‌هایی که متغیر آن کمی‌اند، از آمار پارامتریک استفاده می‌شود. متغیرهای کمی به علت کمی بودن و واحد پذیر بودن از این ویژگی برخوردارند که آن‌ها را میانگین‌پذیر و انحراف معیارپذیر می‌کنند و به دلیل همین ویژگی معمولاً برای استفاده از آزمون‌های پارامتریک، پیش‌فرض‌هایی نظیر بهنجار بودن توزیع جامعه لازم است زیرا در حالتی که توزیع جامعه بهنجار نباشد، میانگین و انحراف معیار، نمایی واقعی از داده‌ها را به تصویر نمی‌کشند.

برای آزمون متغیرهای کیفی و رتبه‌ای از آمار ناپارامتریک استفاده می‌شود. این آزمون‌ها که از آن‌ها با عنوان آزمون‌های بدون پیش‌فرض نیز یاد می‌شود به هیچ پیش‌فرض خاصی نیاز ندارد. درواقع، آماردان‌ها درصدد برآمدند تا روش‌هایی ایجاد کنند که استفاده از آن‌ها مشروط به فرض خاصی دربارهٔ توزیع جامعه نباشد و به‌اصطلاح برای جامعه‌هایی توزیع آزاد[6] نیز کاربرد داشته باشد.

آزمون‌های ناپارامتریک مشروط به مفروضات آمار کلاسیک نیستند و کاربرد اصلی آن‌ها در بررسی جوامع آماری غیربهنجار، جوامع با داده‌های کیفی و نمونه‌های کوچک آماری است. در خصوص تبدیل متغیرها باید یادآور شد که می‌توان متغیرهای کمی را به متغیرهای کیفی تبدیل کرد و آن‌ها را با آزمون‌های ناپارامتریک مورد ارزیابی قرار داد ولی عکس این عمل امکان‌پذیر نیست. بعلاوه، سطح دقت در آزمون‌های آماری پارامتریک از آزمون‌های آماری ناپارامتریک بیشتر است و معمولاً پیشنهاد می‌شود درصورتی‌که استفاده از آزمون‌های پارامتریک امکان‌پذیر باشد از آزمون‌های ناپارامتریک استفاده نشود. در حقیقت بدین منظور روش‌های زیادی ارائه شده است ولی بجای اینکه به روش‌های توزیع آزاد معروف باشند، تسامحاً به روش روش‌های ناپارامتری

1 . Non Parametric Statistics
2 . Parametric Statistics
3 . Sign Test
4 . Frank Whitney
5 . Wolfowitz
6 . Distribution Free

معروف شده‌اند. روش‌های ناپارامتری نسبت به روش‌های پارامتری محاسن و معایبی دارند. اولین حُسن روش‌های ناپارامتری این است که مستلزم فرض خاصی دربارهٔ شکل و توزیع جامعه نیستند. دوم اینکه فهم و استفاده از آن‌ها معمولاً ساده‌تر از روش‌های پارامتری است. علاوه بر این، روش‌های ناپارامتری دو نقص دارند: در این آزمونها فقط از قسمتی از اطلاعات استفاده‌شده و باعث اتلاف اطلاعات می‌شوند و این روش‌ها از کارایی کمتری نسبت به روش‌های پارامتری برخوردارند. مثلاً ممکن است فاصلهٔ اطمینان ۹۵ درصدی روش‌های ناپارامتری دو برابر روش‌های پارامتری باشد. با توجه به مطالب فوق، درواقع ما در استفاده از روش‌های ناپارامتری یا پارامتری یک بده و بستان[1] انجام می‌دهیم چراکه در روش‌های ناپارامتری کمتر فرض می‌کنیم و درنتیجه قدری از دقت و اطلاعات خود را از دست می‌دهیم ولی در عوض، کاربرد روش را گسترش می‌دهیم. بنابراین، به‌طور خلاصه از مزایای روش‌های ناپارامتری این است که فرض‌های کمتری لازم است و در بسیاری از حالات فقط داده‌های اسمی (طبقه‌ای یا ترتیبی، رتبه‌ای) به‌جای داده‌های عددی (فاصله‌ای) کافی است. از معایب روش‌های ناپارامتری این است که معمولاً ترجیح می‌دهیم مدل کاملاً تعریف‌شده داشته باشیم که شامل پارامترهای مهمی نظیر میانگین‌ها و واریانس‌ها باشد تا برای اهداف تعبیر و تفسیر استفاده شود.

نکته

اگر جامعه بهنجار باشد از آزمون‌های پارامتریک و چنانچه غیربهنجار باشد از آزمون‌های غیر پارامتری استفاده می‌نماییم؛ اگر نمونه بزرگ باشد، طبق قضیه حد مرکزی حتی اگر جامعه بهنجار نباشد می‌توان از آزمون‌های پارامتریک استفاده نمود.

پرسش‌های بیشتر، پرسش‌های ۸۱، ۸۶ و ۹۱ را مطالعه کنید.

1 . Trade off

پرسش ۹۱: اندازه اثر چیست و چرا مهم است؟

همان‌گونه که از اطلاعات پرسش ۷۸ به یاد دارید، بین معناداری آماری و معناداری منطقی تفاوت وجود دارد. امّا روشی سودمند دیگری نیز وجود دارد که از طریق آن می‌توان درباره یافته‌های آماری قضاوت کرد. این روش اندازه اثر[1] نام دارد. اندازه اثر شاخصی است که شدت یا بزرگی[2] یک یافته آماری، و نه اندازه مطلق آن را نشان می‌دهد. در واقع، ممکن است در یک پژوهش، بین میانگین دو نمونه تفاوت معنادار زیادی وجود داشته باشد، امّا اندازه اثر کوچک باشد که درواقع بیانگر تقریباً عدم معنا داری تفاوت بین گروه‌ها است. از طرف دیگر، ممکن است در یک پژوهش، بین میانگین دو نمونه تفاوت معنادار کمی وجود داشته باشد، امّا اندازه اثر بزرگ باشد که درواقع بیانگر تفاوت زیاد بین گروه‌ها است.

اندازه اثر عنوانی است که به مجموعه‌ای از شاخص‌ها که بزرگی اثر آزمایش را می‌سنجد، اطلاق می‌شود. برخلاف آزمون‌های معناداری، این شاخص مستقل از حجم نمونه است.

در محاسبه اندازه اثر، از فرمول $ES = \frac{\bar{x}1-\bar{x}2}{SD}$ استفاده می‌شود که در آن ES شاخص اندازه اثر، $\bar{x}1$ میانگین گروه اول، $\bar{x}2$ میانگین گروه دوم و SD انحراف استاندارد یکی از گروه‌ها است.

برای مثال، به اطلاعات زیر که مربوط به نمرات عزت‌نفس دو گروه از افراد است توجه کنید.

گروه اول: اوایل نوجوانی: میانگین ۲۷/۵ و انحراف استاندارد: ۴/۶۵

گروه دوم: اواخر نوجوانی: میانگین ۳۱/۲ و انحراف استاندارد: ۳/۹۸

با جای گذاری اعداد در فرمول اندازه اثر خواهیم داشت

$$ES = \frac{۳۱/۲-۲۷/۵}{۴/۵۶} = ۰/۷۹$$

تفسیر اندازه اثر بسیار ساده است. اگر تفاوت بین گروه ها صفر باشد، اندازه اثر نیز صفر خواهد بود، یعنی بین گروه‌ها صفر درصد مشابهت و صد درصد تفاوت وجود دارد و اگر اندازه اثر ۱ باشد، یعنی بین گروه‌ها، ۴۵ درصد مشابهت و ۵۵ درصد تفاوت وجود دارد؛ درواقع، هر چه میزان اندازه اثر افزایش یابد، نمرات دو مجموعه نمره کمتر به هم شبیه و بیشتر با هم متفاوت خواهند بود و درنتیجه تفاوت مشاهده‌شده منطقی‌تر و واقعی‌تر خواهد بود. هرچند برخی پژوهشگران اندازه اثر را ذکر نمی‌کنند، امّا شاخص بسیار مهمی در نشان دادن معناداری منطقی علارغم عدم معناداری آماری است.

پرسش‌های بیشتر، پرسش‌های ۷۸، ۸۲ و ۸۴ را مطالعه کنید.

1 . effect size
2 . magnitude

بخش ۱۱: نگاهی به روابط بین متغیرها

پرسش ۹۲: روابط بین متغیرها و معناداری آن‌ها چگونه محاسبه می‌شوند؟

بررسی روابط بین متغیرها از طریق آزمون‌های آماری متعددی انجام می‌شود. در این آزمون‌ها ارزش به‌دست‌آمده با ارزش بحرانی موجود در جداول آماری انتهای کتاب مقایسه می‌شود و با سطح خطای ۰/۰۱ یا ۰/۰۵ به تائید یا رد فرضیه صفر اقدام می‌شود. با رجوع به پرسش ۴۵ می توانید به مطالعه چگونگی محاسبه رابطه بین دو متغیر بپردازید.

اختلال اُتیسم یا اختلالات طیف اتیسم[1] شمار رو به افزایشی از کودکان و بزرگسالان را درگیر کرده است. در یک مطالعه، محققان به بررسی رابطه بین شدت اختلال اتیسم و کژکاری چندحسی[2] پرداختند. در این مطالعه بهترین روش آماری، تحلیل همبستگی[3] است. مقیاس حسی مربوط به ۱۰۴ کودک ۳ تا ۵۶ ساله با تشخیص اُتیسم تکمیل شد. تحلیل داده ها همبستگی معناداری بین مولفه های مختلف نشان داد. نتایج نشان داد که اختلالا حسی با شدت اختلال اُتیسم در کودکان همبسته است امّا این همبستگی در نوجوانان و بزرگسالان مشاهده نشد. پژوهشگران مطرح کردند که زمینه های اصلی اختلال اتیسم و پردازشهای چند حسی که در این افراد مشاهده شده است و اختلالی که در پردازشهای حسی آنها وجود دارد، ماهیتاً جهانی هستند.

برای مطالعه بیشتر درباره مثال موجود در این پرسش به مطالعه مرجع زیر بپردازید.

Kern, J. K., Trivedi, M. H., Grannemann, B. D., Garver, C. R., Johnson, D. G ,. Andrews, A. A. . . . Schroeder, J. L. (2007). Sensory correlations in autism .Autism, 11(2), 123ات134.

برای پرسش‌های بیشتر، پرسش‌های ۴۰، ۴۶ و ۴۷ را مطالعه کنید.

1 . autism spectrum
2 . multisensory dysfunction
3 . correlational analysis

پرسش ۹۳: چگونه می‌توانم از طریق اِکسل به بررسی معناداری ضریب همبستگی بپردازم؟

برای پاسخ به این سؤال، از یک مثال بهره می‌بریم. یک محقق به بررسی این سؤال می‌پردازد که آیا بین ساعات صرف شده برای تماشای رسانه‌های جمعی در هفته و موفقیت شغلی[1] رابطه معناداری وجود دارد؟. برای ارزیابی هر دو متغیر از پرسشنامه خودگزارشی با نمره ۰ تا ۱۰ استفاده کرد. فرضیه پژوهش این بود که بین ساعات صرف شده برای تماشای رسانه‌های جمعی در هفته و موفقیت شغلی رابطه معناداری وجود دارد. در ادامه به جدول توزیع داده‌ها اشاره شده است.

ساعات حضور در رسانه اجتماعی	موفقیت شغلی
۲۲	۵
۲۳	۶
۱۵	۵
۷	۳
۲۱	۹
۱۴	۵
۱۵	۶
۲۲	۷
۲۰	۸
۱۹	۹

برای انجام همبستگی در محیط اِکسل مراحل زیر را انجام دهید؟

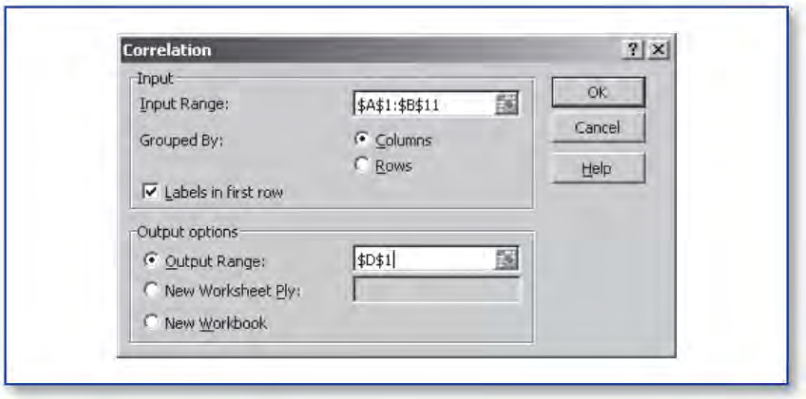
۱. در محیط اِکسل بر روی Data و سپس Data Analysis کلیک کنید؛

۲. در پنجره بازشده، گزینه Correlation را انتخاب کنید؛

۳. دامنه اعداد و سایر اطلاعات خواسته‌شده را تکمیل کنید؛

۴. بر روی OK کلیک کنید؛

1 . job satisfaction



همان‌گونه که در برونداد صفحه بعد ملاحظه می‌کنید، ارزش همبستگی محاسبه‌شده ۰/۶۵ است. امّا برونداد اِکسل و همچنین تابع CORREL= سطح معناداری را نمایه نمی‌کنند. برای اینکه معناداری بین همبستگی دو متغیر مشخص شود، می‌بایست به جدول معناداری آزمون همبستگی رجوع کرد که در ضمیمه قرار داده شده است. در سطح خطای ۰/۰۵، نقطه بحرانی رد فرضیه صفر برابر ۰/۵۴ است. اکنون ازآنجایی‌که میزان همبستگی به‌دست‌آمده (۰/۶۵) بیشتر از میزان بحرانی جدول معناداری است (۰/۵۴)، بنابراین فرضیه صفر مبنی بر عدم معناداری همبستگی بین دو متغیر رد می‌شود و عملاً مشخص می‌شود که همبستگی بین دو متغیر معنادار است.

	A	B	C	D	E	F
1	Social Media Hours	Job Success			*Social Media Hours*	*Job Success*
2	22	5		Social Media Hours	1	
3	23	6		Job Success	0.65	1
4	15	5				
5	7	3				
6	21	9				
7	14	5				
8	15	6				
9	22	7				
10	20	8				
11	19	9				

برای پرسش‌های بیشتر، پرسش‌های ۴۶، ۴۷ و ۹۲ را مطالعه کنید.

پرسش ۹۴: رگرسیون ساده چیست و یک مثال برای چگونگی کاربست آن کدام است؟

در سال ۱۸۷۷ فرانسیس گالتون[1] در مقاله‌ای که درباره بازگشت به میانگین منتشر کرده بود اظهار داشت که متوسط قد پسران دارای پدران قدبلند، کمتر از قد پدرانشان است. به نحو مشابه متوسط قد پسران دارای پدران کوتاه‌قد نیز، بیشتر از قد پدرانشان گزارش شده است. به‌این‌ترتیب گالتون پدیده بازگشت به طرف میانگین را در داده‌هایش مورد تأکید قرارداد. برای گالتون رگرسیون مفهومی زیست‌شناختی داشت، اما کارهای او توسط کارل پیرسون[2] برای مفاهیم آماری توسعه داده شد. گرچه گالتون برای تأکید بر پدیده «بازگشت به سمت مقدار متوسط» از تحلیل رگرسیون استفاده کرد، اما به‌هرحال امروزه واژه تحلیل رگرسیون جهت اشاره به مطالعات مربوط به روابط بین متغیرها به‌کاربرده می‌شود.

اگر بخواهیم واژه رگرسیون[3] را ازلحاظ لغوی تعریف نماییم، این واژه در فرهنگ لغت به معنی پس‌روی، برگشت و بازگشت است. اما اگر آن را از دید آمار و ریاضیات تعریف کنیم اغلب جهت رساندن مفهوم "بازگشت به یک مقدار متوسط یا میانگین" به کار می‌رود. بدین معنی که برخی پدیده‌ها به مرور زمان از نظر کمی به‌طرف یک مقدار متوسط میل می‌کنند. در حقیقت تحلیل رگرسیون، تکنیکی آماری برای بررسی و مدل‌سازی ارتباط بین متغیرها است. رگرسیون تقریباً در هر زمینه‌ای ازجمله مهندسی، فیزیک، اقتصاد، مدیریت، علوم زیستی، بیولوژی و علوم اجتماعی و روانشناسی برای برآورد و پیش‌بینی موردنیاز است. می‌توان گفت تحلیل رگرسیونی، پرکاربردترین روش در بین تکنیک‌های آماری است. شِمایی کلی و خلاصه‌شده از یک تحلیل رگرسیونی ساده به‌صورت زیر می‌باشد: در ابتدا تحلیل‌گر حدس می‌زند که بین دو متغیر نوعی ارتباط وجود دارد، در حقیقت حدس می‌زند که یک رابطه به شکل یک خط بین دو متغیر وجود دارد (رابطه خطی) و سپس به جمع‌آوری اطلاعات کمی از دو متغیر می‌پردازد و این داده‌ها را به‌صورت نقاطی در یک نمودار دو بعدی رسم می‌کند. درصورتی‌که نمودار نشان‌دهنده این باشد که داده‌ها تقریباً (نه لزوماً دقیق) در امتداد یک خط مستقیم پراکنده شده‌اند، حدس تحلیل‌گر تأیید شده و این ارتباط خطی به‌صورت زیر نمایش داده می‌شود:

$$y = a+bx$$

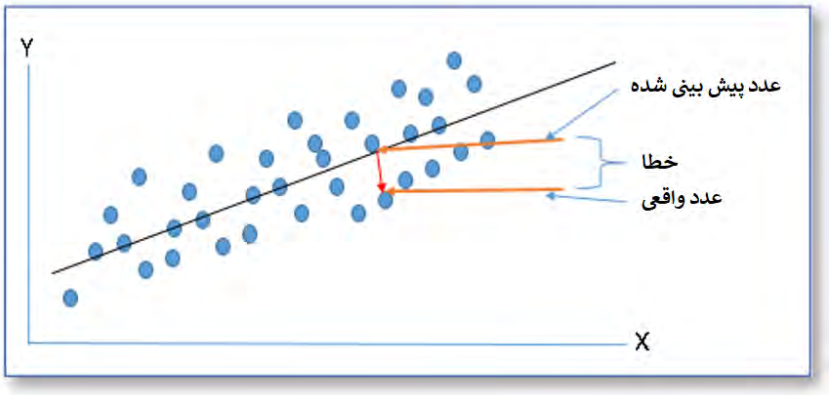
که در آن a عرض از مبدأ و b شیب خط رگرسیون (ضریب زاویه خط) است.

در رابطه خطی، پدیده‌ها یا به عبارتی متغیرها، باهم در رابطه هستند و اگر نموداری برای آن‌ها رسم کنیم به شکل یک خط راست است. مثل $y=x+۲$ در این شرایط می‌گوییم دو متغیر x و y باهم رابطه خطی دارند. در روابط خطی نسبتی که x و y افزایش یا کاهش می‌یابند مقداری ثابت است. در مثالی که زدیم اگر x را یک واحد یک واحد افزایش دهیم y هم یک واحد یک واحد افزایش می‌یابد.

---

1 . Francis Galton
2 . Karl Pearson
3 . Regression

در معادله و خط رگرسیون، بین نقاط واقعی و نقاط یا خط پیش‌بینی‌شده،کمی تفاوت به چشم می‌خورد که از آن به‌عنوان خطای برآورد یاد می‌کنیم. این خطا ممکن است از خطا در اندازه‌گیری، شرایط محیط، تفاوت‌های طبیعی و... ناشی شده باشد. نمونه‌ای از این تفاوت در شکل زیر نشان داده شده است.



بنابراین معادله اولیه را به‌صورت زیر اصلاح می‌کنیم:

$$y = ax + b + \epsilon$$

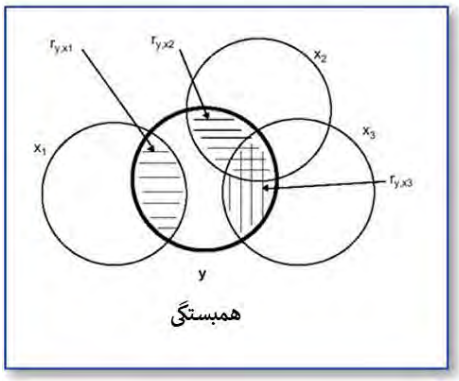
معادله بالا یک مدل رگرسیون خطی نامیده می‌شود. معمولاً به x متغیر مستقل (پیش‌بین) و به y متغیر وابسته (ملاک) گفته می‌شود. که ϵ خطای تصادفی است که برای کامل شدن مدل و نشان دادن این‌که خطا نیز تا حدی وجود دارد در نظر گرفته می‌شود.

پیش‌بینی بر پایه‌ی همبستگی را رگرسیون می‌گویند. یعنی چنانچه رابطه‌ی ما رابطه خطی باشد بر اساس ضریب همبستگی می‌توان پیش‌بینی کرد. پس نباید فراموش شود که تنها در شرایطی که همبستگی بالاست می‌توان از رگرسیون استفاده کرد و سهم متغیر ملاک به‌وسیله‌ی چندین متغیر پیشن بین تعیین یا پیش‌بینی می‌شود.

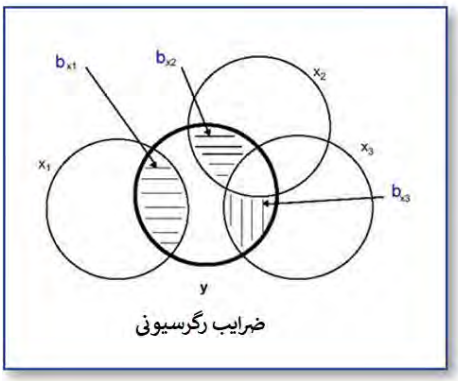
رگرسیون دارای انواع مختلفی نظیر رگرسیون چندمتغیری (معادلات ساختاری)، رگرسیون لجستیک، رگرسیون ترتیبی، رگرسیون منحنی، رگرسیون خطی ساده و رگرسیون خطی چندگانه است، که برای درک بهتر و مفهومی‌تر مطالب ابتدا رگرسیون خطی ساده مطرح شد. مثال: فرض کنید بین نمره ریاضی و فیزیک دانش آموزان رابطه‌ای خطی وجود داشته باشد.در این صورت معادله موردنظر به‌صورت زیر است: نمره فیزیک = مقدار ثابت + (نمره ریاضی×ضریب) + خطا؛ خطا برابر است با مقدار y معادله (پیش‌بینی‌شده) منهای مقدار y که واقعاً هست؛

ساده‌ترین روشی که از ضریب همبستگی پیرسون برای پیش‌بین به کار می‌رود، روش پیش‌بینی نمره‌های استاندارد است که از طریق فرمول $Z_y=(r_{xy})(Z_x)$ محاسبه می‌شود که در آن r_{xy} ضریب همبستگی بین دو متغیر، Z_x نمره استاندارد متغیر x و Z_y نمره پیش‌بینی‌شده برای متغیر y است؛

تا زمانی که دو متغیر به‌صورت کامل همبسته نباشند این گرایش وجود دارد که نمره‌های گروهی در اولین متغیر به میانگین دومین متغیر نزدیک باشد (تأثیر رگرسیون). چنانچه همبستگی بین متغیرها صفر باشد، رگرسیون در اطراف میانگین به‌صورت کامل اتفاق خواهد افتاد و اگر همبستگی بین متغیرها کامل نباشد، نمره‌های ملاک به میانگین نمونه نزدیک‌تر است تا به نمره‌های پیش‌بین. چنانچه همبستگی بین متغیرها بالا باشد گرایش خیلی کمی وجود دارد که میانگین نمره‌های گروه انتخاب‌شده در اولین متغیر به‌طرف میانگین نمره‌های دومین متغیر کشیده شود. اما اگر همبستگی پایین باشد گرایش خیلی زیادی وجود دارد که میانگین نمره‌ها در اولین متغیر به‌طرف میانگین نمره‌های دومین متغیر کشیده شود؛ در شرایطی که همبستگی کامل و برابر ۱ باشد، خط‌های رگرسیون بر هم منطبق‌اند و زاویه بین آن‌ها صفر است؛ در شکل مقابل می‌توان به تفاوت همبستگی ساده و رگرسیون به‌خوبی پی برد.



همبستگی



ضرایب رگرسیونی

اگر همبستگی بین متغیرها صفر باشد قدرت پیش‌بینی وجود ندارد. اما اگر همبستگی ۱ باشد نمره‌ی ملاک برابر نمره‌ی پیش‌بین می‌شود. در این حالت رگرسیون به‌طرف میانگین اصلاً اتفاق نمی‌افتد. بنابراین رگرسیون به‌طرف میانگین با همبستگی بین متغیرها رابطه‌ی معکوس دارد. هرچه همبستگی کامل‌تر رگرسیون به سمت میانگین کمتر است؛

خط برازش یا رگرسیون به‌منظور پیش‌بینی متغیر y از روی متغیر x به‌کاربرده می‌شود. خط رگرسیون موجب می‌شود که خطاهای پیش‌بینی به حداقل برسد، به همین جهت به آن بهترین خط برازش[1] می‌گویند. هدف از به دست آوردن خط پیش‌بینی، یافتن معادله‌ای است که خطای پیش‌بینی را تا حد امکان کاهش دهد. هرگاه خط رگرسیون در نمودار پراکندگی با محور xها یا yها موازی باشد، به این معنا است که همبستگی بین متغیرهای x و y برابر صفر است؛

رگرسیون و همبستگی باهم رابطه معکوس دارند. هر چه همبستگی بیشتر باشد، رگرسیون کمتر می‌شود و هر چه همبستگی کمتر باشد، رگرسیون بیشتر می‌شود. اگر همبستگی بین متغیرها

1 . Best fit line

کامل باشد، رگرسیون اتفاق نمی‌افتند و اگر همبستگی بین متغیرها صفر باشد، رگرسیون در اطراف میانگین کاملاً اتفاق می‌افتند.

برای مطالعه بیشتر درباره مثال موجود در این پرسش به مطالعه مرجع زیر بپردازید.

Chen, S., & Yseng, J. (2010). Body mass index, nutrient intakes, health behaviors and nutrition knowledge: A quantile regression application in Taiwan. Health Education Journal, 69(4), 409ـ426.

برای پرسش‌های بیشتر، پرسش‌های ۴۶، ۹۲ و ۹۵ را مطالعه کنید.

پرسش ۹۵: چگونه می‌توانم از طریق اِکسل معادله رگرسیون ساده را محاسبه کنم؟

برای نشان دادن کاربست اِکسل در محاسبه ضریب رگرسیون ساده، با این مثال آغاز می‌کنیم. پژوهشگری قصد دارد بررسی کند که آیا بین تعداد قهرمانی‌های مدارس در سالهای گذشته در رقابت‌های فوتبال، می‌توان قهرمان شدن آن تیم در سال‌های آینده را پیش‌بینی کرد یا خیر؟ فرض صفر مبتنی بر عدم وجود ارتباط است و در فرضیه تحقیق، مبنا بر این است که قابلیت پیش‌بینی در سطح ۰/۰۵ خطا وجود دارد. داده‌های مربوط به ده مدرسه در جدول زیر نشان داده شده است. در این جدول تعداد قهرمانی‌های قبلی و سال جاری نمایه شده است. ارتباط بین این دو مجموعه داده، می‌تواند سبب شکل‌گیری یک معادله رگرسیونی شود که بر اساس آن می‌توان قهرمانی سال آینده را با ۵ درصد خطا پیش‌بینی کرد.

مدرسه	قهرمان سال قبل	قهرمان امسال
۱	۷	۸
۲	۶	۷
۳	۸	۹
۴	۱۱	۱۰
۵	۱۲	۹
۶	۸	۷
۷	۷	۱۲
۸	۳	۳
۹	۶	۵
۱۰	۵	۶

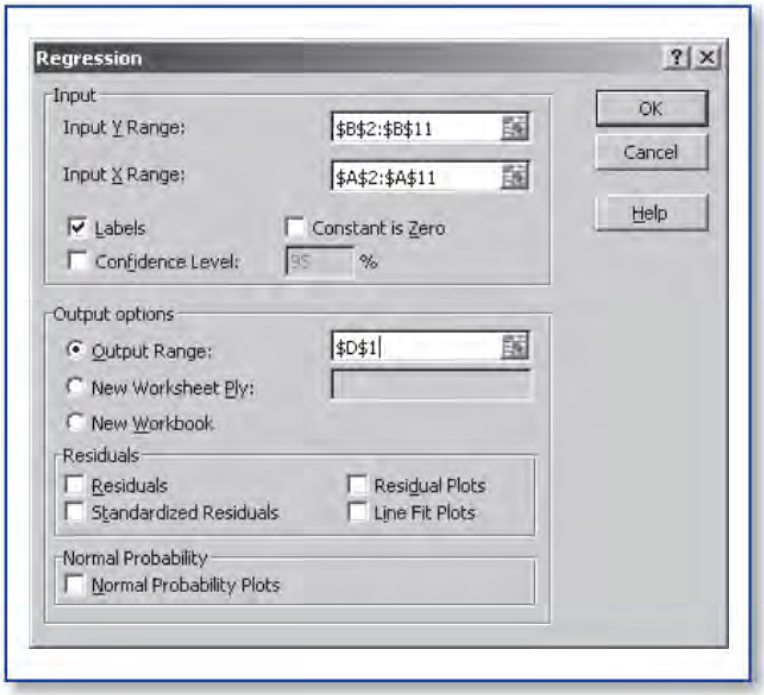
برای انجام آزمون رگرسیون در محیط اِکسل مراحل زیر را انجام دهید:

۱. در محیط اِکسل بر روی Data و سپس Data Analysis کلیک کنید؛

۲. در پنجره بازشده، گزینه Regression را انتخاب کنید؛

۳. دامنه اعداد و سایر اطلاعات خواسته‌شده را تکمیل کنید؛

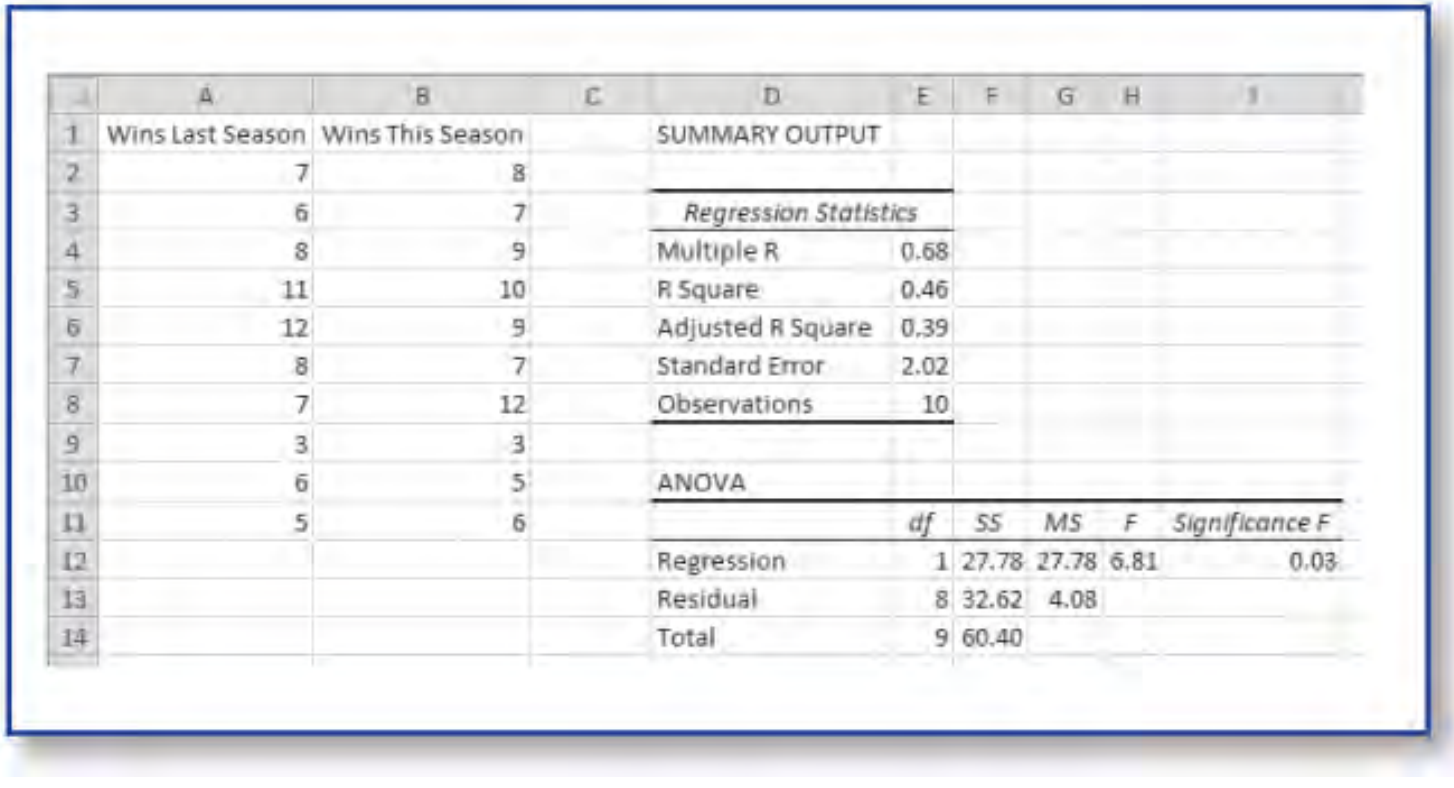
۴. بر روی OK کلیک کنید؛



همان‌گونه که در برونداد صفحه بعد ملاحظه می‌شود میزان F محاسبه‌شده که شاخص قدرت پیش‌بینی متغیر ملاک از روی متغیر پیش‌بین است، ۶/۸۱ به‌دست‌آمده است که در سطح خطای ۰/۰۵ معنی‌دار است (P=۰/۰۳). یعنی حدوداً با ۹۹ درصد اطمینان می‌توان گفت که متغیر پیش‌بین (تعداد سال‌های قهرمان شدن گذشته) می‌تواند متغیر ملاک (قهرمان شدن آینده) را پیش‌بینی کند و در این میان، فقط ۳ مورد از ۱۰۰ مورد پیش‌بینی، ممکن است به دلیل خطا یا تصادف باشد. ضریب R چندگانه[1] یا R^2 مشاهده‌شده در برونداد که ۰/۶۸ به‌دست‌آمده است، دقیقاً همان ضریب همبستگی همبسته (وابسته یا زوجی) بین دو متغیر است زیرا در این مثال فقط دو متغیر وجود دارد.

1 . multiple R

برای پرسش‌های بیشتر، پرسش‌های ۴۰، ۹۳ و ۹۴ را مطالعه کنید.



	A	B	C	D	E	F	G	H	I
1	Wins Last Season	Wins This Season		SUMMARY OUTPUT					
2	7	8							
3	6	7		*Regression Statistics*					
4	8	9		Multiple R	0.68				
5	11	10		R Square	0.46				
6	12	9		Adjusted R Square	0.39				
7	8	7		Standard Error	2.02				
8	7	12		Observations	10				
9	3	3							
10	6	5		ANOVA					
11	5	6			*df*	*SS*	*MS*	*F*	*Significance F*
12				Regression	1	27.78	27.78	6.81	0.03
13				Residual	8	32.62	4.08		
14				Total	9	60.40			

بخش ۱۲: روشهای آماری دیگر

پرسش ۹۶: تحلیل واریانس چند متغیره چیست و چه کاربردی دارد؟

شاید اگر بدانید که چندین نوع مختلف از کاربردهای تحلیل واریانس وجود دارد، متعجب شوید. درواقع انواع مختلفی از تحلیل واریانس متناسب با شرایط متنوع یک پژوهش مقایسه‌ای توسعه‌یافته است. یکی از این موارد، تحلیل واریانس چند متغیره[1] یا مانوا است. مانوا هنگامی به‌کاربرده می‌شود که بیش از یک متغیر وابسته وجود داشته باشد. در واقع با این روش، بجای اینکه اثر گروه‌بندی را بر یک متغیر وابسته مشاهده کنیم، اثر گروه‌بندی را بر چندین متغیر وابسته مشاهده خواهیم کرد. مانوا، شکل بسط داده شده آنوا است. آنچه مانوا انجام می‌دهد این است که اثر متغیر مستقل بر هر کدام از متغیرهای وابسته را به‌روشنی نشان می‌دهد. این متغیرهای وابسته باید به‌نوعی ارتباط داشته باشند، یا باید یک دلیل مفهومی برای کنار هم قرار گرفتن آن‌ها وجود داشته باشد. مانوا گروه‌ها را مقایسه می‌کند و به شما می‌گوید که آیا میانگین تفاوت بین گروه‌ها در ترکیبی از متغیرهای وابسته ناشی از شانس بوده است. برای این کار مانوا یک متغیر وابسته خلاصه جدید ایجاد می‌کند که ترکیب خطی از هریک متغیرهای وابسته اصلی است. سپس تحلیل واریانس را با استفاده از این متغیر وابسته مرکب انجام می‌دهد. مانوا به شما می‌گوید که آیا تفاوت معناداری بین گروه‌ها ازلحاظ این متغیر وابسته مرکب وجود دارد یا خیر؟ همچنین نتایج تک متغیره را برای هر یک از متغیرهای وابسته به‌طور جداگانه ارائه می‌دهد.

برای مثال، پژوهشگران دانشگاه ایدیانا[2] به بررسی اثرات جنسیت، نژاد و سطوح تحصیلی بر چگونگی کنار آمدن نوجوانان با فشار همسالان پرداختند. تحلیل داده‌های جمع‌آوری‌شده از طریق آزمون مانوا انجام شد. این پژوهش شامل یک طرح ۲×۴×۵ بود: طرح ۲ (جنسیت: زن و مرد) × ۴ (نژاد: آسیایی- آمریکایی، آفریقایی، اسپانیولی، قفقازی) × ۵ (سطوح تحصیلی: کلاس ۸ تا کلاس ۱۲). متغیرهای وابسته عبارت بودند از پنج خرده مقیاس از پرسشنامه سبک‌های مقابله‌ای نوجوانان. در این پژوهش با کاربست تحلیل واریانس چند متغیره (مانوا) می‌توان اثرات متغیرهای مستقل (جنسیت، نژاد و سطوح تحصیلی) را بر هر یک از پنج متغیر وابسته ارزیابی کرد.

برای مطالعه بیشتر درباره مثال موجود در این پرسش به مطالعه مرجع زیر بپردازید.

Plucker, J. A. (1998). Gender, race, and grade differences in gifted adolescents coping strategies. Journal for the Education of the Gifted, 21(4), 423ت436.

برای پرسش‌های بیشتر، پرسش‌های ۴۰، ۴۹ و ۹۲ را مطالعه کنید.

1 . multivariate analysis of variance (MANOVA)
2 . Indiana University

پرسش ۹۷: تحلیل کوواریانس چیست و چه کاربردی دارد؟

تحلیل کوواریانس (آنکوا[1]) آزمون بسیار جالبی است زیرا به شما اجازه می‌دهد تفاوت‌های اولیه[2] بین گروه‌ها را یکسان‌سازی[3] کنید. تحلیل کوواریانس کمک می کند که ضمن مقایسه میانگین‌های یک یا چند گروه، اثر یک یا چند متغیر کنترل، مداخله‌گر و هم‌پراش (کمکی)[4] را از معادله خارج می‌شود. مقیاس متغیر هم‌پراش یا کنترل یا کوواریت، باید فاصله‌ای یا نسبی باشد. به عنوان مثال در نظر بگیرید که حامی مالی یک برنامه افزایش سرعت دویدن ۲ گروه از قهرمانان مسابقات هستید. از آنجایی که سرعت دوندگان به قدرت عضلانی آنها ارتباط دارد، می بایست نمهیداتی بیاندیشید تا قدرت عضلانی اولیه دوندگان، تاثیری بر روی سرعت دویدن آنها در انتهای برنامه مداخله ای نداشته باشد. در حقیقت شما می بایست سرعت دویدن دو گروه در انتهای برنامه مداخله ای را با کنترل اثر قدرت عضلانی اولیه آنها مقایسه کنید. برای این کار می بایست قبل از انجام مداخله، قدرت عضلانی اولیه آنها را ارزیابی کنید و از طریق ازمون آنکوا، قدرت عضلانی اولیه آنها را کنترل کنید.

از تحلیل کوواریانس معمولاً در طرح‌های پیش‌آزمون- پس‌آزمون استفاده می‌شود. در این طرح‌ها قبل از اینکه آزمودنی‌ها در شرایط آزمایشی قرار گیرند، یک آزمون بر روی آن‌ها انجام می‌شود و بعد از قرار گرفتن در شرایط آزمایشی، پس آزمون بر روی آن‌ها انجام می‌شود. در اینجا نمرات پیش‌آزمون به‌عنوان متغیر تصادفی کمکی به کار می‌روند. محققی را در نظر بگیرید که دو گروه کنترل و آزمایش را انتخاب کرده و آزمونی را اجرا می‌کند تا دانش افراد را قبل از ورود به دوره کاهش اضطراب بسنجد. در گروه آزمایش مهارت‌های کاهش اضطراب به افراد آموزش داده می‌شود؛ در گروه کنترل مداخله ای صورت نمی گیرد. در پایان دوره آموزشی، آزمون دیگری شبیه به پیش‌آزمون برگزار می‌کند که هدف آن بررسی اثر آموزش مهارت‌های غلبه بر اضطراب است. مقایسه نمره دو گروه در آزمون دوم، همراه با حذف اثر احتمالی دانش قبلی افراد که با پیش‌آزمون اندازه‌گیری شده، بهترین تحلیل آماری برای این نوع طرح تحقیقاتی است؛ در این مثال سه نوع متغیر وجود دارد: ۱. متغیر مستقل (گروه اسمی): آموزش مهارت کنترل اضطراب؛ ۲. متغیر وابسته (پس‌آزمون): پیشرفت در درس سنجش و اندازه‌گیری؛ ۳. متغیر کنترل یا هم‌پراش (پیش‌آزمون): دانش اولیه آزمودنی‌ها.

برای پرسش‌های بیشتر، پرسش‌های ۸۶، ۸۸ و ۹۸ را مطالعه کنید.

---

1 . Analysis of covariance (ANCOVA)
2 . initial differences
3 . equalize
4 . Covariate

پرسش ۹۸: تحلیل واریانس با اندازه‌گیری مکرر چیست و چه کاربردی دارد؟

اکنون با نوع دیگری از تحلیل واریانس آشنا خواهید شد. تحلیل واریانس با اندازه‌گیری مکرر[1] نوعی از تحلیل واریانس است که مانند تمامی آزمون‌های تحلیل واریانس به آزمون تفاوت میانگین‌های دو یا چند گروه می‌پردازد. امّا تفاوتش این است که این مقایسه میانگین‌های دو یا چند گروه، در زمان‌های متفاوت یا به تعبیر درست‌تر با اندازه‌گیری‌های مکرر انجام می‌گیرد. درواقع منظور از اندازه‌گیری مکرر این است که مقایسه میانگین‌های گروه‌های ثابت، در زمان‌های متفاوت صورت می‌پذیرد. اندازه‌گیری مکرر به طرحی گفته می‌شود که در آن هر یک از آزمودنی‌ها در معرض بیش از یک متغیر مستقل قرار می‌گیرند. مورد استفاده مناسب این طرح زمانی است که پژوهشگر علاقه‌مند باشد تغییراتی را که در روند زمان در آزمودنی به وجود می‌آید، مشاهده یا اندازه‌گیری نماید. هدف اساسی این طرح، به حداقل رساندن خطاهای ناشی از تفاوت‌های فردی است.

برای مثال، اگر شما وزن یک گروه را هر هفته به مدت یک سال بررسی کنید و قصدتان این باشد که تفاوت‌های هفتگی را بررسی کنید، می‌بایست از تحلیل واریانس با اندازه‌گیری مکرر استفاده نمایید. برای مثال می‌توان به پژوهشی اشاره کرد که در آن محققان به بررسی تعامل جنسیت (زن و مرد) و شخصیت (درون‌گرا و برونگرا) و سال تحصیلی (اول، دوم، سوم و چهارم) در دانشجویان دوره کارشناسی پرداختند. در این پژوهش ۳ عامل وجود دارد. عامل جنسیت (زن و مرد) که عامل بین آزمودنی[2] است، عامل شخصیت (درون‌گرا و برونگرا) که عامل بین آزمودنی است و عامل سال تحصیلی (اول، دوم، سوم و چهارم) که عامل درون آزمودنی[3] است. روش تحلیل داده‌ها آزمون تحلیل واریانس با اندازه‌گیری مکرر بود. زیرا تفاوت جنسیت (زن و مرد) و تفاوت شخصیت (درون‌گرا و برونگرا) در سال اول، دوم، سوم و چهارم باهم مقایسه شد.

پژوهشگری قصد دارد تأثیر اوقات مختلف روز را روی میزان نشاط تعدادی افراد مورد مطالعه قرار دهد. به این منظور ۲۰ فرد (زن و مرد) انتخاب و میزان نشاط آنان را در اوقات مختلف تعیین شده مورد اندازه‌گیری قرار داد. این پژوهشگر می‌خواهد بداند که آیا اوقات مختلف روز بر میزان نشاط زنان و مردان شاغل و بیکار تأثیر دارد؟ برای این پژوهش این فرضیه قابل تعریف و آزمون است که " میزان نشاط زنان و مردان شاغل و بیکار در اوقات مختلف روز متفاوت است". برای آزمون فرضیه می‌بایست از آزمون تحلیل واریانس با اندازه‌گیری مکرر استفاده کرد. در این مطالعه، زمان، عامل درون آزمودنی و نشاط عامل بین آزمودنی است.

برای پرسش‌های بیشتر، پرسش‌های ۸۶، ۸۸ و ۹۸ را مطالعه کنید.

---

1 . Repeated Measures Analysis of Variance
2 . Within-Subjects Variables
3 . Between Subjects Factor

پرسش ۹۹: رگرسیون چندگانه چیست و چه کاربردی دارد؟

در پرسش ۹۴ آموختید که چگونه با داشتن ارزش یک متغیر، قادر خواهیم بود ارزش متغیر دیگر را پیش‌بینی کنیم. اغلب اوقات، پژوهشگران علوم اجتماعی و رفتاری، به این موضوع علاقه‌مندند که چگونه قادر خواهند بود تا از طریق داشتن ارزش‌های بیش از یک متغیر به‌عنوان متغیرهای پیش‌بین، به پیش‌بینی یک متغیر به‌عنوان متغیر ملاک بپردازند. در رگرسیون چندگانه[1]، متغیرهای بیشتری وارد مطالعه می‌شوند و به لحاظ انجام پیش‌بینی‌های دقیق‌تر، توان بیشتری به دست می‌آید. در تحقیقاتی که از تحلیل رگرسیون استفاده می‌شود، هدف معمولاً پیش‌بینی یک یا چند متغیر ملاک از یک یا چند متغیر پیش‌بین است. چنانچه هدف، پیش‌بینی یک متغیر ملاک از چند متغیر پیش‌بین باشد از مدل رگرسیون چندگانه استفاده می‌شود. درصورتی‌که هدف، پیش‌بینی هم‌زمان چند متغیر ملاک از متغیرهای پیش‌بین یا زیرمجموعه‌ای از آن‌ها باشد از مدل رگرسیون چند گانه متغیری استفاده می‌شود.

برای مثال، مطالعات بسیاری نشان داده‌اند که سواد اطلاعاتی و علاقه تحصیلی والدین نظیر داشتن کتابخانه در منزل و عادت مطالعه در روز در منزل، سبب افزایش و بهبود توانایی خواندن فرزندانشان می‌شود. بنابراین پژوهشگران به بررسی نقش متغیرهای سن، سطح تحصیلی، فعالیت‌های آموزشی و فعالیت‌های درسی مشارکتی والدین در پیش‌بینی مهارت‌های زبانی کودکان پرداختند. این پژوهشگران با استفاده از روش رگرسیون چندگانه دریافتند که فعالیت‌های آموزشی والدین با کودکان و سطح تحصیلی والدین توانستند به‌طور معناداری مهارت‌های زبانی کودکان را پیش‌بینی کننده درحالی‌که سن والدین در این پیش‌بینی نقشی نداشت.

برای مطالعه بیشتر درباره مثال موجود در این پرسش به مطالعه مرجع زیر بپردازید.

Lyytinen, P., Laakso, M. L., & Poikkeus, A. M. (1998). Parental contributions to child،s early language and interest in books. European Journal of Psychology of Education, 13(3), 297ت308.

برای پرسش‌های بیشتر، پرسش‌های ۴۰، ۹۴ و ۹۵ را مطالعه کنید.

1 . multiple regression

پرسش ۱۰۰: تحلیل عاملی چیست و چه کاربردی دارد؟

تحلیل عاملی[1] یک تکنیک آماری است که مبتنی بر چگونگی ارتباط سؤالات با یکدیگر در قالب یک خوشه یا عامل است. هر عامل نماینده چندین متغیر یا سؤال مختلف است. البته هر سؤال نیز می‌تواند نماینده یک متغیر باشد. هر عامل شامل چند سؤال به‌هم‌پیوسته است که عموماً یک موضوع را می‌سنجند. درواقع، یک عامل بسیار مؤثر از یک سؤال یا یک متغیر در نشان دادن یا ارزیابی یک پدیده است. در کاربست این تکنیک، هدف آن است که سؤالات به‌هم‌پیوسته، یک عامل عمومی‌تر را شکل می‌دهند. در حقیقت، به چندین سؤال به هم همبسته، که یک پدیده را می‌سنجند، عامل می‌گویند. این ارتباط بین سؤالات، بر اساس یک منطق و پیشینه پژوهشی فرض می‌شود و نه به‌صورت صرفاً تصادفی و بی‌جهت.

برای مثال، پژوهشگران دانشگاه آنتاریوی غربی[2] تلاش کردند تا دریابند که چگونه تجارب تلخ و بدرفتاری‌های دوران کودکی قبل از ۱۲ سالگی می‌تواند سبک رفتاری و تعاملات اجتماعی آن‌ها در نوجوانی را شکل دهد. بدین منظور، داده‌های مختلفی از متغیرهای مختلف جمع‌آوری و ارتباط بین آن‌ها را ارزیابی کردند. آن متغیرهایی که به نظر می‌رسید باهم همبستگی‌های بالایی دارند و بر اساس نظریات موجود و منطق پژوهشی یک مفهوم عمومی را شکل می‌دهند، به‌عنوان یک عامل مشخص و متمایز شدند. برای مثال، یک عامل را با عنوان عامل بدرفتاری-شرم نام نهادند و عامل دیگر را با عنوان عامل تعاملات مثبت نام نهادند. این پرسشنامه ۱۰ سؤال داشت که ۵ سؤال همبسته به‌عنوان عامل بدرفتاری-شرم و ۵ سؤال همبسته بعدی به‌عنوان عامل ارتباطات مثبت متمایز شدند و تمامی این سوالات با هم مرتبط بوده و نمره کل را تشکیل می دادند.

برای مطالعه بیشتر درباره مثال موجود در این پرسش به مطالعه مرجع زیر بپردازید.

Wolfe, D. A., Wekerle, C., Reitzel-Jaffe, D., & Lefebvre, L. (1968). Factors associated with abusive relationships among maltreated and nonmaltreated youth. Developmental Psychopathology, 10(1), 61ـ85.

برای پرسش‌های بیشتر، پرسش‌های ۴۰، ۴۹ و ۹۲ را مطالعه کنید.

1 . Factor Analysis
2 . The University of Western Ontario

University of Guilan Press

100 Questions (and Answers) About Statistics

Edited by:

Neil J. Salkind

Translated by:

Reza Soltani Shal, Ph. D