

با برخی از اشتباهات رایج در تحلیل‌های آماری آشنا شویم

علی اکبر برومیده^۱

حسن شاهقلیان^۲

چکیده

امروزه کمتر پژوهشی است که از تکنیک‌های آماری بی بهره باشد. در حقیقت استفاده از روشهای آماری جهت بررسی، تجزیه و تحلیل، استنباط و ارائه راهکار در کلیه علوم امری رایج و ضروری است. با توجه به نیاز روز افزون به روشهای آماری توانمند و مناسب، آمارشناسان به معرفی و ارائه روشهای جدید و خلاق آماری می‌پردازند. علی‌رغم پیشرفتهای قابل توجه در این زمینه، متأسفانه هنوز هم برخی اشتباهات رایج در بکارگیری روشهای آماری به چشم می‌خورد. این مهم در ایران با توجه به کمبودهای وافر در هر دو بخش نرم افزار و سخت افزار که بویژه در علوم آماری موجود است، این امر را حادث تر می‌نماید. به عبارت دیگر عدم بهینه کردن منابع آماری با نیازهای کاربردی در سطح جامعه از یک طرف و عدم اخذ مشاوره (مناسب و بموقع) از آمارشناسان خبره و متخصص از طرف دیگر باعث بروز اشتباهاتی فاحش در این زمینه می‌باشد. در برخی موارد این اشکالات تا حدی قابل ملاحظه می‌نماید که حتی از نتایج بدست آمده نمی‌توان به اهداف مورد نظر دست یافت. عدم تعیین و تعریف شفاف از جامعه آماری، نحوه نمونه‌گیری، نحوه تعیین اعضای نمونه، نحوه تعیین برآوردها و دقت آنها از عمده مشکلات مربوط به فاز تولید داده می‌باشد. عدم تشریح روش تجزیه و تحلیل و متدولوژی تحقیق، عدم بکارگیری روشها و فنون آماری مناسب با اهداف پژوهش، عدم بررسی پیش فرضها و مفروضات مربوط به فنون آماری بکار گرفته شده از مشکلات عمده در تحلیل (آماری) داده‌هاست. این نوشتار بر آن است تا به برخی ملاحظات در بکارگیری روشها و فنون آماری بپردازد. با توجه به تجارب نگارنده، سعی خواهد شد تا با ارائه مثالهای واقعی و نقد آنها به بررسی اشکالات و اشتباهات موجود در آنها پرداخته شود.

واژه های کلیدی: تحلیل همبستگی، آنالیز واریانس، حجم نمونه، آزمونهای آماری، P-مقدار (P-value).

۱. مقدمه

علم آمار، مطالعه لذت‌بخشی است در باب این موضوع که، چگونه می‌توان جهان ناشناخته‌ای را با گشودن چند دریچه به روی آن، توصیف کرد. با پرداختن به علم آمار، لذت فکر کردن به یک شیوه‌ای کاملاً جدید کشف خواهد شد. هدف عمده علم آمار، استنباط نسبت به کل جامعه از روی یک نمونه است. مشکاتی (۱۳۷۰) را ببینید.

علم آمار مسأله تصمیم‌گیری را بر اساس داده‌های مشاهده شده و در شرایط عدم قطعیت مورد بررسی قرار می‌دهد، ولی برداشت عام (افکار عمومی و توده مردم) از این علم چنین است که آمار با توده‌ای عظیم از داده‌ها سر و کار داشته و به ارائه فراوانیها، درصدها و میانگینها و یا به نمایش داده‌ها در قالب جداول، نمودارها و توصیف آنها می‌پردازد؛ در حالیکه این موارد، امروزه تنها بخش بسیار کوچکی از علم آمار را تشکیل می‌دهند و این در حالیست که سایر جنبه‌های آمار از قبیل کنترل کیفیت، نمونه‌گیری، تحلیل

^۱ دانشجوی دوره MBA مدیریت بازاریابی و تجارت الکترونیک دانشگاه صنعتی لولتاسوند، دانشگاه صنایع و معادن ایران

^۲ گروه ریاضی، دانشگاه شهر کرد

نمونه پایه، نمونه‌ای است که در نوشتجات آماری «تصادفی» خوانده می‌شود. اجزای نمونه با شانس معین و معلوم (کاملاً تصادفی یا تصادفی ساده) از جامعه ای انتخاب می‌گردند که بخوبی و وضوح کامل تعریف شده باشد. جامعه به زبان آماری، نماینده‌ای از کل است که نمونه بخشی از آن است. نمونه کاملاً تصادفی، تنها نمونه‌ای است که می‌توان آن را با استفاده از نظریه آمار و با اطمینان کامل بررسی کرد. اما این نمونه نیز یک اشکال دارد و آن هم این است که برای بسیاری از مقاصد جمع‌آوری آن پرهزینه و مشکل بوده و این هزینه بالا، استفاده و کاربرد آنرا محدود ساخته است. ولی به هر حال روشهای متنوع و جایگزینی در نظریه نمونه‌گیری معرفی شده است که علاقه‌مندان می‌توانند به کتابهای مختلف در این موضوع مراجعه کنند (برای مثال ارقامی و سایرین (۱۳۸۰) را ببینید). اما علی‌رغم آگاهی از موضوع نمونه‌گیری، باز هم ممکن است با مواردی از آریبی در پژوهشها مواجه شد.

برای مثال، زمانی یک روانشناس گزارش داد که عملاً همه مردم روانی هستند، صرفنظر از این واقعیت که چنین استفاده‌ای از کلمات، مفهوم «روانی» را از بین می‌برد، اولین نکته‌ای که در این گزارش مشکوک به نظر می‌رسد، نحوه نمونه‌گیری و انتخاب نمونه‌هاست که این گزارش بر اساس آن استوار می‌باشد. به عبارت دیگر، این روانشناس چه کسانی را مورد مشاهده قرار داده است؟ با نگاهی عمیق به نمونه وی، آشکار گردید که این نتیجه علمی را با مطالعه بیماران خود بدست آورده است که خیلی متمایز و دورتر از نمونه یک جامعه آماری صحیح می‌باشد. چرا که اگر اینها افراد عادی بودند، روانشناس هیچگاه آنها را ملاقات نمی‌کرد!

در اینجا به مثالی دیگر اشاره می‌گردد. در یک بررسی درباره تعداد خوانندگان مجلات، یکی از پرسشهای اساسی این بود که: خانواده شما چه مجله‌ای را مطالعه می‌کند؟ بعد از خلاصه کردن، جدول‌بندی و تحلیل نتایج به نظر می‌رسید که بسیاری از مردم مجله A را می‌خوانند و تنها گروه اندکی طرفدار مجله B هستند! اما شمارگان این دو مجله در آن زمان دقیقاً بر خلاف این یافته‌ها بود. یعنی شمارگان مجله B در حدود ۶۰۰,۰۰۰ نسخه بوده، در حالی که مجله A شمارگانی در حدود ۶۰,۰۰۰ نسخه داشت.

طراحان این بررسی شاید از خودشان پرسیده باشند که آمارگیران (پرسشگران) احتمالاً با افراد نامناسبی تماس گرفته‌اند. اما چنین نبوده و شواهد نشان دهنده این موضوع بوده است که در این بررسی از همه نوع خانوارها و در تمامی نقاط کشور نظرسنجی به عمل آمده است! پس تنها نتیجه‌گیری منطقی این است که تعداد بسیار زیادی از پاسخ‌دهندگان، حقیقت را نگفته‌اند.

تصمیم و طراحی و تحلیل آزمایشها کمتر مورد توجه مدیران و مهندسين قرار می‌گیرند. محلوجی (۱۳۷۵) را ببینید.

در این راستا آگاهی از مفاهیم اساسی علم آمار لازمه هر نوع مطالعه و پژوهش کمی می‌باشد که به نوعی از تکنیک‌های مرتبط با علم آمار بهره می‌گیرد. از این رو برخی معتقدند که آمار جدید یک نظریه اطلاع^۳ است که هدف غائی آن استنباط می‌باشد. هدف عمده علوم آماری، استنباط در مورد یک جامعه بر اساس اطلاعات موجود در نمونه است، لذا گردآوری داده‌ها و اطلاعات دقیق و صحیح که بعنوان ورودی اصلی در علم آمار قلمداد می‌گردند، اهمیت ویژه‌ای دارد. این داده‌ها بطور معمول حاصل اندازه‌گیری کمیتهای فیزیکی، از قبیل طول، دما، مقاومت، ... می‌باشند که صرفنظر از اینکه مشاهدات و داده‌های گردآوری شده به صورت اعداد، حروف، رده‌ها یا هر چیز دیگر بیان شوند، غالباً روشهای کلی تحلیل آماری به صورت یکسان برای آنها بکار می‌رود که جنبه با اهمیت دیگری را در تبدیل داده‌ها به نکات قابل استفاده به عنوان نتایج نهائی استنباط از اجرای یک طرح و یا تحقیق را نمایان می‌نماید. محلوجی (۱۳۷۵) و ارقامی و سایرین (۱۳۸۰) را ببینید.

بنابراین با توجه به اینکه امروزه بیش از هر زمان دیگر استفاده از روشهای آماری در کلیه حوزه‌های علوم، بویژه علوم پایه، پزشکی و انسانی رایج گردیده است، فراگرفتن روشهای آماری بخصوص برای کسانی که ریاضیات دوره دبیرستان را از یاد نبرده‌اند، کار ساده‌ای است. اما تسلط بر مفاهیم اساسی این روشها به نحوی که شرایط استفاده و موارد بکارگیری هر یک درست و بجا باشد، خود به دقت فراوان و از همه مهمتر به تجربه عملی نیاز دارد. عدم دقت درخصوص این موارد موجبات بروز برخی واقعیات تلخ می‌گردد که موضوع اصلی این مقاله خواهد بود.

مشکل در تحلیل‌های آماری استفاده از روشهای نادرست نیست. گاهی اوقات علی‌رغم اینکه روش کاملاً درست است، نتایج آماری به صورت نادرست بیان می‌شوند. این اشتباهات گاهی در طراحی آزمایش، گاهی در تحلیل داده‌های جمع‌آوری شده و گاهی به صورت بیان جملات نادرست و گاهی هم در اثر نادیده گرفتن مفروضات و متغیرها آشکار می‌شوند.

۲. نمونه با آریبی ذاتی

یک قاعده اساسی در علم آمار بیان می‌دارد که برای اینکه یک گزارش مبتنی بر نمونه‌گیری، ارزش بیشتری داشته باشد، باید از نمونه‌ای که معرف جامعه آماری باشد، استفاده کند. منظور از نمونه معرف جامعه، نمونه‌ای است که تمام منابع ایجاد آریبی از آن حذف شده باشند.

^۳ Information theory

^۴ Biasness

۴. تعریف دقیق جامعه آماری و لزوم نمونه‌گیری از همه سطوح متغیرهای کیفی

□ مثال زیر از ارقامی (۱۳۷۴) اقتباس شده است. در تحقیقی چنین نتیجه‌گیری شده بود که «ازدواج فامیلی احتمال عقب ماندگی ذهنی کودکان را افزایش می‌دهد» هر چند در نگاه اول، این نتیجه‌گیری درست به نظر می‌رسد ولی با بررسی دقیقتر مراحل تولید داده‌ها، مشاهده برخی عدم تطابقات باعث بروز نگرانی و شک در این نتیجه‌گیری شد. هدف این نمونه تحقیقی، بررسی رابطه میان «ازدواج با خویشاوندان نزدیک» و «احتمال عقب ماندگی ذهنی فرزندان» بود. جامعه مورد بررسی این تحقیق را کودکان عقب مانده ذهنی شهر مشهد تشکیل می‌داد. در نمونه اخذ شده، در مورد خویشاوند بودن والدین این کودکان سؤال شده بود و بر اساس آن چنین برآورد شده بود که پدر و مادر ۳۱٪ کودکانی که عقب ماندگی ذهنی دارند، فامیل بوده‌اند. لذا محقق بر اساس این برآورد و بالا بودن میزان درصد آن، نتیجه‌گیری کرده است که «ازدواج فامیلی احتمال عقب ماندگی ذهنی کودکان را افزایش می‌دهد».

چنین نتیجه‌گیری چه از نظر آماری و چه از نظر منطقی نادرست است؛ مگر زمانی که درصد کودکان سالمی که والدین آنها خویشاوند نزدیک هستند نیز در دست باشد. برای مثال اگر معلوم شود که پدر و مادر ۳۵٪ کودکان سالم نیز خویشاوند نزدیک هستند، جای هیچ تردیدی نخواهد بود که استنباط فوق کاملاً نادرست بوده است. ارقامی (۱۳۷۴) متذکر می‌شود که «هنگام تحقیق درباره رابطه میان دو متغیر لازم است که متغیرها را در همه سطوح مشاهده کنیم» برای مثال برای پی بردن به رابطه میان خویشاوند نزدیک بودن زوجها و عقب ماندگی ذهنی کودکان آنها، علاوه بر اطلاع از میزان خویشاوندی پدر و مادر کودکان بیمار، باید از میزان خویشاوندی والدین کودکان سالم نیز مطلع بود.

۵. ملاحظات در تحلیل همبستگی

برومیده (۱۳۸۲) برخی اشکالات و اشتباهات رایج در تحلیل همبستگی را مورد بررسی قرار داده که جمع‌بندی و خلاصه اشتباهات رایج در این موضوع را به صورت زیر می‌توان بیان کرد:

- همبستگی جعلی: نوعی همبستگی است که در آن همبستگی قابل تعریف نیست یعنی (عملاً) از دید منطقی غیر قابل قبول

این بررسی تنها مشخص کرده است که پاسخگویان دچار افاده فروشی شده‌اند.

سرانجام نتیجه‌گیری شد که اگر می‌خواهید بدانید مردم چه می‌خوانند، پرسیدن از آنها بی‌فایده است. شاید نتیجه‌گیری بهتر را بتوان در مراجعه به منزل و درخواست خرید تمام مجلات کهنه - برای مثال یک هفته اخیرشان - و تجزیه و تحلیل آنها بدست آورد. البته این تدبیر زیرکانه نیز مشخص نمی‌کند که مردم واقعاً چه می‌خوانند، بلکه تنها عنوان مجلاتی را مشخص می‌کند که در آن زمان ارائه داده‌اند. برای اطلاعات و مثالهای بیشتر به تقوی (۱۳۷۱) مراجعه نمایید.

۳. تعداد نمونه لازم در مقایسه با حجم جامعه

اولین گام و مهمترین مرحله در آغاز هر طرح و پژوهشی، تعیین تعداد نمونه لازم جهت استنباط و نتیجه‌گیری صحیح می‌باشد. یکی از تصورات نادرست در این مقوله این است که اگر از جامعه‌ای به حجم ۱۰،۰۰۰ نمونه‌ای به اندازه -مثلاً- ۵۰ واحد لازم باشد، از جامعه‌ای به حجم ۱۰۰،۰۰۰ عضو، نمونه‌ای به اندازه ۵۰۰ واحد نیاز خواهد بود!

نظریه آماری ثابت کرده است که چنین نیست. به عبارت دقیقتر، حجم نمونه لازم برای یک طرح متناسب با حجم جامعه افزایش نمی‌یابد. برای مثال، کوکران^۱ فرمول تعداد نمونه لازم را برای برآورد یک پارامتر مجهول در جامعه (مثلاً P نسبت جامعه) با دقت d و با ضریب اطمینان $(1-\alpha)$ را به صورت زیر پیشنهاد کرده است.

$$n = \frac{Z^2 P(1-P) / d^2}{1 + (Z^2 P(1-P) / d^2) / N}$$

که در آن N به حجم جامعه دلالت داشته و Z نیز عددی از جدول توزیع نرمال با سطح زیر منحنی $(1-\alpha)$ می‌باشد. جدول ۱ مقادیر حجم نمونه لازم را به ازای $P=0.5$ ، $d=0.07$ و $1-\alpha=95\%$ و مقادیر مختلف حجم جامعه نشان می‌دهد.

از جدول ۱ به راحتی استنباط می‌شود که برای $N > 100000$ عملاً حجم نمونه لازم تغییر نمی‌کند. بنابراین عباراتی از قبیل «نمونه‌ای به حجم ۱۰٪ اندازه جامعه برای دقت مورد نظر کفایت می‌کند» بدون اینکه اندازه جامعه (لااقل بطور تقریبی) معلوم باشد، بسیار دور از صحت و واقعیت است. ارقامی (۱۳۷۴) را ببینید.

همبستگی ناچیز میان دو متغیر به منزله همبستگی آن دو تلقی می‌شود. برای مواردی مشابه بخش ۱۰ همین مقاله را ببینید.

برای رفع این معضل استفاده از یکی از دو روش زیر پیشنهاد شده است (ارقامی (۱۳۷۴):

- ۱- استفاده از روشهایی (مانند روشهای بیزی) که در آنها امکان استفاده از فرضهای صفر غیر نقطه‌ای (بازه‌ای) با سهولت بیشتر وجود دارد.
- ۲- پائین آوردن سطح آزمون (برای نمونه‌های با حجم به اندازه کافی بزرگ) از ۵ درصد یا ۱ درصد معمول؛ به حدود ۰/۱ درصد یا کمتر از آن به نحوی که احتمال نادرست رد کردن H_0 نیز به اندازه کافی کاهش یابد.

۶. غفلت در استفاده از روشهای آمار چند متغیره

گاهی ماهیت اطلاعات و داده‌های مورد نیاز برای استنباط در یک پژوهش به صورت چندمتغیره^{۱۳} بوده و در نتیجه از هریک از مشاهدات چندین متغیر اندازه‌گیری می‌شود. در چنین حالتی ضروریست تا از روشهای آماری چندمتغیره، بجای روشهای تحلیل یک متغیره^{۱۴} استفاده گردد. به هر حال گاهی برخی پژوهشگران به علت عدم توجه کافی و دقت لازم در تجزیه و تحلیل داده‌ها، اطلاعات و داده‌های چند متغیره را یکی یکی و با استفاده از روشهای یک متغیره بررسی می‌کنند. این کار به معنی دور ریختن بخش مهمی از اطلاعات است و در کوواریانس بین متغیرها نهفته است که نتیجه‌ای گمراه‌کننده نیز به همراه خواهد داشت.

خردمندنیا (۱۳۷۷) مثالی را در این مورد متذکر شده است. فرض کنید که در تحقیقی هدف، بررسی تاثیر یک تیمار روی قد و وزن افراد به صورت همزمان است. برای این کار قد (X_1) و وزن (X_2) ۶۴ نفر را اندازه گرفته و نتایج به صورت زیر بوده است

$$S = \begin{bmatrix} s_{11} & s_{12} \\ s_{21} & s_{22} \end{bmatrix} = \begin{bmatrix} 111/5 & 55 \\ 55 & 113/4 \end{bmatrix}$$

$$\bar{X} = \begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \end{bmatrix} = \begin{bmatrix} 67/9 \\ 171/5 \end{bmatrix}$$

حال فرض کنید که پژوهشگر علاقه‌مند به آزمون فرض

$$H_0: \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} = \begin{bmatrix} 66 \\ 174 \end{bmatrix} \text{ در مقابل } H_1: \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} \neq \begin{bmatrix} 66 \\ 174 \end{bmatrix} \text{ باشد. به}$$

عبارتی فرض صفر اینکه همزمان کسانی که قدی برابر با ۱۷۴ سانتیمتر داشته و

می‌باشد. برای نمونه، همبستگی بین "سن پزشکان مرد" و "سن پزشکان زن" که در یک بیمارستان شاغلند.

□ همبستگی ظاهر فریب^{۱۵}؛ که در آن هیچیک از دو عامل (متغیر) بر دیگری اثر نداشته ولی بین آنها همبستگی واقعی وجود دارد. البته بیشترین سفسطه‌های آماری با استفاده از این نوع همبستگی انجام می‌گیرد. برای مثال همبستگی بین حقوق معلمان شهر لندن و قیمت یک نوع سیگار کوبانی بیانگر وجود یک رابطه نزدیک بین آن دو عامل است!

□ همبستگی کوانداز^{۱۶}؛ مسأله دیگری که در تحلیل‌های همبستگی باید مواظب آن بود، نتیجه‌ای است که در آن تداوم همبستگی فراتر از مقدار (محدوده) آماری که بر اساس آن محاسبه گردیده، استنتاج شده است. برای مثال می‌توان به همبستگی بین «میزان بارش باران» و «رشد گیاهان و محصولات» اشاره کرد؛ که اگر در دوره‌ای باران بسیار بیارد، ممکن است حتی منجر به نابودی محصولات نیز بشود.

□ همبستگی دوسویه^{۱۷}؛ نوع رایج دیگر از این اشتباهات، همبستگی دوسویه است که در آن رابطه مشاهده شده واقعی است، اما نمی‌توان مطمئن شد که کدام متغیر علت و کدام معلول است؛ به عبارتی در این حالت نمی‌توان تعیین کرد که کدامیک بر دیگری تفوق^{۱۸} دارد. همبستگی بین درآمد و مالکیت سهام از این نوع است. آیا مالکیت سهام باعث افزایش درآمد میگردد یا افزایش درآمد باعث ارتقا سطح درآمد می‌گردد؟

در ارتباط با تحلیل همبستگی شایان توجه است که بسیار معنی دار بودن همبستگی بین دو متغیر، لزوماً به معنی بسیار همبسته بودن آن دو متغیر نیست. در عمل نیز مواردی که همبستگی بین دو متغیر دقیقاً برابر با صفر باشد، بسیار نادر است. از طرفی متغیرهای بسیاری وجود دارند که همبستگی آنها به اندازه‌ای ناچیز است که باید عملاً ناهمبسته فرض شوند؛ اما نمونه‌ای به اندازه بسیار بزرگ، میزان همبستگی همین متغیرها را بسیار معنی‌دار نشان می‌دهد. البته این مسأله به سبب نقطه‌ای^{۱۹} بودن فرض صفر پیش می‌آید؛ به این مفهوم که H_0 به صورت تساوی دقیق دو کمیت تعریف می‌شود. در حالیکه معمولاً دقیقاً صفر بودن همبستگی دو متغیر از نظر عملی بسیار نادر است و در عمل

Spurious^{۱۵}Bounded^{۱۶}Bilateral^{۱۷}Dominate^{۱۸}Sharp^{۱۹}Multivariate^{۱۳}Univariate^{۱۴}

۷. تحلیل نادرست طرح تودرتو یا لانه‌ای^{۱۶} به صورت عاملی^{۱۷}

فرض کنید چند نوع مواد خام متفاوت توسط سه کارخانه تولید می‌شود. شرکتی که استفاده‌کننده این مواد خام است نسبت به میزان خلوص تولیدات این سه کارخانه مشکوک است. لذا این شرکت مصمم است تا یک طرح عاملی دو طرفه (سه کارخانه و چهار نوع مواد خام) اجرا نماید. در این طرح، برای هر یک از مواد خام فوق از هر کارخانه سه نمونه (مشاهده) به تصادف انتخاب شده و میزان خلوص آن اندازه‌گیری شده است. اطلاعات و داده‌های فرضی برای این حالت در اسماعیلی (۱۳۷۸) ارائه شده است.

جدول آنالیز واریانس برای آزمون برابری میانگین میزان خلوص مواد خام برای هر سه کارخانه مفروض در جدول ۲ آمده است.

از جدول ۲ به راحتی ملاحظه می‌گردد که بسته‌های مواد خام بطور معنی‌داری متفاوت هستند. به علاوه اثر متقابل (معنی‌دار) به عامل غیر معنی‌دار کارخانجات ربط داده شده که می‌تواند تحلیلگر را به این نتیجه‌گیری سوق دهد که این سه کارخانه واقعاً متفاوتند. اما اثر آنها بوسیله اثر متقابل معنی‌دار، پنهان شده است.

اگر کمی ظریفتر به این مسأله دقت شود اشتباه این تحلیل و نتیجه‌گیری نمایان می‌گردد. آیا مواد خام ۱ که کارخانه اول تولید می‌کند همان بسته مواد خام ۱ است که دومین یا سومین کارخانه تولید می‌کند؟ طبیعتاً چنین نیست. بسته‌های مواد خامی که کارخانه اول تولید می‌کند همان بسته‌هایی که کارخانه دوم و سوم تولید می‌کند، نیست. بنابراین اثر متقابل بین این دو عامل توجیه‌پذیر نخواهد بود. به عبارتی بسته‌های مواد خام هر کارخانه فقط به همان کارخانه منتسب می‌شود، نه به کارخانه‌های دیگر، بنابراین بسته‌های مواد خام در کارخانه‌ها لانه کرده‌اند. پس با توجه به روش تحلیل زیر

(کارخانه × مواد خام) SS = (مواد خام) SS + (کارخانه) SS مواد خام SS و بالاخره جدول آنالیز واریانس جدید، یکسان بودن تولیدات سه کارخانه را تأیید می‌کند. (جدول ۳)

۸. تفسیر نادرست نتایج

بطور معمول آزمونها در سطح ۵٪ انجام می‌شوند؛ به این معنی که مقدار آماره آزمون با عددی از جدول متناظر با احتمال ۵٪ مقایسه می‌شود. در صورتیکه مقدار آماره محاسبه شده از عدد بدست آمده از جدول بزرگتر باشد، فرضیه H_0 رد می‌شود و در غیر این صورت دلیلی بر رد آن وجود

وزنشان ۶۶ کیلوگرم می‌باشد، در مقابل فرض مقابل که چنین نیست. بر اساس این داده‌ها فرض صفر فوق رد می‌شود؛ زیرا،

$$T^2 = n(\bar{X} - \mu)' S^{-1} (\bar{X} - \mu) = ۱۰/۸۴$$

$$> ۱۰/۰۹ = \frac{(n-1)p}{n-p} F_{./,./}(۲,۶۴)$$

حال فرض کنید که کواریانس مشاهده شده را نادیده فرض کرده و به جای آزمون چند متغیری به دو آزمون فرض $\mu_1 = ۶۶$ و H_0 و $\mu_2 = ۱۷۴$ (یک متغیره) بسنده می‌شد در اینصورت بر اساس آزمون t -استیودنت خواهیم داشت؛

$$|t_1| = \frac{۶۷/۹ - ۶۶}{\sqrt{\frac{۱۱۷/۵}{۶۴}}} = ۱/۴۲ < ۲$$

$$|t_2| = \frac{۱۷۱/۵ - ۱۷۴}{\sqrt{\frac{۱۱۳/۴}{۶۴}}} = ۱/۸۸ < ۲$$

که هر دو فرض بر اساس آماره‌های فوق، در مقابل فرضهای دو طرفه متناظرشان رد نمی‌شوند. به عبارت دیگر در این حالت دلیلی برای رد هر دو فرض فوق وجود ندارد.

ارقامی (۱۳۷۴) نیز مثالی دیگر را ارائه کرده است. فرض کنید قرار است بر اساس دو نمونه، میانگین نمرات مدارس شاهد را در ۱۲ درس مختلف با مدارس معمولی مقایسه شود. اگر برای چنین مقایسه‌ای ۱۲ آزمون t انجام گیرد، با فرض این که در هر یک از این آزمونها مثلاً ۵٪ احتمال خطای نوع اول وجود دارد، در این صورت احتمال این که حداقل در یکی از این آزمونها مرتکب خطا شویم بسیار بیشتر از ۵٪ خواهد بود! که با محاسبه‌ای ساده می‌توان نشان داد که احتمال ارتکاب به خطای نوع اول برابر ۴۰٪ است که یقیناً احتمال خطای بسیار بزرگی است. در حالیکه اگر برای مقایسه فوق از آزمون T^2 هتلینگ (که تعمیم آزمون t به حالت چند متغیره است) استفاده می‌شد، احتمال خطای نوع اول، در همان سطح ۵٪ باقی می‌ماند.

در مورد محاسبه بازه‌های اطمینان نیز همین مسأله وجود دارد. در مثال فوق اگر برای مقایسه میانگینهای ۱۲ درس در دو جامعه از ۱۲ بازه اطمینان معمولی ۹۵٪ (به کمک توزیع t) استفاده شود، می‌توان نشان داد که میزان اعتقاد به صحت هر ۱۲ بازه اطمینان به جای ۹۵٪، در حدود ۶۰٪ خواهد بود. در صورتیکه اگر از بازه‌های اطمینان همزمان^{۱۵} استفاده شود، میزان اعتقاد به صحت همه بازه‌های اطمینان بطور همزمان همان ۹۵٪ مطلوب خواهد بود.

^{۱۶} Nested factorial design

^{۱۷} Factorial

^{۱۵} Simultaneously

۱۰. تأثیر حجم نمونه در استنباط

اغلب وقتی که حجم نمونه بسیار بزرگ باشد، احتمال استنباط و استنتاج نادرست نیز افزایش می‌یابد، به عبارتی دیگر در این حالت تفاوت‌های (جزئی) مشاهده شده عملاً اهمیتی ندارند. ارقامی (۱۳۷۴) مثال زیر را عنوان کرده است. در تحقیقی با موضوع «مقایسه میانگین نمرات در نواحی مختلف آموزش و پرورش شهرستان مشهد» برای مقایسه میانگینها از آزمون t استفاده شد. در این تحقیق به علت حجم زیاد نمونه، آماره‌های آزمون محاسبه شده بدون استثناء اعداد بسیار بزرگ (بزرگتر از ۵۰) بدست آمده بودند، در حالیکه که در آزمون t مقادیر بزرگتر از ۳ برای هر آماره، نشان دهنده معنی‌دار بودن بسیار زیاد میانگینهای دو جامعه آماری می‌باشد. از اینرو این سؤال پیش آمد که «چگونه میانگین نمرات نواحی مختلف این اندازه با هم تفاوت دارند؟» وقتی به مقادیر واقعی میانگینها مراجعه شد، ملاحظه گردید که اختلاف میانگینها اکثراً در حدود چند صدم است!

این نوع تعارضات از این جا ناشی می‌شود که اغلب میزان معنی‌داری^{۱۸} یا مقدار احتمال^{۱۹} به عنوان میزان تفاوت دو میانگین مورد استفاده قرار می‌گیرد و بر اساس آن تفسیرهای مربوطه ارائه می‌گردد؛ در حالیکه این کار درست نیست. بنابراین بسیار معنی‌دار بودن دو میانگین، به معنی بسیار متفاوت بودن دو میانگین نیست. ممکن است میانگینهای دو جامعه تفاوتی بسیار ناچیز (و از نظر عملی قابل اغماض) داشته باشند، ولی همین تفاوت ناچیز که در نمونه‌های با حجم متعارف کشف نشده و تفاوت غیر معنی‌دار اعلام می‌شود، در نمونه‌هایی با حجم بسیار بزرگ این تفاوت ناچیز میانگینها (بیشتر ظاهر شده و تفاوتها) بسیار معنی‌دار اعلام گردد که عملاً اهمیتی ندارد.

در مثالی دیگر، در تحقیقی که در مورد یکی از شرکتهای خودروسازی جهت ارزیابی سه مدل طراحی شده با هدف انتخاب مدل برتر بر اساس نظرات مشتریان بود، علی‌رغم اینکه دقت لازم در کلیه مراحل و موارد لازم برای تعریف دقیق و مناسب جامعه آماری، حجم نمونه، نحوه نمونه‌گیری و حتی روش تجزیه و تحلیل داده‌ها رعایت شده بود، ولی اختلاف اندک بین میانگین نظرات مشتریان برای هر سه مدل (جدول ۵ را ببینید)، نتیجه‌گیری را به سمت تأیید معنی‌دار بودن اختلاف میانگین نظرات مشتریان در انتخاب این سه مدل سوق می‌داد که از نظر منطقی محل اشکال و تردید بود؛ چرا که این تفاوتها عملاً اهمیتی ندارند! جدول آنالیز واریانس ۴ نیز تأییدکننده معنی‌داری اختلافهای مشاهده شده می‌باشد. در حالیکه مقادیر میانگین و انحراف معیار مشاهدات در جدول ۵ آورده شده است.

نخواهد داشت. در این حالت در اصطلاح گفته می‌شود آزمون در سطح ۵٪ انجام شده است. مرسوم است که پژوهشگر پس از انجام آزمون، ادعا کند «به صحت نتیجه‌ای که گرفته است (رد یا عدم رد H_0) ۹۵٪ اطمینان دارد»

ارقامی (۱۳۷۴) دلایل نادرستی چنین عباراتی را به خوبی بیان کرده است که خوانندگان علاقه‌مند می‌توانند برای اطلاعات بیشتر به مقاله ایشان مراجعه نمایند. همچنین ذکر این نکته ضروریست که نمونه در حکم یک مثال نقض است. یعنی پژوهشگر فرضی که در مورد آن تردید وجود دارد در فرض صفر قرار داده و سپس با اخذ نمونه‌ای سعی می‌کند که به مثال نقضی دست یابد. فرض صفر، درست بودن گزاره‌ای را به صورت عمومی بیان می‌کند. پس اگر مثال نقضی پیدا شود، درستی این گزاره که با سور عمومی بیان می‌گردد، مورد تردید قرار خواهد گرفت. اما اگر نمونه از ارائه شواهد برای نقض این گزاره عاجز بوده و فرض صفر را تأیید نماید، آیا می‌توان واقعاً آنرا پذیرفت؟

در واقع محقق یک یا چند نمونه از هزاران مثال ممکن را بدست آورده و اگر این مثال گزاره فرض صفر را رد نکند نمی‌توان با اطمینان نتیجه گرفت که گزاره فرض صفر همواره صحیح است. پس عدم رد فرض صفر، به معنای پذیرش آن نیست؛ ممکن است آزمون آنقدر توان نداشته باشد که نادرستی فرض صفر را آشکار نماید. ولی چون در عمل محققین خود را مجبور به پذیرش یکی از دو فرض می‌بینند، بنابراین گاهی اوقات عدم رد فرض صفر به منزله پذیرش آن تلقی می‌گردد.

به هر حال، تنها در صورتی که H_0 رد شود، می‌توان میزان اعتقاد به صحت نتیجه‌گیری را برابر با ۹۵٪ اعلان کرد. در غیر اینصورت (پذیرش H_0) میزان اعتقاد به صحت نتیجه‌گیری، برابر با توان آزمون خواهد بود. به عبارت دیگر درجه اطمینان از صحت پذیرش H_0 را توان آزمون می‌نامند.

۹. بی‌توجهی به مفروضات اولیه

گاهی اوقات محقق، بدون بررسی مفروضات و شرایط اولیه حاکم بر تحقیق اقدام به استفاده از یک تکنیک آماری می‌نماید. برای مثال بسیار مشاهده می‌شود که پژوهشگران از آنالیز واریانس، بدون اینکه فرضیه اساسی همگنی واریانسهای سطوح را بررسی نمایند، استفاده می‌کنند در حالیکه استفاده از آنالیز واریانس معمولی در حالتی که واریانس سطوح برابر نباشد، امری نادرست است. در این حالت باید سعی شود با تبدیلات پایدارسازی مناسب، ابتدا واریانس سطوح را یکسان نموده و سپس آنالیز واریانس را بکار برد. در غیر این صورت باید از آنالیز واریانس ناپارامتری (مثلاً کروسکال-والیس) استفاده گردد.

^{۱۸} Significance level

^{۱۹} P-value

امروزه در بیشتر رشته‌های دانشگاهی لاقابل یکی دو درس اجباری در زمینه علوم آماری وجود دارد. این درسها به دلیل اینکه دانشجو از طریق آنها به اهمیت و لزوم استفاده از روشهای آماری پی برده و زبان تخصصی آمار را می‌آموزد، بسیار مفید است. اما آموختن این مقدار درس آمار باعث می‌شود که دانشجویان (بخصوص کسانی که این درس را خوب فرا گرفته‌اند) هنگام برنامه‌ریزی برای تحقیق یا حتی هنگام تحلیل نتایج، خود را بی‌نیاز مشورت با متخصص آمار بدانند. در نتیجه برای استفاده از روشهای آماری یا اصلاً با شخص صلاحیت‌دار مشورت نمی‌کنند یا این که زمانی اقدام به دریافت مشاوره می‌نمایند که داده‌ها (به نحوی) جمع‌آوری شده‌اند و دست متخصص آماری در بسیاری جهات برای انجام دادن یک تحلیل مناسب بسته است؛ که این مورد اخیر بسیار به چشم می‌خورد. ارقامی (۱۳۷۴) را ببینید.

توصیه نگارنده همچون سایر متخصصین آمار این است که چون استفاده از علم آمار در عمل به نظر سهل و ممتنع می‌آید، امکان دارد تله‌های بیشماری در راه محققانی که در آمار تخصص کافی ندارند و یا تجربیات عملی مناسبی را به عنوان پشتوانه ندارند وجود داشته باشد.

ضعف بسیاری از دانش آموختگان علم آمار در استفاده کاربردی از دانسته‌های خود نیز موجبات برخی مشکلات را در راه استفاده مطلوب از این علم، در کلیه سطوح تصمیم‌گیریهای سازمانی و حتی مملکتی پدید آورده است که لزوم برطرف نمودن این نقیصه را نیز نگارنده تأکید دارد.

بنابراین محققین برای موجه بودن و صحت استنتاجها و استنباطهای خود، باید به ظرافت کار توجه نموده و از ابتدای شروع تحقیق، حتی در زمان تهیه طرح تحقیق، حتماً با افرادی که در این زمینه از نظر علمی و عملی تجربه و دانش کافی دارند مشورت نمایند.

چنین مواردی در تجزیه و تحلیل آماری داده‌ها اغلب مواقعی رخ می‌دهد که حجم نمونه بالا باشد. در این حالت توصیه می‌شود که ضریب تعیین^{۲۰} یا اندازه تأثیر^{۲۱} را محاسبه نمود. در این مثال ضریب تعیین برابر ۰/۰۲ بدست آمد و نتیجه‌گیری منطقی دال بر این موضوع بود که دلیلی بر وجود اختلاف معنی‌دار بین مدلها نیست؛ در حقیقت این سه مدل یکسان ارزیابی شده‌اند ولی به علت حجم بالای نمونه این اختلافات کوچک معنی‌دار گزارش شده بودند.

ارقامی (۱۳۷۴) متذکر می‌شود که اگر یک آزمون آماری را به ذره‌بینی تشبیه کنیم که به کمک آن می‌خواهیم تفاوتها را کشف کنیم، می‌توان گفت که حجم نمونه مورد بررسی، میزان درشتنمایی این ذره‌بین را تعیین خواهد کرد. از این رو وقتی قدرت درشتنمایی این ذره‌بین نیز بسیار زیاد باشد، تصویر درشتی از تفاوتهایی که از نظر عملی کاملاً بی‌اهمیت است، بدست می‌آید. در نتیجه اگر ندانیم این فقط تصویر یک تفاوت ناچیز بی‌جهت بزرگ شده است، استنتاجهای نادرستی در گزارشها درج خواهد شد.

۱۱. جمع‌بندی و پیشنهادها

تا اینجا چنین برداشت می‌شود که اشکالات موجود در برخی تحقیقات، ناشی از قصور و کوتاهی در بکارگیری روشهای توانمند آماری است؛ ولی این قصور عمدتاً از عدم اخذ مشاوره آماری مناسب و بموقع است که متأسفانه در گوشه و کنار این مرز و بوم به چشم می‌خورد. هر چند که ممکن است در برخی موارد تعدی در کار باشد که به علت فراوانی بسیار کم این موارد از آن صرف‌نظر می‌گردد. به هر حال در برخی موارد عدم تسلط کافی بر روشهای مرسوم آماری و کمبود دانش فنی و عملی باعث بروز چنین مشکلاتی در تجزیه و تحلیل آماری داده‌ها می‌شود.

جدول ۱- تعداد نمونه لازم در مقایسه با حجم جامعه

حجم جامعه N	۱	۱۰	۱۰۰	۵۰۰	۱۰۰۰	۱۰۰۰۰	بینهایت
حجم نمونه لازم n	۱	۱۰	۶۶	۱۴۱	۱۶۴	۱۹۳	۱۹۶
درصد (حجم نمونه از حجم جامعه)	۱۰۰	۱۰۰	۶۶	۲۸/۲	۱/۶۴	۱/۹۳	۰/۱۹۶

^{۲۰} Determination coefficient

^{۲۱} Effect size

جدول ۲- جدول آنالیز واریانس برای برابری میانگین میزان خلوص مواد خام برای هر ۳ کارخانه

منبع تغییرات	SS	df	MS	F
کارخانجات	۱۵/۰۶	۲	۷/۵۳	۱/۰۲
بسته‌های مواد خام	۲۵/۶۴	۳	۸/۵۵	۳/۲۴
اثر متقابل (کارخانه x مواد خام)	۴۴/۲۸	۶	۷/۳۸	۲/۸۰
خطا	۴۳/۳۳	۲۴	۲/۶۴	
کل	۱۴۸/۳۱	۳۵		

جدول ۳- آنالیز واریانس بر مبنای طرح لانه‌ای

منبع تغییرات	SS	df	MS	F
کارخانجات	۱۵/۰۶	۲	۷/۵۳	۱/۰۲
بسته‌های مواد خام	۶۹/۹۲	۹	۷/۷۷	۲/۲۴
خطا	۴۳/۳۳	۲۴	۲/۶۴	
کل	۱۴۸/۳۱	۳۵		

جدول ۴- آنالیز واریانس یکطرفه برای آزمون معنی‌دار بودن اختلاف در سه مدل مفروض

منبع تغییرات	SS	df	MS	F	P-Value
بین گروه‌ها	۶۴/۶۸	۲	۳۲/۳۴۱	۱۲/۹۶۴	۰/۰۰۰
درون گروه‌ها	۲۸۳۴/۰۱	۱۱۳۵	۲/۴۹۵		
کل	۲۸۹۸/۶۹	۱۱۳۸			

جدول ۵- آمار توصیفی برای هر سه مدل

مدل ۳	مدل ۲	مدل ۱	
۷/۵۶ (۴۰۴)	۶/۹۷ (۳۳۳)	۷/۲۷ (۴۰۲)	میانگین (تعداد نمونه)
۱/۵۳	۱/۷۶	۱/۴۶	انحراف معیار

مراجع

[۱] ارقامی، ناصر رضا؛ ۱۳۷۳، برخی اشتباهات آماری رایج در تحقیقات علوم تربیتی و اجتماعی، گردهمایی شوراها و تحقیق کشور، تبریز.

- [۲] ارقامی، ناصر رضا؛ سنجرى فارسى پور، ناهید و بزرگنیا، ابوالقاسم؛ ۱۳۸۰، مقدمه‌ای بر بررسی‌های نمونه‌ای، انتشارات دانشگاه فردوسی مشهد.
- [۳] اسماعیلی، حبیب؛ ۱۳۷۸، اشتباهات رایج آماری در پروژه‌های تحقیقاتی، اولین سمینار دانشجویی آمار ایران، دانشگاه شهید بهشتی.
- [۴] برومیده، علی اکبر؛ ۱۳۸۳، برخی ملاحظات در تحلیل همبستگی توأم با نقلی بر یک مقاله، مدیرساز (پذیرفته شده برای چاپ).
- [۵] تقوی، مهدی؛ ۱۳۷۱، چگونه با آمار دروغ می‌گویند، انتشارات میریان.
- [۶] خردمندنی، منوچهر؛ ۱۳۷۷، عواقب نادیده گرفتن کوواریانسها، مجله اندیشه آماری، سال سوم، شماره اول.
- [۷] محلوجی، هاشم؛ ۱۳۷۵، آمار مهندسی، مرکز نشر دانشگاهی.
- [۸] مشکانی، محمد رضا؛ ۱۳۷۰، آمار مقدماتی، جلد اول، چاپ دوم، مرکز نشر دانشگاهی.