مباحث اساسی در آمار کاربردی

مؤلف : علیرضا خوشگویان فرد

بسم الله الرحمن الرحیم

مباحث اساسی در آمار کاربردی

نویسنده

علیرضا خوشگویان‌فرد

مرکز تحقیقات صداوسیما

خوشگویان‌فرد، علیرضا، ۱۳۵۴ –
مباحث اساسی در آمار کاربردی/نویسنده علیرضا خوشگویان‌فرد.
تهران: صداوسیمای جمهوری اسلامی ایران، مرکز تحقیقات، ۱۳۹۱.
۴۰۶ ص.
۱۰۰۰۰۰ ریال: ۰-۲۶-۷۳۷۸-۹۶۴-۹۷۸
فیپا
آمار -- راهنمای آموزشی
آمار -- مسائل، تمرین‌ها و غیره
صداوسیمای جمهوری اسلامی ایران، مرکز تحقیقات
HA۲۹/۵ ۱۳۹۱ ۹۶خ۲ف/
۰۰۱/۴۴۲
۲۸۶۵۴۰۶

مباحث اساسی در آمار کاربردی

نویسنده: علیرضا خوشگویان‌فرد

چاپ اول: ۱۳۹۱
تیراژ: ۱۰۰۰ نسخه
قیمت: ۱۰۰۰۰ تومان
چاپ و صحافی: سروش

شابک: ۰-۲۶-۷۳۷۸-۹۶۴-۹۷۸ ISBN:978-964-7378-26-0

تهران، خیابان ولی‌عصر(عج)، خیابان هتل استقلال، ساختمان اداری جام جم، طبقهٔ دوم،
مرکز تحقیقات صدا و سیمای جمهوری اسلامی ایران
تلفن و نمابر: ۲۲۰۱۳۵۸۵

فهرست

پیشگفتار

پژوهشگران همواره در تلاش‌اند با توجه به محدودیت‌های متعدد، بهترین روش را برای اجرای پژوهش خود برگزینند تا به خوبی به پرسش‌های آنان پاسخ گوید، آنان همچنین می‌کوشند در تحلیل داده‌های کمّی از روش‌های آماری مناسب کمک بگیرند تا یافته‌های حاصل، قابل اعتماد باشند. کتاب حاضر می‌تواند راهنمایی برای استفاده از روش‌های آماری «معمول» در تحلیل داده‌ها باشد؛ روش‌هایی که پژوهشگران در اغلب پردازش‌های آماری کم و بیش با آنها سروکار دارند.

نویسنده، محتوای کتاب را بر اساس چندین سال تجربه پژوهشی، مشاوره آماری و آموزش، بویژه در مرکز تحقیقات تدوین کرده است. از همین رو، مخاطبان اصلی آن، پژوهشگرانی هستند که آشنایی مقدماتی با آمار دارند و مایل‌اند در عمل از آن استفاده کنند. همین رویکرد باعث شده است تا نویسنده ضمن پرداختن به مباحث آمار توصیفی و استنباطی، توجهی خاص نیز به نمونه‌گیری، رابطه آمار با روش پیمایشی و نگارش گزارش آماری داشته باشد تا مجموعه‌ای به نسبت جامع در اختیار خوانندگان کتاب قرار دهد. با وجود این، انتخاب عنوان «مباحث اساسی در آمار کاربردی» برای این کتاب به معنای آن نیست که هر روش آماری را که مورد نیاز پژوهشگران است، در آن می‌توان یافت. در واقع، «مباحث اساسی» شامل آن دسته از روش‌های «آمار کاربردی» است که بیشتر مورد مراجعه کاربرانی در حد متوسط است.

مرکز تحقیقات امیدوار است با انتشار کتاب‌های کاربردی برای مخاطبان پژوهشگر، سهمی کوچک در پیشبرد حرکت علمی کشور داشته باشد. از این رو، این کتاب به پژوهشگران، دانشجویان و آنانکه به ارتقای دانش خود در حوزه آمار علاقه‌مندند، تقدیم می‌شود.

مرکز تحقیقات صداوسیما

مقدمه

امروزه، کاربرد دانش آمار چنان گسترش یافته است که کمتر رشته‌ای را می‌توان یافت که شامل یک یا چند درس آماری نباشد. از همین رو، کتاب‌های آموزشی گوناگونی در بازار در دسترس فراگیران آمار قرار دارد که به‌خوبی به مباحث آماری پرداخته‌اند. اغلب این کتاب‌ها با توجه به اهداف نویسنده به حوزه‌ای خاص از دانش آمار محدود شده‌اند؛ برای مثال، کتاب‌های مفیدی انتشار یافته‌اند که به مباحث آمار توصیفی اختصاص دارند یا کتاب‌های متعدد و راهگشایی را می‌توان یافت که بر آزمون‌های آماری متمرکزند و حتی برخی از آنها به طور خاص به آزمون‌های ناپارامتری می‌پردازند. مباحث نمونه‌گیری نیز عموماً در کتاب‌های مستقلی ارائه می‌شوند. بنابراین، کاربران یا فراگیران آمار در طول کار یا تحصیل، می‌بایست به کتاب‌های مختلفی مراجعه کنند تا به نیازهای معمول خود پاسخ دهند.

انگیزه نویسنده از نگارش این کتاب، ارائه هر سه حوزهٔ آمارِ توصیفی، آمار استنباطی و نمونه‌گیری در یک مجموعه است. بنابراین، خواننده ضمن آشنایی با مبانی آمار، هم در روش‌های آمار توصیفی و هم در روش‌های آمار استنباطی ورزیده می‌شود و در کنار این دو، درکی مناسب از نمونه‌گیری و روش‌های آن به دست می‌آورد. بدیهی است از چنین کتابی نباید انتظار داشت تا تمام نیازهای کاربران یا فراگیران آمار را برآورده سازد، بلکه تأکید آن بر طرح مباحث اساسی‌تر و کاربردی‌تر است.

آنچه نویسنده می‌تواند به خوانندگان کتاب وعده دهد، برآوردنِ نیازهایی است که به طور معمول، فرد در جریان پژوهش‌های کمّی و تحلیل‌های آماری با آنها روبه‌روست. در واقع، اگر فراگیران یا کاربران دانش آمار را به سه سطحِ مقدماتی، میانی و پیشرفته تقسیم کنیم، مخاطبان اصلی این کتاب را بیش از همه، گروهی تشکیل می‌دهند که به تازگی وارد سطح میانی شده‌اند. با وجود این، کسانی که اندک اطلاعی از آمار و حساب داشته باشند نیز قادرند از مطالب این کتاب استفاده کنند.

کتاب شامل ۹ فصل است. مفاهیم بنیادینی که پیش‌نیاز فصل‌های بعد هستند، در فصل اول ارائه می‌شوند. فصل دوم به روش‌های آمار توصیفی اختصاص دارد که اگرچه ساده‌اند، در توصیف داده‌ها و مفهوم بخشیدن به آنها بسیار کارگشا و مؤثرند. آمار استنباطی که به گمان بیشتر کاربران آمار، محور اصلی دانش آمار قلمداد می‌شود، در فصل‌های سوم تا ششم بررسی خواهد شد. این سه فصل پرحجم‌ترین بخش این کتاب را تشکیل می‌دهند و

خواننده در آنها به مطالب متنوعی دست می‌یابد؛ آزمون‌های ساده ولی پرکاربرد نظیر آزمون فرضیه درباره میانگین‌ها همراه با بحث‌های به نسبت پیشرفته‌تر مانند تجزیه واریانس، همبستگی و رگرسیون از آن جمله‌اند. برخی از آزمون‌های ناپارامتری و تحلیل جدول‌های توافقی در ششمین فصل ارائه شده‌اند. برآوردیابی از جمله مباحثی است که برخلاف بسیاری از کتاب‌ها، در مباحث آمار استنباطی این کتاب (فصل چهارم) مورد توجه قرار گرفته است.

فصل هفتم به اجمال به مباحث مختلفی در نمونه‌گیری می‌پردازد. خواننده در این فصل درمی‌یابد که تأکید نمونه‌گیری احتمالی در مقایسه با نمونه‌گیری نااحتمالی بر چیست و با روش‌های هر یک از آنها آشنا می‌شود. فصل هشتم نگاهی کاملاً آماری به روش پیمایشی دارد و برخی مسائل و چالش‌های روش‌شناسی را مطرح می‌کند. خواننده با مطالعه این فصل درخواهد یافت که در پژوهش‌های پیمایشی، با چه پیچیدگی‌ها و سختی‌هایی روبه‌روست و باید انتظار چه مشکلاتی را داشته باشد. این فصل را به نوعی می‌توان مکمل فصل نمونه‌گیری یا «نمونه‌گیری در عمل» به حساب آورد. فصل نهم نیز مانند فصل پیمایش، وجه تمایز دیگر این کتاب از کتاب‌های آماری مشابه با آن است؛ زیرا بحث مستقلی را درباره گزارش یافته‌های آماری ارائه می‌دهد به ویژه آنکه برخی از اشکالات رایج در گزارش یافته‌های را یادآور می‌شود.

به خوانندگان توصیه می‌شود هنگام مطالعه این کتاب به چند نکته توجه کنند تا نتیجه مطلوب‌تری را از مطالعه خود به دست آورند.

۱. پیش از مطالعه هر فصل، ابتدا فهرست مطالب آن را مرور کنید و سپس مقدمه کوتاه نخسین صفحه آن را بخوانید. این کار چشم‌اندازی از آنچه در پیش رو دارید، در اختیار شما قرار می‌دهد. به طور معمول، هر فصل، علاوه بر مقدمه کوتاه، مقدمه مفصل‌تری نیز دارد که مطالعه آن به غنای این چشم‌انداز کمک می‌کند.

۲. انتهای هر فصل شامل اصطلاحات آن فصل همراه با معادل‌های انگلیسی آن است. با به پایان رساندن مطالعه هر فصل، این اصطلاحات را نیز مرور کنید تا اطمینان یابید مفهوم آنها را به خوبی درک کرده‌اید. اگر درباره اصطلاحی تردید داشتید، بار دیگر به قسمت مربوط به آنها باز گردید و مطالب را بازخوانی کنید.

۳. فصل‌های اول و سوم و قسمت ابتدایی فصل پنجم شامل مباحث زیربنایی‌تری هستند که در درک مطالب پس از خود نقش مهمی ایفا می‌کنند. اطمینان یابید که مطالب آنها برای شما روشن و قابل درک هستند.

۴. تمرین‌ها نقشی کلیدی در این کتاب ایفا می‌کنند. آنها نه تنها «تمرین» هستند و به ورزیدگی خواننده برای درک مطالب فصل‌ها کمک می‌کنند، بلکه بخشی از مطالب آموزشی را شامل می‌شوند که برای پرهیز از پیچیدگی و دشواری مطالب در متن ارائه نشده‌اند. بنابراین به خواننده توصیه می‌شود، تمرین‌ها را با دقت و جدیت مطالعه کند.

۵. اگرچه رابطه‌های ریاضی در متن فصل‌ها یا تمرین‌ها بیان شده و در برخی موارد محاسبات مرتبط با آنها نیز در قالب مثال در مطالب فصل‌ها یا در تمرین‌ها گنجانده شده‌اند، ولی تأکید چندانی بر انجام محاسبات و به‌کارگیری رابطه‌های ریاضی نیست؛ زیرا وجود نرم‌افزارهای متنوع در این زمینه که نتایج این محاسبات را با دقت و سرعت به دست می‌دهند، ما را از محاسبات دستی بی‌نیاز می‌سازد. اما هدف از بیان این رابطه‌ها یا انجام محاسبات تنها کمک به درک عمیق‌تر روش‌ها و فرایندهای آماری است تا خواننده به شکلی عمیق آنها را دریابد. البته علاقه‌مندان می‌توانند رابطه‌های ریاضی را با دقت بیشتری بررسی و ابعاد محاسباتی آنها را پیگیری یا به کتاب‌های پیشرفته‌تر مراجعه کنند.

در پایان باید از دوستان عزیز خانم مریم حسینی و آقایان غلامرضا محمدی‌مهر، ابراهیم فتحی، رضا امیراحمدی و بویژه سعادت شیخ، مجید مهدیان و زهرا ندایی برای مطالعه تمام کتاب و آقایان عباس محمدی‌شکیبا و علی قربانپور برای مطالعه برخی از فصل‌ها پیش از چاپ آن سپاسگزاری کنم. همچنین از خانم فریبا یارمحمدی برای صفحه‌آرایی و خانم وجیهه فراهانی برای ویراستاری کتاب قدردانی می‌کنم. نویسنده تلاش کرده است کتاب با کمترین اشتباه آماده طبع شود. با این همه از خوانندگان تقاضا می‌کنم هر گونه خطا یا پیشنهادی را از طریق ای‌میل یا نشانی ناشر با نویسنده در میان بگذارند*.

اللهم صل علی محمد و آل محمد و عجل فرجهم

علیرضا خوشگویان‌فرد

---

*Email: khoshgoo@irib.ir

فصل اول

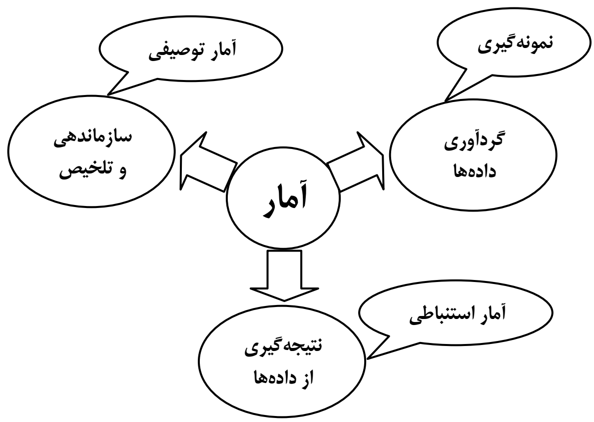
۱

مفاهیم بنیادی

فصل اول به توضیح مفاهیمی اختصاص دارد که در فصل‌های دیگر استفاده خواهند شد و نقشی اساسی در درک سایر مفاهیم این فصل‌ها دارند. بحث با بیان دو برداشت از واژه «آمار» آغاز می‌شود و سپس سه مفهوم کلیدیِ جامعه آماری، متغیر و داده معرفی می‌گردند؛ زیرا روش‌های آماری پیوسته با آنها ارتباط دارند. در ادامه حوزه‌های سه‌گانه دانش آمار شرح داده می‌شوند. شناخت انواع متغیرها از ابعاد گوناگون موضوع بحث بعدی است که در انتخاب روش‌های آماری مناسب برای تحلیل داده‌ها اهمیت ویژه‌ای دارد. در پایان نیز، بحث ابزار گردآوری داده‌ها و مجموعه داده‌ها آمده است.

۱ـ واژه آمار

واژه آمار دارای دو معنی است. در معنی نخست به اعداد و ارقامی اطلاق می‌شود که وضعیت پدیده‌ای را به شکل کمّی بیان می‌کنند. برای مثال، در روزنامه‌ها می‌خوانیم که آمار بزهکاری نسبت به سال گذشته کاهش یافته است؛ یعنی اعداد و ارقامی که میزان بزهکاری را در سال جاری نشان می‌دهند، کمتر از اعداد و ارقام سال گذشته هستند. آمار نیروی انسانی یک سازمان، آمار بیکاری در کشور، آمار مبتلایان به نارسایی کلیه، آمار ریزش باران، آمار پذیرفته‌شدگان در مقطع کارشناسی دانشگاه‌ها یا آمار مراجعه‌کنندگان به یک سایت خبری مثال‌های دیگری از واژه آمار در این معنا هستند.



واژه آمار در معنای دوم خود، عنوانی برای یک شاخه علمی است که در اوایل قرن هجدهم میلادی شکل گرفت. دانش آمار شامل مجموعه‌ای از فنون و روش‌هاست که به منظور «گردآوری داده‌ها»، «سازماندهی و تلخیص» آنها و سرانجام «نتیجه‌گیری» از داده‌ها ابداع شده است. این تعریف بر ابعاد کاربردی دانش آمار اشاره دارد که در سه حوزهٔ آمار توصیفی، آمار استنباطی و نمونه‌گیری جای می‌گیرند.

بحث درباره این سه حوزه را تا بخش سوم این فصل به تعویق می‌اندازیم تا به معرفی برخی از اصطلاحات و واژه‌هایی بپردازیم که از یک سو تعریف دانش آمار و حوزه‌های آن و از سوی دیگر مباحث فصل‌های دیگر به آنها وابسته‌اند. اصطلاح «داده» یکی از این واژه‌هاست که جایگاه ویژه‌ای در تعریف دانش آمار دارد، از این رو بحث بخش بعد به معرفی این اصطلاح اختصاص دارد، ولی برای این منظور، ابتدا لازم است دو اصطلاح «جامعه آماری» و «متغیر» شرح داده شوند

> دوگانگی معنای واژه آمار در معادل انگلیسی آن یعنی Statistics نیز به چشم می‌خورد. گفته می‌شود این واژه برگرفته از واژه لاتین Status به معنی حکومت و کشور (State) است و این امر بر این واقعیت اشاره دارد که پیدایش آمار در نیازهای اطلاعاتی حکومت‌ها ریشه دارد.

تا زمینه تعریف واژه «داده» نیز فراهم شود. گفتنی است اصطلاح متغیر، صـرفنظر از جایگـاهی کـه در تعریف داده دارد، خود از واژه‌هایی است که پیوسته در مباحث آماری به آن اشاره می‌شود.

۲ـ چند اصطلاح

هر لحظه، پدیده‌های بسیار متنوعی پیرامون ما روی می‌دهند که کنجکـاوی پژوهشـگران را برمی‌انگیزند و در ذهن آنان پرسش‌ها و مسائلی را ایجاد می‌کنند. وقوع جُـرم، بارنـدگی در فصل پاییز، مهاجرت از روستا به شهر، تماشای تلویزیون و بیکـاری از جملـه پدیـده‌هـایی هستند که پژوهشگران به بررسی آنها علاقه‌مندند. آنان تمایـل دارنـد پدیـده‌هـا را بـه‌جـای بررسی موردی، به دفعات زیاد بررسی کنند تا بتوانند به نتایج عمومی‌تری دست یابند. برای مثال، بررسی وضعیت بیکاری «آقای X» به عنوان یکی از هزاران بیکار، به دسـتاورد علمـی قابل توجهی درباره بیکاری منجر نمی‌شود، زیرا امری بسیار جزئـی و فـردی اسـت. آنچـه پژوهشگران را ترغیب به پژوهش می‌کند، بررسی پدیده بیکاری در میان انبوهی از بیکـاران است تا از این طریق بتوان به دسته‌بندی انواع دلایل بیکاری، عوارض آن و راهکارهای حل آن در سطحی گسترده پرداخت. بنـابراین، آنـان بـه‌جـای بررسـی‌هـای مـوردی، بـه دنبـال بررسی‌های وسیع‌تر مانند شهر یا حتی یک کشور هستند.

با وجود این، گسترهٔ بررسی پدیده‌ها نمی‌تواند نامحدود و بـدون قلمـرو باشـد، زیـرا از یک سو پژوهشگر باید بداند فعالیت پژوهشی او در میان چه کسانی صـورت مـی‌گیـرد تـا زمان و هزینه تحقیق خود و چگـونگی دسترسـی بـه ایـن افـراد را پـیش‌بینـی و بـرای آن برنامه‌ریزی کند و از سوی دیگر کاربرانِ یافته‌های تحقیق او بدانند این یافته‌ها متعلق به چه گروهی است تا از نتیجه‌گیری‌های افراطی و بی‌منطق مانند تعمیم یافته‌ها به سـایر گـروه‌هـا خودداری کنند. برای نمونه، پژوهشگری که وقوع جرم را بررسی مـی‌کنـد، بایـد در همـان آغاز تحقیق خود مشخص سازد کـه وقـوع جـرم را در میـان چـه گـروه سـنی، جنسـی و جغرافیایی بررسی می‌کند تا به سراغ همان گروه برود و با پیش‌بینی وسعت و پیچیدگی‌های تحقیق خود، در حین اجرای تحقیق دچار سردرگمی نشود. همچنین، اگر پژوهشـگر وقـوع جرم را در میان گروه مشخصی مانند نوجوانـان بررسـی کـرده اسـت، کـاربران یافتـه‌هـای پژوهشی او خواهند دانست که این یافته‌ها لزوماً در میان بزرگسالان صادق نخواهند بـود و باید از تعمیم نتایج آن به این گروه خودداری کنند.

۱ـ۲ـ جامعه آماری

جامعه آماری مجموعه افراد یا اشیایی است که پژوهشگر مطالعه خود را در میان آنها صورت می‌دهد، می‌تواند به تک‌تک آنها دسترسی پیدا کند و یافته‌های پژوهشی خود را به تمامی آنها تعمیم دهد. همه این افراد یا اشیا دارای حداقل یک صفت مشترک هستند که باعث می‌شود در قالب «یک مجموعه» با عنوان جامعه آماری در نظر گرفته شوند. برای مثال، وقتی پژوهشگری محدودهٔ پژوهش خود درباره وقوع جرم را «ساکنان شهر تهران با حداقل ۱۸ سال» تعیین می‌کند، اعضای جامعه آماری تحقیق او دو صفت مشترک دارند: همگی ساکن شهر تهران هستند و همگی حداقل ۱۸ ساله‌اند. بنابراین، تنها افرادی را عضو این جامعه آماری قلمداد می‌کنیم که به طور همزمان هر دو صفت را دارا باشند؛ یعنی یک فرد ۱۵ ساله تهرانی یا یک فرد ۵۰ ساله اصفهانی در جامعه آماری مورد مطالعه او قرار نمی‌گیرند. مثال ۱ نمونه‌های دیگری را از جامعه‌های آماری نشان می‌دهد.

مثال ۱: چند نمونه از جامعه‌های آماری

۱. دانش‌آموزان دختر مقطع دبیرستان کل کشور
۲. فیلم‌های سینمایی اکران‌شده در شش‌ماه نخست سال جاری در سینماهای استان تهران
۳. سرمقاله‌های روزنامه‌های صبح سراسر کشور
۴. خانوارهای مناطق شهری کل کشور
۵. ابیات دیوان شمس
۶. خودروهای پیکان مدل ۱۳۷۰ تا ۱۳۸۰

۲ـ۲ـ متغیر

صرفنظر از صفات مشترکی که برای تعریف جامعه آماری به‌کار می‌رود، اعضای جامعه آماری دارای ویژگی‌های دیگری نیز هستند که پژوهشگران به بررسی آنها علاقه‌مندند. ویژگی‌هایی که تمام اعضای جامعه آماری به طور برابر و یکسان دارا نیستند، متغیر نامیده می‌شوند. در واقع، متغیرها محصول آن دسته از پدیده‌ها هستند که به‌طور یکسان برای تک‌تک اعضای جامعه آماری رخ نداده‌اند و موجب ایجاد وضعیت‌های مختلفی برای آنها شده‌اند. برای مثال، جامعه آماری استان‌های کشور و پدیدهٔ بارندگی را در نظر بگیرید. از آنجا که این پدیده به شکلی یکسان در میان استان‌ها رخ نمی‌دهد، میزان بارندگی در استان‌های مختلف برابر نیست پس «میزان بارش» در این جامعه آماری یک متغیر است. مثال ۲ متغیرهای دیگری را در جامعه‌های آماری مختلف نشان می‌دهد.

مثال ۲: نمونه‌هایی از متغیرها در جامعه‌های آماری مختلف

۱. دانشجویان یک دانشگاه را به عنوان جامعه آماری درنظر بگیرید. «رشته تحصیلی» ویژگی است که می‌تواند متغیر باشد، زیرا دانشجویان دارای رشته‌های تحصیلی مختلفی هستند.

۲. قد ساکنان یک شهر متغیر است؛ زیرا همه مردم هم‌اندازه نیستند.

۳. میزان آلایندگی خودروهای پیکان مدل ۱۳۷۰ تا ۱۳۸۰ را می‌توان یک متغیر دانست؛ زیرا حجم آلایندگی حاصل از خودروها برابر نیست.

یک ویژگی می‌تواند در یک جامعه آماری، متغیر، ولی در جامعه آماری دیگری ثابت باشد. برای مثال، رنگ پوست در جامعه آماری مردم جهان متغیر است (سفید، زرد، سیاه و سرخ پوست). همین ویژگی می‌تواند در جامعه آماری مردم آفریقا به یک ویژگی ثابت تبدیل شود (به آفریقای جنوبی گیر ندهید!). همچنین، اگر تحقیقی به جامعه مردان محدود باشد، ویژگی جنسیت متغیر نخواهد بود؛ زیرا همه اعضای این جامعه آماری، مرد هستند.

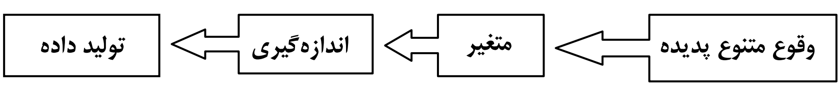
۳ـ۲ـ داده

داده‌ها[1] نتیجهٔ اندازه‌گیری متغیر از اعضای جامعه آماری هستند. به عبارت دیگر، پدیده‌ای در میان اعضای جامعه آماری مفروضی به شکلی متنوع رخ می‌دهد و به این ترتیب، ویژگیِ متغیری را در میان اعضای این جامعه آماری ایجاد می‌کند. اکنون با اندازه‌گیری این ویژگی از اعضای جامعه آماری و ثبت اندازه‌های گردآوری‌شده، داده‌هایی درباره آن متغیر از جامعه آماری تولید می‌شود. بنابراین، داده‌ها «محصولِ» وقوع پدیده‌ای هستند که با اندازه‌گیری متغیر متناظر با آن از اعضای جامعه آماری به دست می‌آید.

> پدیده‌ها منابع بالقوهٔ تولید داده‌ها هستند که در قالب متغیرها به شکلی «نظام‌مند» تعریف می‌شوند و پس از اندازه‌گیری، در قالب داده‌ها نمود پیدا می‌کنند.

برای نمونه، در طول فصل پاییز در استان‌های مختلف بارندگی صورت می‌گیرد (پدیده‌ای رخ می‌دهد). براین اساس، متغیری با عنوان ارتفاع بارندگی در استان‌ها ایجاد می‌شود که با اندازه‌گیری آن، داده‌های ارتفاع بارندگی تولید می‌شوند. شکل ۱ فرایند چهار مرحله‌ای تولید داده را به تصویر کشیده است.

شکل ۱: فرایند تولید داده



۱. معادل «داده» در انگلیسی واژه مفرد Datum است که جمع آن Data است. از آنجا که در مباحث آماری، داده ماهیتی کمّی دارد، از اصطلاح اندازه‌گیری در تعریف آن استفاده شده است. با وجود این، داده به معنای «اطلاعات»، کاربردهای عام‌تری نیز دارد؛ نظیر داده‌های تاریخی یا داده‌ها و شواهد پزشکی در علوم تجربی.

از آنجا که متغیر برخلاف داده پیش از اندازه‌گیری وجود دارد (کافی است پدیده‌ای رخ دهد یا تصور شود، تا بر اساس آن، متغیر یا متغیرهایی تعریف شوند)، این امکان نیز فراهم است که اندازه‌های بالقوه آن را بدون اندازه‌گیریِ عملی فهرست کنید. برای مثال، اندازه متغیر ارتفاع بارندگی می‌تواند از ۰ تا ۱۰۰۰ میلی‌متر باشد، بدون آنکه به اندازه‌گیری آن در بین استان‌ها اقدام کرده باشید. بنابراین، ممکن است برخی از اندازه‌هایی که به‌طور «نظری» برای متغیر برشمرده‌اید، در عمل یعنی پس از اندازه‌گیری، هرگز مشاهده نشوند. برای نمونه، شاید ارتفاع بارش در هیچ‌یک از استان‌ها برابر با ۱۵ میلی‌متر نشود، هرچند وقوع چنین ارتفاع بارشی نامعقول و ناممکن نیست. از این رو، بین مقادیر متغیر و داده‌های مربوط به متغیر تمایز قایل می‌شویم. مقادیر متغیر شامل تمام اندازه‌هایِ ممکن برای آن متغیر است، در حالی که داده‌ها، آن دسته از مقادیر متغیر هستند که در عمل از اعضای جامعه آماری اندازه‌گیری شده‌اند.

۳ـ حوزه‌های آمار

به‌طور معمول، مجموعه داده‌های گردآوری‌شده شامل چند صد تا چندین میلیون داده (عدد) است که بررسی تک‌تک این داده‌ها و خواندن آنها نه تنها گیج‌کننده و دشوار، که بی‌فایده است. برای مثال، جمعیت‌شناسان به بررسی تک‌تک افراد تولد یافته یا فوت‌شده علاقه‌مند نیستند، بلکه به برآیند کلی تولد و مرگ در سراسر جامعه تحت بررسی خود توجه دارند. به عبارت دیگر، نکته با اهمیت، تولد نوزادی در فلان خانواده یا فوت فردی از خانواده دیگر نیست، زیرا این دو، داده‌هایی جزئی (خُرد) از دو خانواده هستند. آنچه برای جمعیت‌شناسان مهم است، اطلاعات کلانی است که از مجموعه خانواده‌ها گردآوری می‌شوند و اطلاعاتی را درباره وضعیت عمومی جامعه، و نه خانواده یا فردی خاص، دراختیار می‌گذارند. بنابراین به روش‌هایی نیاز است که به مجموعه‌ای گسترده از داده‌های انبوه سروسامان بخشد؛ به طوری که به جای اطلاعات انفرادی و جزئی، اطلاعات کلانی از آنها استخراج شود که وضعیت کل داده‌ها را انعکاس دهد. به عبارت دیگر، روش‌هایی لازم است تا با به‌کارگیری آنها بتوان گزارش مختصر و مفیدی از دادها تهیه و ارائه کرد.

> آمار، دانش کار کردن با داده‌ها و معنا بخشیدن به آنهاست و رویکردی کاملاً کمّی دارد. ابزاری است کارامد در خدمت علومی که به مطالعهٔ کمّی پدیده‌ها گرایش دارند و در پی دستیابی به یافته‌هایی برای توصیف، تبیین و پیش‌بینی پدیده‌ها در سطحی کلان هستند.

برای روشن شدن موضوع، کارنامه یک دانش‌آموز دبیرستانی را با نمرات ۱۲ درس، مطابق جدول زیر درنظر بگیرید.

فیزیک	۱۲	جبر	۱۰	فارسی	۱۹	ورزش	۲۰
شیمی	۱۵	انگلیسی	۱۷	عربی	۱۴	کامپیوتر	۱۸
هنر	۱۸	تاریخ	۱۲	هنر	۱۸	تاریخ	۱۲

این کارنامه، اطلاعات انفرادی تک‌تک درس‌ها را به‌دست می‌دهد، ولی در نهایت نمی‌توانیم درباره وضعیت عمومی این دانش‌آموز قضاوت کنیم (مگر در موارد استثنایی که «تمام نمرات» دانش‌آموزی بالا یا پایین باشد). آیا باید نمرات بالای او در درس‌هایی مانند فارسی، ورزش، هنر و کامپیوتر را مبنای قضاوت خود قرار دهیم و او را دانش‌آموزی زرنگ قلمداد کنیم یا نمرات پایین او در درس‌هایی مانند جبر، تاریخ و هندسه را نشانه‌هایی از تنبلی او بدانیم. اگرچه تمرکز بر روی نمره هر درس می‌تواند اطلاعاتی درباره همان درس به‌دست دهد، ولی تک‌تک این نمرات یعنی داده‌های جزئی، اطلاعاتی درباره وضعیت عمومی این دانش‌آموز در اختیار ما نمی‌گذارند، بلکه موجب گیج شدن و سردرگمی ما نیز می‌شوند.

معدل، کمّیتی است که می‌تواند «چکیده‌ای» از این ۱۲ نمره، در قالب تنها یک نمره باشد. به این ترتیب، معدل، به نوعی وضعیت عمومی درس‌های آن دانش‌آموز را نشان می‌دهد، بدون آنکه ما را با تک‌تک ۱۲ نمره درگیر و به مقایسه آنها با یکدیگر مجبور سازد. معرفی شاخص‌هایی نظیر معدل که به تلخیص انبوهی از داده‌ها می‌پردازند، از جمله موضوعاتی است که در حوزه آمار توصیفی در فصل دوم مورد بحث قرار می‌گیرد.

گردآوری داده‌ها و سازماندهی آنها تا زمانی که به تصمیم‌گیری و نتیجه خاصی منجر نشود، کار چندان مفیدی نخواهد بود. در واقع، داده‌ها زمینه‌ای را فراهم می‌آورند تا درباره پدیدهٔ تحت بررسی و قضاوت، تصمیم‌گیری و برنامه‌ریزی کنیم. به عبارت دیگر، هدف اصلی از گردآوری داده‌ها شناخت واقع‌بینانه پدیده‌ها با زبان اعداد و ارقام است؛ به طوری که این اعداد و ارقام انعکاس‌دهنده و آیینه واقعیتِ آن پدیده‌ها باشند و تصویری روشن از آنچه می‌گذرد، به ما نشان دهند. بر این اساس، قادر خواهیم بود عملکرد خود را ارزیابی کنیم و درباره درستی یا نادرستی فعالیت‌ها و اقدامات کنونی خود، دست به قضاوت بزنیم و به تبع آن برای آینده تصمیم بگیریم و برنامه‌هایی را برای بهبود وضعیت کنونی و موفقیت در آینده طراحی کنیم.

برای مثال، اگر یک نامزد انتخاباتی بداند از رقیب خود عقب‌تر است، برنامه‌های تبلیغی وسیعی را تدارک می‌بیند تا این عقب‌افتادگی را جبران کند و حتی از رقیب خود پیشی بگیرد. برای این منظور نیز ابتدا باید به‌درستی درباره وضعیت خود نسبت به رقیبش قضاوت کند، زیرا اگر به اشتباه تصور کند از رقیب خود جلوتر است، فرصت‌های موفقیت را از دست می‌دهد و اگر به اشتباه تصور کند از او عقب‌تر است، زمان و هزینه را صرف تبلیغاتی می‌کند که به آنها نیازی نبوده است و به این ترتیب از کارهای ضروری دیگر باز می‌ماند.

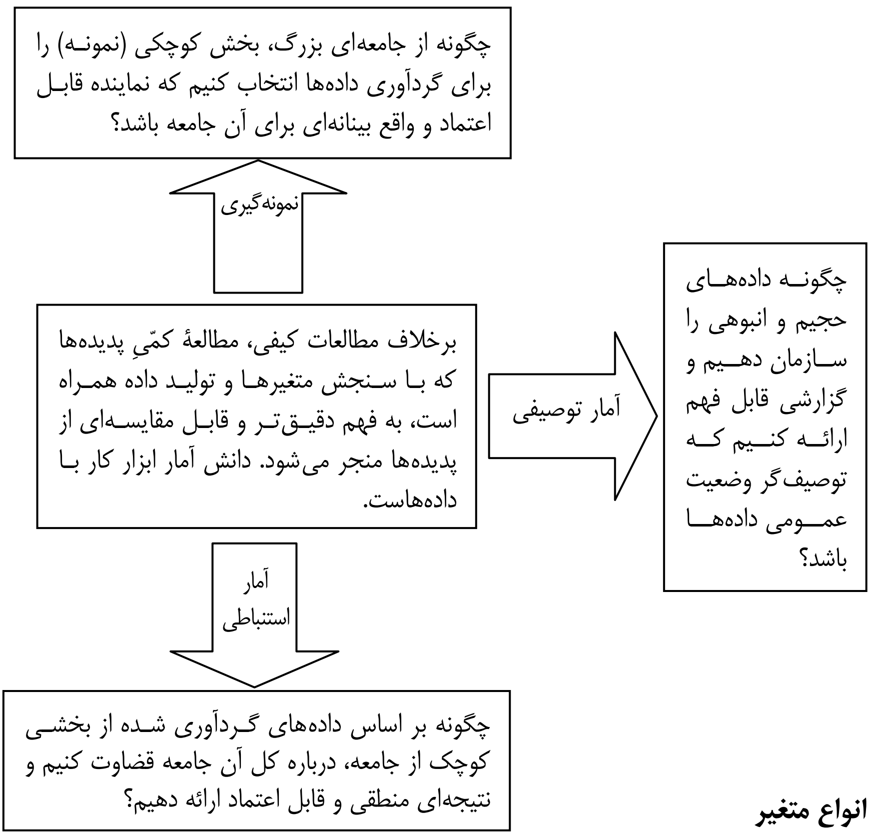
آمار را به دلیل قابلیت روش‌های آمار استنباطی، منطق استقرا (Logic of Induction) نیز می‌نامند، زیرا سازوکاری را در اختیار می‌گذارد تا با استفاده از آن بتوان با اطمینان مشخصی، از داده‌های جزئی (نمونه‌ای از جامعه) به نتیجه‌گیری‌های کلان دست یافت.

نخستین راه گردآوری داده‌ها برای قضاوت درباره یک پدیده، گردآوری داده از تمام اعضای جامعه آماری است که به آن سرشماری می‌گویند. سرشماری‌ها، داده‌های «کاملی» را از جامعه آماری به دست می‌دهند و به این ترتیب می‌توان درباره پدیدهٔ تحت بررسی قضاوتی «قطعی» کرد. برای مثال، اگر رضایت شغلی تمام کارکنان سازمانی سنجیده شود، به‌راحتی می‌توان بر اساس این داده‌ها تعیین کرد که «دقیقاً» چه نسبتی از کارکنان از شغل خود راضی و چه نسبتی از آنان ناراضی هستند. سرشماریِ اعضای یک جامعه آماری کوچک امکان‌پذیر است، زیرا در مدتی کوتاه با هزینه قابل قبولی به انجام می‌رسد، ولی سرشماری از جامعه‌های آماری بزرگ وقت‌گیر و هزینه‌بر است و اغلب توجیه‌پذیر نیست، پس به انتخاب نمونه‌ای از آن اکتفا می‌شود.

از این رو، به سازوکاری نیاز است که بدون صرف زمان و هزینه زیاد، داده‌های «واقع‌بینانه و قابل‌اعتمادی» را از جامعه‌های آماری بزرگ در اختیار بگذارد؛ داده‌هایی که نمایندهٔ واقعی آن جامعه باشند و گمراه‌کننده و غلط نباشند. روش‌های نمونه‌گیری با چنین هدفی ابداع شده‌اند و در پی آن هستند که با گردآوری داده‌ها تنها از بخش کوچکی از جامعه تحت بررسی، و نه تمام آن، اطلاعاتی را به‌دست دهند که به تمام آن جامعه «قابل‌تعمیم» باشد (تصویری واقع‌بینانه و قابل‌اعتماد از آن جامعه ارائه دهد). این بخش کوچک از جامعه را «نمونه» می‌نامند که با سازوکاری علمی انتخاب می‌شود و به این ترتیب، می‌تواند نمایندهٔ قابل اعتمادی برای کل آن جامعه باشد. بحث درباره این سازوکار «علمی» در فصل هفتم کتاب ارائه خواهد شد.

اکنون این پرسش مطرح است که چگونه می‌توان درباره یک پدیده در شرایطی قضاوت کرد که داده‌های گردآوری‌شده از آن تنها به «نمونه» و نه تمام جامعه آماری محدود است؟ موضوع آمار استنباطی که در فصل‌های سوم تا ششم به آن پرداخته می‌شود، ارائه راهکارهایی است که ما را قادر می‌سازد نتایج «قابل‌تعمیمی» را از نمونه برای جامعه آماری به دست آوریم. حوزه‌های مختلف دانش آمار و پرسش‌هایی که هر حوزه به آنها پاسخ می‌دهد، در شکل ۲ به تصویر کشیده شده است.

شکل ۲: پرسش‌های حوزه‌های سه‌گانهٔ آمار



۴ـ انواع متغیر

ابتدا انواع متغیر بر اساس سطح سنجش توضیح داده می‌شود و پس از آن، دسته‌بندی متغیرها به دو نوع پیوسته و گسسته معرفی می‌گردد. متغیر مقوله‌ای، به عنوان یکی از رایج‌ترین متغیرها بویژه در علوم انسانی، موضوع بحث پایانی این بخش است.

۱ـ۴ـ انواع متغیر بر اساس سطح سنجش

در بخش پیشین اشاره شد که متغیر به آن ویژگی‌ای گفته می‌شود که وضعیت‌های متفاوتی در میان اعضای جامعه آماری دارد. برای مثال، ویژگی قد برای فردی از جامعه آماری ۱۵۰ سانتیمتر و برای فرد دیگری ۱۷۵ سانتیمتر است. بنابراین، ۱۵۰ و ۱۷۵ دو وضعیت از این ویژگی هستند که در جامعه وجود دارند. همچنین، فردی از زندگی خود کاملاً ناراضی و فرد دیگری نسبتاً راضی است، پس «کاملاً ناراضی» و «نسبتاً راضی» دو وضعیت از ویژگی رضایت از زندگی هستند.

در آمار، وضعیت‌های متغیر با اعداد بیان و از یکدیگر متمایز می‌شوند، زیرا برخلاف توصیف کیفی پدیده‌ها، توصیف کمّی آنها به فهم دقیق‌تری از پدیده‌ها می‌انجامد و مقایسه پدیده‌ها را با یکدیگر ممکن می‌سازد. از یک رویکرد، سنجش یا اندازه‌گیری شامل فرایند انتساب اعداد به وضعیت‌های مختلف متغیر است که به کمّی کردن آن منجر می‌شود. هر عدد، مقدار یا اندازه وضعیتی از متغیر است. برای مثال، دانش‌آموزان یک کلاس از اطلاعات ریاضی یکسانی برخوردار نیستند، پس ویژگی میزان اطلاعات ریاضی، وضعیت‌های مختلفی در میان دانش‌آموزان کلاس دارد. برای نشان‌دادن تفاوت وضعیت اطلاعات ریاضی دانش‌آموزان از نمرات آزمون ریاضی استفاده می‌شود (این آزمون «وسیلهٔ» سنجش اطلاعات ریاضی دانش‌آموزان است). هر نمره وضعیتی از اطلاعات ریاضی را نشان می‌دهد. به این ترتیب نمرهٔ هر فرد در این آزمون، مقدار یا اندازه اطلاعات ریاضی او خواهد بود. به عبارت دیگر، نمره ۱۷ به وضعیتی از اطلاعات ریاضی اشاره دارد که متفاوت با وضعیت مربوط به نمره ۱۳ است. بنابراین، این اندازه‌ها قادرند وضعیت‌های مختلف اطلاعات ریاضی دانش‌آموزان را از یکدیگر متمایز سازند.

اندازه‌گیری متغیر یا همان انتساب اعداد به وضعیت‌های مختلف آن متغیر به ایجاد چهار نوع متغیرِ اسمی، ترتیبی، فاصله‌ای و نسبتی منجر می‌شود که برآمده از فرایند اندازه‌گیری و به تبع آن، قابلیت‌های محاسباتی این اعداد است. در ادامه به شرح هر یک از این چهار نوع متغیر می‌پردازیم.

۱ـ۱ـ۴ـ متغیر اسمی

متغیر اسمی دارای مقادیری است که ارزش محاسباتی ندارند؛ یعنی عمل‌های چهارگانهٔ حساب را نمی‌توان روی این مقادیر صورت داد. در واقع، فرایند کمّی کردن وضعیت‌های

این متغیر کاملاً قراردادی است؛ به طوری که مقادیر منتسب به وضعیت‌ها صرفاً نام‌ها و برچسب‌های جدیدی برای وضعیت‌های مختلف آن متغیر محسوب می‌شوند. برای مثال، متغیر جنس دارای دو وضعیت مرد و زن است و از این پس قرار می‌گذاریم که عدد ۲ جانشینِ «مرد» که یک وضعیت از متغیر جنس است و عدد ۱ جانشین وضعیت دیگر آن یعنی «زن» باشد. به سخن دیگر، اعداد ۱ و ۲ به عنوان مقادیر متغیر جنس، تنها نام‌ها و برچسب‌های جدیدی برای وضعیت‌های مرد و زن هستند و ارزش ریاضی ندارند.

از آنجا که انتساب اعداد به وضعیت‌ها در متغیر اسمی قراردادی است، پس می‌توان از هر عددی به عنوان مقدار متغیر استفاده کرد؛ مشروط بر آنکه این قرارداد (یعنی چگونگی انتساب) به روشنی بیان شود. برای نمونه، می‌توان به «مرد» و «زن» به ترتیب اعداد ۱۰ و ۵۵ را نسبت داد یا برعکس عدد ۵۵ را به «زن» و عدد ۱۰ را به «مرد» منتسب ساخت. آنچه اهمیت دارد، بیان قراری است که برای انتساب اعداد گذاشته شده است تا هر فرد دیگری نیز مفهوم این اعداد را به درستی درک کند. مثال ۳، نمونه‌های دیگری از متغیر اسمی را نشان می‌دهد.

مثال ۳: نمونه‌هایی از متغیر اسمی

۱. متغیر نژاد دارای چهار وضعیت سرخ پوست، سفید پوست، سیاه پوست و زرد پوست است. برای کمّی کردن این متغیر، می‌توان از چهار عدد ۱=سرخ پوست، ۲=سفید پوست، ۳=سیاه پوست و ۴=زرد پوست استفاده کرد.

۲. متغیر وضعیت اشتغال با دو وضعیت شاغل و بیکار را در نظر بگیرید. می‌توان از دو عدد ۱۰=شاغل و ۲۰=بیکار برای نشان دادن این دو وضعیت استفاده کرد.

۳. متغیر زادگاه با دو وضعیت شهر و روستا می‌تواند به صورت ۰=شهر و ۱=روستا بیان شود.

۴. متغیر نوع خانه مسکونی با وضعیت‌های استیجاری، شخصی و سازمانی می‌تواند به صورت ۵=استیجاری، ۱۰=شخصی و ۱۵=سازمانی کمّی شود.

۵. متغیر استان محل سکونت دارای ۳۰ وضعیت (به تعداد استان‌های موجود در کشور تا سال ۱۳۸۹) است که می‌توان به ترتیب حروف الفبای نام استان‌ها، از اعداد ۱ تا ۳۰ برای آنها استفاده کرد.

۶. رنگ مورد علاقه افراد از میان سه رنگ اصلی قرمز، آبی و زرد متغیری است که می‌تواند به شکل ۱۰۰=آبی، ۲۰۰=قرمز و ۳۰۰=زرد بیان شود.

۲ـ۱ـ۴ـ متغیر ترتیبی

متغیر ترتیبی مشابه متغیر اسمی است، با این تفاوت که وضعیت‌های آن قابل مرتب شدن است. از این رو، برخلاف متغیر اسمی، اعداد به گونه‌ای به وضعیت‌های این متغیر منتسب می‌شوند که به ترتیب این وضعیت‌ها نیز اشاره داشته باشند. برای مثال، رنگ مورد علاقه مردم از میان سه رنگ اصلی قرمز، آبی و زرد متغیری اسمی است، زیرا نمی‌توان برای این سه رنگ هیچ ترتیبی را در نظر گرفت (قرمز پایین‌تر از آبی یا زرد پس از قرمز نیست)، ولی برای متغیری مانند مقطع تحصیلیِ دانشگاه چهار وضعیت کاردانی، کارشناسی، کارشناسی ارشد و دکترا وجود دارد که دارای ترتیب هستند؛ یعنی وضعیت کارشناسی پس از کاردانی، وضعیت کارشناسی ارشد پس از کاردانی و کارشناسی و سرانجام وضعیت دکترا پس از سه وضعیت دیگر قرار دارد. به همین دلیل، در انتساب اعداد به این وضعیت‌ها نیز طوری قرارداد می‌شود که ترتیب اعداد مطابق با ترتیب وضعیت‌ها باشد. بنابراین، متغیر مقطع تحصیلی مثلاً به صورت ۱=کاردانی، ۲=کارشناسی، ۳=کارشناسی ارشد و ۴=دکترا کمّی می‌شود.

متغیر ترتیبی نیز مانند متغیر اسمی به شکل‌های متنوعی قابل کمّی شدن است، مشروط بر آنکه ترتیب وضعیت‌ها در ترتیب اعداد رعایت شده باشد. جدول ۱ پنج شکل دیگر را برای کمّی کردن متغیر مقطع تحصیلی دانشگاه ارائه می‌دهد.

جدول ۱: برخی از شکل‌های کمّی کردن متغیر مقطع تحصیلی

شکل‌ها	وضعیت‌های متغیر مقطع تحصیلی دانشگاه			
	دکترا	کارشناسی ارشد	کارشناسی	کاردانی
شکل اول	۱۲۰	۷۴	۵۰	۱۷
شکل دوم	۲۰۰۰	۱۵۰۰	۱۰۰۰	۵۰۰
شکل سوم	۱۵۰	۸۰	۶۰	۴۰
شکل چهارم	۵	۴	۳	۲
شکل پنجم	۳	۲	۱	۰

اعداد منتسَب در جدول ۱ هماهنگ با ترتیب وضعیت‌ها انتخاب شده‌اند، یعنی با بالا رفتن مقطع تحصیلی، اعداد منتسب نیز افزایش می‌یابند، یعنی اگر وضعیتی پس از وضعیت دیگری قرار دارد، مقدار منتسب به آن وضعیت نیز بیشتر است. توجه به این نکته ضروری است که مقادیر متغیر ترتیبی نیز مانند متغیر اسمی دارای ارزش محاسباتی نیستند، ولی ویژگی مرتب

شدن برای این مقادیر برقرار است. مثال ۴ نمونه‌های دیگری از متغیر ترتیبی را نشان می‌دهد. در این مثال، به قابلیت مرتب شدن وضعیت‌ها و رعایت ترتیب اعداد توجه کنید.

مثال ۴: نمونه‌هایی از متغیر ترتیبی

۱. متغیر رضایت از زندگی با چهار وضعیت کاملاً ناراضی، ناراضی، راضی و کاملاً راضی را می‌توان به صورت ۱=کاملاً ناراضی، ۲=ناراضی، ۳=راضی و ۴=کاملاً راضی کمّی کرد.
۲. متغیر موقعیت سازمانی با سه وضعیت کارمند، مدیر و مدیر کل را می‌توان به شکل ۱۰=کارمند، ۲۰=مدیر و ۵۰=مدیر کل بیان کرد.
۳. متغیر دورهٔ رفتن به مسافرت در طول سال دارای چهار وضعیت هرگز، هفتگی، ماهانه و سالانه است که ۵=هرگز، ۱۰=هفتگی، ۱۵=ماهانه و ۲۰=سالانه، شکلی از کمّی کردن وضعیت‌های آن است.

۳ـ۱ـ۴ـ متغیرهای فاصله‌ای و نسبتی

مقادیر متغیرهای اسمی و ترتیبی جنبهٔ قراردادی دارند و تنها نامگذاری‌ها و برچسب‌های جدیدی برای وضعیت‌های این متغیرها قلمداد می‌شوند. به عبارت دیگر، این اعداد حاصلِ شمارش یا سنجش متغیر با وسیله‌ای خاص نیستند تا آنها را مقداری اندازه‌گیری‌شده به حساب آوریم. بنابراین، عملیات محاسباتی نیز درباره آنها قابل انجام نیست.

برای مثال، اگر وضعیت‌های متغیر اشتغال را، به صورت ۱=شاغل و ۲=بیکار کمّی کنیم، نمی‌توان ادعا کرد که فاصلهٔ شاغل و بیکار برابر با ۱−۲=۱ واحد است، زیرا اندازه شاغل و اندازه بیکار به ترتیب برابر با ۱ و ۲ نیست، بلکه ۱ و ۲ نامگذاری جدیدی برای دو وضعیت شاغل و بیکار هستند. به همین ترتیب، چنین محاسباتی درباره مقادیر متغیر میزان ورزش کردن (۴=همیشه، ۳=اکثر اوقات، ۲=گاهی و ۱=به ندرت) نیز امکان‌پذیر نیست؛ یعنی نمی‌توان ادعا کرد کسی که «اکثر اوقات» ورزش می‌کند، $(۲ \div ۳ =)$ ۱/۵ برابر فردی که «گاهی» ورزش می‌کند، به این امر می‌پردازد؛ زیرا اعداد ۳=اکثر اوقات و ۲=گاهی، تنها جنبهٔ قراردادی دارند و نام‌های جدیدی برای دو گزینهٔ «اکثر اوقات» و «گاهی» محسوب می‌شوند.

برخلاف متغیرهای اسمی و ترتیبی، متغیرهایی وجود دارند که مقدارهای آنها با سنجیدن یا شمارش کردن به دست می‌آیند. برای مثال، متغیر قد با وسیله‌ای مانند متر سنجیده می‌شود و اعدادی که به عنوان مقادیر قد در نظر گرفته می‌شوند، همان چیزی هستند که این وسیلهٔ سنجش نشان می‌دهد. همچنین، تعداد تصادف‌هایی که در طول روز در چهارراه‌های شهری رخ می‌دهند، متغیری است که مقادیر آن با شمارش مشخص

می‌شود. این اعداد جنبهٔ کاملاً قراردادی ندارند، بلکه مقادیری هستند که نتیجهٔ شمارش یا سنجش با ابزارهای خاص هستند. درباره مقادیر چنین متغیرهایی دو پرسش قابل طرح است:

۱. آیا تفاضل دو مقدار از متغیر می‌تواند نشانگر فاصلهٔ آن دو مقدار از یکدیگر باشد؟
۲. آیا با تقسیم مقداری از متغیر بر مقدار دیگر می‌توان گفت آن مقدار چند برابر مقدار دیگر است؟

اگر تنها پاسخ به پرسش اول، برای مقادیر متغیری مثبت باشد، آن متغیر فاصله‌ای است و اگر پاسخ به هر دو پرسش برای مقادیر متغیری مثبت باشد، آن متغیر نسبتی است. در واقع، داشتن یا نداشتن صفر واقعی، معیار فاصله‌ای یا نسبتی بودن متغیرهایی است که نتیجهٔ سنجش با ابزار هستند. متغیرهای نسبتی دارای صفر واقعی، ولی متغیرهای فاصله‌ای فاقد آن هستند[۱].

دما که با وسیله‌ای به‌نام دماسنج اندازه‌گیری می‌شود، متغیری فاصله‌ای است؛ زیرا دمای صفر درجه قراردادی است و به معنای بدون دما بودن نیست (در واقع، دمای صفر درجه نه تنها «بی دما بودن» نیست، درجه‌ای از دماست که آب در آن یخ می‌زند). از این رو، اگر دمای شهری ۴۰ درجه و دمای شهر دیگری ۱۰ درجه باشد، تنها می‌توان گفت دمای شهر اول (۱۰ − ۴۰ =) ۳۰ درجه گرم‌تر از شهر دوم است، ولی نمی‌توان ادعا کرد که شهر اول (۱۰ ÷ ۴۰ =) ۴ برابر گرم‌تر از شهر دوم است، زیرا عمل تقسیم برای متغیر فاصله‌ای که دارای صفر واقعی نیست، صورت نمی‌گیرد.

در مقابل، مدت زمان تماشای تلویزیون متغیری نسبتی است که در آن هم تفاضل میان مقادیر و هم تقسیمِ آنها بر یکدیگر مجاز است. پس اگر کسی ۱۲۰ دقیقه و فرد دیگری ۱۵ دقیقه تلویزیون تماشا کند، هم می‌توان گفت فرد اول (۱۵ − ۱۲۰ =) ۱۰۵ دقیقه بیشتر از فرد دوم تلویزیون تماشا کرده است و هم او (۱۵ ÷ ۱۲۰ =) ۸ برابر فرد اول بینندهٔ تلویزیون بوده است؛ زیرا «صفر دقیقه» صفری واقعی است و معادل اصلاً تماشا نکردن تلویزیون است.

اغلب متغیرهای علوم طبیعی نظیر وزن، ارتفاع، زمان و جِرم، نسبتی هستند، در حالی که متغیرهای علوم انسانی مانند جامعه‌شناسی یا روان‌شناسی به دلیل محدودیت‌های ابزار اندازه‌گیری آنها، از سطح سنجش فاصله‌ای فراتر نمی‌روند (حتی درباره این امر که آیا به سطح فاصله‌ای نیز می‌رسند، تردیدهایی وجود دارد). ویژگی دیگری که می‌توان برای متغیرهای علوم طبیعی برشمرد، بیان مقادیر متغیر با واحد اندازه‌گیری یا به اصطلاح دارا

۱. در واقع، هر چهار عملِ ضرب، تقسیم، جمع و تفریق برای متغیر نسبتی و تنها دو عمل جمع و تفریق برای متغیر فاصله‌ای امکان‌پذیر است.

بودنِ بُعد است. برای مثال، طول با واحد سانتیمتر، دما با واحد سانتیگراد یا وزن با کیلوگرم بیان می‌شود؛ در حالی که مقادیر متغیرهای علوم انسانی، به همان دلیلِ محدودیت ابزار اندازه‌گیری، از چنین ویژگی برخوردار نیستند.

از آنچه گذشت می‌توان نتیجه گرفت که پایین‌ترین سطح سنجش به متغیرهای اسمی تعلق دارد و پس از آن سطح سنجش ترتیبی، فاصله‌ای و نسبتی قرار دارد. در واقع، سطح سنجش نسبتی کامل‌ترین نوع سنجش است؛ زیرا تمام عملیات محاسباتی را می‌توان بر روی مقادیر آن انجام داد.

۲ـ۴ـ متغیر پیوسته و گسسته

مقادیر متغیرها به دو گونه‌اند: پیوسته و گسسته. به بیانی ساده، گاهی مقادیر متغیر به هم چسبیده‌اند؛ یعنی بین هر دو مقدار دلخواه از آنها، مقدار سومی را نیز می‌توان تصور کرد که در این صورت به آن متغیر پیوسته گفته می‌شود. مقادیر متغیر پیوسته از اندازه‌گیری با ابزارهای دقیق به دست می‌آیند که مقدار متغیر را برای عضوی خاص از جامعه آماری اندازه می‌گیرد و به صورت عدد ارائه می‌دهد (یعنی مانند متغیرهای اسمی یا ترتیبی، نسبت اعداد به وضعیت‌ها، به قرارداد نیازی ندارد). مقادیر متغیری مانند دما پیوسته است؛ زیرا بین هر دو مقدار از دما، مقدار دیگری از دما را می‌توان در نظر گرفت (در واقع، نه تنها مقدار دیگری بلکه بی‌شمار مقدار بین دو مقدار از دما قابل تصور است). برای مثال، بین دو مقدار ۲۳ درجه و ۲۴ درجه، مقادیر ۲۳/۱ درجه، ۲۳/۱۱ درجه، ۲۳/۱۱۱ درجه و... قرار دارند. متغیرهای مدت زمان، سن، درآمد و طول، مثال‌های دیگری از متغیرهای پیوسته هستند.

متغیرهای گسسته دارای مقادیر جداجدا و مجزا از یکدیگرند. متغیرهای شمارشی که حاصل شمردن تعداد دفعات رخ‌دادن یک پدیده هستند، متغیرهای گسسته قلمداد می‌شوند. برای مثال، متغیر تعداد روزهایی که فردی در هفته ورزش می‌کند، متغیری گسسته است؛ زیرا مقادیر آن ۰ روز، ۱ روز، ۲ روز،... و ۷ روز در هفته است. توجه کنید که نمی‌توان مقداری را مثلاً بین دو مقدار ۱ روز و ۲ روز برای تعداد روزها در نظر گرفت (برای نمونه، ۱/۵ روز در هفته مفهوم روشنی ندارد).

متغیرهای اسمی و ترتیبی همیشه گسسته‌اند، در حالی که متغیرهای فاصله‌ای و نسبتی می‌توانند گسسته یا پیوسته باشند. تعداد مقادیر متغیر پیوسته همواره نامتناهی است، در حالی که تعداد مقادیر متغیر گسسته می‌تواند متناهی یا نامتناهی باشد.

۳ـ۴ـ متغیر مقوله‌ای

متغیر مقوله‌ای[1] متغیری است که وضعیت‌های آن اعضای جامعه آماری را بخش‌بندی می‌کند؛ به طوری که از یک سو این بخش‌ها دارای هیچ اشتراکی نیستند و از سوی دیگر تمام جامعه آماری را در برمی‌گیرند. بر این اساس، متغیرهای اسمی و ترتیبی متغیرهای مقوله‌ای هستند. برای مثال، متغیر زادگاه با دو وضعیت شهر و روستا می‌تواند جامعه آماری را به دو دسته تقسیم کند. یک دسته شامل کسانی است که در روستا به دنیا آمده‌اند و دستهٔ دیگر شامل کسانی است که در شهر زاده شده‌اند. این دو دسته دارای هیچ اشتراکی نیستند؛ یعنی نمی‌توان فردی را یافت که هم در شهر و هم در روستا به دنیا آمده باشد. همچنین تمام اعضای جامعه را در بر می‌گیرند؛ زیرا زادگاه هر عضو از جامعه، یا در روستا بوده است یا در شهر، پس فردی از جامعه باقی نمی‌ماند که به یکی از این دو تعلق نداشته باشد.

علاوه بر متغیرهای اسمی و ترتیبی، برخی از متغیرهای فاصله‌ای یا نسبتی نیز به طور طبیعی دارای ماهیت مقوله‌ای هستند و برخی دیگر، بویژه متغیرهای پیوسته را می‌توان با دسته‌بندیِ مقادیرشان به متغیر مقوله‌ای تبدیل کرد. متغیر تعداد روزهایی که فرد در هفته ورزش می‌کند، متغیری نسبتی با هشت وضعیت (۰ روز، ۱ روز، ۲ روز، ...، ۷ روز) است که به طور طبیعی افراد جامعه را به هشت دسته تقسیم می‌کند (دسته‌ای که اصلاً ورزش نمی‌کنند، دسته‌ای که ۱ روز در هفته ورزش می‌کنند، دسته‌ای که ۲ روز در هفته ورزش می‌کنند، ... و سرانجام دسته‌ای که هر ۷ روز هفته ورزش می‌کنند).

متغیری مانند سن با مقادیر بسیار متنوع و متعدد را می‌توان با دسته‌بندی مقادیر آن به متغیر مقوله‌ای تبدیل کرد. برای مثال، دسته‌بندی جدول ۲، شکلی از تبدیل متغیر سن به یک متغیر مقوله‌ای را نشان می‌دهد.

جدول ۲: دسته‌بندی متغیر سن به نُه مقوله

۰ تا ۵ سال	۵ سال و ۱ روز تا ۱۰ سال	۱۰ سال و ۱ روز تا ۱۵ سال
۱۵ سال و ۱ روز تا ۲۰ سال	۲۰ سال و ۱ روز تا ۳۰ سال	۳۰ سال و ۱ روز تا ۴۰ سال
۴۰ سال و ۱ روز تا ۵۰ سال	۵۰ سال و ۱ روز تا ۶۰ سال	بیش از ۶۰ سال

جامعه با دسته‌بندی متغیر سن به نُه گروه تقسیم می‌شود که دارای هیچ اشتراکی نیستند؛ یعنی هر عضو از جامعه می‌تواند تنها در یکی از گروه‌های سنی جای بگیرد. از سوی دیگر

۱. متغیر گروهی، رسته‌ای یا طبقه‌ای عنوان‌های دیگری برای متغیر مقوله‌ای (Categorical Variable) هستند.

سن فرد هر چه که باشد، به یکی از این گروه‌ها تعلق خواهد داشت؛ یعنی سنین مختلف با این گروه‌ها پوشش داده می‌شوند و هیچ فردی بیرون از این گروه‌ها باقی نمی‌ماند.

از آنجا که متغیرهای فاصله‌ای یا نسبتیِ دسته‌بندی‌شده به طور طبیعی مقوله‌ای نیستند، رعایت دو شرط «اشتراک نداشتن مقوله‌ها» و «شامل شدن تمام مقادیر متغیر» در دسته‌بندی مقادیر متغیر به چندین مقوله ضروری است؛ زیرا در غیر این صورت متغیر مقوله‌ای نمی‌تواند جامعه را به بخش‌هایی تقسیم کند که از یک سو هیچ عضو مشترکی نداشته باشند و از سوی دیگر تمام اعضای جامعه را شامل شوند. از همین روست که در جدول ۲، گروه دوم با ۵ سال و «۱ روز» آغاز می‌شود تا هیچ‌گونه اشتراکی با گروه قبل خود یعنی ۰ تا ۵ سال نداشته باشد. همچنین گروه آخر به صورت «بیش از» ۶۰ سال تعریف می‌شود تا تمام مقادیر سن خارج از هشت گروه قبل را شامل شود. این دو ویژگی تضمین می‌کنند که هیچ دو گروهی فرد مشترک نداشته باشند و هیچ فردی خارج از این گروه‌ها باقی نماند.

جدول ۳: ویژگی‌های چهار سطح سنجش

نوع متغیر (سطح سنجش)	ویژگی‌ها
اسمی	۱. مقادیر آن قراردادی هستند (نامگذاری جدیدی برای وضعیت‌های متغیر هستند). ۲. مقادیر دارای قابلیت‌های محاسباتی نیستند. ۳. معمولاً وضعیت‌های آن طبیعی هستند. ۴. مقادیر آن گسسته هستند و به طور طبیعی مقوله‌ای به حساب می‌آید.
ترتیبی	۱. مقادیر آن قراردادی هستند (نامگذاری جدیدی برای وضعیت‌های متغیر هستند). ۲. بین وضعیت‌های آن ترتیب و اولویت وجود دارد. ۳. ترتیب مقادیر هماهنگ با ترتیب وضعیت‌های متغیر است. ۴. مقادیر آن گسسته هستند و متغیر مقوله‌ای به حساب می‌آید.
فاصله‌ای	۱. نتیجه اندازه‌گیری با ابزار خاصی هستند. ۲. دارای صفر واقعی نیستند. ۳. می‌تواند دارای واحد اندازه‌گیری باشد. ۴. فاصله بین مقادیر قابل محاسبه است. ۵. دو عمل جمع و تفریق را می‌توان بر روی مقادیر انجام داد. ۶. می‌تواند گسسته یا پیوسته باشد.
نسبتی	۱. نتیجهٔ شمارش یا اندازه‌گیری با ابزار خاصی هستند. ۲. دارای صفر واقعی هستند. ۳. می‌تواند دارای واحد اندازه‌گیری باشد. ۴. هم فاصله بین مقادیر و هم نسبت بین مقادیر در آن قابل محاسبه است. ۵. هر چهار عمل اصلی (×، ÷، + و −) بر روی مقادیر می‌تواند انجام شود. ۶. می‌تواند گسسته یا پیوسته باشد.

از آنچه گذشت می‌توان دریافت که برخی مقوله‌ها طبیعی هستند و برخی ساختگی. به طور معمول، متغیرهای اسمی دارای مقوله‌های طبیعی هستند. برای مثال، متغیر جنس دارای

دو وضعیت طبیعی است که مقوله‌ها را نیز شکل می‌دهند (مرد/زن). در مقابل، متغیری مانند مقطع تحصیلی دانشگاه دارای مقوله‌های ساختگی است (کارشناسی، کارشناسی ارشد، دکترا). همچنین، متغیرهای فاصله‌ای یا نسبتیِ دسته‌بندی‌شده نیز دارای مقوله‌های کاملاً ساختگی هستند. ویژگی‌های انواع متغیرها در جدول ۳ خلاصه شده است.

۵ـ وسیلهٔ گردآوری داده‌ها و مجموعه داده‌ها

اندازه‌گیری متغیرها از اعضای جامعه آماری که به تولید داده‌ها منجر می‌شود، در علوم مختلف با ابزار متفاوتی صورت می‌گیرد. در پزشکی، متغیرهایی مانند فشار خون، ضربان قلب یا کلسترول خون با وسیله‌های دقیقی سنجیده و ثبت می‌شوند؛ به طوری که اندازه‌های حاصل از سنجش اغلب نسبتی هستند. اما وسیله‌های سنجش در علوم انسانی نظیر جامعه‌شناسی یا روان‌شناسی ماهیت دیگری دارند. پرسشنامه یکی از رایج‌ترین وسیله‌هایی است که در این علوم برای سنجش چندین متغیر به کار می‌رود.

پرسشنامه، شامل تعدادی سؤال مکتوب درباره موضوعات مختلف و معمولاً مرتبط است. هر پرسش می‌تواند برای سنجش یک متغیر طراحی شود. برای مثال، متغیر رضایت از زندگی را می‌توان با پرسش زیر سنجید:

تا چه حد از زندگی خود راضی هستید؟ کاملاً راضی، نسبتاً راضی، نسبتاً ناراضی، کاملاً ناراضی

برای این پرسش، چهار پاسخ (گزینه) پیش‌بینی شده است تا پاسخگو یکی از آنها را به دلخواه برای توصیف رضایت خود از زندگی انتخاب کند. این گزینه‌ها دارای نوعی ترتیب هستند؛ بنابراین متغیر رضایت از زندگی با این پرسش در قالب متغیری ترتیبی با چهار وضعیت سنجیده می‌شود. پژوهشگر می‌تواند با طراحی پرسش به شکل زیر، سنجش این متغیر را حتی به سطح اسمی تقلیل دهد:

آیا از زندگی خود راضی هستید؟ بله خیر

در این شکل از پرسش، تنها دو گزینه (بله/خیر) در اختیار پاسخگو قرار دارد، پس متغیر رضایت از زندگی به متغیری اسمی با دو وضعیت تبدیل می‌شود.

گاهی برخی از متغیرها دارای ابعاد گوناگونی هستند که برای سنجش آنها به سنجش متغیرهای دیگری نیاز است. برای مثال، سنجش متغیر حجم از ترکیب سه متغیر طول، عرض و ارتفاع به دست می‌آید. پس ابتدا باید این سه متغیر را سنجید تا در نهایت بتوان از

ترکیب آنها به متغیر حجم رسید. در پرسشنامه‌ها نیز وضعیت مشابهی وجود دارد؛ یعنی برای سنجش متغیری پیچیده، چندین پرسش در پرسشنامه مطرح می‌شود تا ابعاد مختلف آن متغیر سنجیده شود و سپس از ترکیب این پرسش‌ها متغیر اصلی مورد سنجش قرار می‌گیرد. برای نمونه، ممکن است پژوهشگری متغیر رضایت از زندگی را دارای ابعاد گوناگونی نظیر رضایت شغلی، رضایت خانوادگی و رضایت فردی بداند. به این ترتیب، او باید ابتدا این ابعاد را بسنجد تا در نهایت متغیر رضایت از زندگی سنجیده شود. از این رو، برای سنجش هر بعد، پرسشی را طراحی می‌کند:

۱. تا چه حد از شغل خود راضی هستید؟ کاملاً راضی نسبتاً راضی نسبتاً ناراضی کاملاً ناراضی
۲. تا چه حد از خانواده خود راضی هستید؟ کاملاً راضی نسبتاً راضی نسبتاً ناراضی کاملاً ناراضی
۳. تا چه حد از خودتان راضی هستید؟ کاملاً راضی نسبتاً راضی نسبتاً ناراضی کاملاً ناراضی

ترکیب پاسخ‌هایی که یک فرد به این سه پرسش می‌دهد، به سنجش متغیر رضایت از زندگی آن فرد می‌انجامد. برای مثال، اگر گزینه‌های چهارگانه این پرسش‌ها به صورت ۴=کاملاً راضی، ۳=نسبتاً راضی، ۲=نسبتاً ناراضی و ۱=کاملاً ناراضی کمّی شوند و پاسخگویی، گزینهٔ نسبتاً راضی (عدد ۳) را در پاسخ به پرسش اول، گزینه کاملاً راضی (عدد ۴) را در پاسخ به پرسش دوم و سرانجام، گزینه نسبتاً ناراضی (عدد ۲) را در پاسخ به پرسش سوم انتخاب کند، ترکیب سه پاسخ او برای سنجش کلی متغیر رضایت از زندگی می‌تواند با جمع این سه عدد (یعنی ۹ = ۳ + ۴ + ۲) صورت بگیرد. پس مقدار متغیر رضایت از زندگی این فرد برابر با ۹ است (برخی معتقدند مقادیر حاصل از سنجش‌های ترکیبی به ایجاد متغیرهای فاصله‌ای منجر می‌شوند).

همان طور که اشاره شد، در علوم انسانی متغیرها اغلب در قالب پرسش‌هایی سنجیده می‌شوند که در پرسشنامه طراحی می‌گردند. داده‌ها نیز محصول سنجش متغیرها از اعضای جامعه آماری به کمک همین پرسش‌ها هستند. پرسشنامه به سنجش تنها یک متغیر اختصاص ندارد، بلکه پرسش‌های مختلفی را شامل می‌شود. بنابراین می‌تواند به تولید داده‌های متنوعی بینجامد که با متغیرهای مختلفی مرتبط هستند. برای مثال، پرسشنامهٔ یک نظرسنجی حاوی پرسش‌های گوناگونی است که دیدگاه مردم را نسبت به موضوع یا موضوعات مختلفی جویا می‌شود. بنابراین از هر پاسخگو (عضوی از جامعه آماری که به پرسشنامه پاسخ می‌دهد) چندین سؤال پرسیده و به این ترتیب، چندین متغیر از او سنجیده

می‌شود. اکنون تصور کنید پرسشنامه‌ای با ۲۰ سؤال از ۱۰۰۰ نفر پرسیده شده باشد، در این صورت ۲۰۰۰۰=۲۰×۱۰۰۰ داده تولید خواهد شد. همگی این داده‌ها پس از استخراج از پرسشنامه‌های تکمیل‌شده در جدولی به نام «مجموعه داده‌ها» ثبت می‌شوند. هر سطر از این جدول به یک پاسخگو و هر ستون آن به یک پرسش (متغیر) اختصاص دارد.

نمونه‌ای از مجموعه داده‌ها در جدول ۴ نمایش داده شده است. نخستین ستون به شماره پاسخگویان اختصاص دارد. اگر ۱۰۰۰ نفر به پرسشنامه جواب داده باشند، می‌توان از اعداد ۰۰۰۱ تا ۱۰۰۰ برای شماره‌گذاری پاسخگویان استفاده کرد. در این مثال، ستون‌های دوم و سوم مجموعه داده‌ها، نشان‌دهنده متغیرهای سن و جنس هستند. ستون‌های چهارم تا ششم به سه پرسشی اختصاص دارند که پیشتر برای سنجش رضایت از زندگی مطرح شده بودند. پاسخ پرسش‌های (متغیرهای) دیگر نیز می‌توانند در ادامهٔ این شش ستون درج شوند.

هر سطر از این جدول به ثبت پاسخ‌ها و اطلاعات یکی از پاسخگویان اختصاص دارد. برای مثال پاسخگوی اول (شماره ۰۰۰۱) ۲۲ سال دارد و رقم ۱ نشان‌دهنده جنس اوست. توجه کنید که ضروری است تعریف رقم‌های مربوط به متغیرهای اسمی و ترتیبی به عنوان راهنما با مجموعه داده‌ها همراه باشد. بنابراین، باید ذکر شود که رقم ۱ برای مرد و رقم ۲ برای زن انتخاب شده است. همچنین این پاسخگو گزینهٔ ۲ (نسبتاً ناراضی) را در پاسخ به پرسش رضایت شغلی، گزینه ۳ (نسبتاً راضی) را در پاسخ به پرسش رضایت خانوادگی و گزینه ۲ (نسبتاً ناراضی) را در پاسخ به پرسش رضایت فردی انتخاب کرده است. گفتنی است، به جای استفاده از برگه‌های کاغذی، اغلب پژوهشگر مجموعه داده‌ها را در جدولی ثبت و نگهداری می‌کند که نرم‌افزارهای مختلف در اختیار او قرار می‌دهند.

از آنجا که مجموعه داده‌ها شکلی مانند ماتریس در ریاضیات دارد به آن ماتریس داده‌ها نیز گفته می‌شود. از این رو، ابعاد مجموعه داده‌ها مانند شیوهٔ معرفی ابعاد ماتریس‌ها به صورت «تعداد ستون × تعداد سطر» بیان می‌شود. پس مجموعه داده‌های جدول ۴ که حاوی ۱۰۰۰ سطر و ۶ ستون است (ستون مربوط به شماره پاسخگویان نادیده گرفته می‌شود) به عنوان مجموعه داده‌ای ۵×۱۰۰۰ شناخته می‌شود (از چپ به راست ۱۰۰۰ در ۵ بخوانید).

جدول ۴: نمونه‌ای از مجموعه داده‌ها

شماره پاسخگو	سن	جنس	رضایت شغلی	رضایت خانوادگی	رضایت فردی
۰۰۰۱	۲۲	۱	۲	۳	۲
۰۰۰۲	۵۳	۲	۳	۴	۲
۰۰۰۳	۴۹	۱	۳	۳	۳
۰۰۰۴	۳۶	۱	۴	۴	۳
۰۰۰۵	۳۰	۱	۳	۲	۴
۰۰۰۶	۲۹	۲	۱	۲	۱
. . .	. . .	. . .	. . .	. . .	. . .
۱۰۰۰	۶۵	۲	۱	۴	۳

اصطلاحات فصل اول

متغیر فاصله‌ای	Interval Variable	متغیر مقوله‌ای	Categorical Variable
اندازه‌گیری	Measurement	سرشماری	Census
متغیر اسمی	Nominal Variable	ویژگی، خصیصه	Characteristic
متغیر ترتیبی	Ordinal Variable	ثابت، بدون تغییر	Constant
پرسشنامه	Questionnaire	پیوسته	Continuous
متغیر نسبتی	Ratio Variable	ماتریس داده	Data Matrix
نمونه	Sample	مجموعه داده	Data Set
نمونه‌گیری	Sampling	آمار توصیفی	Descriptive Statistics
جامعه آماری	Statistical Population	بُعد	Dimension
مقدار	Value	گسسته	Discrete
متغیر	Variable	آمار استنباطی	Inferential Statistics

تمرین‌های فصل اول

۱. دو مثال از کاربرد واژه آمار در معنای اول آن بزنید.

۲. با انتخاب جامعه آماری مناسب، دو خصیصه مثال بزنید که ثابت باشند نه متغیر.

۳. صفات مشترک در هر یک از جامعه‌های آماری مثال ۱ را بیان کنید.

۴. پژوهشگری مایل است موضوع اعتیاد به اینترنت را بررسی کند. جامعهٔ آماری این مطالعه و برخی از متغیرهای مرتبط با این مطالعه چه می‌توانند باشند؟

۵. متغیری در پرسشنامه‌ای با پرسش و گزینه‌های زیر سنجیده شده است. نوع این متغیر چیست؟

چقدر احتمال می‌دهید در دانشگاه قبول شوید؟ کم/متوسط/زیاد/نمی‌دانم

۶. نوع متغیرهای زیر را مشخص کنید:

آ) تعداد مراجعه‌کنندگان به یک سازمان مفروض در روزهای کاری هفته.

ب) میزان اعتماد به نفس فرد که از ترکیب پاسخ‌های او به ۲۰ پرسش به دست آمده است.

پ) میزان موافقت مردم با یکی از طرح‌های دولت اگر در قالب پرسش «تا چه حد با این طرح موافق هستید؟» با گزینه‌های «اصلاً/تاحدی/زیاد/کاملاً» سنجیده شده باشد.

ت) مدت زمان انتظار در ایستگاه اتوبوس برای سوار شدن.

ث) میزان موافقت مردم با یکی از طرح‌های دولت اگر در قالب پرسش «آیا با این طرح موافق هستید؟» با گزینه‌های «موافقم/مخالفم» سنجیده شده باشد.

ج) وضعیت مهارت کارگران به شکل بی‌مهارت، نیمه ماهر و ماهر.

۷. آیا مقادیر متغیر همواره از اعداد مثبت تشکیل می‌شوند؟

۸. چرا متغیرهای اسمی و ترتیبی گسسته هستند؟

۹. پژوهشگری می‌خواهد بر اساس داده‌هایی که برای تحقیق دیگری گردآوری شده بودند، مطالعه‌ای را صورت دهد. او در مطالعه خود به چه حوزه‌ای از آمار نیاز ندارد؟

۱۰. هر یک از فعالیت‌های آماری زیر با چه حوزه‌ای از آمار مرتبط است؟

آ) ۵۰۰ دانشجوی سال اول و ۵۰۰ دانشجوی سال دوم انتخاب شده‌اند تا بررسی شود که آیا درصد دانشجویان دختر پذیرفته شده در دانشگاه طی ۲ سال افزایش یافته است.

ب) یک شرکت خودروساز جدولی را منتشر کرده است که میزان فروش انواع خودروهای تولیدی آن شرکت را نشان می‌دهد.

پ) پژوهشگری برای مقایسهٔ ارزش‌های فرهنگی سه نسل در یک استان، تصمیم گرفته است مردم را بر اساس سن به سه دستهٔ حداکثر ۳۵ سال، ۳۵ تا ۵۵ سال و بیش از ۵۵ سال طبقه‌بندی کند و سپس از هر دسته ۳۰۰ نفر را انتخاب و داده‌هایی را دربارهٔ ارزش‌های مورد قبول آنان گردآوری کند.

۱۱. کدام‌یک از متغیرهای زیر پیوسته و کدام‌یک گسسته هستند؟

آ) مسافتی که فردی از محل خانه تا محل کار خود طی می‌کند.

ب) میزان رضایت از زندگی اگر از ترکیب سه پرسشی به دست آید که در بخش چهارم این فصل مطرح شد.

پ) تعداد غلط‌های تایپی در یک صفحه که از سوی یک حروف‌نگار صورت می‌گیرد.

ت) میزان درد یک بیمار وقتی پزشک از او می‌خواهد شدت درد خود را از ۰ تا ۲۰ بیان کند.

ث) احتمال شرکت فردی در انتخابات اگر به صورت «حتماً شرکت می‌کنم/ احتمالاً شرکت می‌کنم/ شرکت نمی‌کنم» سنجیده شده باشد.

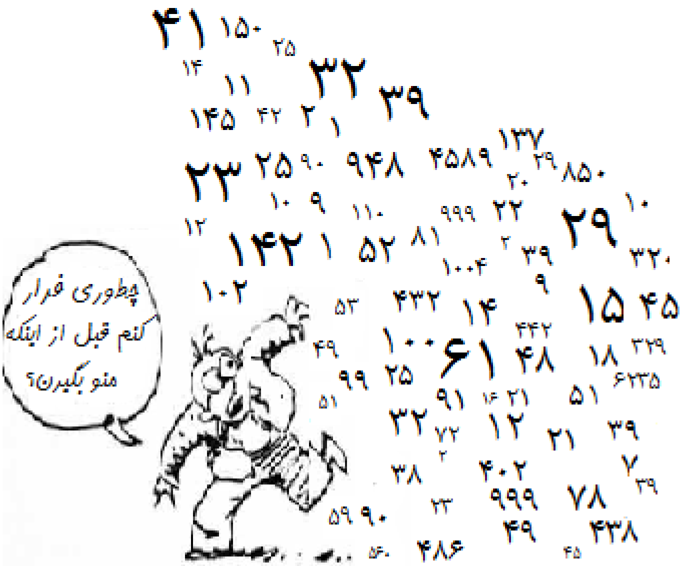
۱۲. بر اساس مجموعهٔ داده‌های جدول ۴، بیان کنید که مشخصات پاسخگوی سوم (شماره ۰۰۰۳) یعنی سن و جنس او و نیز گزینه‌های انتخابی او برای رضایت از زندگی چه بوده‌اند؟

۱۳. ستون جدیدی به مجموعه داده‌های جدول ۴ اضافه کنید که مقدار رضایت کلی از زندگی هر پاسخگو را با ترکیب پاسخ‌های او به سه پرسش رضایت شغلی، خانوادگی و فردی نشان دهد.

فصل دوم

۲

آمار توصیفی



این فصل مباحث مربوط به روش‌های توصیف داده‌ها را پوشش می‌دهد. بحث با نقش و اهمیت آمار توصیفی و دسته‌بندی کلی روش‌های آن آغاز می‌شود. بخش دوم به جدول توزیع فراوانی، نمودارها و منحنی توزیع متغیر می‌پردازد که همگی روش‌هایی برای نمایش توزیع متغیر هستند. مشخصه‌های تمرکز، پراکندگی و شکل توزیع در بخش بعد مورد بحث قرار می‌گیرند. بخش پایانی با عنوان «جمع‌بندی» ضمن آنکه خلاصه‌ای از مطالب مطرح شده در بخش‌های پیشین است، نگاهی کاربردی به آن مطالب نیز دارد تا توجه خواننده را به ارتباط مباحث با یکدیگر و نکات ضروری در به‌کارگیری آنها جلب کند.

۱ـ چشم‌اندازی از آمار توصیفی

یک مطالعه کمّی مانند یک نظرسنجی، شامل سنجش متغیرهای متعدد از افراد زیادی است که در نهایت به مجموعه داده‌ای بزرگ می‌انجامد. برای مثال، مجموعه داده‌ای ۵۰×۱۰۰۰ (یعنی با ۱۰۰۰ سطر و ۵۰ ستون) حاوی داده‌های ۵۰ متغیر از ۱۰۰۰ پاسخگوست و مجموعه داده‌ای کاملاً عادی در نظرسنجی‌هاست که شاید کوچک نیز قلمداد شود. چنین مجموعه داده‌ای پژوهشگر را با ۵۰۰۰۰ عدد و رقم روبه‌رو می‌کند که نه حوصله خواندن تک‌تک آنها را دارد و نه با خواندن آنها می‌تواند به جمع‌بندی و درک خاصی برسد.

برای روشن شدن موضوع، داده‌های جدول ۱ را در نظر بگیرید که سن ۷۲ نفر و میزان رضایت آنان را از زندگی (۱=کم، ۲=متوسط و ۳=زیاد) در خود جای داده است؛ یعنی مجموعه داده‌ای با تنها ۲ متغیر و ۷۲ پاسخگو. چه چیزی درباره سن این ۷۲ نفر می‌توان گفت؟ آنان جوان هستند یا پیر؟ رضایت آنان از زندگی چگونه است؟ روشن است که با نگاه صرف به این جدول و خواندن داده‌های آن نمی‌توان هیچ اطلاعاتی ارائه کرد، زیرا این اعداد داده‌های خُردی هستند که از تک‌تک افراد ثبت شده‌اند و باید مورد پردازش آماری قرار گیرند تا به اطلاعات تبدیل شوند. از این رو به آنها داده‌های خام نیز گفته می‌شود. آمار توصیفی روش‌هایی را برای پردازش داده‌ها بویژه داده‌های خُردی نظیر جدول ۱ در اختیار می‌گذارد تا بتوان به اطلاعاتی دست یافت که این داده‌ها در خود نهفته دارند. به کمک این روش‌ها، توصیفی قابل فهم از مجموعه‌ای بزرگ از داده‌های درهم و انبوه در اختیار خواهیم داشت.

> روش‌های آمار توصیفی به سازماندهی و خلاصه‌سازی داده‌ها برای ارائه گزارشی توصیفی از متغیرها اختصاص دارند. این روش‌ها به دو دسته تقسیم می‌شوند: ۱. توصیف توزیع داده‌ها در قالب جدول‌ها یا نمودارها و ۲. خلاصه کردن خصوصیتی از داده‌ها نظیر میزان پراکندگی آنها در قالب مشخصه‌های توزیع که به صورت کمّی بیان می‌شوند.

طبیعی‌ترین نقطهٔ آغاز در پردازش داده‌ها پاسخ به این پرسش است که چند بار هر داده در مجموعه داده‌ها تکرار شده است؟ برای مثال، شمردن تعداد «۳» در ستون مربوط به رضایت در جدول ۱ به یافتن «اطلاعی» ارزشمند منجر می‌شود؛ یعنی تعداد کسانی که در حد زیاد از زندگی خود راضی هستند. به همین ترتیب می‌توان تعداد ۱ها و تعداد ۲ها را نیز شمارش کرد تا به اطلاعات کامل‌تری درباره رضایت این ۷۲ نفر از زندگی دست یافت و آنها را در قالب جدولی نمایش داد تا به‌راحتی در دسترس باشند. فرایند شمارش ۱ها،

۲ها و ۳ها و تعیین تعداد افرادی که در حد کم، متوسط یا زیاد از زندگی خود راضی هستند، پردازش آماری متغیر رضایت در این مجموعه داده‌ها نامیده می‌شود.

جدول ۱: داده‌های مربوط به متغیر سن و رضایت از زندگی ۷۲ نفر

رضایت	سن	رضایت	سن	رضایت	سن	رضایت	سن	رضایت	سن	رضایت	سن
۳	۱۶	۱	۱۳	۲	۱۴	۱	۲۳	۳	۵۹	۱	۲۵
۱	۴۰	۲	۱۴	۲	۴۰	۲	۱۵	۲	۲۷	۲	۱۵
۲	۲۹	۳	۵۹	۲	۱۷	۲	۳۳	۱	۳۲	۳	۲۴
۲	۴۸	۱	۵۰	۳	۱۹	۳	۳۹	۲	۱۳	۱	۱۴
۳	۳۹	۱	۱۳	۱	۱۲	۲	۷۴	۲	۱۴	۱	۵۹
۱	۸۵	۲	۱۴	۱	۱۳	۱	۴۸	۲	۵۹	۲	۲۱
۳	۱۴	۲	۵۹	۲	۱۴	۲	۲۰	۱	۵۰	۳	۳۲
۲	۳۷	۱	۵۹	۳	۵۹	۲	۷۸	۳	۴۰	۳	۳۹
۱	۶۳	۲	۲۴	۳	۱۷	۳	۶۶	۳	۱۴	۲	۸۵
۱	۴۷	۳	۱۴	۱	۳۲	۳	۴۰	۳	۴۰	۲	۱۴
۲	۲۳	۲	۴۰	۲	۳۹	۳	۱۸	۳	۲۹	۱	۳۷
۳	۷۰	۳	۱۴	۳	۸۵	۱	۸۱	۱	۴۸	۱	۶۳

هنگامی که پس از شمارش (پردازش) مشخص می‌شود که ۲۲ نفر در حد کم، ۲۷ نفر در حد متوسط و ۲۳ نفر در حد زیاد از زندگی خود راضی هستند، به سه قلم اطلاعاتی درباره متغیر رضایت در میان این ۷۲ نفر دست یافته‌ایم (هر یک از اعداد ۲۲، ۲۷ و ۲۳ یک قلم اطلاعاتی محسوب می‌شوند) و توصیفی از این ۷۲ داده ارائه کرده‌ایم که بسیار مختصرتر و مفیدتر از ۷۲ دادهٔ خُرد است. پردازش‌هایی از این نوع و شیوهٔ ارائه اطلاعات حاصل از آنها، با عنوان «توزیع متغیر» در بخش دوم این فصل بررسی خواهند شد. این بخش روش‌های به‌کارگیری جدول‌ها و نمودارها را برای ارائه گزارشی توصیفی از وضعیت متغیرها به دست می‌دهد.

اکنون به سراغ متغیر سن در جدول ۱ می‌رویم. یکی از ساده‌ترین پرسش‌های پژوهشگر می‌تواند به میزان تفاوت سنی این ۷۲ نفر مربوط باشد. از آنجا که کمترین سن در مجموعه داده‌ها ۱۲ و بیشترین سن ۸۵ سال است، با تفاضل این دو عدد از هم به قلمرو تغییرات

سن پاسخگویان یعنی رقم (۱۲– ۸۵ =) ۷۳ سال می‌رسیم. عدد ۷۳ «مشخصه‌ای» برای متغیر سن است که پس از پردازش داده‌ها (یافتن کمترین و بیشترین داده و محاسبه تفاضل آنها) اطلاعی درباره «دامنهٔ» تغییرات این متغیر به دست می‌دهد. بحث درباره مشخصه‌هایی از این نوع در بخش سوم این فصل مطرح خواهد شد.

۲ـ توزیع متغیر

هنگامی که ویژگی‌ای در یک جامعه آماری متغیر شد، یعنی اعضای جامعه به اشکال گوناگونی آن را دارا بودند، یکی از پرسش‌های اساسی درباره این ویژگی آن است که چگونه این اشکال گوناگون یا همان مقادیر متغیر در میان اعضای جامعه آماری پخش شده است؟ توزیع متغیر، نحوهٔ پخش شدن مقادیر متغیر را در میان اعضایِ جامعه آماری نشان می‌دهد.

یافتن توزیع متغیر بر اساس داده‌های گردآوری‌شده و نمایش این توزیع، از مهم‌ترین مباحث آمار توصیفی است. معمولاً توزیع‌ها در قالب جدول یا نمودار ارائه می‌شوند که در این بخش بررسی خواهند شد. در آغاز، از توزیع متغیر مقوله‌ای بحث خواهیم کرد و سپس به چگونگی توزیع متغیرهای فاصله‌ای و نسبتی خواهیم پرداخت.

۱ـ۲ـ توزیع متغیر مقوله‌ای (جدول توزیع فراوانی)

توزیع متغیر مقوله‌ای سهم هر مقوله (وضعیت) از متغیر را در جامعه آماری نشان می‌دهد. به عبارت دیگر، توزیع بیانگر آن است که چه نسبتی از اعضای جامعه به هر مقوله تعلق دارند. معمولاً از درصد برای نشان‌دادن این سهم استفاده می‌شود. برای مثال، متغیر اسمیِ نوع خانهٔ مسکونی را با سه وضعیت ۱=استیجاری، ۲=شخصی و ۳=سازمانی در جامعه‌ای شامل ۱۰۰۰ نفر در نظر بگیرید. هدف از یافتن توزیع این متغیر در این جامعه، بررسی این امر است که چه نسبتی از این ۱۰۰۰ نفر دارای خانه استیجاری، چه نسبتی دارای خانه شخصی و چه نسبتی دارای خانه سازمانی هستند.

برای این منظور، ابتدا فراوانی هر مقوله را با شمارش تعداد اعضای متعلق به هر مقوله به دست آورید. اگر پس از بررسی جامعه آماری مشخص شود که تعداد ۳۵۰ نفر از ۱۰۰۰ نفر دارای خانه استیجاری، تعداد ۶۰۰ نفر دارای خانه شخصی و تعداد ۵۰ نفر دارای خانه سازمانی هستند، اعداد ۳۵۰، ۶۰۰ و ۵۰ را فراوانی‌های این سه مقوله می‌نامند. در مرحله بعد، فراوانی

نسبی هر مقوله را با تقسیم فراوانی آن مقوله بر جمعیت جامعه آماری به دست آورید. می‌توان با ضرب فراوانی‌های نسبی در عدد ۱۰۰، آنها را به «درصد» نیز بیان کرد. محاسبات زیر، شیوه یافتن فراوانی نسبی و درصد را برای متغیر نوع خانهٔ مسکونی نشان می‌دهد:

نسبتِ (یا فراوانی نسبی) خانه استیجاری = ۳۵۰÷۱۰۰۰= ۰/۳۵ یا ۳۵٪

نسبتِ (یا فراوانی نسبی) خانه شخصی = ۶۰۰÷۱۰۰۰= ۰/۶ یا ۶۰٪

نسبتِ (یا فراوانی نسبی) خانه سازمانی = ۵۰ ÷ ۱۰۰۰= ۰/۰۵ یا ۵٪

این محاسبات نشان می‌دهند که ۳۵ درصد از این جامعه را مستأجران، ۶۰ درصد را ساکنان خانه‌های شخصی و ۵ درصد را ساکنان خانه‌های سازمانی تشکیل می‌دهند. این بدان معنی است که این ۱۰۰۰ نفر، به نسبتِ ۳۵ درصد، ۶۰ درصد و ۵ درصد، بین سه مقولهٔ متغیر نوع خانه مسکونی «توزیع» شده‌اند. معمولاً توزیع متغیر در قالب جدولی به نام جدول توزیع فراوانی مانند جدول ۲ نمایش داده می‌شود که ستونی از آن به مقوله‌ها، ستونی به فراوانی‌ها و ستونی به فراوانی‌های نسبی یا درصدها اختصاص دارد.

جدول ۲: توزیع متغیر نوع خانه

نوع خانه	فراوانی	فراوانی نسبی	درصد
استیجاری	۳۵۰	۰/۳۵	۳۵
شخصی	۶۰۰	۰/۶	۶۰
سازمانی	۵۰	۰/۰۵	۵
جمع	۱۰۰۰	۱	۱۰۰

هنگامی که جدول توزیع فراوانی مانند جدول ۲ برای متغیر اسمی رسم می‌شود، اهمیتی ندارد که مقوله‌ها به چه ترتیبی در ستون خود قرار می‌گیرند. در جدول ۲، ستون مقوله‌ها با مقوله «استیجاری» آغاز شده است، در حالی که می‌توانست با مقوله «شخصی» یا «سازمانی» نیز آغاز شود. ولی اگر جدول توزیع فراوانی به متغیری ترتیبی اختصاص دارد، بهتر است مقوله‌های متغیر از کمترین به بیشترین در ستون خود قرار بگیرند. جدول ۳، نمونه‌ای از جدول توزیع فراوانی را برای متغیری ترتیبی با عنوان «رضایت از شغل» نشان می‌دهد. همان طور که دیده می‌شود، ستون مربوط به مقوله‌ها با پایین‌ترین وضعیت رضایت یعنی «کاملاً کم» آغاز می‌شود و به ترتیب، وضعیت‌های بعدی قرار گرفته‌اند تا به بالاترین وضعیت رضایت یعنی «کاملاً زیاد» برسیم.

جدول ۳: توزیع متغیر رضایت از شغل

مقوله‌های متغیر	فراوانی	درصد	درصد تجمعی
کاملاً کم	۱۰۰	۲	۲
کم	۵۵۰	۱۱	۱۳
نسبتاً کم	۱۲۰۰	۲۴	۳۷
نسبتاً زیاد	۷۰۰	۱۴	۵۱
زیاد	۲۰۰۰	۴۰	۹۱
کاملاً زیاد	۴۵۰	۹	۱۰۰
جمع	۵۰۰۰	۱۰۰	–

با اولین نگاه به ستون فراوانی جدول ۳ می‌توان دریافت که توزیع این متغیر از داده‌های گردآوری‌شده از ۵۰۰۰ نفر تهیه شده است. این جدول فراوانی نسبی ندارد، زیرا پیشتر اشاره شد که می‌توان از فراوانی نسبی یا درصد در جدول توزیع فراوانی استفاده کرد، هر چند درصد رایج‌تر از فراوانی نسبی است.

جدول ۳ دارای ستونی با عنوان «درصد تجمعی[1]» است که تنها در جدول‌های مربوط به متغیرهای ترتیبی به کار می‌رود (به همین دلیل در جدول ۲ دیده نمی‌شد). درصد تجمعی هر مقوله، حاصلِ جمعِ درصد آن مقوله و درصدهای مقوله‌های پیش از آن است. برای مثال، درصد تجمعی مقوله «نسبتاً کم» از جمع درصد این مقوله یعنی ۲۴٪ با درصدهای دو مقوله پیش از آن یعنی مقوله کم (۱۱٪) و مقوله کاملاً کم (۲٪) به دست می‌آید، پس درصد تجمعی آن برابر با ۳۷=۲+۱۱+۲۴ است. توجه کنید که درصد تجمعی مقوله اول، برابر با درصد آن مقوله است، زیرا مقوله‌ای پیش از آن وجود ندارد تا با آنها جمع شود. درصد تجمعی مقوله آخر نیز همواره برابر با ۱۰۰ است[2].

فراموش نکنید که در جدول توزیع فراوانی، تمام وضعیت‌های متغیر فهرست شوند؛ چه آنهایی که مشاهده شده و چه آنهایی که مشاهده نشده‌اند. برای مثال اگر رضایت از زندگی، متغیری ترتیبی با سه وضعیت «زیاد، متوسط و کم» است، ولی هیچ‌کس رضایت خود را از زندگی در حد «کم» اعلام نکرده است، این وضعیت در جدول آورده و فراوانی آن برابر با صفر در نظر گرفته می‌شود.

درصد تجمعی به شناخت چگونگی توزیع متغیر ترتیبی در جامعه آماری کمک می‌کند. برای مثال، اگر پرسیده شود که چه نسبتی از جامعه از شغل خود «حداکثر» در حد «زیاد» راضی هستند، بلافاصله با مشاهده درصد تجمعی مربوط به مقوله «زیاد» به عدد ۹۱٪ خواهیم رسید، زیرا عدد ۹۱ از جمع درصد مقوله «زیاد» و مقوله‌های پیش از آن به دست آمده است، پس نسبتی از جامعه را نشان می‌دهد که در حد «زیاد» یا کمتر از آن از شغل خود رضایت دارند.

رسم جدول توزیع فراوانی برای متغیرهای دسته‌بندی‌شده (فاصله‌ای یا نسبتی) مانند متغیر ترتیبی است، زیرا دسته‌بندیِ مقادیر متغیرهای فاصله‌ای یا نسبتی به ایجاد متغیری ترتیبی منجر می‌شود (متغیر دسته‌بندی‌شده، دیگر فاصله‌ای یا نسبتی نیست بلکه متغیری ترتیبی است).

---

۱. به آن، درصد «تراکمی» نیز گفته می‌شود.

۲. همین بحث را می‌توان درباره فراوانی نسبی نیز مطرح کرد تا فراوانی نسبیِ تجمعی به دست آید.

۲ـ۲ـ نمایش توزیع متغیر مقوله‌ای در قالب نمودار

اگرچه جدول توزیع فراوانی می‌تواند اطلاعات کافی را درباره توزیع متغیر مقوله‌ای در اختیار بگذارد، ولی کاربران یافته‌های آماری اطلاعات مورد نیاز خود را از نمودارها ساده‌تر از جدول‌ها به دست می‌آورند. همچنین، نمودارها می‌توانند برخی نکات نهفته در یافته‌ها را به شکلی برجسته به نمایش بگذارند، در حالی که جدول نمی‌تواند توجه کاربر را مستقیم و بدون صرف دقت به آن یافته‌ها جلب کند. از این رو نمودارها به تنهایی یا به عنوان مکمل جدول توزیع فراوانی در ارائه یافته‌های آماری به خدمت گرفته می‌شوند.

نمودار دایره‌ای و نمودار میله‌ای دو نمودار رایج برای نمایش توزیع متغیرهای مقوله‌ای هستند. نمودار دایره‌ای نسبت‌های هر مقوله را در قالب بُرش‌هایی از یک دایره نمایش می‌دهد. شکل ۱ نمونه‌ای از یک نمودار دایره‌ای برای نمایش توزیع متغیر اسمی با ۷ مقوله است.

شکل ۱: نمودار دایره‌ای برای توزیع متغیر انواع فیلم سینمایی مورد علاقه

فرض کنید شکل ۱، توزیع متغیر موضوع سینمایی مورد علاقه مردم را نشان می‌دهد؛ به این ترتیب که از مردم پرسیده شده است: «چه نوع فیلم سینمایی را بیشتر دوست دارید؟» و آنان می‌بایست یکی از ۷ گزینهٔ «فیلم‌های جنایی، فیلم‌های سیاسی، فیلم‌های تاریخی، فیلم‌های ترسناک، فیلم‌های علمی‌ـتخیلی، فیلم‌های خانوادگی یا فیلم‌های حادثه‌ای» را انتخاب کنند. برای نمایش توزیع این متغیر در قالب نمودار دایره‌ای شکل ۱، ابتدا باید توزیع

جدول ۴: اطلاعات برای رسم نمودار دایره‌ای

انواع فیلم سینمایی	فراوانی	درصد	زاویه (درجه)
جنایی	۸۰۰	۸	۲۸/۸
سیاسی	۱۲۰۰	۱۲	۴۳/۲
تاریخی	۱۲۰۰	۱۲	۴۳/۲
ترسناک	۸۰۰	۸	۲۸/۸
علمی‌ـتخیلی	۹۰۰	۹	۳۲/۴
خانوادگی	۲۹۰۰	۲۹	۱۰۴/۴
حادثه‌ای	۲۲۰۰	۲۲	۷۹/۲
جمع	۱۰۰۰۰	۱۰۰	۳۶۰

آن را با یافتن فراوانی هر مقوله و سپس، فراوانی نسبی یا درصد مقوله‌ها بـه دسـت آوریـد. این اطلاعات در ستون فراوانی و درصد جدول ۴ ارائه شده‌اند.

در گام بعد، باید محاسبه کنید که چه برشی از دایره به یک مقولـه اختصـاص دارد. از آنجـا که کل دایره ۳۶۰ درجه است، برش (زاویه) مربوط به یک مقوله از رابطهٔ زیر به دست می‌آید:

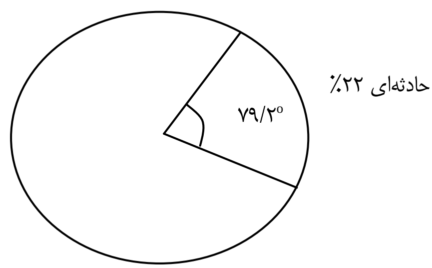
۳۶۰×درصد=زاویه

برای مثال، زاویه مربوط به مقوله فیلم تاریخی بر اساس اطلاعات جدول ۴ به صورت زیـر به دست می‌آید:

زاویهٔ مربوط به فیلم‌های تاریخی $= ۴۳/۲^{\circ} = \%۱۲\times ۳۶۰ = \frac{۱۲}{۱۰۰}\times ۳۶۰$

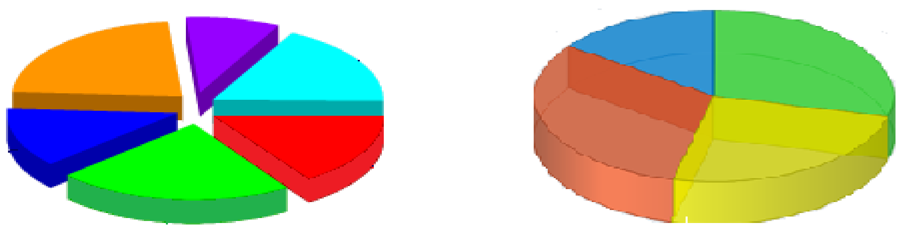
جدول ۴ دارای ستونی به نام زاویه است که نتیجهٔ محاسبات فوق را برای تک‌تک مقوله‌هـا نشان می‌دهد. اکنون اطلاعات لازم برای رسم نمودار دایره‌ای در دسـت اسـت و مـی‌تـوان برای هر مقوله، برشی از دایره را به اندازه زاویه آن انتخاب کـرد تـا آن مقولـه روی سـطح دایـره نمـایش داده شـود. شـکل ۲، انتخـاب برش را برای فیلم حادثه‌ای نشان می‌دهد کـه بر اساس جـدول ۴، زاویـه‌ای برابـر بـا ۷۹/۲ درجه دارد. به همین ترتیـب، بـرش بعـد در ادامه این برش قرار می‌گیرد تا تمام مقولـه‌هـا در نمودار نمایش داده شـوند و شـکل ۱ بـه دست آید.

شکل ۲: تعیین بُرش‌ها در نمودار دایره‌ای



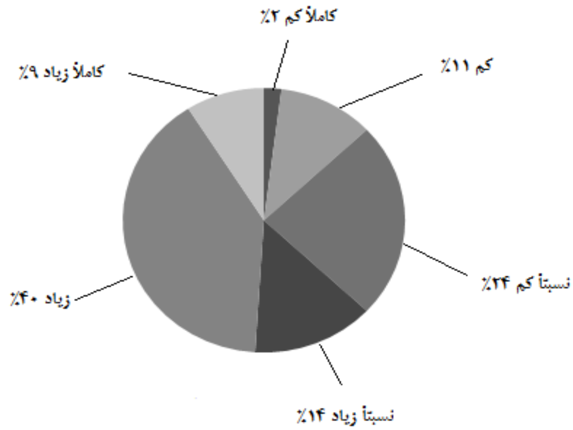
پژوهشگر می‌تواند مانند شکل ۳، به نمودار دایره‌ای حجم دهد یـا بـرش‌هـای آن را از یکدیگر جدا سازد. در هر صورت این نمودارها نیز دایره‌ای هستند، زیرا سهم هر مقولـه بـا زاویه در آنها نشان داده می‌شود، هر چند سلیقه پژوهشگر صـورت‌هـای متفـاوتی بـه آنهـا بخشیده است.

شکل۳: دو نمونه از نمودار دایره‌ای



مانند جدول توزیع فراوانی، اهمیتی ندارد که برش‌های مقوله‌های متغیر اسمی با چه ترتیبی در سطح دایره ظاهر شوند، ولی درباره متغیرهای ترتیبی، مناسب‌تر آن است که در جهت حرکت عقربه‌های ساعت، برش‌ها به ترتیب از پایین‌ترین مقوله به بالاترین آنها قرار گیرند.

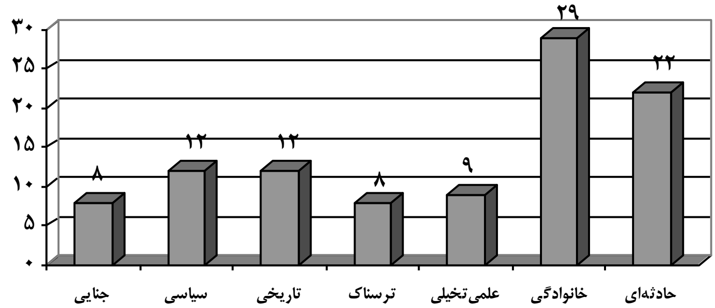
شکل ۴: نمودار دایره‌ای برای متغیر رضایت از شغل جدول ۳



نمودار دیگری که برای نمایش توزیع متغیر مقوله‌ای به‌کار می‌رود، نمودار میله‌ای است که به آن نمودار ستونی نیز گفته می‌شود. در این نمودار، درصدهای مربوط به مقوله‌ها در قالب مستطیل‌هایی (میله‌هایی) نمایش داده می‌شوند که بر روی محور مختصات افقی قرار دارند؛ ارتفاع مستطیل اختصاص‌یافته به هر مقوله برابر با درصد آن مقوله است که به کمک محور مختصات عمودی نشان داده می‌شود.

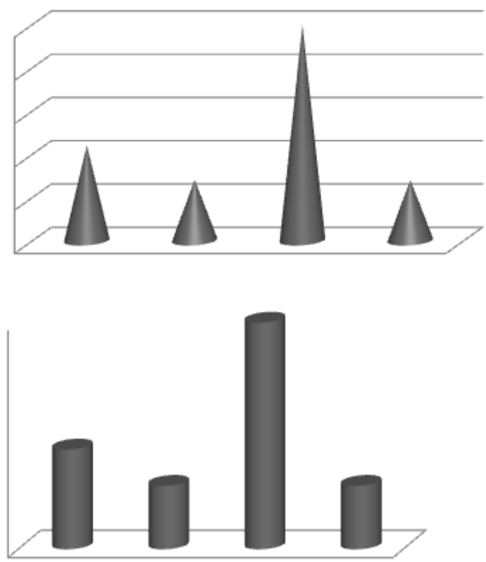
نمودار میله‌ای در شکل ۵ بر اساس اطلاعات جدول ۴ رسم شده است. همان طور که دیده می‌شود، ۷ مقولهٔ متغیر فیلم سینمایی مورد علاقه بر روی محور افقی فهرست شده‌اند. به هر فیلم، مستطیلی اختصاص یافته است که ارتفاع آن برابر با درصد آن مقوله در جدول ۴ است (محور عمودی سمت چپ بر اساس درصد درجه‌بندی شده است تا ارتفاع مستطیل‌ها با آن متناسب باشد). برای مثال، ارتفاع مستطیل مربوط به فیلم «جنایی» و فیلم «ترسناک» هم‌ارتفاع و برابر با ۸ است. می‌توان برای راحتی کاربران، این درصدها را بالای هر مستطیل نمایش داد.

شکل ۵: نمودار میله‌ای برای نمایش توزیع متغیر فیلم سینمایی مورد علاقه



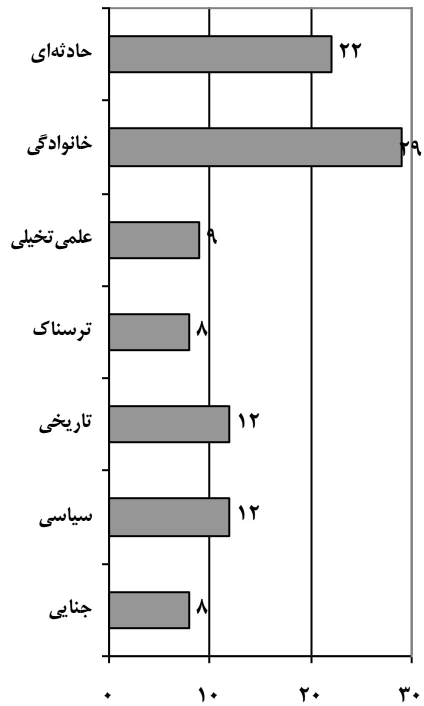
نمودار میله‌ای نیز مانند نمودار دایره‌ای، بر حسب سلیقه پژوهشگر به اشکال گوناگونی نمایش داده می‌شود، در حالی که ماهیت همه آنها همان نمودار میله‌ای است. دو نمونه دیگر از نمودار میله‌ای در شکل ۶ دیده می‌شود.

شکل۶: دو نمونه از نمودار میله‌ای



گاهی نیز جای محورهای عمودی و افقی جا به جا می‌شود؛ یعنی مستطیل‌ها به صورت افقی و نه عمودی رسم می شوند. شکل ۷ تبدیل‌یافتهٔ نمودار میله‌ای شکل ۵ است.

شکل ۷: حالت دیگری از نمایش نمودار میله‌ای شکل ۵



مانند نمودار دایره‌ای و جدول توزیع فراوانی، اولویتی میان مقوله‌های متغیر اسمی در محل قرار گرفتن مستطیل‌های آنها وجود ندارد، ولی برای متغیرهای ترتیبی، منطقی‌تر آن است که مستطیل‌ها از چپ به راست، از پایین‌ترین مقوله به بالاترین آنها مرتب شوند. نمودار میله‌ای رضایت از شغل بر اساس اطلاعات جدول ۳ در شکل ۸ ارائه شده است؛ مقوله‌ها از چپ با «کاملاً کم» آغاز می‌شوند و با «کاملاً زیاد» پایان می‌یابند.

شکل ۸: نمودار میله‌ای برای متغیر رضایت از شغل در جدول ۳

مثال‌هایی که تاکنون از توزیع متغیرها ارائه شد، به متغیرهای مقوله‌ای اختصاص داشت که اسمی یا ترتیبی بودند، در حالی که یافتن و نمایش توزیع متغیرهای فاصله‌ای و نسبتی نیز از اهمیت ویژه‌ای در توصیف داده‌ها برخوردار است. برخی از متغیرهای فاصله‌ای یا نسبتی متغیرهایی گسسته با تعداد مقادیر محدود و برخی دیگر متغیرهایی پیوسته یا متغیرهایی گسسته با مقادیری نامتناهی یا بسیار متعدد هستند. ارائه توزیع متغیرهای فاصله‌ای یا نسبتیِ گسسته با تعداد مقادیر محدود با همان روش ارائه توزیع متغیرهای ترتیبی صورت می‌گیرد؛ یعنی به کمک جدول توزیع فراوانی یا نمودارهای دایره‌ای و میله‌ای. برای مثال، متغیر تعداد روزهایی که فرد در هفته ورزش می‌کند، متغیر نسبتی گسسته‌ای با تنها ۸ وضعیت است که می‌توان توزیع آن را مانند یک متغیر ترتیبی ارائه کرد. در دو بخش بعد به بررسی شیوه نمایش توزیع متغیرهایی پرداخته می‌شود که یا پیوسته هستند یا گسسته، ولی با مقادیر نامتناهی یا بسیار متعدد.

۳ـ۲ـ دسته‌بندی

به‌کارگیری جدول توزیع فراوانی یا نمودارها برای نشان‌دادن توزیع متغیرهای پیوسته یا متغیرهای گسسته‌ای که مقادیر نامتناهی یا متعددی دارند، به ساده شدن کار با داده‌ها و یافتن اطلاعات روان کمکی نمی‌کند، زیرا تنوع مقادیر این متغیرها به ایجاد جدول‌های توزیع فراوانی مختصر و مفیدی منجر نمی‌شوند تا بتوان اطلاعات مشخص و سرراستی از آن به دست آورد.

فرض کنید پس از پخش سریالی تلویزیونی با ۲۴ قسمت، از مردم پرسیده می‌شود که چند قسمت از این سریال را تماشا کرده‌اند؟ تعداد ۲۵ پاسخ برای این پرسش وجود دارد (۰ قسمت، ۱ قسمت، ۲ قسمت،... و ۲۴ قسمت). این پرسش به ایجاد متغیری گسسته منجر می‌شود که وضعیت‌های (مقادیر) متعددی دارد. برای ارائه توزیع این متغیر در قالب جدول توزیع فراوانی، جدولی با ۲۵ سطر خواهیم داشت که خود مانند یک مجموعه داده، شلوغ است و به تلخیص نیاز دارد.

دسته‌بندی برای چه متغیرهایی مناسب است؟

۱. متغیرهایی فاصله‌ای نسبتی پیوسته
۲. متغیرهای فاصله‌ای یا نسبتی گسسته با تعداد مقادیر نامتناهی
۳. متغیرهای فاصله‌ای یا نسبتی گسسته با تعداد مقادیر محدود، ولی متنوع و متعدد

در مثالی دیگر، جدول ۱ را در نظر بگیرید که شامل ۳۱ رقم مختلف برای سن است که از ۷۲ نفر «مشاهده» شده است. از آنجا که سن، متغیری پیوسته است، از نگاه نظری علاوه بر این ۳۱ رقم،

مقادیر بی‌شمار دیگری را نیز می‌توان برای آن برشمرد که در عمل مشاهده نشده‌اند. بنابراین، نمی‌توان جدول توزیع فراوانی‌ای تشکیل داد که تمام مقادیر «ممکن» برای متغیر سن در آن فهرست شوند. حتی ارائه جدول برای ۳۱ مقدار مشاهده‌شده نیز به رسم جدولی با ۳۱ سطر می‌انجامد که به دلیل بزرگ‌بودن قابل توصیف نیست و تنها پژوهشگر را سردرگم می‌سازد.

دسته‌بندیِ مقادیر این نوع متغیرها شیوه‌ای رایج است که برای اطلاع از توزیع آنها بسیار راهگشاست. هدف از دسته‌بندی، تقسیم مقادیر «ممکنِ» متغیر به دسته‌های «دوبه‌دو مجزا» و «فراگیر» است (تبدیل متغیر فاصله‌ای یا نسبتی به یک متغیر مقوله‌ای). دسته‌های دوبه‌دو مجزا، دارای هیچ اشتراکی نیستند؛ یعنی نمی‌توان عضوی از جامعه آماری را یافت که به بیش از یک دسته تعلق داشته باشد. از سوی دیگر، فراگیر بودن این دسته‌ها حاکی از آن است که همگی آنها تمام مقادیر متغیر را شامل می‌شوند؛ یعنی هیچ عضوی از جامعه آماری باقی نمی‌ماند که به یکی از این دسته‌ها تعلق نداشته باشد.

با دسته‌بندی، ستونی طولانی از داده‌های مرتبط با یک متغیر، در چند دسته مرتب می‌شوند و سازمان می‌یابند و به تعبیری، فشرده می‌شوند. تعیین تعداد دسته‌ها و طول دسته‌ها، دو مسئله مهم در دسته‌بندی مقادیر متغیرهاست. حداقل تعداد دسته‌ها می‌تواند دو دسته باشد. برای نمونه، اگر متغیر گسستهٔ تعداد سال‌های تحصیل را به صورت «۰ سال یا همان بی‌سواد، ۱ سال، ۲ سال، ...، ۲۴ سال یا همان دکترا» در نظر بگیریم، مقادیر این متغیر را در ساده‌ترین شکل می‌توان به دو دستهٔ بدون تحصیلات (۰ سال) و دارای تحصیلات (بزرگ‌تر از ۰ سال) دسته‌بندی کرد. بنابراین، تمام افراد بدون تحصیلات در دسته اول و سایرین در دسته دوم جای می‌گیرند؛ به این ترتیب، متغیری با ۲۵ وضعیت، به متغیری با تنها دو وضعیت تبدیل می‌شود.

تعداد دسته‌ها نباید نه آن قدر زیاد باشد که به ایجاد جدول توزیع فراوانی پیچیده و شلوغی منجر شود و نه آن قدر کم که وضعیت‌های «ناهمگونی» را در یک دسته جای دهد. به این ترتیب، دو دستهٔ بدون تحصیلات و دارای تحصیلات در مثال دسته‌بندی تعداد سال‌های تحصیل نیز نامناسب است؛ زیرا دستهٔ «دارای تحصیلات» شامل طیف ناهمگونی از وضعیت‌هاست که از فردی با تحصیلات دبستانی تا فردی با تحصیلات عالی دانشگاهی را شامل می‌شود. بنابراین، بهتر است بیش از دو دسته برای این متغیر در نظر گرفته شود؛

مانند بدون تحصیلات (۰ سال)، تحصیلات کم (۱ تا ۵ سال یعنی تحصیلات دبستانی)، تحصیلات عمومی (۶ تا ۱۱ سال)، دیپلم (۱۲ سال)، کارشناسی (۱۳ تا ۱۶ سال)، کارشناسی ارشد و دکترا (۱۷ تا ۲۴ سال). توجه کنید که با توجه به گسسته بودن متغیر تعداد سال‌های تحصیل، یک تک مقدار مانند ۱۲ سال تحصیل یا همان دیپلم نیز می‌تواند یک دسته قلمداد شود.

دسته‌بندی متغیر و تعیین طول دسته‌ها می‌تواند برگرفته از نظریه‌ها و تعاریف علوم مختلف باشد. برای مثال، متغیر سن را می‌توان بر اساس طبقه‌بندی روان‌شناسان از سن به نوجوان، جوان، میانسال، پیر و کهنسال دسته‌بندی کرد. بنابراین، پژوهشگری که رویکرد روان‌شناسی را برمی‌گزیند، دیگر به دلخواه خود سن را دسته‌بندی نمی‌کند، بلکه باید با مراجعه به تعاریف علم روان‌شناسی طول هر دسته را تعیین کند. برای مثال اگر گروه سنی ۱۲ تا ۱۷ سال به عنوان نوجوان، ۱۷ تا ۳۵ سال به عنوان جوان، ۳۵ تا ۵۵ سال به عنوان میانسال، ۵۵ تا ۷۰ سال به عنوان پیر و بیش از ۷۰ سال به عنوان کهنسال قلمداد شود، او نیز دسته‌بندی خود را بر همین اساس صورت می‌دهد و سن را به ۵ دسته با طول‌های پیشگفته دسته‌بندی می‌کند.

اگر متغیری را بر اساس نظریه‌ها و تعاریف علمی دسته‌بندی نمی‌کنید و خود نیز معیار خاصی برای دسته‌بندی در نظر ندارید، ساده‌ترین شکلِ دسته‌بندی مقادیر متغیر، ایجاد دسته‌هایی با طول برابر (متساوی‌الفاصله) است و معمولاً حداکثر ۱۰ دسته کافی است. با وجود این توجه کنید که دسته‌بندی به ایجاد دسته‌هایی منجر نشود که دارای فراوانی ناچیز یا فراوانی صفر هستند. برای مثال، دسته‌بندی سن به ۱۰ دسته، به گونه‌ای که ۹۰ تا ۱۰۰ سال نیز یکی از دسته‌های آن باشد نامناسب به نظر می‌رسد، زیرا تعداد افراد ۹۰ تا ۱۰۰ ساله در جامعه آماری چنان اندک است که یا فراوانی این دسته صفر خواهد بود یا بسیار ناچیز.

معیارهای دسته‌بندی (تعیین تعداد و طول دسته‌ها) کدامند؟

۱. دسته‌ها باید دوبه‌دو مجزا و فراگیر باشند.
۲. تعداد دسته‌ها به گونه‌ای باشد تا جدول فراوانی مختصر و مفیدی به دست دهد.
۳. معمولاً حداکثر ۱۰ دسته کافی است.
۴. در تعیین طول دسته‌ها نباید وضعیت‌های ناهمگون در یک دسته جای بگیرند.
۵. نظریه‌ها و تعاریف علمی می‌توانند مبنایی برای دسته‌بندی باشند.
۶. در شرایطی که هیچ معیاری در دست نباشد، دسته‌هایی با طول برابر گزینه مناسبی است.
۷. به دسته‌هایی با فراوانی ناچیز یا فراوانی صفر توجه کنید.

متغیرهای فاصله‌ای و نسبتی پس از دسته‌بندی به متغیر ترتیبی تبدیل می‌شوند (تنها در موارد استثنایی ممکن است به متغیر اسمی تبدیل شوند). در واقع دسته‌ها، وضعیت‌های (مقوله‌های) جدیدی هستند که از مقادیر یک متغیر فاصله‌ای یا نسبتی ایجاد شده‌اند و قابل مرتب شدن هستند. برای نمونه، وقتی مقادیر متغیر سن در جدول ۱ را به صورت ۱۰ تا ۲۰ سال، ۲۰ تا ۳۰ سال، ۳۰ تا ۵۰ سال، و بیش از ۵۰ سال دسته‌بندی می‌کنید، متغیر ترتیبی جدیدی با ۴ وضعیت خواهید داشت که وضعیت اول ۱۰ تا ۲۰ سالگی، وضعیت دوم ۲۰ تا ۳۰ سالگی، وضعیت سوم ۳۰ تا ۵۰ سالگی و وضعیت چهارم بیش از ۵۰ سالگی است. اگر قرار باشد این وضعیت‌ها کمّی شوند، برای مثال به آنها اعداد ۱=۱۰ تا ۲۰، ۲=۲۰ تا ۳۰، ۳=۳۰ تا ۵۰ و ۴=بیش از ۵۰ را نسبت می‌دهیم، زیرا نوعی ترتیب بین این ۴ دسته وجود دارد.

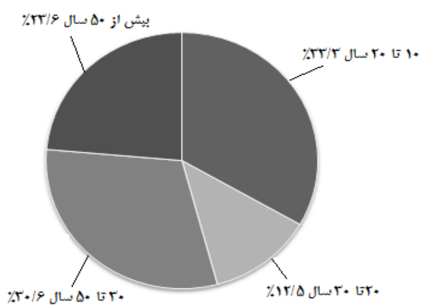
جدول ۵: متغیر سنِ دسته‌بندی‌شده از جدول ۱

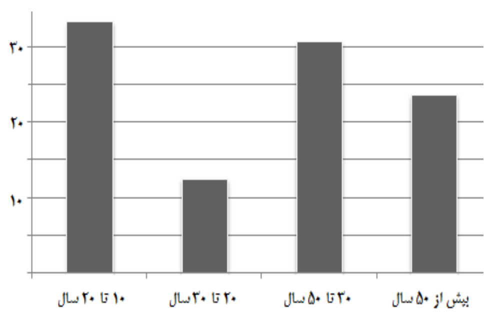
دسته‌ها	فراوانی	درصد	درصد تجمعی
۱۰ تا ۲۰ سال	۲۴	۳۳/۳	۳۳/۳
۲۰ تا ۳۰ سال	۹	۱۲/۵	۴۵/۸
۳۰ تا ۵۰ سال	۲۲	۳۰/۶	۷۶/۴
بیش از ۵۰ سال	۱۷	۲۳/۶	۱۰۰
جمع	۷۲	۱۰۰	–

اکنون که دسته‌بندی متغیر فاصله‌ای یا نسبتی به ایجاد متغیری ترتیبی منجر شد، از روش‌هایی که برای یافتن و نمایش توزیع متغیرهای ترتیبی به کار می‌رود، برای توزیع مقادیر متغیر حاصل از دسته‌بندی نیز استفاده می‌شود. بنابراین، جدول توزیع فراوانی متغیر سن با ۴ دسته، بر اساس داده‌های ۷۲ نفر پاسخگوی جدول ۱، مطابق جدول ۵ به دست می‌آید.

بر اساس اطلاعات این جدول، نمودارهای دایره‌ای و میله‌ای این توزیع نیز مطابق شکل ۹ رسم می‌شود.

شکل ۹: نمودار دایره‌ای و میله‌ای برای نمایش توزیع متغیر سن در جدول ۵





۴ـ۲ـ منحنی توزیع متغیر (بافت‌نگار)

در بخش قبل به کمک دسته‌بندی، به اطلاعاتی درباره توزیع یک متغیر پیوسته دست یافتیم. در این بخش، مفهوم توزیع را برای متغیرهای پیوسته با یافتن «منحنی توزیع متغیر» بیان می‌کنیم که دقیق‌تر از دسته‌بندی کردن است. برای یافتن منحنی توزیع متغیر از ترسیم بافت‌نگار (هیستوگرام) توزیع آن متغیر استفاده می‌شود. برای این منظور، مراحل زیر صورت می‌گیرد:

۱. متغیر پیوسته را با دسته‌بندی مقادیر آن به یک متغیر مقوله‌ای تبدیل کنید (اغلب ایجاد دسته‌هایی با طول برابر برای رسم بافت‌نگار مناسب است[۱]). توجه کنید که نقطهٔ انتهایی هر دسته، نقطهٔ ابتدایی دسته دیگر باشد. بنابراین، متغیری مانند سن به صورت زیر به پنج گروه دسته‌بندی می‌شود: ۰ تا ۱۰، ۱۰ تا ۲۰، ۲۰ تا ۴۰، ۴۰ تا ۶۰، بیش از ۶۰. همان‌طور که دیده می‌شود، عدد ۱۰ نقطهٔ انتهایی دسته اول است، بنابراین به عنوان نقطهٔ ابتدایی دستهٔ دوم انتخاب شده است. همچنین، عدد ۲۰ نقطهٔ انتهایی دستهٔ دوم است که به عنوان نقطهٔ ابتدایی دستهٔ سوم انتخاب شده است.

۲. مشابه جدول ۶، فراوانی نسبی هر مقوله (دسته) را مانند فرایند یافتن جدول توزیع فراوانی برای متغیرهای مقوله‌ای به دست آورید (توجه کنید که لازم نیست فراوانی‌های نسبی را به درصد تبدیل کنید).

جدول ۶: اطلاعات لازم برای ترسیم بافت‌نگار متغیر سن

شماره دسته	دسته (مقوله)	فراوانی	فراوانی نسبی	طول دسته	چگالی دسته
۱	۰ تا ۱۰	۲۰۰	۰/۲	۱۰	۰/۰۲=۰/۲÷۱۰
۲	۱۰ تا ۲۰	۲۵۰	۰/۲۵	۱۰	۰/۰۲۵=۰/۲۵÷۱۰
۳	۲۰ تا ۴۰	۳۵۰	۰/۳۵	۲۰	۰/۰۱۱۷=۰/۳۵÷۲۰
۴	۴۰ تا ۶۰	۱۲۰	۰/۱۲	۲۰	۰/۰۰۶=۰/۱۲÷۲۰
۵	۶۰ تا ۹۰	۸۰	۰/۰۸	۳۰	۰/۰۰۲۶=۰/۰۸÷۳۰
جمع		۱۰۰۰	۱	–	–

۳. طول هر دسته را با تفاضل نقطهٔ ابتدایی دسته از نقطهٔ انتهایی آن به دست آورید. برای مثال در دسته‌بندی متغیر سن، طول دستهٔ اول ۱۰=۰−۱۰ یا طول دستهٔ چهارم

---

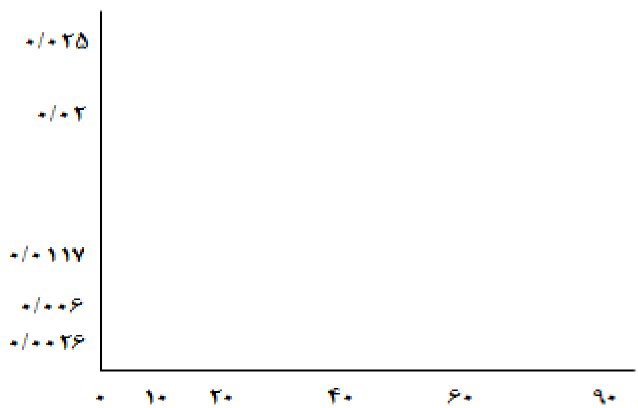
۱. بر اساس تعداد داده‌ها (n)، از رابطه استرژ (Sturges) نیز به صورت ۱+ ۳/۳۲۲ log(n) برای تعیین تعداد دسته‌هایی با طول برابر استفاده می‌شود. این رابطه، حداقل به ۴ و حداکثر به ۲۰ دسته منجر می‌شود.

۶۰ − ۴۰ = ۲۰ است. اما درباره دستهٔ پنجم که نقطهٔ انتهایی ندارد، چه باید کرد؟ برای این منظور می‌توان بالاترین سنی را که ممکن است به طور معمول در جامعه آماری وجود داشته باشد مثلاً ۱۲۰ سال، به عنوان نقطهٔ انتهایی در نظر گرفت. پیشنهاد دیگر، انتخاب بالاترین سنی است که در مجموعه داده‌ها «مشاهده» شده است؛ برای مثال ممکن است این سن برابر با ۹۰ سال باشد.

۴. فراوانی نسبی هر دسته را بر طول آن دسته تقسیم کنید تا به «چِگالی» آن دسته برسید. جدول ۶، این محاسبات را برای پنج گروه متغیر سن در مجموعه داده‌ای با ۱۰۰۰ عضو نمایش می‌دهد.

۵. اکنون اطلاعات کافی برای رسم بافت‌نگار متغیر پیوسته فراهم است. دستگاه مختصاتی با دو محور عمود بر هم تشکیل دهید. محور افقی به دسته‌ها و محور عمودی به چگالی اختصاص دارد. محور افقی را متناسب با طول دسته‌ها مدرج کنید. در شکل ۱۰، محور افقی بر اساس اطلاعات جدول ۶ مدرج شده است. همان طور که دیده می‌شود، فاصلهٔ ۰ تا ۱۰ و ۱۰ تا ۲۰ برابر ۱۰ واحد، ولی فاصله ۲۰ تا ۴۰ یا ۴۰ تا ۶۰ دو برابر آن یعنی ۲۰ واحد در نظر گرفته شده است. فاصله ۶۰ تا ۹۰ نیز ۳۰ واحد یعنی سه برابر فاصله ۰ تا ۱۰ است.

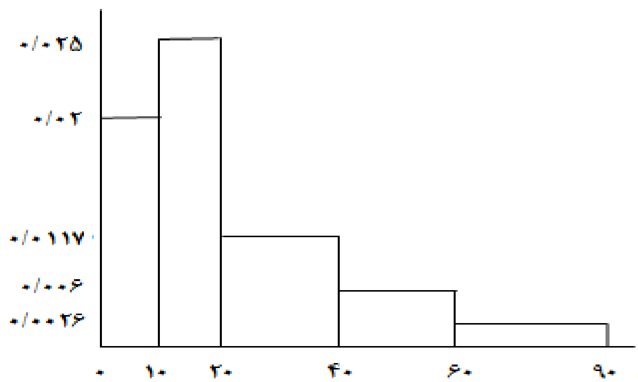
شکل ۱۰: دستگاه مختصات برای رسم بافت‌نگار مرتبط با اطلاعات جدول ۶



۶. هر دسته، مستطیلی بر روی نمودار خواهد داشت که عرض آن برابر با طول دسته و ارتفاع آن برابر با چگالی آن دسته است. بافت‌نگار متغیر سن، بر اساس اطلاعات جدول ۶ در شکل ۱۱ نمایش داده شده است. همان طور که دیده می‌شود، نخستین مستطیل به

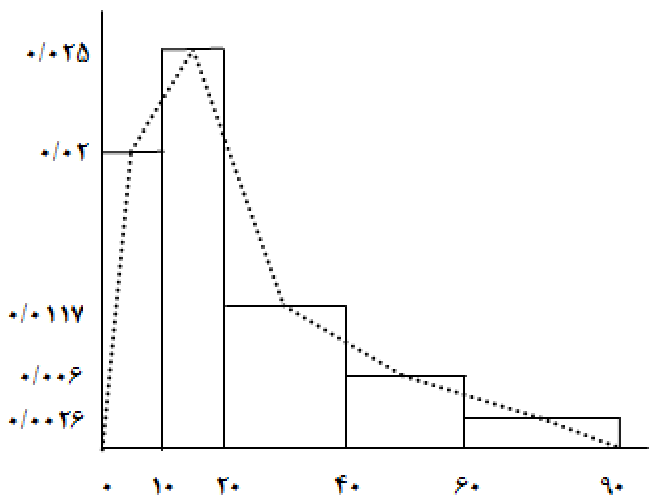
نخستین دسته اختصـاص دارد و عـرض آن از ۰ تـا ۱۰ و ارتفـاع آن برابـر بـا ۰/۰۲ یعنـی چگالی دستهٔ اول است.

شکل ۱۱: بافت‌نگار متغیر سن بر اساس اطلاعات جدول ۶



پس از رسم بافت‌نگار، شکل تقریبی منحنی توزیع از اتصال نقاط میـانی اضـلاع بـالایی مستطیل‌ها به یکدیگر نمایان می‌شود. خط‌چین‌ها در شکل ۱۲ همان شـکل تقریبـی منحنـی توزیع متغیر سن هستند که نمودار چند ضلعی نامیده می‌شود.

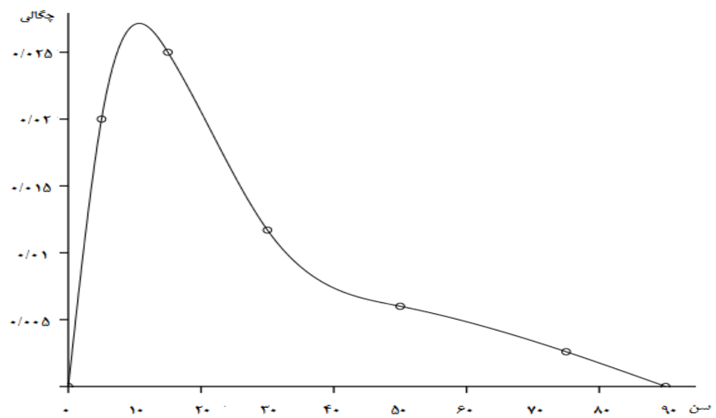
شکل ۱۲: شکل تقریبی منحنی توزیع متغیر سن



در شکل ۱۲، مقادیر متغیر سن به ۵ گروه دسته‌بندی شده است، بنابراین ۵ مستطیل و به تبع آن شکل تقریبی منحنی، از ۶ پاره‌خط تشکیل شده است (همیشه تعداد پاره‌خط‌ها یکـی

بیشتر از تعداد گروه‌هاست). فرض کنید متغیر سن را به ۱۰ گروه دسته‌بندی کنیم. برای مثال، هر یک از گروه‌های سنی جدول ۶ را با نصف کردن، به ۲ گروه تبدیل و نمودار بافت‌نگار جدیدی را برای آنها رسم کنیم؛ یعنی ۰–۵ و ۵–۱۰به جای ۰–۱۰، ۱۰–۱۵ و ۱۵–۲۰ به جای ۱۰–۲۰، ۲۰–۳۰ و ۳۰–۴۰ به جای ۲۰–۴۰، ۴۰–۵۰ و ۵۰–۶۰ به جای ۴۰–۶۰ و سرانجام ۶۰–۷۵ و ۷۵–۹۰ به جای ۶۰–۹۰. اکنون ۱۰ مستطیل خواهیم داشت و نمودار چند ضلعی آن از ۱۱ پاره‌خط تشکیل می‌شود که ظریف‌تر از نمودار شکل ۱۲ است. اگر این فرایند را با ۲۰ گروه انجام دهیم (مثلاً ۱۰ گروه قبل را با نصف کردن به ۲۰ گروه تبدیل و بافت‌نگار آنها را رسم کنیم)، به نمودار چندضلعی ظریف‌تری با ۲۱ پاره‌خط خواهیم رسید. رفته‌رفته با افزایش تعداد گروه‌ها از ۲۰ به ۴۰ و از ۴۰ به ۸۰ و... نمودار چند ضلعی، به منحنی پیوسته‌ای مانند شکل ۱۳ تبدیل می‌شود که به آن منحنی توزیع متغیر سن می‌گویند (دایره‌های کوچک بر این منحنی همان نقاط میانی مستطیل‌های شکل ۱۲ هستند).

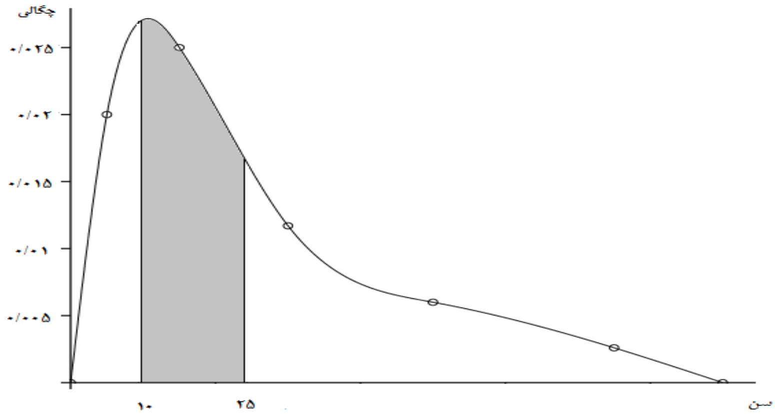
شکل ۱۳: منحنی توزیع متغیر سن



اکنون این پرسش مطرح می‌شود که چگونه از منحنی توزیع برای یافتن توزیع متغیر استفاده می‌شود. برای پاسخ، به یاد بیاورید که توزیع متغیر مقوله‌ای، نسبت‌هایی از اعضای جامعه آماری را ارائه می‌کرد که به مقوله خاصی از متغیر تعلق داشتند. مساحت زیر منحنی توزیع نیز می‌تواند همین نسبت‌ها را به دست دهد. به عبارت دیگر، مساحتی از زیر منحنی توزیع که بین دو مقدار متغیر قرار دارد، نسبتی از اعضای جامعه آماری را نشان می‌دهد که بین این دو مقدار از متغیر قرار گرفته‌اند. برای مثال، برای یافتن نسبت افراد ۱۰ تا ۲۵ ساله،

یعنی کسانی که در مقوله ۱۰ تا ۲۵ سال جای می‌گیرند، کافی است مساحتی را از زیر منحنی توزیع متغیر سن به دست آوریم که بین دو مقدار ۱۰ و ۲۵ قرار دارد (یعنی مساحت ناحیهٔ مشخص‌شده در منحنی شکل ۱۴).

شکل ۱۴: ناحیه‌ای از منحنی توزیع متغیر سن که نسبت ۱۰ تا ۲۵ ساله‌ها را نشان می‌دهد



اغلب، محاسبه مساحت ناحیه‌های خاصی از زیر منحنی توزیع به کمک رایانه و با نرم‌افزارهای آماری یا ریاضی صورت می‌گیرد؛ هر چند با محاسبات ساده‌ای روی نمودار چندضلعی نیز می‌توان به مساحت تقریبی این ناحیه‌ها دست یافت. همچنین، برخی از منحنی‌های توزیع، به دلیل کاربرد زیاد، دارای نام‌های اختصاصی هستند؛ نظیر منحنی توزیع نرمال، منحنی توزیع t یا منحنی توزیع F، و جدول‌هایی برای آنها تهیه شده است که مساحت زیر منحنی توزیع آنها را به ازای مقادیر مختلف متغیر به دست می‌دهند. بنابراین، اگر منحنی توزیع متغیری مانند منحنی یکی از این توزیع‌های شناخته‌شده باشد، می‌توان بدون استفاده از نرم‌افزار یا محاسبات خاص و تنها با مراجعه به این جدول‌ها، به مساحت ناحیه مورد نظر دست یافت.

همواره مساحت کل زیر منحنی توزیع، زیر منحنی چندضلعی و زیر بافت‌نگار، برابر با ۱ است؛ صرفنظر از اینکه برای چه متغیری رسم شده باشند.

۳ـ مشخصه‌های کمّی توزیع

مانند یک خودرو که دارای مشخصاتی نظیر مدل، تعداد سیلندر، نوع سوخت، حداکثر سرعت و... است و این مشخصات به شناخت ما از خودرو کمک می‌کند، توزیع متغیر نیز مشخصاتی دارد که هر یک از آنها «اطلاعی» را درباره خصوصیت و کیفیتی از آن به دست می‌دهد. در این

بخش، از مشخصه‌های توزیع تحت سه عنوانِ تمرکز، پراکندگی و شکل بحث می‌شود، ولی پیش از آن، برخی از نمادهای آماری شرح داده می‌شوند که به ارائه بهتر این بحث کمک می‌کنند.

۱ـ۳ـ نمادگذاری

نمادهایی که در این بخش معرفی می‌شوند، در فصل‌های بعد نیز استفاده خواهند شد. نمادها، بیان شیوه‌های محاسبه را در بحث‌های آماری، ساده و دقیق می‌سازند، از این رو به خواننده توصیه می‌شود در به‌کارگیری آنها مهارت یابد.

متغیرها، با حرف‌های بزرگ انگلیسی مانند X، Y یا Z و داده‌های مربوط به آنها با حرف‌های کوچک اَندیس‌دار نمایش داده می‌شوند. برای مثال، اگر نام متغیری X باشد داده‌های مربوط به آن با $x_۱$، $x_۲$، ...، x_n نشان داده می‌شوند. اندیس‌ها نشان می‌دهند که داده‌ها نتیجهٔ اندازه‌گیری متغیر از کدام عضو هستند. برای مثال، اندیسِ ۱ در $x_۱$، مقدار متغیر X را برای اولین عضو و اندیسِ ۲ در $x_۲$، مقدار متغیر X را برای دومین عضو نشان می‌دهد. بنابراین، عدد n به عنوان آخرین اندیس به آخرین فرد اندازه‌گیری شده اشاره دارد و برابر با تعداد کل داده‌های مربوط به متغیر است. پس اگر متغیری از ۱۰ نفر اندازه‌گیری شده باشد، $n = ۱۰$ خواهد بود.

از حروف انگلیسی برای ارائه بحث‌های آماری در حالتی کلی مثلاً نمایش یک فرمول استفاده می‌شود، زیرا به هیچ عدد و رقم خاصی اشاره نمی‌کنند. برای مثال، $x_۱$، $x_۲$، $x_۳$، $x_۴$ و $x_۵$ تنها به ۵ داده از متغیری به نام X اشاره دارد، بدون آنکه عدد یا رقم خاصی را برای این داده‌ها معین کرده باشد. هنگامی که مایل باشیم محاسباتی را بر روی داده‌های مشخصی صورت دهیم، اعداد به جای حروف قرار داده می‌شوند. برای نمونه، اگر متغیر سن را با حرف X نشان دهیم و مقدار این متغیر برای ۵ عضو از جامعه آماری ۲۵، ۴۰، ۳۹، ۱۲ و ۴۷ سال باشد، در این صورت $n = ۵$ و $x_۱ = ۲۵$، $x_۲ = ۴۰$، $x_۳ = ۳۹$، $x_۴ = ۱۲$ و $x_۵ = ۴۷$ است.

گاهی از حروف کوچک انگلیسی به عنوان اندیس استفاده می‌شود تا به داده‌ای دلخواه اشاره شود. برای مثال، از نماد x_i برای نشان دادن دادهٔ i اُم استفاده می‌شود، ولی به درستی مشخص نمی‌شود که i چه مقداری دارد؛ پس iاُمین داده می‌تواند دادهٔ اول، دادهٔ دوم، ... یا حتی دادهٔ آخر باشد. پس تفاوت x_i با مثلاً $x_۲$ در آن است که $x_۲$ دقیقاً به دادهٔ دوم اشاره می‌کند، در حالی که x_i می‌تواند به هر یک از داده‌های متغیر X مربوط باشد.

عملیات ریاضی نظیر جمع، تفریق، ضرب، تقسیم یا توان را نیز می‌توان با این نمادها به کار برد. مواردی که در ادامه می‌آید، نمونه‌هایی از آنها هستند.

$x_i^۲$ دادهٔ i اُم به توان ۲ می‌رسد (مربع دادهٔ i اُم)

$x_i - a$ عدد a از دادهٔ i اُم کم می‌شود

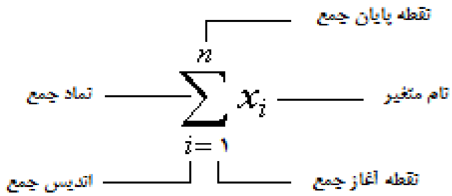
$a\,x_i$ عدد a در دادهٔ i اُم ضرب می‌شود

$x_۱ + x_۲ + \cdots + x_n$ تعداد n داده با هم جمع می‌شوند

یکی از نمادهای بسیار رایج در آمار، نمادِ جمع $\sum$ است که شیوهٔ نوشتن مجموع تعدادی داده را به صورت زیر خلاصه‌تر می‌سازد (بخوانید سیگمایِ x_i، i از ۱ تا n):

$$\sum_{i=۱}^{n} x_i = x_۱ + x_۲ + \cdots + x_n$$

اجزای این نماد به شرح زیر هستند:



این نماد می‌تواند با عملیات‌های دیگر نیز ترکیب شود، مثلاً جمع توان‌های دوم داده‌ها که به شکل زیر است:

$$\sum_{i=۱}^{n} x_i^۲ = x_۱^۲ + x_۲^۲ + \cdots + x_n^۲$$

یا جمع تفاضل‌های داده‌ها از عدد a که به صورت زیر نمایش داده می‌شود:

$$\sum_{i=۱}^{n} (x_i - a) = (x_۱ - a) + (x_۲ - a) + \cdots + (x_n - a)$$

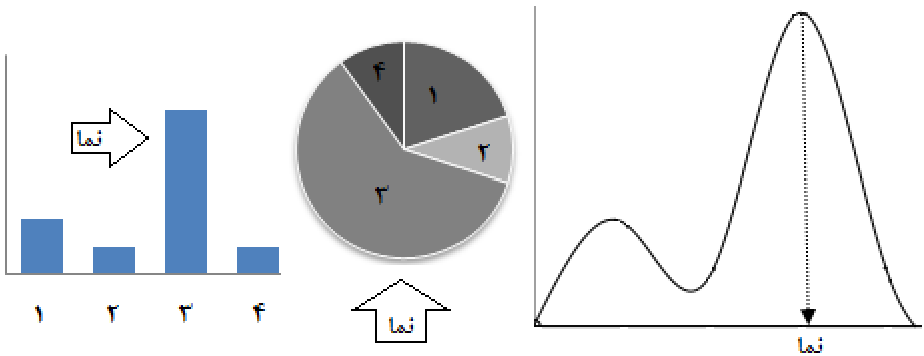
۲ـ۳ـ مشخصه‌های تمرکز توزیع

در فصل پیش بیان شد که متغیرها ویژگی‌هایی هستند که تمام اعضای جامعه آماری به طور یکسان آن را دارا نیستند. مقادیر متغیرها، بویژه متغیرهای پیوسته می‌تواند بسیار متنوع باشد. یافتن توزیع متغیر شیوه‌ای برای سامان دادن به این تنوع و گوناگونی است، به طوری که پژوهشگر دریابد هر مقدار از متغیر را چه نسبتی از اعضای جامعه دارا هستند. مشخصه‌های تمرکز توزیع، اطلاعات توزیع را در تنها یک عدد خلاصه می‌سازند، به طوری که آن عدد تا حدی «نمایندهٔ عمومی» توزیع مقادیر متغیر در جامعه آماری باشد و وضعیت عمومی توزیع مقادیر را نشان دهد.

۱ـ۲ـ۳ـ نما

نخستین معیاری که به عنوان نماینده توزیع، مناسب به نظر می‌رسد، مقداری از متغیر است که دارای بیشترین فراوانی است؛ یعنی در مقایسه با سایر مقادیر متغیر، تعداد بیشتری از اعضای جامعه آماری دارای آن هستند (اکثریت با این مقدار است). این مقدار، نمای توزیع نامیده می‌شود. برای مثال، اگر ۶۰ درصد جامعه از میان سه وضعیت خانه استیجاری، خانه شخصی و خانه سازمانی، دارای خانه شخصی باشند، پس وضعیت خانه شخصی دارای بیشترین فراوانی بوده و اکثریت را به خود اختصاص داده است (به جدول ۲ مراجعه کنید). از این رو، خانه شخصی بهتر از دو وضعیت دیگر می‌تواند وضعیت عمومی توزیع متغیر نوع خانه مسکونی را در جامعه آماری نشان دهد.

شکل ۱۵: یافتن نما از نمودار یا منحنی توزیع

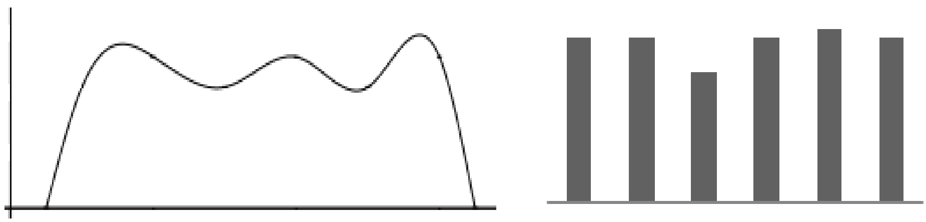


درصدها در جدول توزیع فراوانی برای تشخیص نما به کار می‌روند. مانند شکل ۱۵، بزرگ‌ترین برش در نمودار دایره‌ای و بلندترین مستطیل در نمودار میله‌ای، نشان‌دهنده نمای توزیع متغیر است. مقداری از متغیر در منحنی توزیع، نما به حساب می‌آید که قلّه این منحنی را نشان می‌دهد (در واقع، آن مقداری از متغیرِ «پیوسته» نما قلمداد می‌شود که بیشترین چگالی را دارا باشد، پس قلهٔ توزیع، نما را نشان می‌دهد، زیرا به بیشترین چگالی اشاره دارد).

ارزش نما به عنوان نماینده عمومی توزیع به بزرگی فراوانی آن بستگی دارد. هر چه فراوانی نما چشمگیرتر باشد، نمایندهٔ درصد بیشتری از جامعه آماری خواهد بود و داده‌ها بر مقدار آن متمرکزتر هستند. بنابراین، در توزیع‌های هموار مانند شکل ۱۶ که تفاوت درصدهای وضعیت‌های مختلف متغیر چندان زیاد نیست، یعنی فراوانی‌ها تقریباً به یک

نسبت و به شکلی یکنواخت بین وضعیت‌ها تقسیم شده‌اند، نما نمی‌تواند نماینده قابل قبولی برای توزیع متغیر باشد، زیرا فراوانی نما اختلاف چندانی با فراوانی سایر مقادیر متغیر ندارد تا آن را نسبت به مقادیر دیگر برجسته‌تر سازد.

شکل ۱۶: دو نمونه از توزیع‌های تاحدی هموار



همچنین، به این نکته توجه کنید که توزیع می‌تواند چندنمایی باشد؛ یعنی بیش از یک نما داشته باشد که در این صورت نمی‌توان تنها یک مقدار را به عنوان مشخصه تمرکز توزیع ارائه کرد (نمونه‌هایی از منحنی‌های توزیع چندنمایی در شکل ۲۵ نمایش داده شده است).

۲ـ۲ـ۳ـ چَندَک‌ها و میانه

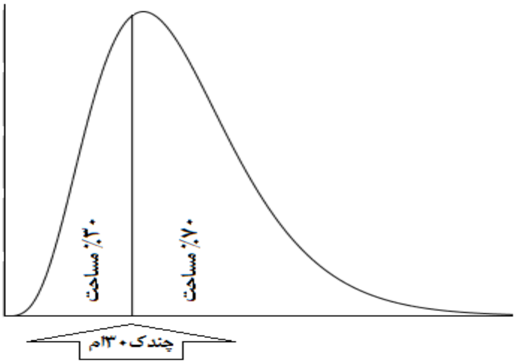
میانه به خانواده‌ای از مشخصه‌های توزیع تعلق دارد که با عنوان چندک‌ها شناخته می‌شوند. در اینجا، ابتدا تعریفی از چندک‌ها و شیوه محاسبه آنها ارائه می‌شود و سپس به طور خاص درباره میانه به عنوان یکی از مشخصه‌های تمرکز توزیع بحث خواهد شد.

برای درک مفهوم چندک، ابتدا متغیر «پیوستهٔ» X و عدد p را با مقداری بین ۰ تا ۱۰۰ در نظر بگیرید. مقدار Q_p را چندک p اُم توزیع متغیر پیوستهٔ X می‌نامیم؛ اگر p درصد از جامعه آماری کمتر از Q_p و $(۱۰۰-p)$ درصد از جامعه بیشتر از آن باشند. به این ترتیب، چندک، اعضای جامعه آماری را به دو بخش تقسیم می‌کند: اعضایی که دارای مقداری کمتر از مقدار چندک هستند؛ یعنی p درصد از جامعه و اعضایی که دارای مقداری بیش از مقدار چندک هستند؛ یعنی $(۱۰۰-p)$ درصد از جامعه. برای مثال، اگر $p=۳۰$ باشد، چندک ۳۰ اُم یعنی $Q_{۳۰}$ عبارت از مقداری است که ۳۰ درصد از جامعه دارای مقداری کمتر از آن و ۷۰ درصد از جامعه دارای مقداری بیشتر از آن هستند.

مقدار چندک در منحنی توزیع با خطی عمودی مشخص می‌شود که سطح زیر منحنی را به دو بخش تقسیم می‌کند؛ به طوری که p درصد از سطح زیر منحنی در سمت چپ این

خط و درصد باقیمانده در سمت راست آن قرار بگیرد. این خط برای چندک ۳۰ اُم در شکل ۱۷ نمایش داده شده است.

شکل ۱۷: محل چندک ۳۰ ام در منحنی توزیع



اکنون تعریفی از چندک ارائه می‌دهیم که هم برای متغیرهای پیوسته و هم برای متغیرهای گسسته به کار می‌رود و در یافتن چندک برای داده‌های گردآوری‌شده و انجام محاسبات مربوط به آن نیز راهگشاست. مقدار Q_p را چندک p اُم می‌گویند؛ اگر از یک سو مقدار متغیر برای حداقل p درصد از جامعه آماری حداکثر برابر با Q_p و از سوی دیگر مقدار متغیر برای حداکثر $(۱۰۰-p)$ درصد از جامعه آماری بیش از Q_p باشد. این تعریف به زبان آماری به شکل زیر بیان می‌شود[1]:

$$P(X \le Q_p) \ge p \quad \text{و} \quad P(X > Q_p) \le ۱۰۰ - p$$

> Q_p چندک p اُم توزیع است؛ اگر p درصد از جامعه بزرگ‌تر از Q_p نباشد و $(۱۰۰-p)$ درصد نیز کوچک‌تر از آن نباشد.

برای مثال، اگر $p = ۳۰$ باشد، چندک ۳۰ اُم یعنی $Q_{۳۰}$ عبارت از مقداری است که مقدار متغیر برای حداقل ۳۰ درصد از جامعه آماری حداکثر برابر با آن و برای حداکثر ۷۰ درصد از جامعه آماری بیش از آن باشد.

برای محاسبه چندک p اُم توزیع متغیر X از روی داده‌ها، ابتدا آنها را از کوچک‌ترین داده به بزرگ‌ترین مرتب کنید. فرض کنید تعداد n داده به صورت $x_۱$، $x_۲$، ...، x_n در دست است. این داده‌ها را پس از مرتب شدن به صورت $x_{(۱)}$، $x_{(۲)}$، ...، $x_{(n)}$ نشان می‌دهیم. فرض کنید حاصلِ $(n+۱)\times\frac{p}{۱۰۰}$ عددی به صورت a/d باشد (علامت / نشان‌دهنده ممیز است، پس a بخش صحیح و d بخش اعشاری این عدد است). در این صورت، چندک p اُم داده‌ها از رابطه زیر به دست می‌آید:

$$Q_p = x_{(a)} + [x_{(a+۱)} - x_{(a)}] \times ۰/d$$

---

۱. نماد $P(X \le Q_p)$ را این گونه بخوانید: درصدی از جامعه آماری که مقدار متغیر X برای آنها حداکثر برابر با عدد Q_p است (یعنی بیش از Q_p نیست). به همین ترتیب، $P(X > Q_p)$ به معنی درصدی از جامعه آماری است که مقدار متغیر X برای آنها بیش از عدد Q_p است.

برای مثال، فرض کنید می‌خواهید چندک‌های بیست‌وسوم $(p=۲۳)$ و هفتادوپنجم $(p=۷۵)$ را برای ۱۲ داده زیر محاسبه کنید.

$$x_۱=۵,\ x_۲=۳,\ x_۳=۹,\ x_۴=۱,\ x_۵=۴,\ x_۶=۱۲$$
$$x_۷=۲۳,\ x_۸=۸,\ x_۹=۱,\ x_{۱۰}=۹,\ x_{۱۱}=۵,\ x_{۱۲}=۶$$

ابتدا باید آنها را به شکل زیر از کمترین به بیشترین مرتب کنید:

$$x_{(۱)}=۱,\ x_{(۲)}=۱,\ x_{(۳)}=۳,\ x_{(۴)}=۴,\ x_{(۵)}=۵,\ x_{(۶)}=۵$$
$$x_{(۷)}=۶,\ x_{(۸)}=۸,\ x_{(۹)}=۹,\ x_{(۱۰)}=۹,\ x_{(۱۱)}=۱۲,\ x_{(۱۲)}=۲۳$$

سپس مقدار $(n+۱)\times\frac{p}{۱۰۰}$ را برای $p=۲۳$ و $p=۷۵$ محاسبه کنید. از آنجا که $n=۱۲$ است، این مقدار به‌ازای $p=۲۳$ برابر با ۲/۹۹ و به‌ازای $p=۷۵$ برابر با ۹/۷۵ است. بنابراین، برای $p=۲۳$ داریم $a=۲$ و $d=۰/۹۹$ و برای $p=۷۵$ داریم $a=۹$ و $d=۰/۷۵$. پس چندک بیست‌وسوم و چندک هفتادوپنجم به صورت زیر به دست می‌آیند:

$$Q_{۲۳}=x_{(۲)}+[x_{(۳)}-x_{(۲)}]\times۰/۹۹=۱+[۳-۱]\times۰/۹۹=۲/۹۸$$
$$Q_{۷۵}=x_{(۹)}+[x_{(۱۰)}-x_{(۹)}]\times۰/۷۵=۹+[۹-۹]\times۰/۷۵=۹$$

میانه، همان چندک پنجاهم $(p=۵۰)$ توزیع است. به زبان ساده، مقداری است که توزیع را به دو نیمه با فراوانی برابر تقسیم می‌کند. از همین رو به عنوان مشخصه تمرکز توزیع، یعنی نماینده عمومی آن انتخاب شده است. در منحنی توزیع، میانه با خطی عمودی به دست می‌آید؛ به طوری که مساحت زیر منحنی در دو طرف این خط برابر باشد.

باید به این نکته اساسی توجه داشت که لزوماً مقدار وسطِ مقادیر متغیر همان میانه نیست. برای مثال، اگر مقادیر متغیری از ۰ تا ۱۳۰ باشد، مقدار وسط برابر با ۶۵ است، ولی میانه می‌تواند مقدار دیگری باشد؛ زیرا میانه، توزیع جمعیت جامعه و نه مقادیر متغیر را به دو نیمه تقسیم می‌کند، پس برای یافتن میانه باید به دنبال مقداری گشت که مقادیر متغیر برای نیمی از «جامعه» کمتر از آن و برای نیمی دیگر بیشتر از آن باشد.

برای محاسبه میانه، ابتدا مقادیر مشاهده شده از متغیر یعنی همان داده‌ها را از کوچک‌ترین به بزرگ‌ترین مرتب کنید. فرض کنید $x_{(۱)}$، $x_{(۲)}$، ...، $x_{(n)}$ داده‌های مرتب

شده باشند. اگر تعداد داده‌ها یعنی n فرد باشد، از رابطهٔ (۱) و اگر زوج باشد از رابطهٔ (۲) برای محاسبه میانه استفاده می‌شود.

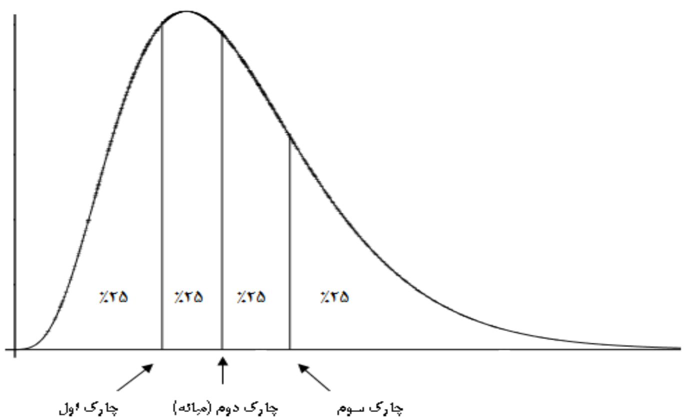
$$Q_{۵۰} = \begin{cases} x_{(\frac{n+۱}{۲})} & (۱) \\ \dfrac{x_{(\frac{n}{۲})} + x_{(\frac{n}{۲}+۱)}}{۲} & (۲) \end{cases}$$

بنابراین، میانهٔ ۱۲ داده‌ای که دو چندک آن را پیشتر محاسبه کردید، به صورت زیر به دست می‌آید (توجه کنید که $n=۱۲$ زوج است):

$$Q_{۵۰} = \frac{x_{(۶)} + x_{(۷)}}{۲} = \frac{۵+۶}{۲} = ۵/۵$$

برخی از چندک‌های دیگر نیز مانند میانه دارای نام‌های خاصی هستند؛ نظیر چارَک‌ها، دَهَک‌ها و صَدَک‌ها. هر توزیع دارای سه چارک است: چارک اول یعنی $Q_{۲۵}$، چارک دوم که همان میانه است و چارک سوم یعنی $Q_{۷۵}$. سه چارک با هم می‌توانند توزیع را به چهار بخش با حجم برابر تقسیم کنند، به طوری که هر بخش شامل ۲۵ درصد از جامعه آماری (داده‌ها) باشد. شکل ۱۸ این تقسیم بندی را بر روی منحنی توزیع نشان می‌دهد.

شکل ۱۸: نمایش چارک‌ها بر روی منحنی توزیع



مساحت زیر منحنی، در سمت چپ چارک اول، برابر با ۲۵ درصد کل مساحت است. از آنجا که مساحت سمت چپ چارک دوم برابر با ۵۰ درصد کل مساحت است، پس مساحت زیر منحنی که بین چارک اول و دوم قرار دارد، برابر با ۲۵ درصد کل مساحت

است. به همین ترتیب، مساحت بین چارک دوم و سوم و مساحت سمت راست چارک سوم نیز برابر با ۲۵ درصد کل مساحت است.

هر توزیع دارای نُه دهک به صورت دهک اول $Q_{.1}$، دهک دوم $Q_{.2}$، دهک سوم $Q_{.3}$، ... و دهک نهم $Q_{.9}$ است که با هم می‌توانند توزیع را به ۱۰ قسمت تقسیم کنند که هر قسمت شامل ۱۰ درصد از جامعه آماری (داده‌ها) باشد. به همین ترتیب، هر توزیع دارای ۹۹ صدک به صورت صدک اول Q_1، صدک دوم Q_2، ... و صدک نودونهم Q_{99} است که با هم می‌توانند توزیع را به ۱۰۰ قسمت تقسیم کنند، به طوری که هر قسمت شامل ۱ درصد از جامعه آماری (داده‌ها) باشد.

۳ـ۲ـ۳ـ میانگین

نما، یکی از مقادیر مشاهده شده از متغیر است که بیشترین فراوانی را در مقایسه با سایر مقادیر مشاهده شده دارد. میانه و به طور کلی چارک‌ها نیز بر اساس فراوانی به دست می‌آیند. برای مثال، میانه کمّیتی است که داده‌ها را به دو نیمه با فراوانی برابر تقسیم می‌کند. به عبارت دیگر، این دو مشخصه تمرکز تنها بر اساس فراوانی‌ها تعیین می‌شوند و بزرگی یا کوچکی مقادیر نقشی در محاسبه آنها ندارد.

برای مثال، میانه و نمای دو مجموعه دادهٔ جدول ۷ یکسان و به ترتیب برابر با ۵ و ۱ است، با وجود آنکه این دو مجموعه داده متفاوت هستند، ولی در مقدار نما و میانه اشتراک دارند. این بدان معناست که مثلاً میانه، تنها داده‌ها را به دو بخش با فراوانی برابر تقسیم می‌کند و با داشتن آن تنها می‌دانیم که بخشی از مقادیر از آن تجاوز نمی‌کنند، ولی نمی‌دانیم که تا چه حد از آن کوچک‌ترند، یا می‌دانیم نیمی از داده‌ها از آن بزرگ‌ترند، ولی نمی‌دانیم که تا چه حد از آن بزرگ‌ترند. میانگین، مشخصه‌ای از توزیع است که تحت تأثیر تمام مقادیر است.

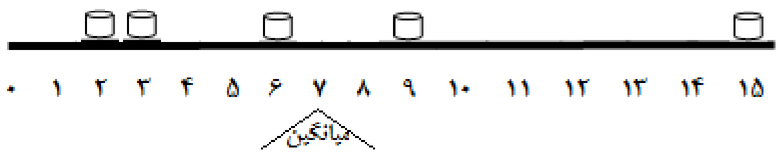
جدول ۷: مثالی از دو مجموعه دادهٔ متفاوت، ولی با میانه و نمای یکسان

مجموعه اول	۱، ۱، ۱، ۱، ۲، ۵، ۶، ۷، ۷، ۷ و ۸
مجموعه دوم	۱، ۱، ۱، ۳، ۴، ۵، ۱۰۰، ۱۰۱، ۱۰۲، ۱۰۳ و ۱۰۴

میانگین، نقطه تعادل داده‌هاست. برای ارائه تعبیری فیزیکی، تصور کنید میله‌ای مدرج مانند شکل ۱۹ در اختیار دارید و ۵ دادهٔ ۲، ۳، ۶، ۹ و ۱۵ در دست است. ۵ وزنه هم‌وزن را در محل مربوط به هر داده، روی میله قرار دهید. پس وزنه‌ای در محل درجه ۲، وزنه‌ای در

محل درجه ۳، وزنه‌ای در محل درجه ۶، وزنه‌ای در محل درجه ۹ و سرانجام وزنه‌ای در محل درجه ۱۵ قرار می‌گیرد. اکنون می‌خواهید این میله را با ریسمانی (مانند ترازوهای قدیمی) بلند کنید، به طوری که تعادل آن حفظ شود و میله به هیچ سمتی کج نشود. میانگین این ۵ عدد یعنی عدد ۷ همان نقطه‌ای از میله مدرج است که می‌توانید ریسمان را به آن ببندید و میله را بدون از دست دادن تعادل از آن نقطه بلند کنید.

شکل ۱۹: میانگین، نقطهٔ تعادل داده‌ها



عدد ۷، از جمع این ۵ داده که برابر با ۳۵ است و تقسیم آن بر تعداد داده‌ها یعنی ۵ به دست آمده است. محاسبه میانگین n دادهٔ x_1، x_2،...، x_n با رابطه زیر صورت می‌گیرد (نمادِ $\bar{x}$ را «ایکس بار» بخوانید که در آمار نشان‌دهنده میانگین متغیر X است):

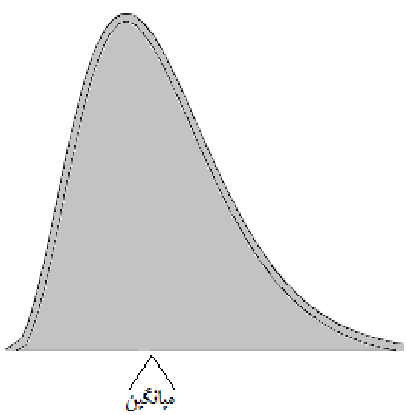
$$\bar{x} = \frac{x_1 + x_2 + \cdots + x_n}{n} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

ویژگی آشکار میانگین، وابستگی آن به تمام داده‌هاست؛ زیرا در محاسبه آن «همگی» داده‌ها با هم جمع می‌شوند. انجام این محاسبات برای دو مجموعه داده جدول ۷، دو رقم ۴/۱۸ برای مجموعه اول و ۴۷/۷۲ برای مجموعه دوم را به دست می‌دهد که به‌خوبی تفاوت داده‌های دو مجموعه را نمایان می‌سازد.

میانگین مجموعه دوم اختلاف زیادی با مجموعه اول دارد، زیرا شامل داده‌های بزرگ‌تری است که مقدار میانگین را به سمت خود می‌کشند.

میانگین، نقطهٔ تعادل بافت‌نگار و منحنی توزیع نیز قلمداد می‌شود. اگر بافت‌نگار یا منحنی توزیع را یک جسم فیزیکی در نظر بگیرید، میانگین، نقطه‌ای از آن است که این جسم می‌تواند با تعادل روی آن بایستد. شکل ۲۰، محل میانگین را به عنوان نقطه تعادل منحنی توزیع شکل ۱۸ نشان

شکل ۲۰: میانگین نقطه تعادل منحنی توزیع



می‌دهد. در واقع، اگر این منحنی توده‌ای دارای جِرم باشد، میانگین، مرکز ثقل آن است که با قرار گرفتن بر روی آن به شکلی متعادل می‌ایستد.

میانگین، دارای انواعی است که یکی از رایج‌ترین آنها به نام میانگین حسابی یا همان متوسط، شناخته می‌شود. هر کجا واژه «میانگین» به تنهایی به کار رود، منظور میانگین حسابی (متوسط یا معدل) است. میانگین هندسی (ریشه n اُم حاصل‌ضرب n داده) و میانگین سازوار (معکوسِ میانگینِ حسابیِ وارون داده‌ها) دو نوع دیگر از میانگین هستند که به صورت زیر تعریف می‌شوند:

$$\bar{x}_g = (x_۱ \times x_۲ \times \cdots \times x_n)^{\frac{۱}{n}} = \sqrt[n]{x_۱ \times x_۲ \times \cdots \times x_n}$$

$$\bar{x}_h = \frac{n}{\frac{۱}{x_۱} + \frac{۱}{x_۲} + \cdots + \frac{۱}{x_n}} = \frac{n}{\sum_{i=۱}^{n} \frac{۱}{x_i}}$$

۳ـ۳ـ مشخصه‌های پراکندگی توزیع

پیشتر بیان شد که مشخصه‌های تمرکز در پی یافتن نقطه تعادل توزیع هستند؛ نقطه‌ای که بتوان به عنوان نماینده عمومی داده‌ها معرفی کرد. اما یافتن نقطه‌ای که داده‌ها به آن گرایش دارند، تنها مشخصه‌ای نیست که برای شناخت ویژگی‌های توزیع کافی باشد، بلکه میزان دوری یا نزدیکی داده‌ها از یکدیگر و به عبارت دیگر میزان «پخش‌شدگی» و «پهنای» آنها نیز از مشخصه‌های مهم توزیع قلمداد می‌شود. این دسته از مشخصه‌های توزیع را مشخصه‌های پراکندگی می‌نامند.

برای روشن شدن اهمیت مشخصات پراکندگی، معلمی را در نظر بگیرید که ادعا می‌کند میانگین نمرات آزمون ۳۰ دانش‌آموز برابر با ۱۷ است و از این رو، خود را در تدریس فرد موفقی می‌داند. عدد ۱۷ نمره نسبتاً مطلوبی است و به عنوان مشخصه تمرکز نمرات ۳۰ دانش‌آموز کلاس، از گرایش عمومی نمرات (داده‌ها) به عدد ۱۷ حکایت دارد. ولی این پرسش مطرح می‌شود که آیا نمره ۱۷ به معنای آن است که اکثر دانش‌آموزان نمره‌ای نزدیک به آن (نمراتی مثل ۱۶/۵، ۱۸ یا ۱۷/۲۵) داشته‌اند؟ یا «تنوع و گوناگونی» نمراتی که دانش‌آموزان این کلاس در آزمون کسب کرده‌اند تا آن حد زیاد است که افرادی با نمرات ۱۱، ۱۳ یا ۸ نیز در کلاس به چشم می‌خورند؟ در صورتی که نمرات کسب شده در کلاس با یکدیگر اختلاف زیادی داشته باشند، دیگر نمی‌توان آن معلم را با وجود میانگین ۱۷ دانش‌آموزان وی موفق دانست.

مشخصه‌های پراکندگی، اطلاعی درباره «تنوع و گوناگونی» داده‌ها در اختیار می‌گذارند و مکمل اطلاعات مشخصه‌های تمرکز هستند. سه مشخصه دامنه، واریانس و ضریب تغییرات به عنوان مشخصه‌های پراکندگی توزیع در این بخش بررسی می‌شوند.

۱ـ۳ـ۳ـ دامنه

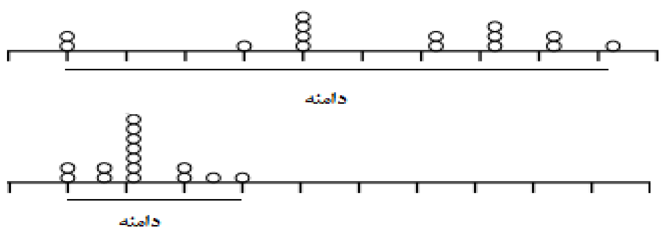
دامنه یکی از منطقی‌ترین و طبیعی‌ترین مشخصه‌های پراکندگی است. هنگامی که مایلیم اطلاعی از پهنا و گسترهٔ توزیع به دست آوریم، ساده‌ترین کار یافتن تفاضل کمترین مقدار از بیشترین مقدار مشاهده شده است. این کمّیت، حداکثر میزان اختلاف بین داده‌ها را نشان می‌دهد. فرض کنید $x_{(1)}$، $x_{(2)}$، ...، $x_{(n)}$ داده‌های مرتب شده باشند. بر این اساس، دامنه عبارت است از:

$$R = x_{(n)} - x_{(1)}$$

بر این اساس، مقدار دامنه برای دو مجموعه داده جدول ۷ به ترتیب برابر با ۱−۸=۷ برای مجموعه اول و ۱−۱۰۴=۱۰۳ برای مجموعه دوم است. همان طور که انتظار می‌رفت، پراکندگی مجموعه داده دوم بزرگ‌تر از مجموعه اول است.

شکل ۲۱ به درک بهتر پراکندگی و دامنه کمک می‌کند. این شکل دو مجموعه دادهٔ ۱۵تایی را بر روی یک محور افقی به تصویر کشیده است. هر دایره کوچک نشان‌دهنده یک داده است (نمودارهای شکل ۲۱ را نمودار نقطه‌ای می‌نامند، زیرا هر داده با یک نقطه یا دایره کوچک نشان داده می‌شود). این شکل به‌خوبی رابطه دامنه با پهنای توزیع (پراکندگی) را نمایان می‌سازد. هر چه داده‌ها در گسترهٔ پهن‌تری توزیع شده باشند، دامنه نیز بزرگ‌تر و حاکی از پراکندگی بیشتر آنهاست.

شکل ۲۱: مقایسه پراکندگی دو مجموعه داده با توجه به دامنه

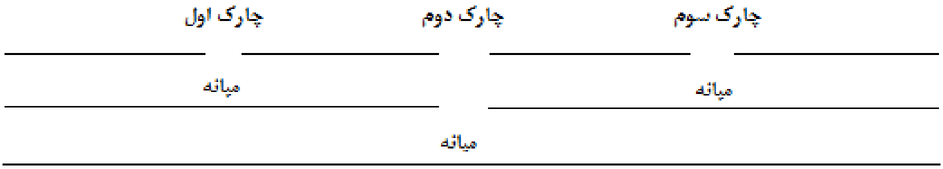


۲ـ۳ـ۳ـ دامنهٔ میان‌چارکی

استفاده از چارک‌ها و به طور کلی چندک‌ها برای اطلاع از پراکندگی داده‌ها گزینه‌ای مناسب و در واقع مکملی برای دامنه است. پیشتر اشاره شد که میانه، داده‌ها را به دو نیمه

با حجم (فراوانی) برابر تقسیم می‌کند. اکنون می‌توان هر نیمه را نیز مطابق شکل ۲۲ به دو نیمه تقسیم کرد تا به این ترتیب، داده‌ها به چهار قسمت با حجم (فراوانی) برابر تقسیم شوند. هر قسمت شامل ۲۵ درصد از تعداد کل داده‌هاست. از همین رو، ۵۰ درصد از داده‌ها بین چارک اول و سوم قرار دارند.

شکل ۲۲: چارک‌ها، داده‌ها را به چهار بخش با حجم (فراوانی) برابر تقسیم می‌کنند

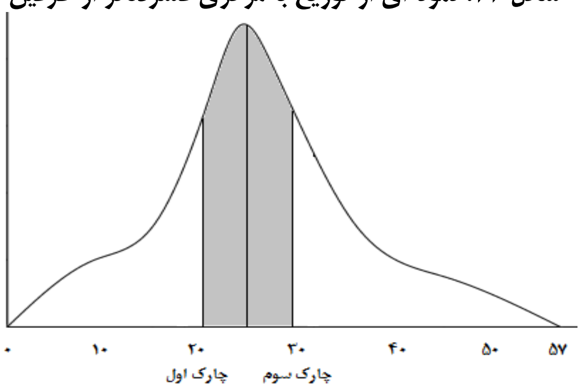


دامنه، پهنایی را نشان می‌دهد که کل داده‌ها در آن جای دارند؛ زیرا از تفاضل کمترین و بیشترین مقدار به دست می‌آید. دامنهٔ میان‌چارکی (IQR) که حاصل تفاضل چارک اول از چارک سوم است، پهنایی را نشان می‌دهد که میان دو چارک است و ۵۰ درصد از داده‌ها در آن قرار دارند.

$$IQR = Q_{75} - Q_{25}$$

این داده‌ها ویژگی خاصی دارند؛ زیرا از یک سو داده‌های پایین چارک اول یعنی داده‌هایی با مقدار کوچک را شامل نمی‌شوند و از سوی دیگر داده‌های بالای چارک سوم یعنی داده‌هایی با مقدار بزرگ را در بر نمی‌گیرند. به عبارت دیگر، دامنه میان‌چارکی نشان می‌دهد که داده‌های معمولی (داده‌های نه چندان بزرگ و نه چندان کوچک) در چه پهنایی از توزیع قرار دارند.

شکل ۲۳: نمونه‌ای از توزیع با مرکزی فشرده‌تر از طرفین



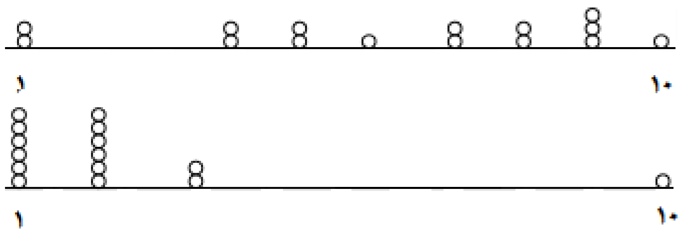
اشاره شد که دامنه چارکی و میان‌چارکی مکمل‌هایی برای دامنه هستند و در کنار هم اطلاع کامل‌تری از چگونگی پراکندگی به دست می‌دهند. برای مثال، مانند شکل ۲۳، اگر دامنه داده‌هایی ۵۷ و دامنه چارکی آنها ۱۰ باشد، به راحتی می‌توان

نتیجه گرفت که ۵۰ درصد از داده‌ها تنها در ۱۰ واحد از فاصله ۵۷ واحدی دامنه جای گرفته‌اند؛ یعنی توزیع در میان خود فشرده‌تر از طرفین خود است.

۳ـ۳ـ۳ـ واریانس

محاسبه دامنه تنها بر اساس کمترین و بیشترین مقدار داده‌ها صورت می‌گیرد و توجهی به پراکندگی داده‌ها در میان این دو مقدار ندارد. از این رو، می‌توان دو مجموعه داده مانند شکل ۲۴ یافت که دامنهٔ برابری دارند، در حالی که داده‌ها به طور متفاوتی پراکنده شده‌اند. واریانس مشخصه‌ای از توزیع است که میزان دور بودن داده‌ها را از نقطه تمرکزشان اندازه می‌گیرد. به عبارت دیگر، پراکندگی از نگاه واریانس ناشی از اختلافی است که داده‌ها از نقطه تعادل خود یعنی میانگین دارند.

شکل ۲۴: دو مجموعه داده با دامنه برابر و پراکندگی متفاوت



فرض کنید داده‌های ۵، ۴، ۸، ۱۰ و ۳ مقادیر متغیری باشند که از ۵ نفر اندازه‌گیری شده‌اند. اگر این ۵ نفر نسبت به این متغیر دارای پراکندگی نبودند، مقدار متغیر برای همگی آنها ثابت و برابر با میانگین آنها بود، در حالی که آنچه مشاهده شده است، ۵ مقدار مختلف است که در داده‌ها پراکندگی ایجاد کرده‌اند. میانگین این داده‌ها به عنوان نقطه تعادل آنها برابر با ۶ است. اختلاف داده اول با میانگین ۱- = ۶ - ۵، اختلاف داده دوم ۲- = ۶ - ۴، اختلاف داده سوم ۲+ = ۶ - ۸، اختلاف داده چهارم ۴+ = ۶ - ۱۰ و سرانجام اختلاف داده پنجم ۳- = ۶ - ۳ است. پس پراکندگی، نتیجهٔ این اختلاف‌ها یعنی ۱-، ۲-، ۲+، ۴+ و ۳- است. دورترین داده از میانگین، داده چهارم با اختلاف ۴+ و نزدیک‌ترین داده به میانگین، داده اول با اختلاف ۱- است. توجه کنید که علامت مثبت نشان می‌دهد داده از میانگین بزرگ‌تر و علامت منفی نشان می‌دهد داده از میانگین کوچک‌تر است.

این اختلاف‌ها می‌توانند مبنای مناسبی برای اندازه‌گیری پراکندگی باشند. نخستین گزینه‌ای که به نظر می‌رسد، محاسبه مجموع این اختلاف‌هاست. جمع این اختلاف‌ها برابر

با ۰ است؛ زیرا اختلاف‌های مثبت و منفی به شکل زیر همدیگر را خنثی می‌کنند و «همواره» حاصل جمع را برابر با ۰ می‌سازند:

۰ = (۱−) + (۲−) + (۲+) + (۴+) + (۳−) = مجموع اختلاف داده‌ها از میانگین

پس نمی‌توان از این شیوه برای اندازه‌گیری پراکندگی استفاده کرد. از این رو، مربع (توان دوم) اختلاف‌ها به صورت زیر با هم جمع می‌شوند که به آن اندازه تغییرپذیری داده‌ها گفته می‌شود:

۳۴ = ^{۲}(۱−) + ^{۲}(۲−) + ^{۲}(۲+) + ^{۲}(۴+) + ^{۲}(۳−) = تغییرپذیری یا مجموع مربعات اختلاف داده‌ها از میانگین

عدد ۳۴ مقدار کل مربع اختلاف همه داده‌ها از میانگین را نشان می‌دهد، در حالی که می‌خواهیم بدانیم به طور «متوسط» داده‌ها چقدر با میانگین خود اختلاف دارند، بنابراین، میانگین مربع اختلاف‌ها را محاسبه می‌کنیم؛ یعنی:

$$\text{واریانس} = \text{میانگین مجموع مربعات اختلاف‌ها} = \frac{۳۴}{۵} = ۶/۸$$

عدد ۶/۸ بیان می‌کند که به طور متوسط، داده‌ها مربع اختلافی برابر با ۶/۸ از میانگین خود دارند.

بر اساس محاسبات فوق، واریانسِ n دادهٔ $x_۱$، $x_۲$، ...، x_n باید از رابطه زیر محاسبه شود:

$$\frac{(x_۱-\bar{x})^۲+(x_۲-\bar{x})^۲+\cdots+(x_n-\bar{x})^۲}{n}=\frac{۱}{n}\sum_{i=۱}^{n}(x_i-\bar{x})^۲$$

ولی در عمل برای محاسبه واریانسِ n دادهٔ $x_۱$، $x_۲$، ...، x_n، مجموع مربعات اختلاف‌ها به جای n، بر $n-۱$ تقسیم می‌شود. بنابراین، واریانس داده‌ها که با $S^۲$ نمایش داده می‌شود، از رابطه زیر به دست می‌آید. بر این اساس، واریانس ۵ دادهٔ فوق برابر با $\frac{۳۴}{۴}=۸/۵$ است[۱].

$$S^۲=\frac{(x_۱-\bar{x})^۲+(x_۲-\bar{x})^۲+\cdots+(x_n-\bar{x})^۲}{n-۱}=\frac{۱}{n-۱}\sum_{i=۱}^{n}(x_i-\bar{x})^۲$$

مربع کردن اختلاف‌ها باعث شد علامت‌های مثبت و منفی اختلاف‌ها به خنثی شدن آنها با یکدیگر منجر نشود، ولی از سوی دیگر اختلاف‌ها با به توان ۲ رسیدن به شکل مبالغه‌آمیزی بزرگ جلوه می‌کنند. بنابراین، از جذر واریانس که «انحراف معیار» نامیده می‌شود نیز برای بیان مشخصهٔ پراکندگی استفاده می‌شود.

$$S=\sqrt{\tfrac{۱}{n-۱}\sum_{i=۱}^{n}(X_i-\bar{X})^۲}$$

---

۱. در آمار توصیفی، همواره از n در مخرج کسرِ واریانس استفاده می‌شود، ولی در فصل برآوردیابی خواهیم دید که تحت چه شرایطی باید از n-۱ استفاده کرد.

بر اساس این رابطه، انحراف معیار ۵ دادهٔ فوق با گرفتن جذر عدد ۸/۵ برابر با ۲/۹۱ است. مزیت دیگری که انحراف معیار در مقایسه با واریانس دارد، یکسانی واحد اندازه‌گیری آن با واحد اندازه‌گیری داده‌هاست. برای مثال، اگر ۵ دادهٔ فوق مربوط به مدت زمانی بر حسب «ثانیه» باشند که فرد به محرک عصبی خاصی پاسخ می‌دهد، واحد اندازه‌گیریِ واریانس «ثانیه به توان ۲» است؛ زیرا اختلاف‌ها به توان ۲ رسیده‌اند، در حالی که واحد اندازه‌گیریِ انحراف معیار همان ثانیه است؛ یعنی انحراف معیار این ۵ داده ۲/۹۱ ثانیه است.

هرچه مقدار واریانس به صفر نزدیک‌تر باشد داده‌ها در اطراف میانگین خود متمرکزتر هستند زیرا کوچکی واریانس نشان‌دهنده اختلاف کم داده‌ها با میانگین است.

واریانس با به توان ۲ رساندن اختلاف‌ها باعث می‌شود اختلاف‌های بزرگ، بزرگ‌تر و اختلاف‌های کوچک، کوچک‌تر جلوه کنند. برای مثال، اگر اختلاف داده‌ای از میانگین برابر با ۲ باشد، مربع آن بزرگ‌تر یعنی برابر با ۴ خواهد شد، در حالی که اختلاف کوچکی مانند ۰/۰۲ با به توان ۲ رسیدن از این هم کوچک‌تر، یعنی برابر با ۰/۰۰۰۴ می‌شود. از این رو، میانگینِ قدرمطلقِ اختلاف‌ها[1] که به صورت زیر محاسبه می‌شود نیز به عنوان مشخصه پراکندگی پیشنهاد شده است.

$$\text{MAD}=\frac{1}{n}\sum_{i=1}^{n}\left|x_i-\bar{x}\right|$$

در واقع، میانگین قدرمطلق اختلاف‌ها، به جای مربع اختلاف‌ها، از قدرمطلق آنها برای حل مشکل خنثی شدن علامت‌های مثبت و منفی کمک می‌گیرد. با وجود اینکه MAD در مقدار اختلاف‌ها اغراق نمی‌کند، ولی استفاده از آن مانند واریانس رواج ندارد، زیرا واریانس دارای خواص ریاضی مطلوبی است که MAD فاقد آن است.

۴ـ۳ـ۳ـ ضریب تغییرات

هنگامی که مقایسه پراکندگی توزیع دو متغیر مطرح است، واحد اندازه‌گیری متغیرها می‌تواند مسئله‌ساز باشد. برای مثال، فرض کنید ۸، ۵، ۶، ۶ و ۹ مدت زمانی بر حسب ساعت باشد که ۵ نفر در شبانه روز می‌خوابند. میانگین و انحراف معیار این ۵ داده به ترتیب برابر با ۶/۸ و ۱/۶۴ «ساعت» است. اکنون فرض کنید این ۵ داده به جای ساعت بر

1. Mean of Absolute Deviations (MAD)

حسب دقیقه ثبت شده باشند، یعنی ۴۸۰، ۳۰۰، ۳۶۰، ۳۶۰ و ۵۴۰. در این صورت، میانگین و انحراف معیار داده‌ها به ترتیب برابر با ۴۰۸ و ۹۸/۴ «دقیقه» است. این تفاوت نشان می‌دهد مشخصه پراکندگی، به‌شدت تحت تأثیر واحد اندازه‌گیری داده‌هاست. از این رو، اگر پژوهشگری بخواهد پراکندگی مدت زمان استراحت دو گروه را با هم مقایسه کند، در حالی که داده‌های یک گروه بر حسب دقیقه و داده‌های گروه دیگر بر حسب ساعت ثبت شده باشند، پیشاپیش می‌توان گفت پراکندگی داده‌هایی که بر حسب دقیقه هستند، بیش از داده‌هایی است که بر حسب ساعت گردآوری شده‌اند، حتی اگر طول زمان استراحت هر دو گروه، درست مانند هم باشد! پس پیش از مقایسهٔ پراکندگی توزیع دو متغیر، باید چاره‌ای برای این مشکل اندیشید.

دو راهکار برای حل این مشکل وجود دارد. اگر واحد اندازه‌گیری یک متغیر قابل تبدیل به واحد اندازه‌گیری دیگری باشد، ابتدا آنها را هم‌واحد و سپس پراکندگی را برای داده‌های تغییر یافته محاسبه می‌کنیم. برای مثال، واحد اندازه‌گیری داده‌هایی را که بر حسب ساعت هستند، با ضرب کردن تک‌تک آنها در عدد ۶۰ به دقیقه تبدیل می‌کنیم تا هر دو مجموعه داده بر حسب دقیقه باشند. اکنون می‌توان انحراف معیار را برای آنها محاسبه و با هم مقایسه کرد. این راهکار هنگامی مناسب است که دو متغیر دارای ماهیت یکسانی باشند تا واحد اندازه‌گیری آنها قابل تبدیل به یکدیگر باشد.

در شرایطی که مقایسهٔ پراکندگیِ دو متغیر کاملاً متفاوت مد نظر است، تبدیل واحد اندازه‌گیری نیز امکان‌پذیر نیست. برای مثال، مقایسه پراکندگی سن افراد با قد آنها با تبدیل واحد اندازه‌گیری ممکن نیست؛ زیرا سن بر حسب سال و قد بر حسب متر اندازه‌گیری می‌شود و این دو واحد را نمی‌توان به یکدیگر تبدیل کرد. در این صورت باید از نوعی مشخصهٔ پراکندگی استفاده کرد که به واحد اندازه‌گیری وابسته نباشد. ضریب تغییرات چنین مشخصه‌ای است.

ضریب تغییرات از تقسیم انحراف معیار متغیر بر میانگین آن به صورت زیر به دست می‌آید:

$$CV = \frac{S}{\bar{X}}$$

از آنجا که واحد اندازه‌گیری انحراف معیار و میانگین، همان واحد اندازه‌گیری متغیر است، تقسیم آنها باعث خنثی شدن واحد صورت با واحد مخرج می‌شود و به این ترتیب، کمّیت حاصل (ضریب تغییرات) بدون واحد اندازه‌گیری و خنثی است. برای نمونه، ۵ داده مربوط

به مدت زمان خواب در شبانه روز را به یاد بیاورید، محاسبه ضریب تغییرات آنها بر حسب دقیقه و ساعت به صورت زیر است:

$$\text{ضریب تغییرات} = \frac{۹۸/۴}{۴۰۸} = ۰/۲۴ \quad \text{اندازه‌گیری بر حسب دقیقه}$$
$$\text{ضریب تغییرات} = \frac{۱/۶۴}{۶/۸} = ۰/۲۴ \quad \text{اندازه‌گیری بر حسب ساعت}$$

همان طور که دیده می‌شود، با وجود تفاوت انحراف معیارها، ضریب تغییرات داده‌ها برابر است؛ یعنی پراکندگی یکسانی در دو مجموعه داده وجود دارد و دلیل تفاوت انحراف معیارها در اختلاف واحدهای اندازه‌گیری آنها نهفته است.

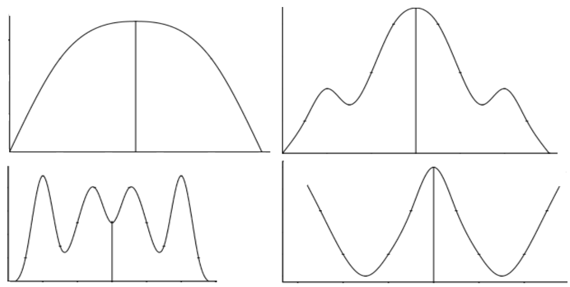
ضریب تغییرات، یک ضعف اساسی نیز دارد؛ اگر هم انحراف معیار و هم میانگین به صفر نزدیک و انحراف معیار بیش از میانگین باشد، مقدار ضریب تغییرات می‌تواند بسیار بزرگ باشد. برای نمونه، اگر انحراف معیار برابر با ۰/۰۲ و میانگین برابر با ۰/۰۰۱ باشد، ضریب تغییرات برابر با ۲۰ است.

۴ـ۳ـ مشخصه‌های شکل توزیع

از نظر شکل، دو ویژگی در منحنی توزیع متغیر مهم قلمداد می‌شود: چولگی و برجستگی. چولگی یا کَجی، اطلاعی درباره متقارن یا نامتقارن بودن شکل توزیع در اختیار می‌گذارد. منحنی توزیعِ متقارن را می‌توان با خطی عمودی موسوم به محور تقارن طوری به دو نیمه تقسیم کرد که یکی، تصویر آینه‌ای دیگری باشد (اگر منحنی بر روی کاغذی رسم شده باشد و این کاغذ را از محل محور تقارن تا کنید، دو نیمه بر هم منطبق می‌شوند). شکل ۲۵، چند نمونه از منحنی‌های توزیع متقارن را با محورهای تقارنشان نشان می‌دهد.

ویژگی اصلی منحنی متقارن، یکسانی توزیع مقادیر متغیر در دو سوی منحنی است (دو سوی منحنی توزیع را دُم‌های توزیع می‌نامند). به عبارت دیگر، مقادیر بزرگ (سمت راست منحنی) و مقادیر کوچک (سمت چپ منحنی) به یک میزان هستند. تقارن باعث می‌شود مقدار میانگین و میانه توزیع همیشه برابر و همان نقطهٔ قرارگرفتن محور تقارن باشند. در برخی از منحنی‌های توزیع، نما نیز با میانگین و میانه برابر است. در شکل ۲۵، میانگین، میانه و نما در سه منحنی توزیع برابر است (تنها استثنا، منحنی پایین سمت چپ است که میانگین و میانهٔ آن برابرند، ولی نما جای دیگری است).

شکل ۲۵: نمونه‌هایی از منحنی توزیع متقارن

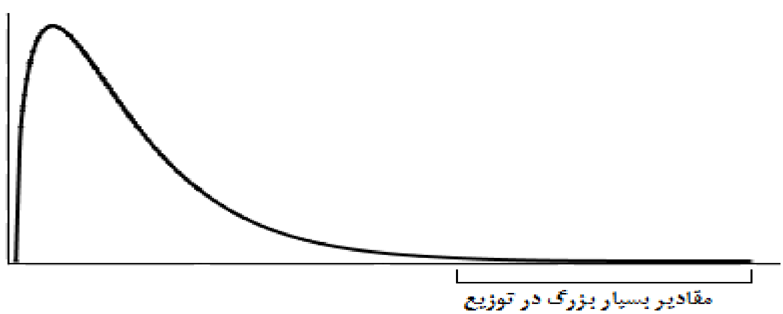


منحنی توزیع‌های چوله نامتقارن است، به طوری که یکی از دُم‌های منحنی توزیع، به سمت راست یا چپ تمایل بیشتری از دُم دیگر دارد (شکل ۲۶).

شکل ۲۶: منحنی توزیع چوله به راست و چوله به چپ

چولگی در مقابل توازنی قرار دارد که منحنی متقارن دارای آن است. چولگی باعث می‌شود نسبت مقادیر بزرگ و مقادیر کوچک در توزیع برابر نباشد و بر حسب اینکه چولگی به چه سمت و تا چه میزانی باشد، مقادیر بسیار بزرگ یا بسیار کوچکی در توزیع وجود خواهد داشت که متفاوت با سایر مقادیر توزیع متغیر هستند. شکل ۲۷ این ویژگی را برای توزیع کشیده به راست نمایش می‌دهد.

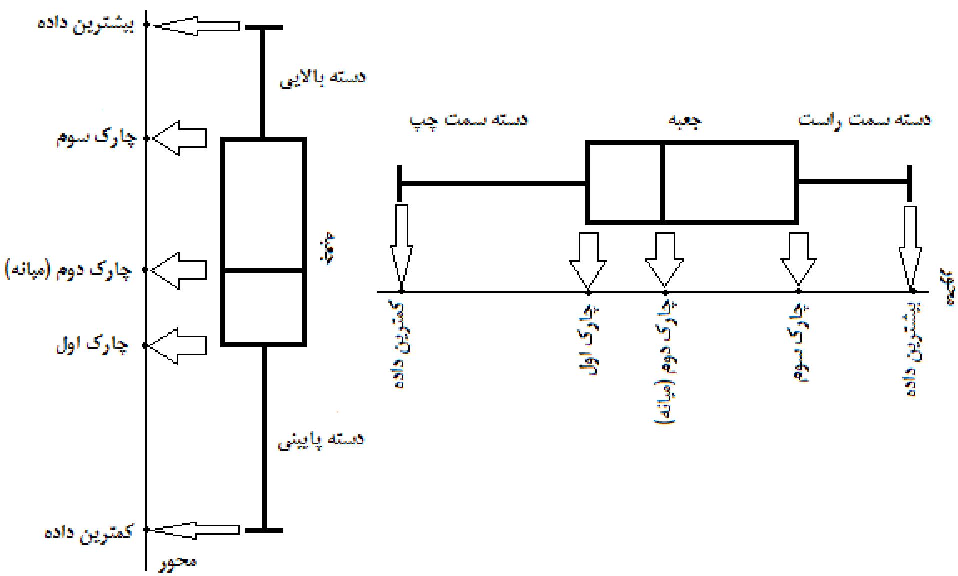
شکل ۲۷: وجود مقادیر بسیار بزرگ در منحنی توزیعی با تمایل (چولگی) زیاد به راست



دقیق‌ترین شیوه برای تشخیص متقارن یا نامتقارن بودن منحنی توزیع، ترسیم این منحنی است. با وجود این می‌توان با داشتن برخی از مشخصات توزیع نظیر چارک‌ها و رسم نمودار جعبه‌ای، اطلاعاتی تقریبی درباره تقارن توزیع به دست آورد.

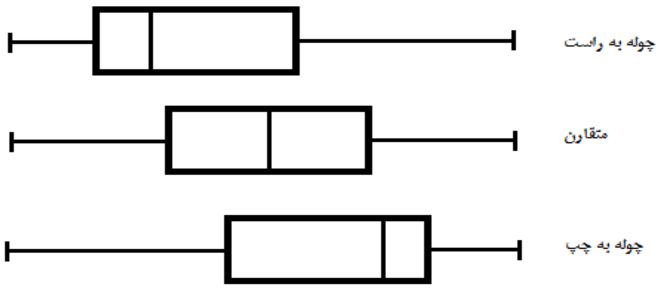
نمودار جعبه‌ای از یک مستطیل به نام جعبه و دو دسته تشکیل شده است که به دو طرف این مستطیل متصل هستند. کنار این نمودار محور مختصاتی وجود دارد که به درک کمّی نمودار کمک می‌کند. این نمودار در دو حالت افقی و عمودی در شکل ۲۸ نمایش داده شده است. همان‌طور که دیده می‌شود، رسم این نمودار با یافتن محل پنج نقطه بر روی محور مختصات امکان‌پذیر است که محل اضلاع مستطیل و میزان امتداد دسته‌ها را مشخص می‌کنند.

شکل ۲۸: نمودار جعبه‌ای در دو حالت افقی و عمودی



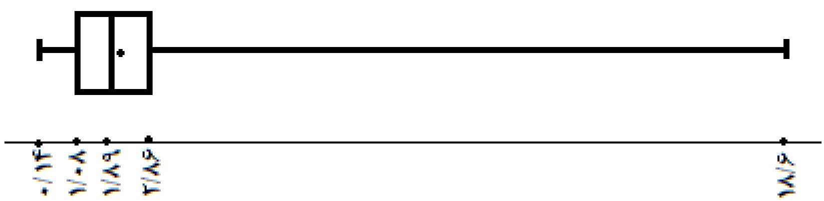
این نمودار به کمک این پنج مشخصه توزیع می‌تواند درباره شکل و میزان پراکندگی توزیع، اطلاعات قابل ملاحظه‌ای به دست دهد. منحنی متقارن دارای نمودار جعبه‌ای متعادلی مانند شکل ۲۹ است؛ به این معنی که خط مربوط به میانه در وسط مستطیل (جعبه) قرار می‌گیرد و دسته‌ها نیز به یک اندازه در دو سوی آن امتداد می‌یابند. در توزیع چوله به راست، خط مربوط به میانه به ضلع سمت چپ نزدیک‌تر و دسته سمت راست طولانی‌تر از دسته سمت چپ است. عکس این حالت برای توزیع چوله به چپ وجود دارد.

شکل ۲۹: تشخیص شکل توزیع به کمک نمودار جعبه‌ای



برای مثال، کمترین و بیشترین مقدار توزیع شکل ۲۷ به ترتیب ۰/۱۴ و ۱۸/۶ و چارک‌های آن نیز به ترتیب ۱/۰۸، ۱/۸۹ و ۲/۸۶ هستند، بنابراین نمودار جعبه‌ای آن مطابق شکل ۳۰ است. دسته طولانی سمت راست بر چولگی زیاد منحنی این توزیع دلالت دارد، اگر چه خط مربوط به میانه (۱/۸۹) تنها کمی به سمت چپ گرایش دارد. گاهی در نمودار جعبه‌ای مانند شکل ۳۰، محل قرار گرفتن میانگین نیز نشان داده می‌شود تا برای بررسی تقارن، انطباق آن با میانه بررسی شود. میانگین توزیع شکل ۳۰ که با نقطه نشان داده شده است، برابر با ۲/۱ و متفاوت با مقدار میانه است.

شکل ۳۰: نمودار جعبه‌ای برای منحنی توزیع شکل ۲۷ (چولگی به سمت راست)



تاکنون برای تشخیص چولگی از نمودار جعبه‌ای یعنی روشی تصویری استفاده کرده‌ایم و اکنون می‌خواهیم از مشخصهٔ توزیعی به نام ضریب چولگی چارکی برای این منظور استفاده کنیم. با توجه به شکل ۲۹، رابطه زیر برای توزیع متقارن برقرار است:

$$Q_{۷۵} - Q_{۵۰} = Q_{۵۰} - Q_{۲۵}$$

زیرا در توزیع متقارن، میانه (چارک دوم) در وسط مستطیل قرار دارد و آن را به دو نیمه برابر تقسیم می‌کند؛ یعنی فاصله میانه تا چارک اول برابر با فاصله میانه تا چارک سوم

است. طرفین این تساوی را بر دامنه میان‌چارکی تقسیم می‌کنیم تا به واحد اندازه‌گیری داده‌ها بستگی نداشته باشد:

$$\frac{Q_{۷۵}-Q_{۵۰}}{Q_{۷۵}-Q_{۲۵}}=\frac{Q_{۵۰}-Q_{۲۵}}{Q_{۷۵}-Q_{۲۵}}$$

ضریب چولگی چارکی که با CS_Q نشان داده می‌شود، از تفاضل دو طرف تساوی فوق به صورت زیر به دست می‌آید. از آنجا که این تساوی در توزیع متقارن همواره برقرار است، مقدار ضریب چولگی چارکی برای توزیع‌های متقارن برابر با ۰ است.

$$CS_Q=\frac{Q_{۷۵}-Q_{۵۰}}{Q_{۷۵}-Q_{۲۵}}-\frac{Q_{۵۰}-Q_{۲۵}}{Q_{۷۵}-Q_{۲۵}}=\frac{Q_{۷۵}-۲Q_{۵۰}+Q_{۲۵}}{Q_{۷۵}-Q_{۲۵}}$$

اگر توزیع، چوله به راست باشد، انتظار می‌رود فاصله چارک اول از میانه کوتاه‌تر از فاصله میانه از چارک سوم باشد، بنابراین مقدار ضریب چولگی چارکی مثبت خواهد بود، در حالی که اگر توزیع چوله به چپ باشد، انتظار می‌رود فاصله چارک اول از میانه بلندتر از فاصله میانه از چارک سوم باشد و به این ترتیب، مقدار ضریب چولگی چارکی منفی خواهد بود.

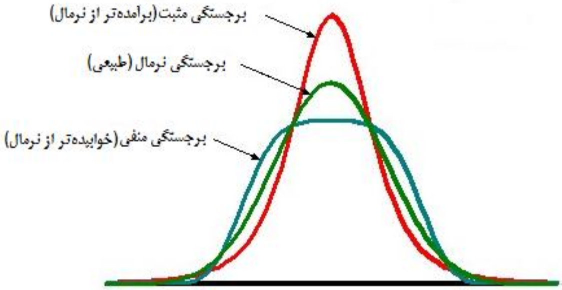
ضریب چولگی چارکی نمی‌تواند سنجش کاملی از چولگی ارائه دهد، زیرا ممکن است فاصله چارک‌ها در توزیعی برابر باشد، یعنی میانه در وسط مستطیل نمودار جعبه‌ای قرار داشته باشد، ولی دم‌های توزیع هم‌اندازه نباشند و ضخامت متفاوتی داشته باشند. در این صورت ضریب چولگی چارکی برابر با ۰ است، در حالی که منحنی توزیع متقارن نیست، بنابراین ضریب چولگی گشتاوری که در زیر تعریف شده است، برای بررسی انحراف از تقارن، مناسب‌تر از ضریب چولگی چارکی است (مانند واریانس، در آمار توصیفی از n به جای n−۱ استفاده می‌شود).

$$g_{۱}=\frac{۱}{n-۱}\sum_{i=۱}^{n}\left[\frac{x_i-\bar{x}}{S}\right]^{۳}$$

برجستگی، ویژگی دیگر شکل منحنی توزیع است که به میزان برآمدگی یا خوابیدگی منحنی توزیع اشاره دارد. همان طور که معیار چولگی، انحراف از حالت تقارن است، برای برجستگی نیز به معیاری نیاز است. به عبارت دیگر، چه حدی از برجستگی طبیعی است تا میزانی بیش از آن را برآمده و میزانی کمتر از آن را خوابیده قلمداد کنیم؟ این معیار در آمار، منحنی توزیع نرمال است که در شکل ۳۱ مشاهده می‌شود. این شکل نشان می‌دهد که قلهٔ

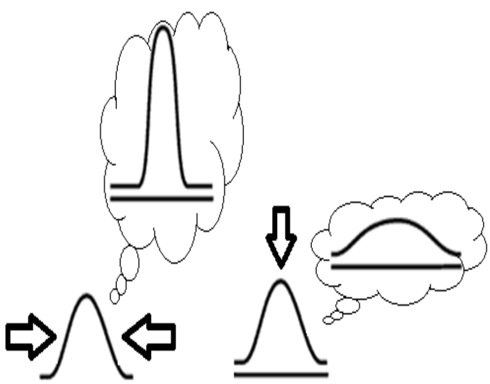
توزیع در توزیعی با برجستگی مثبت بلندتر از توزیع نرمال است، درحالی که قلهٔ توزیعی با برجستگی منفی پایین‌تر از قلهٔ توزیع نرمال قرار دارد.

شکل ۳۱: مقایسه برجستگی منحنی توزیع با برجستگی طبیعی



برجستگی از آن رو اهمیت پیدا می‌کند که اطلاعی درباره پراکندگی توزیع، بویژه توزیع‌های متقارن به دست می‌دهد. منحنی متقارنی که برآمدگی‌ای بیش از برآمدگی نرمال دارد، داده‌هایی را با فراوانی زیاد در زیر قله خود یعنی اطراف میانگین جا می‌دهد، پس انتظار داریم دُم‌های توزیع نازک‌تر از منحنی نرمال باشند. بنابراین، پراکندگی موجود در این توزیع بیش از همه ناشی از داده‌های موجود در دُم‌های آن است. به عبارت دیگر، انتقال از توزیع نرمال به توزیعی برآمده‌تر از آن، به معنی فشردن طرفین قلهٔ توزیع نرمال است؛ به طوری که برآمدگی آن بیشتر و داده‌ها در زیر قله متمرکزتر شوند. در مقابل، انتقال از توزیع نرمال به توزیعی خوابیده‌تر از آن به معنی فشارآوردن بر روی قله آن است، به طوری که پهن‌تر و خوابیده‌تر گردد و از تمرکز داده‌ها کاسته شود. این موضوع در شکل ۳۲ به تصویر کشیده شده است.

شکل ۳۲: تغییر منحنی نرمال به منحنی برآمده‌تر یا خوابیده‌تر

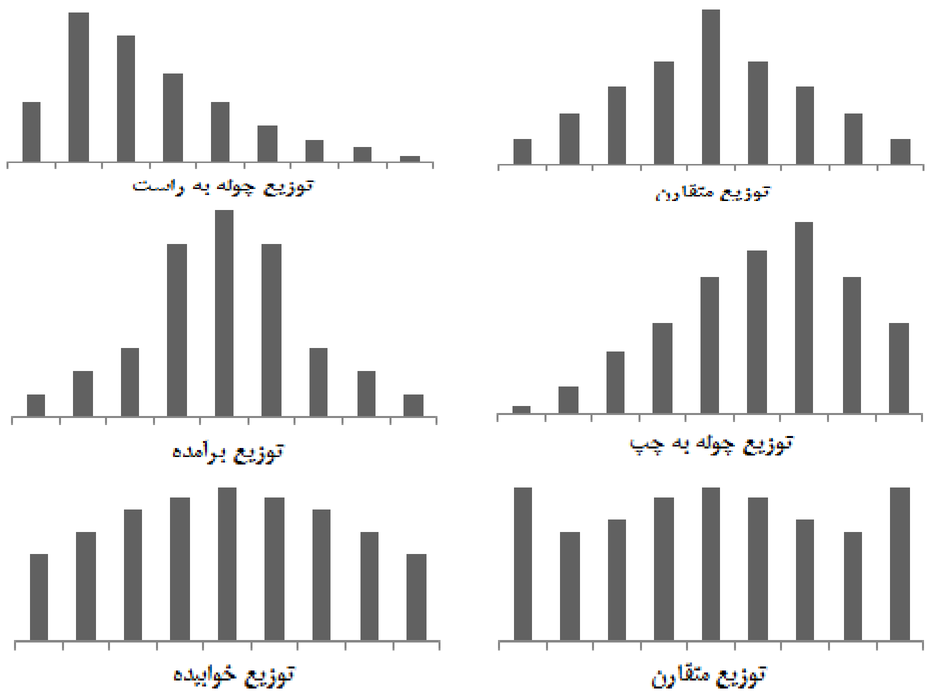


ضریب برجستگی که به صورت زیر تعریف می‌شود، مشخصه‌ای برای توزیع است کـه میـزان برجستگی آن را نسبت به توزیع نرمال می‌سنجد. مقداری برابر با صفر نشان می‌دهد که برجسـتگی توزیع مانند توزیع نرمال است. مقدار مثبت حاکی از آن است کـه توزیـع برآمـده‌تـر از برجسـتگی توزیع نرمال و مقدار منفی بیانگر آن است که توزیع خوابیده‌تر از برجسـتگی توزیـع نرمـال اسـت (مانند واریانس، در آمار توصیفی از n به جای ۱−n استفاده می‌شود).

$$g_2 = \frac{1}{n-1}\sum_{i=1}^{n}\left[\frac{x_i - \bar{x}}{S}\right]^4 - 3$$

اگرچه بحث چولگی و برجستگی بر اساس منحنی توزیع، یعنی متغیرهای پیوسته بیان شد، ولی می‌توان آن را برای متغیرهای گسستهٔ فاصله‌ای یا نسبتی نیز ماننـد شـکل ۳۳ مطـرح کـرد. ضـریب چولگی و کشیدگی نیز برای چنین متغیرهایی از طریق همان رابطه‌های پیشین محاسبه می‌شود.

شکل ۳۳: نمودارهای میله‌ای مختلف برای متغیر گسسته



۴ـ محاسبه مشخصه‌های توزیع برای داده‌های دسته‌بندی‌شده

رابطه‌هایی که تاکنون برای محاسبه مشخصه‌های توزیع ارائه شده‌اند، به داده‌های خُرد اختصاص دارند. گاهی آنچه از داده‌ها در دست است، جدول توزیع فراوانی برای داده‌های دسته‌بندی‌شده است. به عبارت دیگر، تنها اطلاعات موجود از داده‌ها نقاط ابتدایی و انتهایی هر دسته و فراوانی‌های متناظر با دسته‌هاست. به این ترتیب، نمی‌توان مشخصه‌های توزیع را محاسبه کرد، زیرا داده‌ای در اختیار ندارید تا آنها را با هم جمع و بر تعدادشان تقسیم کنید و میانگین آنها را به دست آورید. در چنین حالتی برای هر دسته نماینده‌ای تعیین می‌شود و مشخصه‌های توزیع بر اساس رابطه‌های دیگری محاسبه می‌شوند.

فرض کنید داده‌های مربوط به متغیر X به k دسته مطابق جدول ۸ طبقه‌بندی شده‌اند. در این دسته‌بندی L_i و U_i به ترتیب نقطهٔ ابتدایی و نقطهٔ انتهایی دستهٔ i اُم را نشان می‌دهد. برای مثال بر اساس اطلاعات جدول ۵، در دستهٔ اول یعنی به ازای $i=۱$، نقاط ابتدایی و انتهایی دسته، $L_۱=۱۰$ و $U_۱=۲۰$ یا در دستهٔ سوم یعنی به ازای $i=۳$، نقاط ابتدایی و انتهایی دسته، $L_۳=۳۰$ و $U_۳=۵۰$ هستند.

جدول ۸: اطلاعات داده‌های دسته‌بندی‌شده برای محاسبه مشخصه‌های توزیع

دسته‌ها	فراوانی	نمایندهٔ دسته
$L_۱$ تا $U_۱$	$f_۱$	$x_۱$
$L_۲$ تا $U_۲$	$f_۲$	$x_۲$
⋮	⋮	⋮
L_k تا U_k	f_k	x_k
جمع	$n=\sum_{i=۱}^{k} f_i$	

بر این اساس، نماینده دسته i اُم که معمولاً نقطهٔ میانی آن دسته است، به صورت زیر به دست می‌آید:

$$x_i = L_i + \frac{U_i - L_i}{۲}$$

برای مثال، نقطهٔ میانی دسته اول در جدول ۵ برابر با $۱۰+\frac{۲۰-۱۰}{۲}=۱۵$ است. اکنون می‌توان مشخصه‌های توزیع را با توجه به رابطه‌های جدول ۹ محاسبه کرد.

جدول ۹: رابطه‌های محاسبه مشخصه‌های توزیع برای داده‌های دسته‌بندی شده

$\bar{x}=\frac{1}{n}\sum_{i=1}^{k}f_i x_i$	میانگین حسابی
$s^2=\frac{1}{n-1}\sum_{i=1}^{k}f_i(x_i-\bar{x})^2$	واریانس
$\bar{x}_h=\dfrac{n}{\sum_{i=1}^{k}\frac{1}{f_i x_i}}$	میانگین سازوار
$\bar{x}_g=(\prod_{i=1}^{k}f_i x_i)^{\frac{1}{n}}$	میانگین هندسی
$g_1=\frac{1}{n-1}\sum_{i=1}^{k}f_i\left[\frac{x_i-\bar{x}}{s}\right]^3$	ضریب چولگی
$g_2=\frac{1}{n-1}\sum_{i=1}^{k}f_i\left[\frac{x_i-\bar{x}}{s}\right]^4-3$	ضریب کشیدگی

مشخصه دیگری که ممکن است از اطلاعات جدول توزیع فراوانی برای داده‌های دسته‌بندی‌شده محاسبه شود چندک p اُم داده‌هاست. برای این منظور، ابتدا «اولین» دسته‌ای را پیدا کنید که فراوانی تجمعی آن کمتر از $n\times\frac{p}{100}$ نباشد. این همان دسته‌ای است که چندک p اُم در آن قرار دارد. فرض کنید دسته m اُم شامل چندک p اُم است، نقطهٔ ابتدایی و انتهایی این دسته را به ترتیب با L_m و U_m و فراوانی آن را با f_m نشان می‌دهیم. در این صورت، چندک p اُم از رابطهٔ زیر به دست می‌آید:

$$Q_p=L_m+\frac{\frac{p}{100}\times n+\sum_{i=1}^{m-1}f_i}{f_m}\times(U_m-L_m)$$

توجه کنید که در رابطهٔ فوق، $\sum_{i=1}^{m-1}f_i$ برابر با فراوانی تجمعی تا پیش از دستهٔ m اُم است؛ زیرا حد بالای سیگما $m-1$ انتخاب شده است.

نما نیز از مشخصه‌هایی است که برای داده‌های دسته‌بندی‌شده قابل محاسبه است. برای یافتن نما، فرض کنید دسته m اُم دارای بیشترین فراوانی در میان سایر دسته‌ها باشد. نقطهٔ ابتدایی و انتهایی این دسته را به ترتیب با L_m و U_m و فراوانی آن را با f_m نشان

می‌دهیم. همچنین، فراوانی دستهٔ پیش از آن و دستهٔ پس از آن را نیز بـه ترتیـب بـا f_{m-1} و f_{m+1} نمایش می‌دهیم. بر این اساس، نما از رابطهٔ زیر به دست می‌آید:

$$Mode = L_m + \frac{f_m - f_{m-1}}{(f_m - f_{m-1}) + (f_m - f_{m+1})} \times (U_m - L_m)$$

تاکنون بایـد ایـن پرسـش بـرای خواننـده مطـرح شـده باشـد کـه آیـا نتـایج محاسـبه مشخصه‌های توزیع از طریق داده‌های دسته‌بندی‌شده با نتایج محاسبه آنها از طریق داده‌هـای خُرد متفاوت است؟ برای مثال، آیا واریانس حاصل از داده‌های خُرد با واریـانس حاصـل از دسته‌بندی همان داده‌ها برابر است؟ پاسخ این پرسش منفی است، زیرا مشخصـهٔ حاصـل از داده‌های خُرد بر اساس داده‌های «واقعی» محاسبه می‌شود، در حالی که مشخصهٔ حاصـل از داده‌های دسته‌بندی‌شده بر اساس نمایندهٔ دسته‌ها محاسبه می‌گردد.

برای روشن شدن دلیل این تفاوت، ۷ داده زیر را در نظر بگیرید کـه تنهـا بـه دو دسـته تقسیم شده‌اند. دسته اول شامل ۳ دادهٔ ۱، ۷ و ۱۰ و دسـته دوم شـامل ۴ دادهٔ ۱۱، ۱۸، ۱۹ و ۲۰ است. نماینده دستهٔ اول و دستهٔ دوم، بر اساس رابطه‌ای که پیشتر بیـان شـد، بـه ترتیـب برابر با ۵/۵ و ۱۵/۵ است.

دسته	داده‌ها	نمایندهٔ دسته	فراوانی
۱ تا ۱۰	۱، ۷ و ۱۰	۵/۵	۳
۱۱ تا ۲۰	۱۱، ۱۸، ۱۹ و ۲۰	۱۵/۵	۴

میانگین بر اساس داده‌های خُرد و داده‌های دسته‌بندی شده عبارت است از:

میانگین داده‌های خُرد $= \frac{1+7+10+20+19+18+11}{7} = \frac{86}{7} = 12/28$

میانگین داده‌های دسته‌بندی‌شده $= \frac{(5/5\times3+15/5\times4)}{7} = \frac{78/5}{7} = 11/23$

همان طور که در محاسبات دیده می‌شـود، در محاسـبه میـانگین داده‌هـای خُـرد، ۷ داده بـا یکدیگر جمع و بر عدد ۷ تقسیم شده‌اند. میانگین داده‌های دسته‌بندی‌شده نیز بـا بازنویسـی محاسبات فوق به شکل زیر از جمع ۷ داده و تقسیم آن بر عدد ۷ به دست آمده است:

$$\frac{(5/5\times3+15/5\times4)}{7} \quad \text{یا} \quad \frac{5/5+5/5+5/5+15/5+15/5+15/5+15/5}{7}$$

در واقع در محاسبات فوق، نماینده دستهٔ اول جانشین ۳ دادهٔ این دسته و نماینده دستهٔ دوم جانشین ۴ دادهٔ دسته دوم شده است، پس ۳ بار عدد ۵/۵ و ۴ بار عدد ۱۵/۵ در محاسبه جمع به کار رفته است. این بدان معناست که میانگین حاصل از داده‌های دسته‌بندی‌شده تمایزی بین داده‌های موجود در یک دسته قایل نمی‌شود و تمام داده‌های یک دسته را برابر با نماینده آن دسته قلمداد می‌کند. برای مثال، وجود ۳ دادهٔ متمایز ۱، ۷ و ۱۰ را در دسته اول نادیده می‌گیرد و به جای این ۳ عدد، ۳ بار عدد ۵/۵ را به کار می‌برد. این همان منبع بروز تفاوتی است که بین مشخصه‌های حاصل از داده‌های خُرد و مشخصه‌های حاصل از داده‌های دسته‌بندی‌شده به چشم می‌خورد.

اگرچه دسته‌بندی، یکی از روش‌های کارآمد در تلخیص داده‌هاست، ولی خواننده باید بداند مشخصه‌های محاسبه‌شده از دسته‌بندی داده‌ها با خطا همراه است، زیرا با یکسان پنداشتن داده‌های هر دسته، اطلاعات موجود در داده‌های خُرد را از دست می‌دهد. بنابراین، اگر داده‌های خُرد در دسترس پژوهشگر باشد، استفاده از آنها برای محاسبه مشخصه‌های توزیع، بر داده‌های دسته‌بندی‌شده برتری دارد، زیرا داده‌های خُرد واقعی‌تر از داده‌های دسته‌بندی‌شده هستند.

۵ـ داده‌های دورافتاده

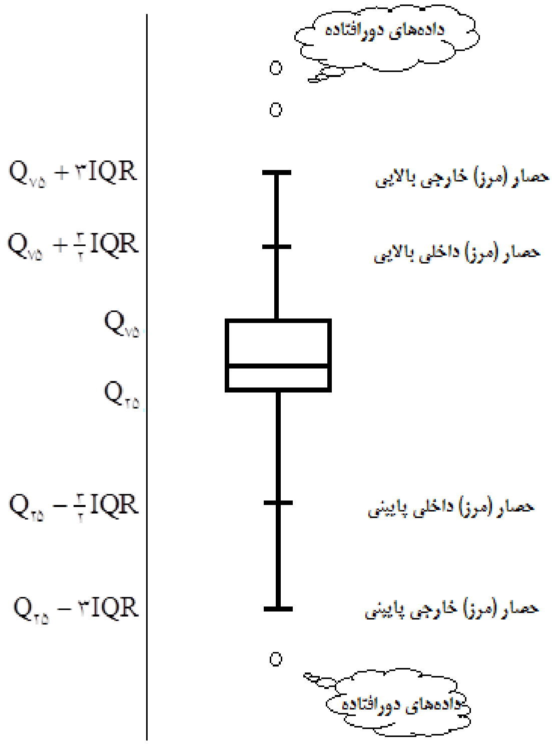
گاهی داده‌هایی که از یک متغیر گردآوری می‌شود، شامل مقادیر نادری است که در مقایسه با سایر داده‌ها بسیار بزرگ یا بسیار کوچک هستند؛ یعنی از عموم داده‌ها فاصله دارند. این داده‌ها با وجود فراوانی ناچیز خود می‌توانند مشخصه‌های توزیع را به شدت تحت تأثیر قرار دهند و تصویر گمراه‌کننده‌ای از توزیع ارائه کنند. برای مثال، ممکن است عموم مردم درآمدی بین ۵۰۰ تا ۸۰۰ هزار تومان داشته باشند، ولی عده کمی از جامعه دارای درآمدی بیش از ۱ میلیون تومان باشند. این عده باعث می‌شوند میانگین درآمد آن جامعه به شکلی کاذب بزرگ جلوه کند و تصویری نادرست از تمرکز توزیع درآمد در جامعه به دست دهد. داده‌هایی از این نوع را داده‌های دورافتاده می‌نامند.

داده‌های دورافتاده می‌توانند در توزیع متغیر، چولگی شدیدی ایجاد کنند و مشخصه‌های مختلف توزیع را تحت تأثیر خود قرار دهند. برای مثال، تنها یک مقدار بسیار بزرگ در میان داده‌ها می‌تواند دُم سمت راست توزیع را به سمت خود بکشد و دامنه توزیع را به شدت بزرگ نشان دهد، در حالی که ممکن است بخش عمده‌ای از داده‌ها به هم نزدیک و

متمرکز باشند. محاسبه مشخصه‌های توزیع با حذف داده‌های فَرین، یعنی داده‌هایی که در دو انتهای دُم توزیع قرار دارند، از راهکارهایی است که در صورت وجود داده‌های دورافتاده پیشنهاد می‌شود. توجه کنید که داده‌های فرین لزوماً دورافتاده نیستند، بلکه می‌توانند داده‌هایی عادی در توزیع باشند، پس باید شیوه‌ای برای تشخیص داده‌های دورافتاده یافت.

همان طور که اشاره شد، شناسایی داده‌های دورافتاده در حین توصیف مجموعه‌ای از داده‌ها برای جلوگیری از ارائه اطلاعات نادرست اهمیت ویژه‌ای دارد. نمودار شِمایی که با اعمال تغییراتی بر نمودار جعبه‌ای به دست می‌آید، وسیله مناسبی برای شناسایی داده‌های دورافتاده است. شکل ۳۴ جزئیات ترسیم نمودار شمایی را نشان می‌دهد. جعبه در نمودار شمایی مانند نمودار جعبه‌ای است، ولی طول دسته‌ها به شکلی متفاوت محاسبه می‌شود (شکل ۲۸ را ببینید).

شکل ۳۴: نمودار شِمایی برای شناسایی داده‌های دورافتاده



داده‌های دورافتاده از مرز خارجی بالایی، بیشتر یا از مرز خارجی پایینی، کمتر هستند. به عبارت دیگر، داده‌های بیشتر از مرز خارجی بالایی در مقایسه با سایر داده‌ها بسیار بزرگ و داده‌های کمتر از مرز خارجی پایینی در مقایسه با سایر داده‌ها بسیار کوچک قلمداد می‌شوند. داده‌هایی که بین مرزهای داخلی و خارجی قرار دارند، داده‌های دورافتادهٔ معتدل نامیده می‌شوند؛ زیرا در مقایسه با داده‌های دورافتاده، کمتر بر مشخصه‌های

توزیع تأثیر می‌گذارنـد. در ادامـه، بـا جزئیـات بیشـتری بـه تـأثیر داده‌هـای دورافتـاده بـر مشخصه‌های توزیع می‌پردازیم.

از میان مشخصه‌های تمرکز، میانگین حسابی از رایج‌ترین و کامـل‌تـرین مشخصـه‌هـایی است که نقطه تمرکز توزیع متغیر را به دست می‌دهد، زیرا تمام داده‌ها در محاسبه آن نقـش دارند. با این همه، این مشخصه در توزیع‌های به شدت چوله کـه معمـولاً شـامل داده‌هـای دورافتاده نیز هستند، چندان مناسب نیست، زیرا مقدار میانگین بـه سـمت دُم چولـه توزیـع گرایش پیدا می‌کند، هر چند فراوانی این داده‌ها ناچیز باشد. برای مثال در منحنی شکل ۲۷، دُم توزیع به سمت راست چوله است، پس میانگین تحـت تـأثیر ایـن مقـادیر بـزرگ قـرار می‌گیرد و به آنها گرایش پیدا می‌کند، در حالی کـه عمـده داده‌هـا در سـمت چـپ توزیـع تمرکز دارند.

در شرایطی که با توزیع‌هایی با چولگی شدید و داده‌های دورافتـاده روبـه‌رو هسـتید، از دو مشخصهٔ میانه و نما برای بیان نقطه تمرکز توزیع کمک بگیرید؛ زیرا بزرگی یا کـوچکی مقادیر، نقشی در محاسبه آنها ندارند و به این ترتیب تحت تـأثیر داده‌هـای دورافتـاده قـرار نمی‌گیرند. استفاده از میانگین پیراسته راهکار دیگری اسـت کـه در زمـان وجـود داده‌هـای دورافتاده پیشنهاد می‌شود. از آنجا که داده‌های دورافتاده در دو دُم توزیع جای دارند، انتظـار می‌رود با حذف ۲/۵ درصـد از کوچـک‌تـرین داده‌هـا (دُم سـمت چـپ) و ۲/۵ درصـد از بزرگ‌ترین داده‌ها (دُم سمت راست)، دیگر داده دورافتاده‌ای باقی نمانـد؛ در واقـع، بـا ایـن حذف، داده‌ها از داده‌های دورافتاده پاک می‌شوند. اکنون می‌توان میانگین داده‌هـای پیراسـته را محاسبه کرد که ۹۵ درصد از کل داده‌ها را تشکیل می‌دهند. میانگین پیراسته، تحت تـأثیر داده‌های دورافتاده قرار ندارد، زیرا آنها در محاسبه آن به کار گرفته نمی‌شوند.

دامنه در میان مشخصه‌های پراکندگی، اطلاع مفیدی از گسترهٔ تغییرات داده‌ها بـه دسـت می‌دهد، ولی به شدت تحت تأثیر داده‌های دورافتاده است، در حالی که دامنهٔ میـان‌چـارکی، با داده‌های دورافتاده تهدید نمی‌شود، زیرا داده‌های کمتر از چارک اول (داده‌هـای کوچـک) و داده‌های بیشتر از چارک سوم (داده‌های بزرگ) نقشی در محاسـبه آن ایفـا نمـی‌کننـد. بـا وجود این، همواره این ضعف برای دامنهٔ میـان‌چـارکی وجـود دارد کـه ۵۰ درصـد داده‌هـا (۲۵درصد پایینی و ۲۵درصد بالایی) را نادیده می‌گیرد. از این رو، دامنه میان‌صدکی[1] نیز بـه

1. Inter-Percentile Range

عنوان مشخصه پراکندگی پیشنهاد شده است. دامنه میان‌صدکی به شکل زیر از تفاضل صدک دهم از صدک نودم به دست می‌آید (یا تفاضل دهک اول و نهم):

$$IPR = Q_{۹۰} - Q_{۱۰}$$

دامنه میان‌صدکی، ۲۰ درصد از داده‌ها را نادیده می‌گیرد که باعث می‌شود مانند دامنه میان‌چارکی تحت تأثیر داده‌های دورافتاده قرار نگیرد.

واریانس و بویژه انحراف معیار، مشخصه‌های کاملی برای سنجش پراکندگی داده‌ها هستند؛ زیرا نخست آنکه دوری داده‌ها را از نقطه تمرکز (میانگین) می‌سنجند و دوم آنکه تک‌تک داده‌ها در محاسبه آنها به کار می‌روند. ولی همین امر باعث می‌شود واریانس و انحراف معیار نیز نسبت به داده‌های دورافتاده حساس باشند. ضریب تغییرات به عنوان مشخصه دیگر پراکندگی توزیع، کمتر از انحراف معیار تحت تأثیر داده‌های دورافتاده، بویژه داده‌های دورافتادهٔ بزرگ قرار دارد؛ زیرا اگر داده‌های دورافتاده‌ای در میان داده‌ها وجود داشته باشند و مثلاً باعث بزرگ شدن میانگین و انحراف معیار شوند، تقسیم انحراف معیار بر میانگین برای محاسبه ضریب تغییرات می‌تواند به کاستن تأثیر این داده‌های دورافتاده منجر شود.

۶ـ جمع‌بندی

در این فصل، به ابزارهای توصیف داده‌ها پرداخته شد که با خلاصه‌سازی و سازماندهی داده‌های مربوط به یک متغیر، گزارشی از توزیع آن متغیر ارائه می‌کنند. این ابزارها در دو دسته جای می‌گیرند:

- ابزارهای تصویری ارائه توزیع متغیر
  - ✓ جدول توزیع فراوانی (اگر داده‌ها به متغیر مقوله‌ای مربوط نباشند، جدول توزیع فراوانی پس از دسته‌بندی به کار می‌رود)
  - ✓ نمودارهای مختلف نظیر نمودار دایره‌ای، نمودار میله‌ای، بافت‌نگار، منحنی توزیع و نمودار جعبه‌ای
- ابزارهای کمّی برای ارائه مشخصه‌های مختلف توزیع
  - ✓ مشخصه‌های تمرکز مانند میانگین حسابی، نما و چندک‌ها (میانه و چارک‌ها)
  - ✓ مشخصه‌های پراکندگی مانند دامنه، دامنه میان‌چارکی، دامنه میان‌صدکی، واریانس (انحراف معیار)، میانگین قدرمطلق اختلاف‌ها و ضریب تغییرات
  - ✓ مشخصه‌های شکل توزیع شامل ضریب چولگی و ضریب کشیدگی

داشتن توزیع متغیر، چه در قالب جدول توزیع فراوانی و چه در قالب نمودارها می‌تواند «اطلاعات کاملی» از متغیر در اختیار کاربر قرار دهد؛ زیرا تمام مشخصه‌های توزیع از اطلاعات موجود در توزیع متغیر قابل محاسبه هستند. با وجود این، گاهی نیاز کاربر تنها با اطلاع از یک یا چند مشخصه توزیع برآورده می‌شود، بنابراین پردازش‌های آماری به محاسبه این مشخصه‌ها محدود می‌گردد و نیازی به ارائه توزیع متغیر نیست.

هنگامی که نمایش توزیع برای کاربر اهمیت دارد، نمودارها بر جدول‌های توزیع فراوانی اولویت دارند؛ زیرا از یک سو ساده‌تر هستند و کاربر را کمتر با اعداد و ارقام درگیر می‌کنند و از سوی دیگر به شکلی بصری می‌توانند توجه کاربر را به نکات مهم و برجستهٔ توزیع جلب کنند. برای مثال، وضعیتی از متغیر با بیشترین یا کمترین فراوانی به‌خوبی در نمودار به چشم کاربر می‌آید و توجه او را جلب می‌کند، در حالی که یافتن همین اطلاع از طریق جدول به کمی دقت و مقایسه اعداد و ارقام با یکدیگر نیاز دارد. همچنین، با دیدن نمودار یا منحنی توزیع به‌راحتی می‌توان به هموار بودن آن و به طور کلی ویژگی‌های مرتبط با شکل، نظیر تقارن یا چولگی پی‌برد.

سطح سنجش متغیر در انتخاب روش مناسب برای توصیف داده‌ها مؤثر است. جدول توزیع فراوانی و نمودارهای دایره‌ای و میله‌ای برای متغیرهای مقوله‌ای (متغیرهای اسمی، ترتیبی یا شمارشی) کاربرد دارند، در حالی که بافت‌نگار، منحنی توزیع، نمودار جعبه‌ای و نمودار شمایی برای متغیر پیوسته مناسب است. استفاده از جدول توزیع فراوانی یا نمودارهای دایره‌ای و میله‌ای برای متغیرهای پیوسته تنها پس از دسته‌بندی مقادیر آنها، یعنی تبدیل شدن به یک متغیر مقوله‌ای مناسب است.

مشخصه‌های توزیع به طور عمده به متغیرهای فاصله‌ای و نسبتی اختصاص دارند. در این میان دو استثنا وجود دارد. نما می‌تواند برای تمام متغیرها از جمله متغیرهای اسمی و ترتیبی نیز به کار رود و میانه برای متغیر ترتیبی نیز قابل استفاده است. دلیل این امر به فرایند کمّی شدن متغیرهای اسمی و ترتیبی باز می‌گردد. مشخصه‌های توزیع به جز نما و میانه، به محاسبات ریاضی نظیر جمع و تفریق یا به توان رساندن نیاز دارند، در حالی که مقادیر متغیرهای اسمی و ترتیبی دارای قابلیت‌های محاسباتی نیستند. مشخصه نما به محاسبه خاصی نیاز ندارد، پس برای متغیرهای اسمی و ترتیبی نیز به کار می‌رود. همچنین، یافتن میانه به مرتب کردن داده‌ها و تقسیم آنها به دو نیمه با «فراوانی» برابر وابسته است، از

همین رو کافی است مقادیر متغیر قابل مرتب شدن باشند تا بتوان میانه را برای آن به دست آورد. جدول ۱۰ خلاصه‌ای از پیوند روش‌های آمار توصیفی با نوع متغیرها ارائه می‌کند.

جدول ۱۰: چکیده‌ای از رابطه نوع متغیر با روش‌های توصیف داده‌ها

کاربرد	ابزارهای آمار توصیفی	نوع متغیر				ملاحظات
		اسمی	ترتیبی	فاصله‌ای	نسبتی	
ابزارهای نمایش توزیع متغیر	جدول توزیع فراوانی	✓	✓			برای متغیرهای فاصله‌ای و نسبتی پس از دسته‌بندی می‌تواند به کار برود.
	نمودارهای دایره‌ای و میله‌ای	✓	✓			برای متغیرهای فاصله‌ای و نسبتی پس از دسته‌بندی می‌تواند به کار برود.
	منحنی توزیع			✓	✓	به متغیرهای پیوسته اختصاص دارد.
	نمودارهای جعبه‌ای و شمایی			✓	✓	شکل توزیع با نمودار جعبه‌ای و داده‌های دورافتاده با نمودار شمایی شناسایی می‌شود.
مشخصه‌های تمرکز	نما	✓	✓	✓	✓	برای توزیع‌های هموار مناسب نیست. توزیع می‌تواند چندنمایی باشد.
	میانه		✓	✓	✓	به داده‌های دورافتاده حساس نیست.
	میانگین			✓	✓	بر اساس تمام داده‌ها محاسبه می‌شود. به داده‌های دورافتاده حساس است. میانگین پیراسته می‌تواند راهگشا باشد.
مشخصه‌های پراکندگی	دامنه و دامنهٔ میان‌چارکی			✓	✓	دامنه برخلاف دامنه میان‌چارکی، تحت تأثیر داده‌های دورافتاده قرار می‌گیرد. دامنه میان‌صدکی می‌تواند جایگزین مناسبی باشد.
	واریانس			✓	✓	به داده‌های دورافتاده حساس است. دوری از میانگین توزیع را می‌سنجد. میانگین قدرمطلق اختلاف‌ها رقیبی برای آن است.
	ضریب تغییرات			✓	✓	بدون واحد اندازه‌گیری است.
مشخصه‌های شکل	ضریب چولگی			✓	✓	میزان انحراف از تقارن را در منحنی توزیع می‌سنجد.
	ضریب کشیدگی			✓	✓	برآمدگی یا خوابیدگی منحنی توزیع را نسبت به توزیع نرمال می‌سنجد.

این پرسش درباره مشخصه‌های توزیع مطرح است که آیا می‌توان درباره بزرگی یا کوچکی مشخصه‌های توزیع قضاوت کرد؟ برای مثال، آیا ضریب چولگی ۲/۵ مقداری بزرگ محسوب می‌شود و از چولگی «شدید» توزیع به راست حکایت می‌کند یا واریانس ۱۰۰ بیانگر پراکندگی چشمگیری در داده‌هاست؟ قاعده کلی برای پاسخ به این پرسش چنین است: اگر مشخصه‌ای در فاصلهٔ معینی تغییر کند، با توجه به آن فاصله قابل ارزیابی است. برای مثال، معدل ۱۷ دانش‌آموز می‌تواند دورنمایی از وضعیت تحصیلی او به دست دهد؛ زیرا حداقل و حداکثر مقدارِ «ممکن» برای معدل ۰ یا ۲۰ است و در واقع، معدلِ ۱۷ در فاصلهٔ ۰ تا ۲۰ تفسیر می‌شود و همین امر به ارزیابی عدد ۱۷ کمک می‌کند.

گاهی نمی‌توان فاصلهٔ تغییرات مشخصه‌ای را به‌دقت مشخص کرد، ولی در عمل می‌توان محدوده‌ای برای آن در نظر گرفت، به طوری که داشتن مقداری خارج از آن محدوده برای آن مشخصه نادر باشد. در این حالت نیز ارزیابی مقدار مشخصه با توجه به آن محدوده امکان‌پذیر است. برای نمونه، حداقل درآمد می‌تواند ۰ باشد، ولی برای حداکثر آن نمی‌توان سقفی قطعی تعیین کرد، ولی با توجه به شرایط جامعه ممکن است درآمدی بین a تا b تومان معمول باشد، به این ترتیب متوسط درآمد در این فاصله قابل ارزیابی خواهد بود. اغلب، می‌توان برای مشخصه‌های تمرکز توزیع چنین فاصله‌ای را در نظر گرفت و مقدار آنها را در این فاصله ارزیابی کرد.

نگاه دیگر به مشخصه‌های توزیع نگاه مقایسه‌ای است. برای مثال، پژوهشگر می‌خواهد بداند توزیع درآمد در سال جاری نسبت به سال گذشته چوله‌تر شده است. پس ضریب چولگی امسال را با ضریب چولگی سال گذشته مقایسه می‌کند تا درباره بزرگی آنها قضاوت کند. همچنین، پژوهشگری مایل است بداند پراکندگی توزیع متغیر انسجام ملی که از ترکیب مجموعه‌ای از پرسش‌ها به دست می‌آید، در دو کشور چگونه است، پس واریانس این متغیر را در دو کشور محاسبه و مقایسه می‌کند تا به پاسخ پرسش خود دست یابد.

در هر حال پژوهشگری که از ابزار آمار توصیفی استفاده می‌کند، باید تحلیل اکتشافی داده‌ها[1] را مد نظر قرار دهد و بداند داده‌ها اطلاعات متعدد و متنوعی را از متغیر در خود پنهان کرده‌اند که باید با وارسی جامع، کشف و ارائه شوند؛ گو اینکه پژوهشگر با داده‌ها آن‌قدر بازی می‌کند و کلنجار می‌رود تا اطلاعات مختلف نهفته در آنها را از ابعاد گوناگون

---

1. Explanatory Data Analysis (EDA)

بیرون بکشد. این بدان معناست که مشخصه‌های توزیع در کنار هم، و نه به تنهایی می‌توانند تصویری از توزیع متغیر ارائه دهند. از این رو، بیان مشخصه تمرکز توزیع به تنهایی، یعنی بدون دادن اطلاعی درباره پراکندگی آن یا برعکس، بـه تصـویری مـبهم و نـاقص از ابعـاد مختلف توزیع می‌انجامد.

اصطلاحات فصل دوم

Arithmetic Mean	میانگین حسابی	Micro-data	داده‌های خُرد
Asymmetric	نامتقارن	Midpoint	نقطه میانی (نماینده دسته)
Average	متوسط	Mode	نما
Bar Chart	نمودار میله‌ای (ستونی)	Moment Coefficient of Skewness	ضریب گشتاوری چولگی
Box Plot	نمودار جعبه‌ای	Multi-modal	چند نمایی
Central Tendency	گرایش مرکزی	Normal Distribution Curve	منحنی توزیع نرمال
Classification	طبقه‌بندی (دسته‌بندی)	Outliers	داده‌های دورافتاده
Coefficient of Variation	ضریب تغییرات	Peak	قله منحنی توزیع
Cumulative Percent	درصد تجمعی (تراکمی)	Percent	درصد
Decile	دهک	Percentile	صدک
Density	چگالی	Pie Chart	نمودار دایره‌ای
Dispersion	پراکندگی	Polygon	نمودار چندضلعی
Distribution	توزیع	Proportion	نسبت
Dot Plot	نمودار نقطه‌ای	Quantile	چندک
Explanatory Data Analysis	تحلیل اکتشافی داده‌ها	Quartile Coefficient of Skewness	ضریب چارکی چولگی
Extreme	فَرین	Quartile	چارک
Flat	هموار، صاف	Range	دامنه
Frequency	فراوانی	Raw Data	داده‌های خام
Frequency Distribution Table	جدول توزیع فراوانی	Schematic Plot	نموداری شِمایی
Geometric Mean	میانگین هندسی	Skewness	چولگی (کجی)
Harmonic mean	میانگین سازوار	Standard Deviation	انحراف معیار
Histogram	بافت‌نگار، هیستوگرام	Statistical Process	پردازش آماری
Information	اطلاعات	Symmetric	متقارن
Inter-Percentile Range	دامنه میان‌صدکی	Distribution Tail	دُم توزیع
Inter-Quartile Range	دامنه میان‌چارکی	Trimmed Mean	میانگین پیراسته
Kurtosis	برجستگی	Variability	تغییرپذیری
Mean	میانگین	Variable Distribution Curve	منحنی توزیع متغیر
Mean of Absolute Deviations	قدرمطلق انحراف از میانگین	Variance	واریانس
Median	میانه		

تمرین‌های فصل دوم

۱. جدول زیر توزیع فراوانی کارمندان دو بخش از یک اداره را بر حسب گروه سنی نشان می‌دهد. برای مثال، ۸ نفر از کارمندان بخش «آ» ۲۰ تا ۳۰ ساله هستند، در حالی که ۱۵ نفر از کارمندان بخش «ب» در این گروه سنی قرار دارند.

آ) برای هر یک جدول توزیع فراوانی رسم کنید.

ب) بر اساس درصد تجمعی جدول‌های توزیع فراوانی، سن چند درصد از کارمندان بخش «آ» کمتر از ۵۰ سال است؟

پ) سن چند درصد از کارمندان بخش «ب» بین ۳۰ تا ۶۰ سال است؟

ت) توزیع سن کارمندان بخش «آ» را با نمودار دایره‌ای و توزیع سن کارمندان بخش «ب» را با نمودار میله‌ای نمایش دهید.

گروه سنی (سال)	بخش «آ»	بخش «ب»
۳۰–۲۰	۸	۱۵
۴۰–۳۰	۱۷	۳۲
۵۰–۴۰	۱۱	۲۰
۶۰–۵۰	۸	۴
۷۰–۶۰	۲	۰

۲. در تمرین ۱، ابتدا دو طبقه ۵۰ تا ۶۰ سال و ۶۰ تا ۷۰ سال را ادغام کنید تا طبقه ۵۰ تا ۷۰ سال ایجاد شود. اکنون بافت‌نگار توزیع متغیر سن را برای کارمندان بخش «آ» رسم کنید. مساحت زیر این بافت‌نگار را با جمع کردن مساحت چهار مستطیل آن محاسبه کنید.

۳. نمودار میله‌ای را برای تمرین ۲ رسم کنید. این نمودار چه تفاوتی با بافت‌نگاری که پیشتر رسم کردید، دارد؟

۴. جمع‌های زیر را با نماد $\sum$ بیان کنید:

۱) $x_1 + x_2 + x_3 + x_4 + x_5$

۲) $x_3 + x_4 + x_5 + x_6 + x_7$

۳) $x_1^3 + x_2^3 + x_3^3 + x_4^3 + x_5^3$

۴) $\left[x_1 + x_2 + x_3 + x_4 + x_5\right]^3$

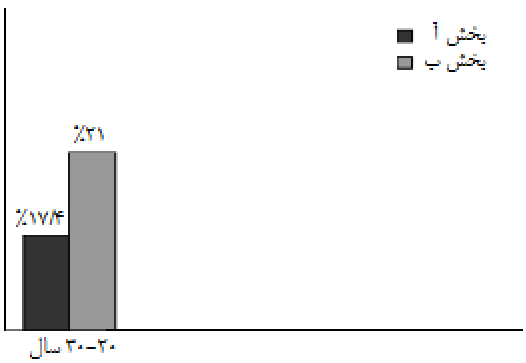
۵) $x_1y_1 + x_2y_2 + x_3y_3 + x_4y_4 + x_5y_5$

۶) $x_1^3y_1 + x_2^3y_2 + x_3^3y_3 + x_4^3y_4 + x_5^3y_5$

۵. پس از تکمیل جدول زیر، بافت‌نگار داده‌ها را رسم کنید. مانند تمرین ۲، مساحت زیر این بافت‌نگار چقدر است؟ چه نتیجه‌ای از یافته این تمرین و تمرین ۲ می‌توان گرفت؟

دسته	طول دسته	فراوانی	(طول دسته ÷ فراوانی نسبی) = چگالی
۰ تا ۱۰	۱۰	۱۰۰	۰/۰۲۴
۱۰ تا ۱۵	۵	۲۵	۰/۰۱۲
۱۵ تا ۵۰		۷۰	
۵۰ تا ۷۰	۲۰	۸۰	
۷۰ تا ۱۲۰	۵۰		۰/۰۰۷
جمع	–	۴۲۵	–

۶. در تمرین ۱، می‌توان برای مقایسه وضعیت گروه سنی یک بخش با گروه سنی متناظر آن در بخش دیگر، نمودار میله‌ای هر دو بخش را در قالب یک نمودار نمایش داد. برای این منظور، میله‌های دو بخش را که مربوط به یک گروه سنی هستند، در کنار هم قرار می‌دهند. این کار برای اولین گروه سنی در نمودار زیر صورت گرفته است. سایر گروه‌های سنی را به آن اضافه کنید.



۷. اگر $x_1 = ۳$، $x_2 = ۵$ و $x_3 = ۱۰$ باشد، حاصل جمع‌های زیر را به دست آورید:

$$۱)\sum_{i=1}^{3} x_i \quad ۲)\left[\sum_{i=1}^{3} x_i\right]^2 \quad ۳)\sum_{i=1}^{3}(x_i - ۶)^2 \quad ۴)\sum_{i=1}^{3} x_i^2 - \left[\sum_{i=1}^{3} x_i\right]^2 \quad ۵)\sum_{i=1}^{2} x_i x_{i+1}$$

۸. جدول زیر اطلاعات دهک‌های توزیع سه متغیر را نشان می‌دهد که مقادیر آنها بین ۰ تا ۱۰۰ است. برای مثال، دهک سوم متغیر X برابر با ۱۵ و دهک هفتم متغیر Z برابر با ۴۵

است. شکل تقریبی منحنی توزیع هر یک از این متغیرها را با استفاده از اطلاعات دهک‌های آنها رسم کنید.

دهک / متغیر	۱	۲	۳	۴	۵	۶	۷	۸	۹
X	۳	۸	۱۵	۳۰	۵۰	۷۰	۸۵	۹۲	۹۷
Y	۲۰	۳۵	۴۵	۴۹	۵۰	۵۱	۵۵	۶۵	۸۰
Z	۵	۷	۹	۱۰	۲۵	۲۹	۴۵	۶۵	۸۳

۹. اگر داده‌ها به یک اندازه افزایش یا کاهش یابند، میانگین آنها نیز به همان اندازه افزایش یا کاهش می‌یابد. همچنین اگر داده‌ها در مقدار ثابتی ضرب شوند، میانگین آنها نیز در همان مقدار ضرب می‌شود. این موضوع را با داده‌های جدول زیر بررسی کنید (داده‌های سطر دوم از افزودن ۳ به داده‌های سطر اول و داده‌های سطر سوم از ضرب کردن ۴ در داده‌های سطر اول به دست آمده‌اند). اکنون فکر می‌کنید اگر داده‌ها در مقدار ثابتی ضرب شوند و سپس مقدار ثابت دیگری نیز به همه آنها افزوده شود، چه اتفاقی برای میانگین داده‌ها رخ می‌دهد؟ به نظر شما، این اتفاق چه تأثیری بر سایر مشخصه‌های تمرکز مانند نما یا میانه می‌گذارد؟

داده‌ها	میانگین
۲، ۵، ۳، ۸، ۹، ۶	۵/۵
۵، ۸، ۶، ۱۱، ۱۲، ۹	
۸، ۲۰، ۱۲، ۳۲، ۳۶، ۲۴	

۱۰. کدام جمله درباره واریانس اشتباه است؟

آ) نشان می‌دهد که متوسط مربع انحراف از میانگین داده‌ها چقدر است.

ب) بیان می‌کند که چقدر داده‌ها متنوع و دور از هم هستند.

پ) مقدار آن را می‌توان به تنهایی تفسیر کرد.

ت) هیچ‌گاه منفی نیست.

ث) کمتر از انحراف معیار تحت تأثیر داده‌های دورافتاده قرار دارد.

۱۱. شیوهٔ ساده‌تر برای محاسبهٔ دستی واریانس n داده، استفاده از رابطهٔ زیر است:

$$S^2 = \frac{1}{n-1}\left(\sum_{i=1}^{n} X_i^2 - n\overline{X}^2\right)$$

در این صورت، به محاسبهٔ تفاضل هر داده از میانگین نیازی نیست. واریانس دو مجموعه داده زیر را یک بار با این رابطه و بار دیگر با رابطهٔ معرفی‌شده در فصل دوم محاسبه کنید. چه نتیجه‌ای از محاسبات خود می‌گیرید؟

مجموعه داده اول	۲، ۶، ۳، ۸، ۱
مجموعه داده دوم	۳، ۹، ۵، ۶، ۲

۱۲. جدول زیر، دو مجموعه داده ۱۵ تایی را نشان می‌دهد.

متغیر X	متغیر Y
۱۸، ۱۲، ۱۰، ۳۰، ۳۰، ۳۸، ۲۰، ۲۸، ۱۸، ۲، ۲۴، ۱۴، ۶، ۴۰، ۱۶	۸، ۲۰، ۳، ۷، ۱۲، ۱، ۹، ۱۴، ۱۰، ۱۹، ۱۵، ۱۵، ۵، ۶، ۹

آ) واریانس آنها را محاسبه کنید. چه رابطه‌ای بین واریانس‌های دو مجموعه داده وجود دارد؟ چرا؟

ب) به نظر شما پراکندگی کدام مجموعه بیشتر است؟

پ) ضریب تغییرات دو مجموعه را محاسبه کنید. آیا نتایج قبلی شما را درباره پراکندگی دو مجموعه تأیید می‌کند؟

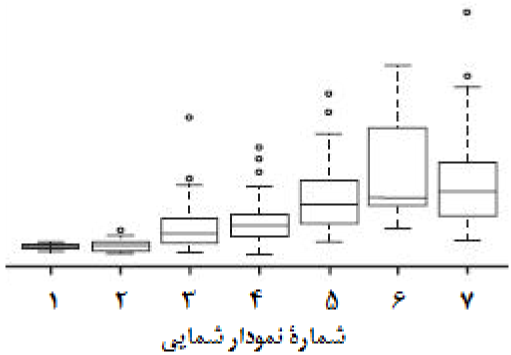
۱۳. بر اساس اطلاعات جدول زیر به پرسش‌ها پاسخ دهید:

مشخصه‌ها / متغیر	X	Y	Z
میانگین	۴۰	۲۰۰	۱۰
میانه	۴۵	۵۰	۱۵
انحراف معیار	۱۲۰	۶۰۰	۱۲۰

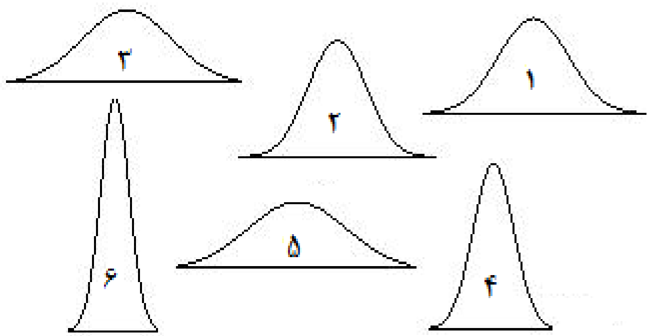
آ) پراکندگی توزیع متغیرها در مقایسه با یکدیگر چگونه است؟

ب) انتظار دارید کدام‌یک از توزیع‌ها دارای نقاط دورافتاده باشد؟

۱۴. بر اساس نمودارهای شمایی مقابل، درباره شکل و پراکندگی توزیع متغیرهای متناظر با آنها بحث کنید.



۱۵. منحنی‌های زیر همگی متقارن و تک نمایی هستند. تفاوت آنها در چیست؟



۱۶. با توجه به اطلاعات جدول تمرین ۱ به پرسش‌های زیر پاسخ دهید.

آ) میانگین سنی کدام بخش کمتر است؟

ب) تنوع سنی در کدام بخش بیشتر است؟

پ) بر اساس پاسخ‌های دو پرسش فوق، آیا تفاوتی بین توزیع سن کارمندان دو بخش وجود دارد؟

۱۷. انتظار دارید چولگی کدام‌یک از توزیع‌های زیر بیشتر باشد؟

آ) توزیعی که میانگین آن پس از دهک نهم است.

ب) توزیعی که میانگین آن بین میانه و دهک ششم است.

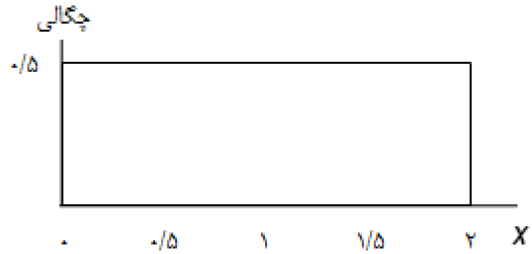
۱۸. کدام‌یک از عبارت‌های زیر درباره متغیر اشتباه است؟

آ) تمام اطلاعات مربوط به متغیر از توزیع آن به‌دست می‌آید.

ب) توزیع‌های متقارن بهتر از توزیع‌های چوله هستند.

پ) روش یافتن توزیع متغیرهای اسمی مانند متغیرهای فاصله‌ای پیوسته است.
ت) اگر واریانس توزیع متغیری بیش از ۴۰۰ باشد، پراکندگی آن توزیع خیلی زیاد است.
ث) چندک ۲۷ توزیع متغیر یکی از مشخصه‌های آن است.
ج) شکل توزیع متغیر به جامعه آماری بستگی دارد که توزیع از آن به‌دست می‌آید.
چ) یکسان‌بودن توزیع دو متغیر به آن معنی است که ماهیت دو متغیر نیز یکسان است.
ح) ممکن است توزیع متغیری در طول زمان تغییر کند.
خ) اگر مشخصه‌های دامنه و نمای دو توزیع برابر باشد، آن دو توزیع نیز یکسان هستند.
د) هر توزیعی دارای نماست.

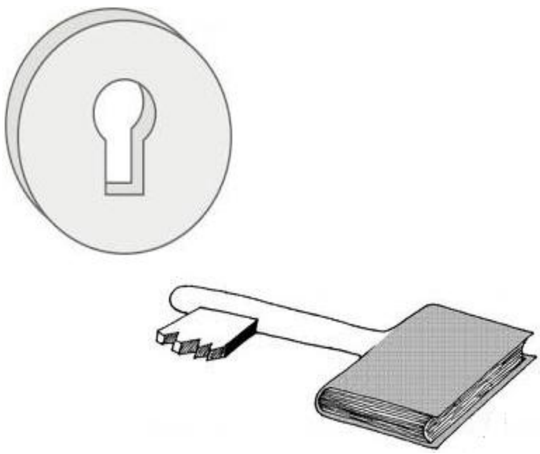
۱۹. منحنی توزیع متغیر پیوسته X به صورت زیر است.
آ) نمای توزیع چیست؟ از پاسخ خود چه نتیجه‌ای درباره این منحنی می‌گیرید؟
ب) آیا این توزیع متقارن است؟
پ) میانگین توزیع چیست؟
ت) چند درصد از مقادیر متغیر X بیشتر از ۱/۵ است؟



فصل سوم

۳

آمار استنباطی مفاهیم پایه



مباحث آمار استنباطی بیشترین حجم این کتاب را تشکیل می‌دهند. در چند فصل آینده، خواننده با روش‌هایی آشنا خواهد شد که به او در نتیجه‌گیریِ قابل تعمیم به جامعه از داده‌های نمونه کمک می‌کنند. این فصل به معرفی مفاهیم پایه در آمار استنباطی اختصاص دارد. آشنایی با توزیع نرمال و شیوه استفاده از جدول چندک‌های آن و توزیع نمونه‌گیری یک آماره از جملهٔ این مفاهیم است. قضیه حد مرکزی به عنوان یکی از مشهورترین و راهگشاترین قضیه‌ها در آمار استنباطی پایان‌بخش مطالب این فصل است. صرف وقت برای درک مفاهیم این فصل می‌تواند به درک بهتر موضوعات فصل‌های بعد کمک مؤثری کند.

۱ـ چشم‌اندازی از آمار استنباطی

در فصل اول بیان شد که آمار، دانش کسب اطلاعات درباره متغیرها بر اساس داده‌هایی است که از جامعه آماری مشخصی گردآوری شده‌اند. در واقع، هدف اصلی از گردآوری داده‌ها شناخت آن متغیرها در «جامعه آماری» است. برای مثال، پژوهشگری مایل است یک شهر را از نگاه متغیر وضعیت سواد بررسی کند. بنابراین می‌کوشد ضمن شناسایی توزیع این متغیر، دریابد چه درصدی از مردم آن «شهر» باسواد و چه درصدی بی‌سواد هستند.

آنچه اطلاعات آماری را ارزشمند می‌سازد، قابلیت آنها در شناخت «تمام» جامعه آماری تحت مطالعه است. نخستین راهکار برای شناخت «کل» جامعه آماری، گردآوری داده‌ها از «تمام» اعضای آن جامعه یعنی سرشماری است. به این ترتیب، توزیع متغیر و مشخصه‌هایی که از داده‌ها به دست می‌آیند، وضعیت آن متغیر را به طور دقیق در کل جامعه آماری نشان می‌دهند؛ زیرا متغیر از تک‌تک اعضای جامعه آماری اندازه‌گیری شده و داده‌ای از هر عضو جامعه در دست است. بنابراین، به‌راحتی می‌توان تنها با استفاده از روش‌های آمار توصیفی توزیع متغیر و مشخصه‌های آن را به دست آورد و گزارشی از آنها ارائه کرد.

اگرچه راهکار سرشماری برای جامعه‌های آماری کوچک، مانند جامعه‌ای با چند هزار عضو، مناسب به نظر می‌رسد، ولی برای جامعه‌های آماری بزرگ با چند میلیون عضو عملی نیست؛ زیرا محدودیت‌های زمانی و مالی اجازه گردآوری داده‌هایی را با چنین وسعت و گستره‌ای به ما نمی‌دهند. از این رو، بخش کوچکی از جامعه آماری به نام «نمونه» با سازوکاری علمی انتخاب می‌شود تا داده‌ها تنها از اعضای آن بخش گردآوری شوند. داده‌های حاصل از نمونه را می‌توان با روش‌های آمار توصیفی سازماندهی و خلاصه کرد و اطلاعات توزیع متغیر و مشخصه‌های آن را برای اعضای «نمونه» به دست آورد. با وجود این، هدف اصلی، کسب اطلاعات درباره «کل» جامعه آماری است، نه نمونه‌ای از آن. از این رو، اگر بتوان اطلاعات نمونه را با سازوکاری به جامعه آماری که نمونه از آن انتخاب شده است، «تعمیم» داد، آنگاه به هدف اصلی یعنی کسب اطلاعات درباره کل جامعه آماری رسیده‌ایم. آمار استنباطی شامل روش‌هایی است که سازوکار تعمیم اطلاعات نمونه را به جامعه آماری فراهم می‌کند.

برای مثال، فرض کنید یک نظرسنجی با هدف یافتن میانگین روزانه مدت زمان انتظار برای استفاده از اتوبوس‌های درون شهری تهران صورت گرفته است (به یاد بیاورید که

میانگین، یک مشخصهٔ توزیع است). برای این منظور، نمونه‌ای ۱۰۰۰ نفری از مسافران اتوبوس‌های درون‌شهری انتخاب و از آنان پرسیده می‌شود، به طور معمول چند دقیقه در ایستگاه اتوبوس در انتظار می‌مانید؟ اگر میانگین مدت زمان انتظار این ۱۰۰۰ نفر برابر با ۶ دقیقه باشد، این پرسش مطرح است که آیا می‌توان ادعا کرد که میانگین مدت زمان انتظار «تمام» مسافران اتوبوس‌های درون شهری تهران نیز ۶ دقیقه است؟ یعنی آیا میانگین ۶ دقیقه که به نمونه ۱۰۰۰ نفری از مسافران اختصاص دارد، به تمام مسافران تهرانی استفاده‌کننده از اتوبوس‌های درون شهری قابل تعمیم است؟ پاسخ پرسش‌هایی از این دست به آمار استنباطی مربوط می‌شود.

مباحث آمار استنباطی به دو حوزه برآوردیابی و آزمون فرضیه تقسیم می‌شوند. بحث برآوردیابی به این پرسش می‌پردازد که چگونه می‌توان مشخصه‌های توزیع متغیر را بر اساس داده‌های نمونه با احتمال معین برای کل جامعه آماری تخمین زد؟ این پرسش دارای دو نکته اساسی است: «احتمال معین» و «تخمین زدن». تخمین زدن به این واقعیت اشاره دارد که اطلاعاتی که از داده‌های نمونه درباره مشخصه‌های توزیع متغیر در جامعه آماری ارائه می‌شوند، اطلاعاتی «تقریبی» هستند و از همین رو «برآورد» نامیده می‌شوند. بنابراین، برآورد مشخصه‌های توزیع از داده‌های نمونه با «خطا» همراه است و اطلاعی قطعی قلمداد نمی‌شود. تقریبی بودن برآورد مشخصه‌ها با نگرانی درباره قابل اعتماد بودن آنها یعنی میزان راستی‌نمایی آنها همراه است. پژوهشگر همواره از خود می‌پرسد که تا چه حد می‌توانم به برآوردی اعتماد کنم که از نمونه برای مشخصه توزیع در «کل» جامعه به دست آورده‌ام؟ «احتمال معین» بیان می‌کند که تا چه حد برآوردها قابل اعتماد هستند؛ یعنی چقدر احتمال دارد که برآورد حاصل از نمونه قابل تعمیم به کل جامعه آماری باشد؟ به این ترتیب، پژوهشگر به روشنی می‌داند احتمال تطبیق برآورد او با واقعیت چقدر است.

> هنگامی که یافته‌ای را بر اساس داده‌های نمونه برای جامعه آماری ارائه می‌کنیم، دست به استنباط آماری زده‌ایم. ارائه برآوردی از مشخصه توزیع در جامعه یا بررسی درستی یک ادعا درباره مشخصهٔ توزیع، مثال‌هایی از استنباط آماری هستند.

بحث دیگر آمار استنباطی، یعنی آزمون فرضیه، به بررسی درستی ادعاهایی می‌پردازد که درباره مشخصه‌های توزیع متغیر مطرح می‌شوند. برای مثال، این ادعا که میانگین مدت زمان انتظار برای اتوبوس در شهر تهران بیش از شهر کرج است، به مقایسه مشخصه

میانگین در دو جامعه آماری می‌پردازد. این ادعا نیز بر اساس داده‌های دو نمونه صورت می‌گیرد که یکی از شهر تهران و دیگری از شهر کرج انتخاب شده‌اند. در صورتی که میانگین مدت زمان انتظار برای نمونه شهر تهران برابر با ۶ دقیقه و برای نمونه شهر کرج برابر با ۵ دقیقه باشد، آیا می‌توان بزرگ‌تر بودنِ ۶ نسبت به ۵ را دلیلی بر بزرگ‌تر بودن میانگین مدت زمان انتظار در شهر تهران نسبت به شهر کرج به حساب آورد؟ به عبارت دیگر، تفاوت میانگین دو نمونه به معنای وجود تفاوت بین میانگین دو «شهر» است؟

روش‌های آزمون فرضیه سازوکاری را در اختیار می‌گذارند تا ادعا درباره مشخصه‌های توزیع را بر اساس داده‌های نمونه بیازماییم و درباره درستی آن ادعا در «جامعه آماری» قضاوت کنیم. نتیجه بررسی چنین ادعاهایی نیز با خطا روبه‌روست؛ یعنی نمی‌توان اثبات ادعایی را درباره مشخصه توزیع بر اساس داده‌های نمونه در جامعه آماری قطعی دانست. به عبارت دیگر، نتیجه آزمون یک ادعا نیز مانند برآوردیابی با احتمال مشخصی قابل اعتماد است.

سه فصل بعد به موضوع استنباط آماری یعنی برآوردیابی و آزمون فرضیه می‌پردازند، ولی پیش از آن، با مفهوم آماره، پارامتر، فضای نمونه، توزیع نمونه‌گیری و توزیع نرمال در این فصل آشنا خواهیم شد که مبنایی برای طرح مباحث آمار استنباطی فراهم می‌کنند.

۲ـ توزیع نمونه‌گیری

ابتدا با معرفی دو اصطلاح «آماره» و «پارامتر»، بین مشخصهٔ حاصل از داده‌های نمونه با مشخصهٔ حاصل از سرشماری (داده‌های تمام اعضای جامعه آماری) تمایز قایل می‌شویم. آماره برآوردی از مشخصهٔ توزیع است که بر اساس داده‌های نمونه محاسبه می‌شود، در حالی که پارامتر مقدار واقعی مشخصهٔ توزیع است که از داده‌های متعلق به تمام اعضای جامعه آماری به دست می‌آید. بنابراین، بر حسب اینکه داده‌ها در سطح نمونه یا جامعه آماری در دسترس باشند، هر یک از مشخصه‌های توزیع می‌تواند آماره یا پارامتر نامیده شود[۱].

> پارامتر مقدار یکتایی دارد که اغلب نامعلوم است، زیرا تنها از طریق سرشماری به دست می‌آید. مقدار آماره می‌تواند از نمونه‌ای به نمونه دیگر تغییر کند، بنابراین از توزیع نمونه‌گیری آماره‌ها صحبت می‌شود.

برای مثال، درآمد خانوارهای تهرانی دارای منحنی توزیعی است که تنها هنگامی در دست

۱. به زبان ساده، آماره به هر کمّیتی گفته می‌شود که بر اساس داده‌های نمونه به دست می‌آید و پارامتر کمّیتِ متناظر با آماره است که حاصل از داده‌های سرشماری است. بنابراین، آماره‌ها و پارامترها تنها به مشخصه‌هایی از توزیع محدود نمی‌شوند که در فصل دوم معرفی شده‌اند. در فصل‌های بعد با آماره‌ها و پارامترهای دیگری آشنا خواهید شد.

خواهد بود که درآمد تک‌تک خانوارهای تهرانی از طریق سرشماری گردآوری شده باشد. به همین ترتیب، مشخصه‌های توزیع درآمد مانند میانگین، واریانس، دامنه، ضریب چولگی و... نیز در صورت وجود داده‌هایی از تمام خانوارهای تهرانی قابل محاسبه است. بنابراین، تنها هنگامی می‌توان به مقدار «واقعی» مشخصه‌ای مانند میانگین درآمد دست یافت که داده‌ها از تمام خانوارهای تهرانی گردآوری شده باشند. مشخصه‌هایی را که از توزیع واقعی به دست می‌آیند، پارامترهای توزیع درآمد می‌نامند.

اکنون تصور کنید نمونه کوچکی از کل خانوارهای تهرانی انتخاب و داده‌های مربوط به درآمد آنها گردآوری شده است. بر اساس داده‌های نمونه نیز می‌توان منحنی توزیع درآمد را به دست آورد و به این ترتیب، مشخصه‌های این توزیع را هم محاسبه کرد. از آنجا که این منحنی توزیع به داده‌های نمونه تعلق دارد، نه کل جامعه خانوارهای تهرانی، به آن، منحنی توزیع نمونه می‌گویند و مشخصه‌هایی را که از آن به دست می‌آیند نظیر میانگین نمونه، واریانس نمونه و...، آماره می‌نامند. برای تأکید بر تمایز بین آماره و پارامتر از نمادهای متفاوتی نیز مانند جدول ۱ استفاده می‌شود.

جدول ۱: نمادهای به‌کاررفته برای جامعه و نمونه

مشخصه	جامعه	نمونه
اندازه (تعداد)	N	n
میانگین	$\mu = \frac{1}{N}\sum_{i=1}^{N} X_i$	$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$
واریانس	$\sigma^2 = \frac{1}{N}\sum_{i=1}^{N}(X_i - \overline{X})^2$	$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$
انحراف معیار	σ	S

اغلب، پارامترهای توزیع متغیر نامعلوم هستند؛ زیرا تنها بر اساس داده‌های گردآوری‌شده از «کل» جامعه آماری یعنی سرشماری قابل محاسبه‌اند، در حالی که پیشتر اشاره شد، محدودیت‌های زمانی و مالی اجازه گردآوری داده‌ها را از تمام اعضای یک جامعه آماری بزرگ نمی‌دهند. بنابراین در عمل به انتخاب نمونه‌ای از آن اکتفا می‌شود و به جای محاسبه پارامترها، آماره‌های متناظر با آنها به دست می‌آیند.

۱ـ۲ـ تکرار نمونه‌گیری

از یک جامعه آماری نمونه‌های متعددی را می‌توان انتخاب کرد که می‌توانند به آماره‌هایی با مقادیر مختلف منجر شوند؛ زیرا هر نمونه شامل بخشی از اعضای جامعه آماری است که می‌تواند متفاوت با نمونه دیگر باشد. برای مثال، جامعه آماری ۵ عضوی و درآمدهای آنها را مطابق جدول ۲ در نظر بگیرید (عدد مقابل هر نام، درآمد اوست؛ مثلاً، درآمد سعید ۱۰ تومان است). اگر قرار باشد از این ۵ عضو، ۳ نفر با قرعه‌کشی انتخاب شوند، ۱۰ حالت مختلف برای این انتخاب ۳ تایی امکان‌پذیر است که در جدول ۲ فهرست شده‌اند. برای مثال، انتخاب امید، امیر و مجید یعنی نمونه ۹، یکی از این انتخاب‌هاست.

میانگین درآمد جامعه آماری به عنوان یک پارامتر که بر اساس درآمد هر ۵ عضو جامعه محاسبه می‌شود، برابر با ۱۰ است، ولی میانگین نمونه‌ها به عنوان آماره بر حسب اینکه چه افرادی با قرعه‌کشی در نمونه انتخاب شده باشند، از نمونه‌ای به نمونه دیگر متفاوت است. برای مثال، میانگین درآمد در نمونه اول که شامل سعید، امید و امیر است، برابر با $۱۱ = \frac{۱۰+۱۵+۸}{۳}$ و در نمونه هشتم که شامل امیر، حامد و مجید است برابر با $۸/۳۳ = \frac{۸+۱۱+۶}{۳}$ است. این بدان معناست که پارامتر دارای مقداری یکتاست، در حالی که مقدار آماره در نمونه‌های مختلف لزوماً یکسان نیست و می‌تواند برای یک نمونه متفاوت با نمونه دیگر باشد.

این تفاوت که «تغییرپذیری» آماره نامیده می‌شود، تعریف متغیر را به یاد خواننده می‌آورد (ویژگی‌ای که اعضای جامعه آماری به طور یکسان دارا نباشند). گو اینکه می‌توان مجموعه تمام نمونه‌های ۳ تایی از جامعه آماری جدول ۲ را یک جامعه آماری جدید با ۱۰ عضو و آماره را ویژگی‌ای برای این اعضا در نظر گرفت که مقدار آن برای اعضای این جامعه جدید یکسان نیست. پس آماره در جامعه آماریِ نمونه‌های ۳ تایی، متغیری است که می‌توان مانند هر متغیر دیگری از توزیع آن صحبت کرد.

> توزیع نمونه‌گیری آماره ناشی از تغییرات نمونه‌ای است؛ یعنی تفاوت‌های بین نمونه‌ها که به مجموعه داده‌های گوناگون و به تبع آن آماره‌های متفاوت می‌انجامد.

معمولاً به جای آنکه از اصطلاح «جامعه آماریِ نمونه‌های ۳ تایی» استفاده شود، مجموعه تمام نمونه‌های ۳ تایی را که از جامعه آماریِ ۵ عضویِ جدول ۲ می‌توان انتخاب

کرد، «فضای نمونه» آن جامعه آماری می‌نامند که شامل ۱۰ عضو مختلف است[1]. توجه کنید که اگر به جای نمونه‌های ۳ تایی، نمونه‌های ۲ تایی یا ۴ تایی را از جامعه آماری انتخاب می‌کردیم، دو فضای نمونه دیگر برای آن ایجاد می‌شد؛ یعنی یک جامعه آماریِ مفروض، بر حسب تعداد اعضای نمونه، فضاهای نمونه متعددی دارد. اعضای فضای نمونه را نمونه‌ها تشکیل می‌دهند و هر آماره نیز در این جامعه یک متغیر است. توزیع آماره در فضای نمونه را توزیع نمونه‌گیری آن آماره می‌نامند تا بر این امر تأکید شود که توزیع ایجادشده ناشی از تغییرات نمونه‌ای است.

جدول ۲: نمونه‌های ۳ تایی ممکن از یک جامعه آماری ۵ عضوی

اعضای جامعه آماری و درآمد آنها									
سعید	۱۰	امید	۱۵	امیر	۸	حامد	۱۱	مجید	۶
۱۰ = ۵ ÷ (۶+۱۱+۸+۱۵+۱۰) = میانگین درآمد جامعه آماری									
فهرست ۱۰ نمونه ۳ تایی ممکن از جامعه آماری ۵ عضوی فوق									
نمونه	اعضای نمونه			میانگین نمونه	نمونه	اعضای نمونه			میانگین نمونه
۱	سعید	امید	امیر	۱۱	۶	سعید	امید	حامد	۱۲
۲	سعید	امیر	حامد	۹/۶۷	۷	امید	حامد	مجید	۱۰/۶۷
۳	سعید	امیر	مجید	۸	۸	امیر	حامد	مجید	۸/۳۳
۴	سعید	امید	مجید	۱۰/۳۳	۹	امید	امیر	مجید	۹/۶۷
۵	امید	امیر	حامد	۱۱/۳۳	۱۰	سعید	حامد	مجید	۹

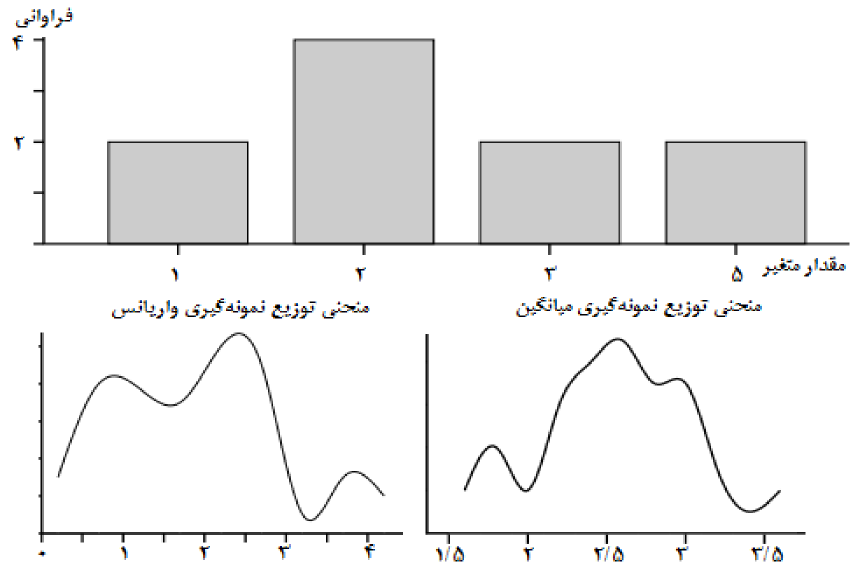
خواننده باید تاکنون دریافته باشد که توزیع متغیر، متفاوت با توزیع نمونه‌گیریِ آماره است. برای مثال، توزیع درآمد خانوارهای تهرانی متفاوت با توزیع میانگین‌های درآمد نمونه‌های ۱۰۰۰تایی مختلفی است که از خانوارهای تهرانی انتخاب می‌شوند. برای روشن شدن موضوع از شکل ۱ کمک می‌گیریم. این شکل شامل نمودار میله‌ای توزیع فراوانی یک متغیر در جامعه آماری ۱۰ عضوی است (مقدار متغیر برای تک‌تک اعضا نیز در قالب جدولی در شکل ۱ ارائه شده است). همچنین، منحنی توزیع نمونه‌گیری میانگین و

۱. این مثال حالت ساده‌ای را در نظر گرفته است که به فضای نمونه‌ای تنها با ۱۰ عضو منجر می‌شود. برای درک وسعت یک فضای نمونه توجه کنید که فضای نمونه‌های ۵ تایی ممکن از یک جامعه آماری ۱۰۰ عضوی شامل ۷۵۲۸۷۷۵۲۰ نمونه مختلف است.

واریانس ۵۱ نمونه ۵ تایی نیز نمایش داده شده است که از این جامعه ۱۰ عضـوی انتخـاب شده‌اند؛ یعنی ۵۱ بار از این جامعه نمونه‌های ۵ تـایی انتخـاب و میـانگین و واریـانس هـر نمونه محاسبه شده است. به این ترتیب ۵۱ میانگین و ۵۱ واریانس در دست اسـت کـه بـر اساس آنها منحنی توزیع نمونه‌گیری میانگین و منحنی توزیع نمونـه‌گیـری واریـانس رسـم شده است.

شکل ۱: مشخصات جامعه و منحنی توزیع نمونه‌گیری میانگین و واریانس نمونه‌های ۵ تایی آن

جامعه آماری ۱۰ عضوی و مقدار متغیر برای هر یک از آنها									
مریم	۵	رضا	۲	لیلا	۱	سیما	۳	فربد	۲
مانی	۲	سارا	۱	فرید	۲	ندا	۵	آرزو	۳

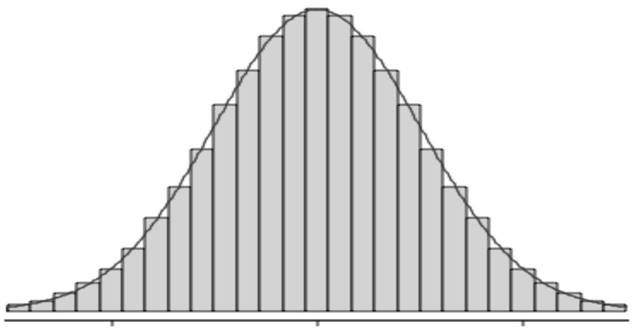


این مثال دو نکته را درباره توزیع نمونه‌گیری آماره در اختیار می‌گـذارد. نخسـت آنکـه، توزیع آماره می‌تواند هیچ شباهتی به توزیع متغیر نداشته باشد. در شکل ۱، توزیع آماره‌ها با منحنی‌های توزیعی نمایش داده شده است که هیچ شباهتی بـه نمـودار میلـه‌ای مربـوط بـه متغیر ندارند. دوم آنکه، دو منحنی توزیع متفاوت برای میانگین و واریانس بـه دسـت آمـده است، پس آماره‌های مختلف می‌توانند توزیع‌های متفاوتی داشته باشند، حتـی اگـر بـه یـک متغیر مربوط باشند.

۳ـ توزیع نرمال

یکی از منحنی‌های توزیع مشهور با کاربردهای فراوان، منحنی توزیع نرمال با شکلی ناقوسی مانند شکل ۲ است. پدیده‌های (متغیرهای) متعدد و متنوعی در جهان واقعی وجود دارند که توزیع آنها مانند منحنی توزیع نرمال است. اگر نشانه‌هایی در دست باشد که متغیری از توزیع نرمال پیروی می‌کند، کار پژوهشگر بسیار ساده خواهد شد؛ زیرا مشخصات مختلف این توزیع کاملاً شناخته شده‌اند و می‌توان به‌راحتی از اطلاعات موجود در کتاب‌های آماری برای توصیف آن متغیر استفاده کرد.

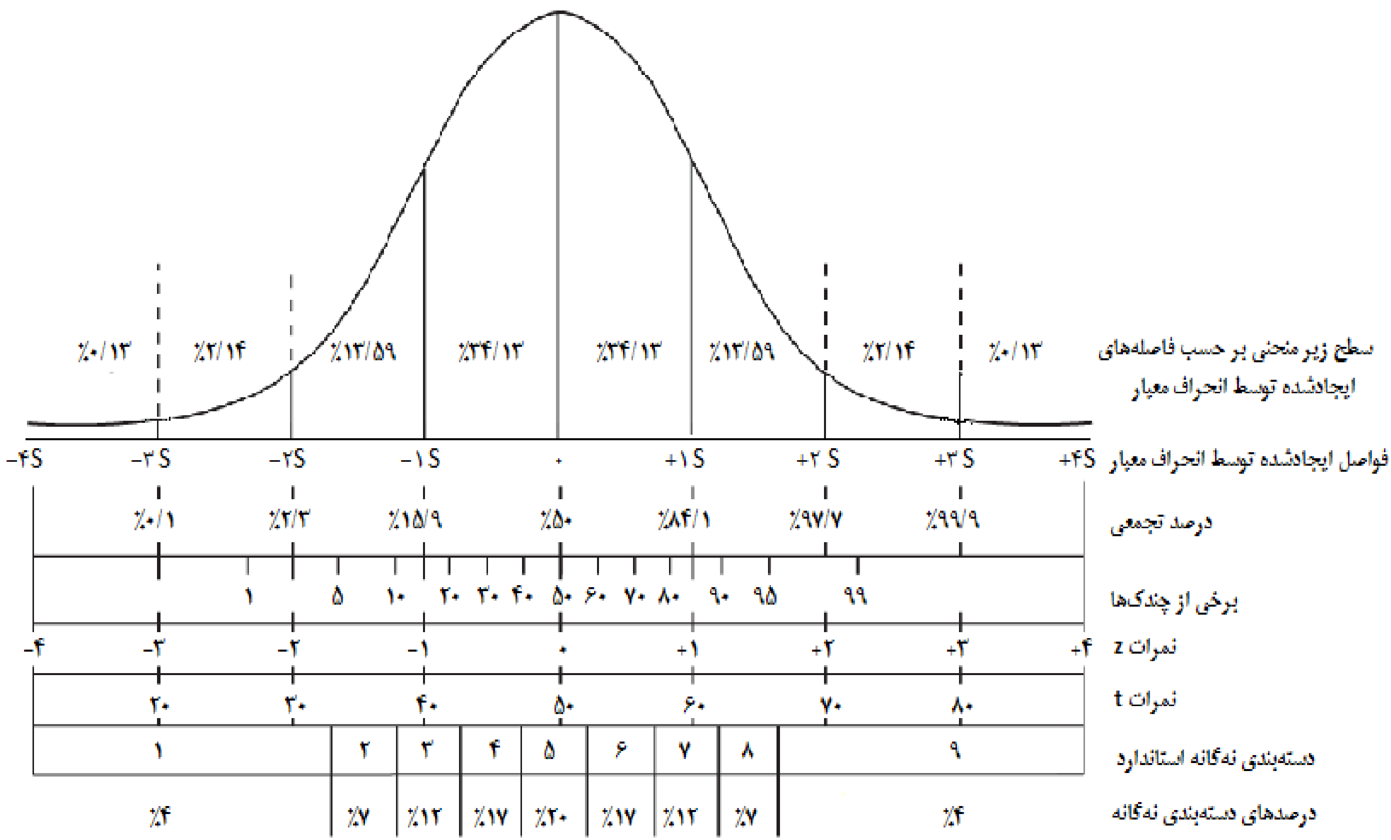
شکل ۲: منحنی توزیع نرمال



منحنی توزیع نرمال، متقارن، با برجستگی صفر است. میانگین، میانه و نمای توزیع برابر هستند. همان طور که در شکل ۳ دیده می‌شود، توزیع از تمرکز نسبتاً بالایی در اطراف میانگین خود برخوردار است؛ زیرا ۶۸ درصد از سطح زیر منحنی در محدودهٔ ۱ انحراف معیار از میانگین و ۹۵/۴٪ درصد از آن در محدودهٔ ۲ انحراف معیار از میانگین قرار دارند و کمتر از ۰/۰۱ درصد از مساحت زیر منحنی در محدوده‌ای خارج از ۴ برابر انحراف معیار از میانگین جای دارند.

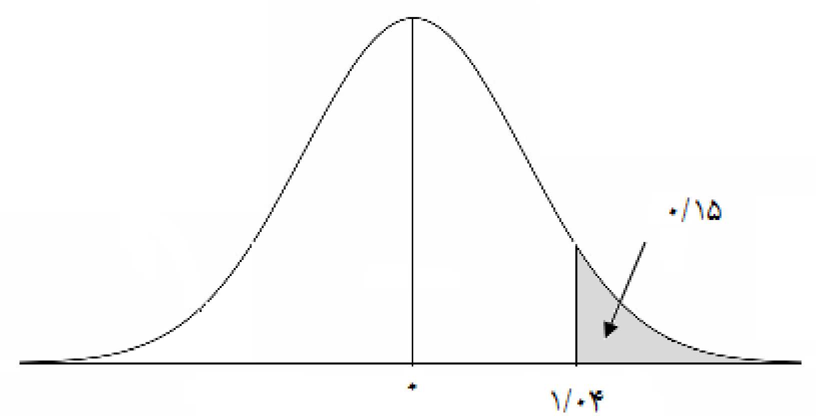
نوع خاصی از توزیع نرمال با انحراف معیارِ ۱ و میانگینِ ۰ را توزیع نرمال استاندارد می‌نامند. در بیشتر کتاب‌های آمار کاربردی، جدولی مانند جدول ۳ مشاهده می‌شود که چندک‌های مختلف توزیع نرمال استاندارد از آن به دست می‌آید. این جدول مساحت سمت راست نقطه‌ای را نشان می‌دهد که از ترکیب سطر و ستون خاکستری جدول مشخص می‌شوند. برای مثال، مساحت پس از مقدار ۰/۰۹ برابر با ۰/۴۶۴۱ است (تلاقی سطر آخر و ستون آخر جدول را مشاهده کنید). در ادامه، روش یافتن چندک‌های مختلف توزیع نرمال استاندارد را با استفاده از این جدول بررسی می‌کنیم.

شکل۳: برخی از مشخصه‌های توزیع نرمال



فرض کنید به دنبال یافتن چندک هشتادوپنجم هستید، یعنی مقداری از متغیر که ۱۵ درصد از داده‌ها بیشتر از آن و ۸۵ درصد از داده‌ها کمتر از آن باشد؛ بر حسب سطح زیر منحنی، این چندک نقطه‌ای در منحنی شکل ۴ است که مساحت پس از آن ۰/۱۵ و مساحت پیش از آن ۰/۸۵ باشد.

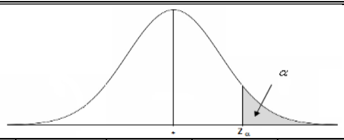
شکل ۴: محل چندک هشتادوپنجم منحنی توزیع نرمال استاندارد



برای یافتن چندک هشتادوپنجم در جدول ۳، خانه‌ای از این جدول را پیدا کنید که نزدیک‌ترین مقدار را به ۰/۱۵ دارد. با کمی جستجو عدد ۰/۱۴۹۲ به عنوان نزدیک‌ترین مقدار به ۰/۱۵ مشخص می‌شود. سطر مربوط به این خانه در جدول ۳، عدد ۱/۰ و ستون مربوط به آن عدد ۰/۰۴ را نشان می‌دهد، پس نقطهٔ ۱/۰۴ همان مقداری در منحنی توزیع نرمال استاندارد است که مساحت پس از آن برابر با ۰/۱۴۹۲ است و می‌تواند به عنوان چندک هشتادوپنجم انتخاب شود (البته ۱/۰۴ درست برابر با چندک هشتادوپنجم نیست، زیرا مساحت ۰/۱۴۹۲ دقیقاً برابر با ۰/۱۵ نیست. با وجود این، ۱/۰۴ به اندازه کافی به مقدار واقعی چندک هشتادوپنجم نزدیک است).

تقارن توزیع نرمال باعث می‌شود اطلاع از چندک‌های نیمی از منحنی توزیع برای یافتن تمام چندک‌های آن کافی باشد. جدول ۳ نیز تنها شامل چندک‌هایی است که پس از میانگین قرار دارند، پس تمام این اعداد مثبت و بزرگ‌ترین مساحت جدول برابر با ۰/۵ است که به میانگین یعنی عدد ۰ اختصاص دارد (به یاد بیاورید که تمام مساحت زیر منحنی توزیع برابر با ۱ است و میانگین در توزیع نرمال همان محور تقارن است که توزیع را به دو نیمهٔ برابر با مساحت ۰/۵ تقسیم می‌کند).

جدول ۳: مساحت سمت راست نقطهٔ z در زیر منحنی توزیع نرمال استاندارد



Z	۰/۰۰	۰/۰۱	۰/۰۲	۰/۰۳	۰/۰۴	۰/۰۵	۰/۰۶	۰/۰۷	۰/۰۸	۰/۰۹
۳/۴	۰/۰۰۰۳	۰/۰۰۰۳	۰/۰۰۰۳	۰/۰۰۰۳	۰/۰۰۰۳	۰/۰۰۰۳	۰/۰۰۰۳	۰/۰۰۰۳	۰/۰۰۰۳	۰/۰۰۰۲
۳/۳	۰/۰۰۰۵	۰/۰۰۰۵	۰/۰۰۰۵	۰/۰۰۰۴	۰/۰۰۰۴	۰/۰۰۰۴	۰/۰۰۰۴	۰/۰۰۰۴	۰/۰۰۰۴	۰/۰۰۰۳
۳/۲	۰/۰۰۰۷	۰/۰۰۰۷	۰/۰۰۰۶	۰/۰۰۰۶	۰/۰۰۰۶	۰/۰۰۰۶	۰/۰۰۰۶	۰/۰۰۰۵	۰/۰۰۰۵	۰/۰۰۰۵
۳/۱	۰/۰۰۱۰	۰/۰۰۰۹	۰/۰۰۰۹	۰/۰۰۰۹	۰/۰۰۰۸	۰/۰۰۰۸	۰/۰۰۰۸	۰/۰۰۰۸	۰/۰۰۰۷	۰/۰۰۰۷
۳/۰	۰/۰۰۱۳	۰/۰۰۱۳	۰/۰۰۱۳	۰/۰۰۱۲	۰/۰۰۱۲	۰/۰۰۱۱	۰/۰۰۱۱	۰/۰۰۱۱	۰/۰۰۱۰	۰/۰۰۱۰
۲/۹	۰/۰۰۱۹	۰/۰۰۱۸	۰/۰۰۱۸	۰/۰۰۱۷	۰/۰۰۱۶	۰/۰۰۱۶	۰/۰۰۱۵	۰/۰۰۱۵	۰/۰۰۱۴	۰/۰۰۱۴
۲/۸	۰/۰۰۲۶	۰/۰۰۲۵	۰/۰۰۲۴	۰/۰۰۲۳	۰/۰۰۲۳	۰/۰۰۲۲	۰/۰۰۲۱	۰/۰۰۲۱	۰/۰۰۲۰	۰/۰۰۱۹
۲/۷	۰/۰۰۳۵	۰/۰۰۳۴	۰/۰۰۳۳	۰/۰۰۳۲	۰/۰۰۳۱	۰/۰۰۳۰	۰/۰۰۲۹	۰/۰۰۲۸	۰/۰۰۲۷	۰/۰۰۲۶
۲/۶	۰/۰۰۴۷	۰/۰۰۴۵	۰/۰۰۴۴	۰/۰۰۴۳	۰/۰۰۴۱	۰/۰۰۴۰	۰/۰۰۳۹	۰/۰۰۳۸	۰/۰۰۳۷	۰/۰۰۳۶
۲/۵	۰/۰۰۶۲	۰/۰۰۶۰	۰/۰۰۵۹	۰/۰۰۵۷	۰/۰۰۵۵	۰/۰۰۵۴	۰/۰۰۵۲	۰/۰۰۵۱	۰/۰۰۴۹	۰/۰۰۴۸
۲/۴	۰/۰۰۸۲	۰/۰۰۸۰	۰/۰۰۷۸	۰/۰۰۷۵	۰/۰۰۷۳	۰/۰۰۷۱	۰/۰۰۶۹	۰/۰۰۶۸	۰/۰۰۶۶	۰/۰۰۶۴
۲/۳	۰/۰۱۰۷	۰/۰۱۰۴	۰/۰۱۰۲	۰/۰۰۹۹	۰/۰۰۹۶	۰/۰۰۹۴	۰/۰۰۹۱	۰/۰۰۸۹	۰/۰۰۸۷	۰/۰۰۸۴
۲/۲	۰/۰۱۳۹	۰/۰۱۳۶	۰/۰۱۳۲	۰/۰۱۲۹	۰/۰۱۲۵	۰/۰۱۲۲	۰/۰۱۱۹	۰/۰۱۱۶	۰/۰۱۱۳	۰/۰۱۱۰
۲/۱	۰/۰۱۷۹	۰/۰۱۷۴	۰/۰۱۷۰	۰/۰۱۶۶	۰/۰۱۶۲	۰/۰۱۵۸	۰/۰۱۵۴	۰/۰۱۵۰	۰/۰۱۴۶	۰/۰۱۴۳
۲/۰	۰/۰۲۲۸	۰/۰۲۲۲	۰/۰۲۱۷	۰/۰۲۱۲	۰/۰۲۰۷	۰/۰۲۰۲	۰/۰۱۹۷	۰/۰۱۹۲	۰/۰۱۸۸	۰/۰۱۸۳
۱/۹	۰/۰۲۸۷	۰/۰۲۸۱	۰/۰۲۷۴	۰/۰۲۶۸	۰/۰۲۶۲	۰/۰۲۵۶	**۰/۰۲۵۰**	۰/۰۲۴۴	۰/۰۲۳۹	۰/۰۲۳۳
۱/۸	۰/۰۳۵۹	۰/۰۳۵۱	۰/۰۳۴۴	۰/۰۳۳۶	۰/۰۳۲۹	۰/۰۳۲۲	۰/۰۳۱۴	۰/۰۳۰۷	۰/۰۳۰۱	۰/۰۲۹۴
۱/۷	۰/۰۴۴۶	۰/۰۴۳۶	۰/۰۴۲۷	۰/۰۴۱۸	۰/۰۴۰۹	۰/۰۴۰۱	۰/۰۳۹۲	۰/۰۳۸۴	۰/۰۳۷۵	۰/۰۳۶۷
۱/۶	۰/۰۵۴۸	۰/۰۵۳۷	۰/۰۵۲۶	۰/۰۵۱۶	۰/۰۵۰۵	۰/۰۴۹۵	۰/۰۴۸۵	۰/۰۴۷۵	۰/۰۴۶۵	۰/۰۴۵۵
۱/۵	۰/۰۶۶۸	۰/۰۶۵۵	۰/۰۶۴۳	۰/۰۶۳۰	۰/۰۶۱۸	۰/۰۶۰۶	۰/۰۵۹۴	۰/۰۵۸۲	۰/۰۵۷۱	۰/۰۵۵۹
۱/۴	۰/۰۸۰۸	۰/۰۷۹۳	۰/۰۷۷۸	۰/۰۷۶۴	۰/۰۷۴۹	۰/۰۷۳۵	۰/۰۷۲۱	۰/۰۷۰۸	۰/۰۶۹۴	۰/۰۶۸۱
۱/۳	۰/۰۹۶۸	۰/۰۹۵۱	۰/۰۹۳۴	۰/۰۹۱۸	۰/۰۹۰۱	۰/۰۸۸۵	۰/۰۸۶۹	۰/۰۸۵۳	۰/۰۸۳۸	۰/۰۸۲۳
۱/۲	۰/۱۱۵۱	۰/۱۱۳۱	۰/۱۱۱۲	۰/۱۰۹۳	۰/۱۰۷۵	۰/۱۰۵۶	۰/۱۰۳۸	۰/۱۰۲۰	۰/۱۰۰۳	۰/۰۹۸۵
۱/۱	۰/۱۳۵۷	۰/۱۳۳۵	۰/۱۳۱۴	۰/۱۲۹۲	۰/۱۲۷۱	۰/۱۲۵۱	۰/۱۲۳۰	۰/۱۲۱۰	۰/۱۱۹۰	۰/۱۱۷۰
۱/۰	۰/۱۵۸۷	۰/۱۵۶۲	۰/۱۵۳۹	۰/۱۵۱۵	۰/۱۴۹۲	۰/۱۴۶۹	۰/۱۴۴۶	۰/۱۴۲۳	۰/۱۴۰۱	۰/۱۳۷۹
۰/۹	۰/۱۸۴۱	۰/۱۸۱۴	۰/۱۷۸۸	۰/۱۷۶۲	۰/۱۷۳۶	۰/۱۷۱۱	۰/۱۶۸۵	۰/۱۶۶۰	۰/۱۶۳۵	۰/۱۶۱۱
۰/۸	۰/۲۱۱۹	۰/۲۰۹۰	۰/۲۰۶۱	۰/۲۰۳۳	**۰/۲۰۰۵**	۰/۱۹۷۷	۰/۱۹۴۹	۰/۱۹۲۲	۰/۱۸۹۴	۰/۱۸۶۷
۰/۷	۰/۲۴۲۰	۰/۲۳۸۹	۰/۲۳۵۸	۰/۲۳۲۷	۰/۲۲۹۶	۰/۲۲۶۶	۰/۲۲۳۶	۰/۲۲۰۶	۰/۲۱۷۷	۰/۲۱۴۸
۰/۶	۰/۲۷۴۳	۰/۲۷۰۹	۰/۲۶۷۶	۰/۲۶۴۳	۰/۲۶۱۱	۰/۲۵۷۸	۰/۲۵۴۶	۰/۲۵۱۴	۰/۲۴۸۳	۰/۲۴۵۱
۰/۵	۰/۳۰۸۵	۰/۳۰۵۰	۰/۳۰۱۵	۰/۲۹۸۱	۰/۲۹۴۶	۰/۲۹۱۲	۰/۲۸۷۷	۰/۲۸۴۳	۰/۲۸۱۰	۰/۲۷۷۶
۰/۴	۰/۳۴۴۶	۰/۳۴۰۹	۰/۳۳۷۲	۰/۳۳۳۶	۰/۳۳۰۰	۰/۳۲۶۴	۰/۳۲۲۸	۰/۳۱۹۲	۰/۳۱۵۶	۰/۳۱۲۱
۰/۳	۰/۳۸۲۱	۰/۳۷۸۳	۰/۳۷۴۵	۰/۳۷۰۷	۰/۳۶۶۹	۰/۳۶۳۲	۰/۳۵۹۴	۰/۳۵۵۷	۰/۳۵۲۰	۰/۳۴۸۳
۰/۲	۰/۴۲۰۷	۰/۴۱۶۸	۰/۴۱۲۹	۰/۴۰۹۰	۰/۴۰۵۲	۰/۴۰۱۳	۰/۳۹۷۴	۰/۳۹۳۶	۰/۳۸۹۷	۰/۳۸۵۹
۰/۱	۰/۴۶۰۲	۰/۴۵۶۲	۰/۴۵۲۲	۰/۴۴۸۳	۰/۴۴۴۳	۰/۴۴۰۴	۰/۴۳۶۴	۰/۴۳۲۵	۰/۴۲۸۶	۰/۴۲۴۷
۰/۰	۰/۵۰۰۰	۰/۴۹۶۰	۰/۴۹۲۰	۰/۴۸۸۰	۰/۴۸۴۰	۰/۴۸۰۱	۰/۴۷۶۱	۰/۴۷۲۱	۰/۴۶۸۱	**۰/۴۶۴۱**

اکنون چگونه می‌توان مثلاً چندک بیست‌وپنجم را از این جدول به دست آورد؟ توجه کنید که باید نقطه‌ای از منحنی را بیابید که مساحت سمت راست آن ۰/۷۵ و مساحت سمت چپ آن ۰/۲۵ است، در حالی که بیشترین مساحت در جدول ۳، برابر با ۰/۵ است.

برای این منظور کافی است چندک مکمل چندک بیست‌وپنجم یعنی چندک هفتادوپنجم را پیدا و آن را قرینه کنید تا خودِ چندک بیست‌وپنجم به دست آید. از آنجا که چندک هفتادوپنجم برابر با ۱/۹۶ است، پس چندک بیست‌وپنجم برابر با قرینهٔ آن یعنی ۱/۹۶- خواهد بود.

همان طور که اشاره شد، جدول ۳ به چندک‌های توزیع نرمال استاندارد اختصاص دارد، پس این پرسش مطرح می‌شود که چگونه می‌توان چندک‌های توزیع نرمالی را که استاندارد نیست، مانند توزیع نرمالی با میانگین ۱۰ و انحراف معیار ۴، به دست آورد؟ خوشبختانه تمام توزیع‌های نرمال با توزیع نرمال استاندارد ارتباط دارند و به این ترتیب می‌توان چندک‌های آنها را به کمک جدول ۳ محاسبه کرد. در واقع، اگر متغیر X دارای توزیع نرمال با میانگین μ و انحراف معیار σ باشد، آنگاه از رابطهٔ زیر می‌توان آن را به متغیر Z با توزیع نرمال استاندارد تبدیل کرد. از همین رو به این کار استاندارد کردن مقادیر متغیر گفته می‌شود.

$$Z = \frac{X - \mu}{\sigma}$$

برای مثال، مقدار ۱۲ در توزیع نرمالی با میانگین ۱۰ و انحراف معیار ۴، معادل مقدار $\frac{۱۲-۱۰}{۴} = ۰/۵$ در توزیع نرمال استاندارد است. در واقع، عدد ۰/۵، مقدار استاندارد عدد ۱۲ است[1]. پس برای اینکه بدانید مساحت زیر منحنی پس از عدد ۱۲ چقدر است، کافی است این مساحت را برای مقدار استاندارد آن از جدول ۳ به دست آورید. از تقاطع سطر ۰/۵ و ستون ۰ در جدول ۳، به رقم ۰/۳۰۸۵ می رسید که همان مساحت سمت راست نقطهٔ ۰/۵ در زیر منحنی توزیع نرمال استاندارد است. به این ترتیب، مساحت تا نقطهٔ ۱۲ نیز برابر با ۰/۳۰۸۵ است.

---

۱. یافتن مقدار «معادل» موضوع جدیدی نیست. برای مثال با یک تناسب ساده می‌توان نمرات درسی را که در مقیاس ۰ تا ۲۰ است، به مقیاس ۰ تا ۱۰۰ یا بر عکس تبدیل کرد. برای نمونه نمرهٔ ۱۰ در مقیاس ۰ تا ۲۰ «معادل» نمرهٔ ۵۰ در مقیاس ۰ تا ۱۰۰ است یا نمرهٔ ۱۳ در مقیاس ۰ تا ۲۰ معادل نمرهٔ ۶۵ در مقیاس ۰ تا ۱۰۰ است. برای تبدیل نمرات ۰ تا ۲۰ به نمرات ۰ تا ۱۰۰ کافی است با یک تناسب ساده، این نمرات در عدد ۵ ضرب شوند. تبدیل مقادیر توزیع‌های نرمال به توزیع نرمال استاندارد نیز با منطق به نسبت مشابهی صورت می‌گیرد.

اکنون فرض کنید به جای یافتن مساحت زیر منحنی در جستجوی چندکی از توزیع نرمال با میانگین μ و انحراف معیار σ هستید. برای این منظور، ابتدا این چندک را در توزیع نرمال استاندارد پیدا کنید و سپس چندک معادل آن را برای توزیع نرمال دلخواه خود از رابطهٔ زیر به دست آورید:

$$X = \mu + Z\sigma$$

برای مثال، اگر در پی یافتن چندک بیستم در توزیع نرمالی با میانگین ۱۰ و انحراف معیار ۴ هستید، ابتدا چندک بیستم را در توزیع نرمال استاندارد به دست آورید. نزدیک‌ترین خانه به عدد ۰/۲ در جدول ۳، عدد ۰/۲۰۰۵ است که در تقاطع سطر ۰/۸ و ستون ۰/۰۴ قرار دارد. بنابراین، چندک بیستم در توزیع نرمال استاندارد $z = ۰/۸۴$ و مقدار معادل آن در توزیع نرمالی با میانگین ۱۰ و انحراف معیار ۴ به کمک رابطهٔ فوق برابر با $۱۳/۳۶ = ۱۰ + (۰/۸۴) \times ۴$ است.

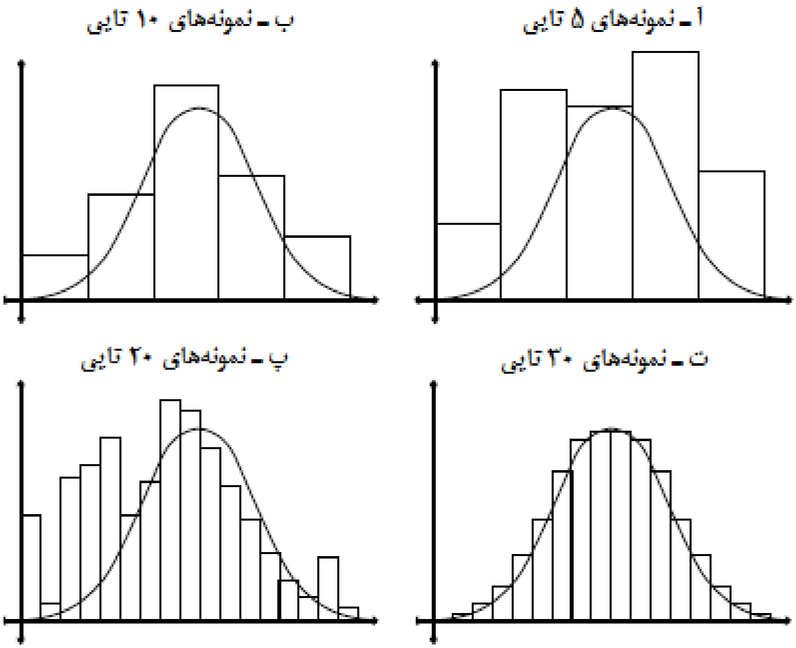
۴ـ توزیع نمونه‌گیری میانگین

یکی از آماره‌های بسیار پرکاربرد، میانگین نمونه است. قانونی در آمار با عنوان «قضیهٔ حد مرکزی» وجود دارد که بیان می‌کند اگر اندازه نمونه «کافی» باشد، توزیع نمونه‌گیری آمارهٔ میانگین، نرمال است. به این ترتیب، بدون نیاز به تشکیل فضای نمونه و محاسبه میانگین تک‌تک نمونه‌های آن و یافتن توزیع این میانگین‌ها، به‌راحتی می‌توان توزیع نمونه‌گیری این آماره را به دست آورد.

کافی بودن اندازه نمونه در قضیهٔ حد مرکزی، تضمین می‌کند که توزیع نمونه‌گیری میانگین‌های نمونه‌ها نرمال باشد. شکل ۵ به روشن شدن نقش اندازه نمونه در توزیع نمونه‌گیری این میانگین‌ها کمک می‌کند. این شکل چهار بافت‌نگار توزیع نمونه‌گیری را نشان می‌دهد که به ازای اندازه نمونه‌های ۵، ۱۰، ۲۰ و ۳۰ تایی رسم شده‌اند. برای مثال، شکل «آ» بافت‌نگار توزیع نمونه‌گیری میانگین‌هایی است که از تمام نمونه‌های ۵ تایی به دست آمده‌اند. همان طور که دیده می‌شود با افزایش اندازه نمونه، بافت‌نگار به منحنی توزیع نرمال شبیه‌تر می‌شود؛ به طوری که بافت‌نگار «ت» که از میانگین‌های تمام نمونه‌های ۳۰ تایی به دست آمده، بر منحنی توزیع نرمال منطبق است.

قضیهٔ حد مرکزی علاوه بر نوع منحنی توزیع میانگین‌های نمونه‌ها، درباره مشخصهٔ تمرکز آنها نیز بیان می‌کند که میانگین توزیع نمونه‌گیری میانگین‌ها همواره برابر با میانگین جامعه‌ای است که نمونه‌ها از آن انتخاب شده‌اند (این موضوع را بر اساس ۱۰ میانگین مثال جدول ۱ بررسی کنید). برای مثال، اگر میانگین توزیع متغیری در جامعه آماری برابر با ۱۰ باشد، آنگاه میانگینِ میانگین‌های حاصل از تمام نمونه‌های ۵ تایی انتخاب‌شده از این جامعه آماری نیز برابر با ۱۰ است.

شکل ۵: نقش اندازه نمونه در توزیع نمونه‌گیری میانگین



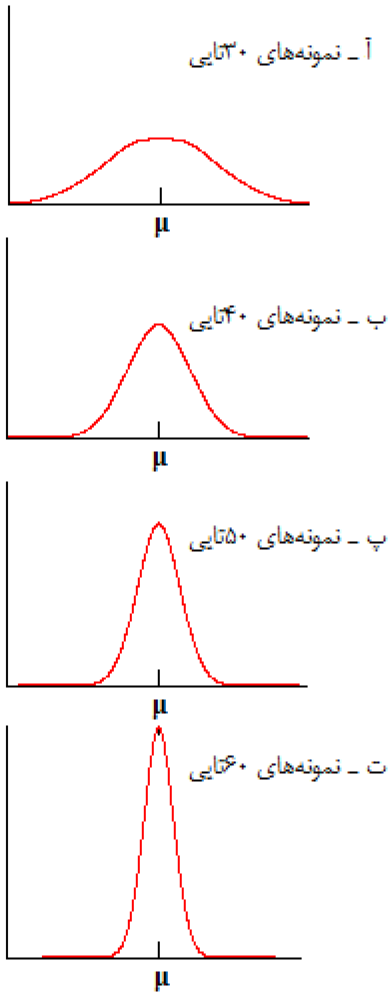
همچنین در قضیهٔ حد مرکزی نشان داده می‌شود که واریانس توزیع نمونه‌گیری میانگین از رابطهٔ زیر به دست می‌آید (واریانس توزیع نمونه‌گیری را با V نشان می‌دهیم):

$$V_{\bar{X}} = \frac{\sigma_X^2}{n}$$

این رابطه بیان می‌کند واریانس میانگین‌های حاصل از نمونه‌های مختلف از تقسیم واریانس متغیر در جامعه بر اندازه نمونه به دست می‌آید. بنابراین، هر چه اندازه نمونه بزرگ‌تر باشد،

واریانس این میانگین‌ها کوچک‌تر خواهد شد، زیرا مخرج کسر یعنی n بزرگ می‌شود. این بدان معناست که هر چه اندازه نمونه‌ها بزرگ‌تر باشند، میانگین‌های حاصل از آنها به هم نزدیک‌تر هستند.

شکل ۶: کاهش واریانس توزیع نمونه‌گیری میانگین با افزایش اندازه نمونه

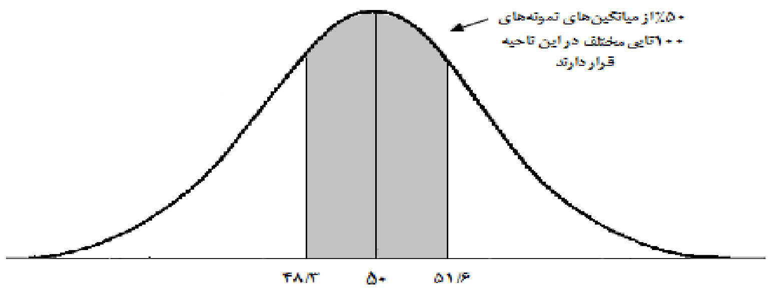


پس در مجموع، قضیهٔ حد مرکزی درباره توزیع نمونه‌گیری میانگین بیان می‌کند که اگر اندازه نمونه «کافی» باشد، آنگاه ۱. توزیع نمونه‌گیری میانگین نرمال است، ۲. میانگین این توزیع برابر با میانگین جامعه آماری است و ۳. واریانس این توزیع از تقسیم واریانس جامعه بر اندازه نمونه به دست می‌آید.

از آنجا که توزیع نمونه‌گیری میانگین‌های نمونه‌های حاصل از جامعه‌ای با انحراف معیار σ و میانگین μ، نرمال با انحراف معیار $\frac{\sigma}{\sqrt{n}}$ و میانگین μ است، پس می‌توان چندک‌های آن را نیز از جدول ۳ به دست آورد؛ زیرا تمام توزیع‌های نرمال به توزیع نرمال استاندارد قابل تبدیل هستند. برای مثال اگر توزیع متغیری در «جامعه» دارای میانگین ۵۰ و انحراف معیار ۲۵ باشد، پیشاپیش خواهید دانست که توزیع نمونه‌گیری میانگین‌های نمونه‌های ۱۰۰ تایی از این جامعه، توزیعی نرمال با میانگین ۵۰ و انحراف معیار $\frac{۲۵}{\sqrt{۱۰۰}}=۲/۵$ است. همچنین با استفاده از فرایند استاندارد کردن به‌راحتی چندک‌های مختلف آن را به دست خواهید آورد؛ برای مثال، چندک بیست‌وپنجم (چارک اول) و هفتادوپنجم (چارک دوم) آن به ترتیب $۵۰+۲/۵\times(-۰/۶۷)=۴۸/۳$ و $۵۰+۲/۵\times۰/۶۷=۵۱/۶$ هستند؛ زیرا چندک بیست‌وپنجم و هفتادوپنجم توزیع نرمال استاندارد با توجه به جدول ۳ به ترتیب ۰/۶۷− و

۰/۶۷ است. بنابراین، ۵۰ درصد از مساحت زیر منحنی توزیع نمونه‌گیری مطابق شکل ۷ بین دو مقدار ۴۸/۳ و ۵۱/۶ قرار دارد.

شکل ۷: توزیع نمونه‌گیری میانگین‌های نمونه‌های ۱۰۰ تایی
از جامعه‌ای با میانگین ۵۰ و انحراف معیار ۲۵

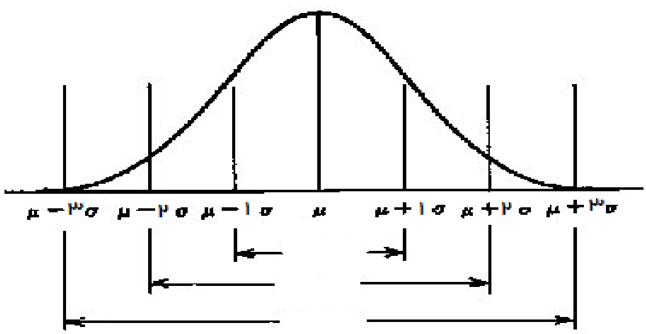


اصطلاحات فصل سوم

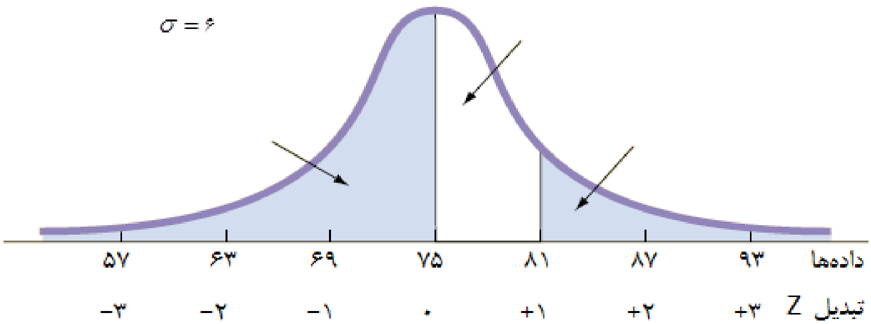
Bell-shape	ناقوسی شکل	Sample Space	فضای نمونه
Central Limit Theorem	قضیه حد مرکزی	Sampling Distribution	توزیع نمونه‌گیری
Error	خطا	Standard Normal Distribution	توزیع نرمال استاندارد
Estimation	برآورد	Statistic	آماره
Parameter	پارامتر	Test of Hypothesis	آزمون فرضیه

تمرین‌های فصل سوم

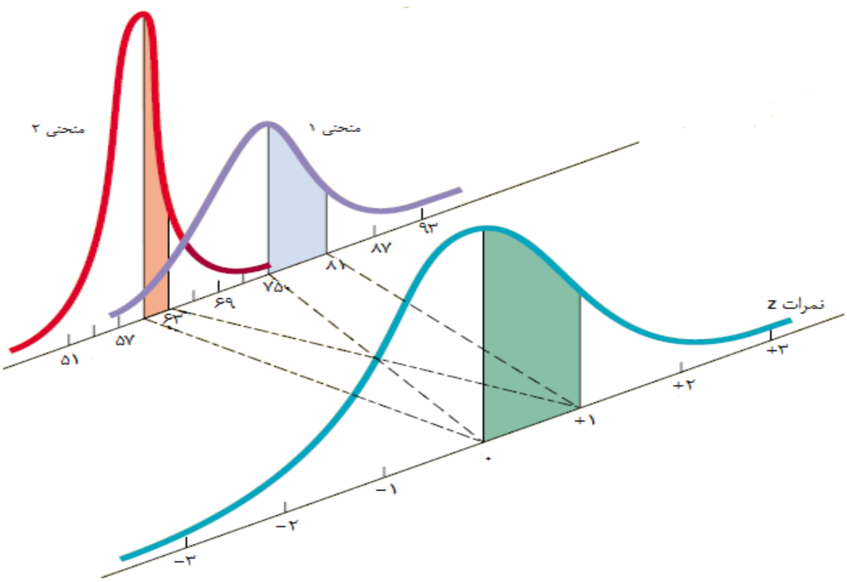
۱. اگر شکل زیر منحنی توزیع نرمالی با میانگین μ و انحراف معیار σ باشد، چند درصد از مساحت زیر منحنی در فاصله‌های تعیین‌شده قرار دارد؟ اکنون فرض کنید منحنی زیر مربوط به توزیع نرمالی با میانگین ۵ و واریانس ۱۰۰ است. شش مقدار متناظر با انحراف از میانگین را در شکل زیر محاسبه کنید.



۲. منحنـی نرمـال زیـر را در نظـر بگیریـد. چنـد درصـد از مسـاحت در ناحیـه‌هـای مشخص‌شده قرار دارد؟



۳. در زیر سه منحنی نرمال نمایش داده شده است. دو نقطـه از منحنـی نمـرات Z (منحنـی نرمـال استاندارد) با خط‌چین به دو منحنی دیگر وصل شده است. بر این اساس، میانگین و انحراف معیـار هر یک از این منحنی‌های توزیع را به دست آورید. مساحت ناحیه‌های مشخص‌شده در هر یـک از این دو منحنی چقدر است؟



۴. در توزیع نرمال استاندارد:

آ) مساحت بین ۰ و ۱/۲ چقدر است؟

ب) مساحت بین ۰/۴۶− و ۲/۲۱ چقدر است؟

پ) مساحت سمت راست ۱/۲۸ چقدر است؟

۵. توزیـع نرمـالی را بـا میـانگین ۷۵ و انحـراف معیـار ۶ در نظـر بگیریـد (مـی‌توانیـد قسمت‌های «چ» و «ح» را پس از یادگیری مفهوم چارک‌ها و میانه حل کنید).

آ) مقدارهای ۶۹ و ۸۱ از این توزیع متناظر با چه مقدارهایی در توزیع نرمال استاندارد هستند؟

ب) چند درصد از سطح زیر منحنی این توزیع کمتر از مقدار ۶۹ است؟

پ) چند درصد از سطح زیر منحنی این توزیع بین ۶۹ و ۸۱ است؟

ت) چند درصد از سطح زیر منحنی این توزیع بزرگ‌تر از ۸۱ است؟

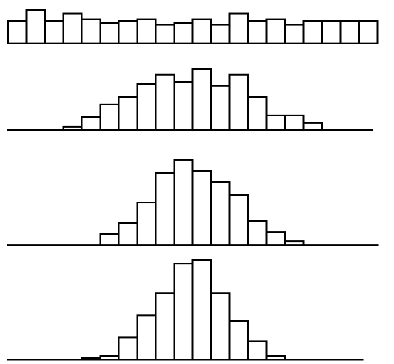
ث) مقدار ۲/۰۲ از توزیع نرمال استاندارد متناظر با چه مقداری از این توزیع است؟

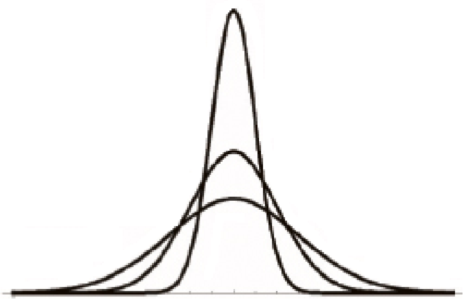
ج) مقدار ۲/۰۲− از توزیع نرمال استاندارد متناظر با چه مقداری از این توزیع است؟

چ) چارک سوم این توزیع چه مقداری است؟

ح) میانه این توزیع چه مقداری است؟

۶. شکل‌های سمت راست و چپ زیر به چه مفاهیم آماری اشاره دارند؟





۷. فرض کنید از سرشماری سال جاری می‌دانیم میانگین روزانه مطالعه مـردم ۵۰ دقیقـه با انحراف معیار ۸ دقیقه است. میانگین و انحراف معیـار توزیـع میـانگین‌هـای نمونـه‌هـای تصادفی ۶۴ نفری چیست؟ بـر ایـن اسـاس، چنـد درصـد از نمونـه‌هـای ۶۴ نفـری دارای میانگینی بین ۴۸ و ۵۲ هستند؟

فصل چهارم

۴ آمار استنباطی برآوردیابی

برآورد برخی مشخصه‌های مهم توزیع متغیر مانند میانگین، درصد و واریانس، موضوع اصلی این فصل است. از این رو، خواننده هم مطالبی درباره برآورد نقطه‌ای و هم برآورد فاصله‌ای این مشخصه‌ها فرا خواهد گرفت. معیارهای یک برآوردگر خوب و کاربرد قضیه حد مرکزی در برآوردیابی، از مباحث مورد تأکید این فصل است. همچنین، دو توزیع پرکاربردِ t و χ^2 و چگونگی استفاده از جدول‌های آنها معرفی خواهند شد.

۱ـ برآوردیابی

ارائه برآوردی قابل اعتماد از پارامترهای نامعلوم جامعه آماری، مثل مشخصه‌های توزیع متغیر بر اساس داده‌های نمونه‌ای از آن جامعه، مسئله‌ای است که در بحث برآوردیابی بررسی می‌شود. برآوردگر همان ابزاری است که به کمک داده‌های نمونه، تخمینی از پارامتر نامعلوم به دست می‌دهد. بنابراین، هر پارامتر برآوردگر خاص خود را دارد؛ برآوردگری که مشخصه‌ای مانند میانگین توزیع جامعه را تخمین می‌زند، متفاوت با برآوردگری است که واریانس این توزیع را تخمین می‌زند.

برآوردگرها دو گونه‌اند: نقطه‌ای و فاصله‌ای. برآوردگر نقطه‌ای مقدار مشخصی را به عنوان تخمین پارامتر نامعلوم جامعه ارائه می‌کند، در حالی که برآوردگر فاصله‌ای مجموعه‌ای از مقادیر را در قالب یک فاصله به عنوان تخمینی از پارامتر به دست می‌دهد. برای مثال، اگر گفته شود برآورد میانگین مدت زمان انتظار در ایستگاه اتوبوس برابر با ۶ دقیقه است، از برآورد نقطه‌ای استفاده شده است، در حالی که اگر گفته شود این میانگین دارای مقداری بین ۳ تا ۹ دقیقه است، از برآورد فاصله‌ای استفاده شده است، زیرا به رقم مشخصی اشاره نشده بلکه طیفی از مقادیر به عنوان تخمین ارائه شده است.

از آنجا که برآوردگر نیز بر اساس داده‌های نمونه محاسبه می‌شود، یک آماره به حساب می‌آید. برآوردگرها خواه نقطه‌ای باشند خواه فاصله‌ای، به دلیل اتکا بر داده‌های نمونه در تخمین پارامتر نامعلومِ جامعه، دور از خطا نیستند و تنها با احتمال معینی می‌توان به تخمین آنها اعتماد کرد. در ادامه، برآوردگر نقطه‌ای و فاصله‌ای را به طور جداگانه بررسی خواهیم کرد.

۲ـ برآوردگر نقطه‌ای

نخستین راهکار برای برآورد پارامترهای نامعلومی مانند میانگین، واریانس یا چولگی توزیع متغیر در جامعه، استفاده از آماره‌های متناظر با آنهاست. برای مثال، میانگین جامعه آماری با میانگین داده‌های یک نمونه ۵۰۰ تایی از آن جامعه تخمین زده شود. با وجود این، مثال جدول ۲ در فصل قبل نشان داد که تخمین حاصل از آماره‌ها را باید با احتیاط به کار برد، زیرا از داده‌های نمونه و نه سرشماری، به دست می‌آید و با خطا همراه است؛ یعنی با مقدار واقعی پارامتر اختلاف دارد.

در آن مثال به‌خوبی اختلاف بین میانگین جامعه و مقدار آماره به عنوان برآوردگر آن مشاهده می‌شود. مثلاً اگر از میانگین نمونه اول برای برآورد میانگین جامعه استفاده شود،

۱واحد، میانگین جامعه را «بیشتر» تخمین زده‌ایم و اگر از میانگین نمونه سوم برای برآورد آن استفاده شود، ۲ واحد، میانگین جامعه را «کمتر» تخمین زده‌ایم. پراکندگی در مقادیر برآوردگر یا همان تغییرپذیری امری اجتناب‌ناپذیر است؛ زیرا این مقادیر برآمده از نمونه‌های مختلفی هستند که هر یک تنها می‌تواند بخشی از اطلاعات جامعه آماری را در خود منعکس سازد. از این رو باید معیارهایی را ارائه کرد که قابل اعتماد بودن یک برآوردگر را تضمین کند. دو معیار رایج برای برآوردگر «خوب»، «نااُریبی» و «کمترین واریانس» نامیده می‌شوند.

اگر برآوردگری نااریب (بدون سوگیری) باشد، نقطهٔ تمرکز توزیع آن درست برابر با همان پارامتری است که برآورد می‌کند. به یاد بیاورید که بهترین مشخصهٔ تمرکز توزیع، میانگین است، پس برآوردگری نااریب قلمداد می‌شود که میانگین تمام برآوردهای حاصل از آن برابر با مقدار پارامتری باشد که برآورد می‌کند. در مثال جدول ۲ در فصل قبل، ۱۰ نمونه ۳ تایی به ۱۰ برآورد برای میانگین جامعه منجر شدند. اکنون بررسی می‌کنیم که آیا میانگین نمونه، برآوردگر نااریبی برای میانگین جامعه است؟

$$\text{میانگینِ میانگین‌ها} = \frac{۱۱+۹/۶۷+۸+۱۰/۳۳+۱۱/۳۳+۱۲+۱۰/۶۷+۸/۳۳+۹/۶۷+۹}{۱۰} = ۱۰$$

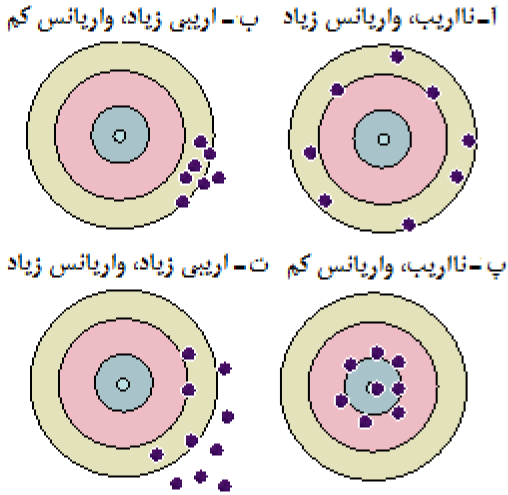
همان طور که دیده می‌شود، میانگینِ این ۱۰ میانگین برابر با میانگینِ جامعه است، پس میانگین را می‌توان برآوردگری نااریب از میانگین جامعه قلمداد کرد؛ زیرا نقطهٔ تمرکز توزیع نمونه‌گیری آن برابر با میانگین جامعه (مقدار پارامتر) است و این با آنچه قضیهٔ حد مرکزی بیان می‌کند، کاملاً سازگاری دارد.

پیشتر اشاره شد که برآوردهای حاصل از یک برآوردگر به دلیل نمونه‌گیری، دارای پراکندگی هستند. حال اگر برآوردگری نااریب باشد، پراکندگی چه نقشی بر دقت تخمین‌های آن خواهد داشت؟ نااریبی تضمین می‌کند که نقطهٔ تمرکز برآوردها برابر با همان پارامتری باشد که برآورد می‌کنند. اکنون اگر پراکندگی برآوردها کوچک باشد، می‌توان نتیجه گرفت که برآوردهای مختلف در اطراف مقدار پارامتر جمع شده‌اند و زیاد با آن اختلاف ندارند. در مقابل، پراکندگی زیاد بیانگر آن است که حداقل برخی برآوردها اختلاف چشمگیری با پارامتر دارند و با خطای زیادی آن را تخمین می‌زنند. از آنجا که

واریانس (انحراف معیار) بهترین مشخصهٔ سنجش پراکندگی است، پس برآوردگر نااریبی «خوب» است که واریانس (انحراف معیار) کوچکی نیز داشته باشد.

در شکل ۱، فرایند برآورد کردن یک پارامتر به شلیک به سوی هدف تشبیه شده است.

شکل ۱: نقش نااریبی و واریانس در دقت برآوردگر



شلیکِ نااریب (بدون سوگیری) به معنای اصابت تیر به اطراف هدف و واریانس کوچک به معنی نزدیک بودن محل‌های اصابت تیر به یکدیگر است. شکل‌های «ب» و «ت» نشان‌دهنده شلیک‌هایی هستند که به سمت خاصی از هدف گرایش (سوگیری) دارند، پس نمی‌توان آنها را نااریب قلمداد کرد (شاید لوله سلاح کج بوده است!). از سوی دیگر، بین این دو، وضعیت شکل «ب» بهتر از شکل «ت» است؛ زیرا لااقل محل‌های اصابت در آن به یکدیگر نزدیکند؛ یعنی واریانس کوچک‌تری دارند. شکل‌های «آ» و «پ» شلیک‌های نااریبی را نشان می‌دهند، زیرا محل‌های اصابت به سمت خاصی از هدف گرایش ندارند و در اطراف هدف هستند، ولی از آنجا که محل‌های اصابت در شکل «پ» به یکدیگر نزدیک‌ترند، وضعیت آن بهتر از وضعیت شکل «آ» است. بنابراین، دقیق‌ترین شلیک‌ها در شکل «پ» و نادقیق‌ترین‌ها در شکل «ت» مشاهده می‌شوند.

تاکنون دریافتیم برآوردگری با واریانس زیاد، به معنای برآوردهای پراکنده و دور از هم برای پارامتر است. اُریبیِ برآوردگر نیز حاکی از گرایش برآوردها به مقداری متفاوت با پارامتر است. ترکیب این دو معیار با هم که با عنوان «میانگینِ مربعِ خطا[1]» یا MSE شناخته می‌شود، می‌تواند معیار کامل‌تری را برای سنجش میزان خطای برآوردگر به دست دهد. پیش از تعریف MSE برآوردگر، نمادهایی را معرفی می‌کنیم که تعریف MSE را ساده‌تر می‌سازد.

1. Mean Squared Error (MSE)

فرض کنید θ (تِتا) پارامتر جامعه‌ای باشد که می‌خواهید برآورد کنید (θ می‌تواند یکی از مشخصه‌های توزیع متغیر مانند میانگین یا واریانس جامعه باشد). فضای نمونه را شامل K نمونه در نظر بگیرید که برآورد θ از نمونه اول $\hat{\theta}_۱$، برآورد آن از نمونه دوم $\hat{\theta}_۲$،... و برآورد آن از نمونه K اُم $\hat{\theta}_k$ است[1]. بر اساس مثال جدول ۲ در فصل قبل، θ میانگین جامعه است، پس $\theta = ۱۰$؛ فضای نمونه نیز شامل ۱۰ نمونه ۳ تایی است، پس $k = ۱۰$ و $\hat{\theta}$ها نیز همان میانگین‌های نمونه‌ها هستند یعنی:

$$\hat{\theta}_۱ = ۱۱,\ \hat{\theta}_۲ = ۹/۶۷,\ \hat{\theta}_۳ = ۸,\ \ldots,\ \hat{\theta}_۹ = ۹/۶۷,\ \hat{\theta}_{۱۰} = ۹$$

با توجه به این نمادگذاری، میانگین برآوردها که با $\overline{\theta}$ نشان داده می‌شود، عبارت است از:

$$\overline{\theta} = \frac{\hat{\theta}_۱ + \hat{\theta}_۲ + \cdots + \hat{\theta}_k}{k} = \frac{۱}{k}\sum_{i=۱}^{k}\hat{\theta}_i$$

پس رابطهٔ زیر برای برآوردگر نااریب برقرار است (برابری میانگین برآوردها با مقدار پارامتر):

$$\overline{\theta} - \theta = ۰ \quad \text{یا} \quad \overline{\theta} = \frac{۱}{k}\sum_{i=۱}^{k}\hat{\theta}_i = \theta$$

به این ترتیب اگر برآوردگر نااریب نباشد، مقدار اریبی (سوگیری) آن که با b نشان می‌دهیم، حاصلِ تفاضل میانگین برآوردها از مقدار پارامتر به صورت زیر است:

$$b = \overline{\theta} - \theta = \overline{\theta} - \frac{۱}{k}\sum_{i=۱}^{k}\hat{\theta}_i = \frac{(\hat{\theta}_۱ - \theta) + (\hat{\theta}_۲ - \theta) + \cdots + (\hat{\theta}_k - \theta)}{k}$$

واریانس برآوردها که با V نشان می‌دهیم نیز از رابطهٔ زیر قابل محاسبه است:

$$V_{\hat{\theta}} = \frac{۱}{k}\sum_{i=۱}^{k}(\hat{\theta}_i - \overline{\theta})^۲$$

از آنجا که بزرگی یا کوچکی انحراف معیار برآوردگر، یعنی جذرِ V، نشانه‌ای از میزان خطای برآوردگر است، در بحث برآوردیابی، به انحراف معیارِ برآوردگر، خطای معیار[2] نیز می‌گویند و آن را با SE نشان می‌دهند.

$$SE_{\hat{\theta}} = \sqrt{V_{\hat{\theta}}} = \sqrt{\frac{۱}{k}\sum_{i=۱}^{k}(\hat{\theta}_i - \overline{\theta})^۲}$$

---

۱. $\hat{\theta}$ را تِتا هَت بخوانید. علامت ^ در آمار برای نشان دادن برآورد و برآوردگر به کار می‌رود.

2. Standard Error (SE)

اکنون به سراغ تعریف MSE برآوردگر و ارتباط آن با اریبی و واریانس برآوردگر می‌رویم. همان طور که از عنوان «میانگینِ مربعِ خطا» برمی‌آید، برای یافتن MSE باید خطای تک‌تک برآوردها محاسبه شود و سپس به توان ۲ برسد و در نهایت، میانگین آنها محاسبه گردد. منظور از خطای برآورد، اختلافی است که برآورد با مقدار پارامتر دارد. به این ترتیب، خطای برآورد نمونه اول، $\hat{\theta}_1 - \theta$، خطای برآورد نمونه دوم، $\hat{\theta}_2 - \theta$، ... و خطای برآورد نمونه K اُم، $\hat{\theta}_k - \theta$ است. بنابراین میانگین مربع این خطاها یا همان MSE به صورت زیر خواهد بود (به تفاوت آن با واریانس برآوردگر توجه کنید):

$$MSE = \frac{(\hat{\theta}_1 - \theta)^2 + (\hat{\theta}_2 - \theta)^2 + \cdots + (\hat{\theta}_k - \theta)^2}{k} = \frac{1}{k}\sum_{i=1}^{k}(\hat{\theta}_i - \theta)^2$$

به‌راحتی می‌توان دریافت که MSE برآوردگر، از جمعِ واریانس برآوردگر با مربع اریبی آن به شکل زیر نیز قابل محاسبه است:

$$MSE = V + b^2$$

بنابراین، اگر برآوردگر نااریب باشد، یعنی $b = 0$، در این صورت MSE همان واریانس برآوردگر خواهد بود.

بحث را با یک مثال بر اساس داده‌های جدول ۲ در فصل قبل پایان می‌دهیم. در آمار توصیفی، بیان شد که واریانس همان «میانگین» مربع اختلاف داده‌ها از میانگین آنهاست، ولی به جای آنکه برای محاسبه این میانگین، مربع اختلاف‌ها را بر تعداد آنها یعنی n تقسیم کنیم، آن را بر $n-1$ تقسیم می‌کنیم؛ یعنی به جای رابطهٔ (۱) در زیر، از رابطهٔ (۲) استفاده می‌شود:

$$(1) \qquad \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2 \qquad\qquad (2) \qquad \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2$$

دلیل این امر در این واقعیت نهفته است که اگر از این دو رابطه به عنوان برآوردگر واریانس جامعه استفاده شود، رابطهٔ (۲) برخلاف رابطه (۱)، برآوردگری نااریب است. ما این ویژگی رابطهٔ (۲) را به کمک مثال جدول ۲ از فصل قبل نشان می‌دهیم.

جدول ۱، واریانس ۱۰ نمونه آن جدول را بر اساس دو رابطۀ (۱) و (۲) ارائه مـی‌دهـد. برای مثال، اگر واریانس داده‌های نمونه اول یعنی ۱۰، ۱۵ و ۸ را با رابطۀ (۱) محاسبه کنیـد، به عدد ۸/۶۷ و اگر با رابطۀ (۲) محاسبه کنید، به عدد ۱۳ خواهید رسید.

در مرحلۀ بعد، میانگین واریانس‌هـای حاصـل از رابطـۀ (۱) و میـانگین واریـانس‌هـای حاصل از رابطۀ (۲) را محاسبه می‌کنیم تا بتوان آن را با واریانس جامعه آماری مقایسه کرد.

جدول ۱: مقایسه اریبی دو روش محاسبه واریانس

رابطه	شماره نمونه در جدول ۲ از فصل قبل										میانگین
	۱	۲	۳	۴	۵	۶	۷	۸	۹	۱۰	
(۱)	۸/۶۷	۱/۵۶	۲/۶۷	۱۳/۵۶	۸/۲۲	۴/۶۷	۱۳/۵۶	۴/۲۲	۱۴/۸۹	۴/۶۷	۷/۶۷
(۲)	۱۳	۲/۳۳	۴	۲۰/۳۳	۱۲/۳۳	۷	۲۰/۳۳	۶/۳۳	۲۲/۳۳	۷	۱۱/۵

همان طور که در ستون آخر جدول ۱ دیده می‌شود، این دو کمّیـت بـرای رابطـۀ (۱) و (۲) به ترتیب ۷/۶۷ و ۱۱/۵ است. اگر میانگین واریانس‌ها با آنچه بـرآورد مـی‌کننـد، یعنـی واریانس جامعه، برابر باشد، برآوردگر واریانس نااریب خواهد بود. اکنون واریـانس جامعـه را بر اساس داده‌های جدول ۲ در فصل قبل به صورت زیر محاسبه می‌کنیم:

$$\sigma^۲ = \frac{(۱۰-۱۰)^۲ + (۱۵-۱۰)^۲ + (۸-۱۰)^۲ + (۱۱-۱۰)^۲ + (۶-۱۰)^۲}{۵-۱} = ۱۱/۵$$

پس رابطۀ (۲) برآوردگری نااریب از واریانس جامعه به دست می‌دهد، در حالی که رابطۀ (۱) دارای $(۱۱/۵ - ۷/۶۷ =) \; ۳/۸۳ -$ اریبی است.

۳ـ برآوردگر فاصله‌ای

پیشتر اشاره شد که مقادیر حاصل از یک برآوردگر به عنوان برآورد پارامتر دارای پراکندگی است. در عمل، تنها یکی از نمونه‌های فضـای نمونـه در دسـت اسـت و بـه ایـن ترتیـب، پژوهشگر تنها به یکی از برآوردهای ممکن دسترسی دارد. ایـن بـرآورد مـی‌توانـد یکـی از برآوردهای نزدیک به پارامتر یا یکی از برآوردهای دور از آن باشد.

جدول ۲، برآوردهای میانگین و برآوردهای واریانس حاصل از ۱۰ نمونـه جامعـه مثـال جدول ۲ در فصل قبل را همراه با خطای آنها بر اساس اطلاعـات ایـن جـدول و جـدول ۱

ارائه کرده است. همان طور که دیده می‌شود، در میان برآوردهای میانگین، برآوردی مانند ۹/۶۷ از نمونه دوم و نهم به چشم می‌خورد که تنها ۰/۳۳− خطا دارد و نیز برآوردی مانند ۸ از نمونه سوم وجود دارد که ۲− خطا دارد. برای برآوردهای واریانس نیز وضعیت مشابهی مشاهده می‌شود؛ برآورد حاصل از نمونه نهم دارای ۱۰/۳۸ خطاست، در حالی که برآورد حاصل از نمونه پنجم ۰/۸۳ خطا دارد. پس بر حسب اینکه کدامیک از نمونه‌ها انتخاب شوند، ممکن است برآوردی کم‌خطا یا پرخطا برای پارامتر جامعه به دست آید.

سردرگمی در تشخیص اینکه برآورد نقطه‌ای حاصل از نمونه، از برآوردهای کم‌خطای پارامتر است یا از برآوردهای پرخطای آن، استفاده از فاصله‌ای از برآوردها را به جای تنها یک برآورد قابل توجه می‌سازد. به عبارت دیگر، هنگامی که فردی نمی‌تواند با تنها یک شلیک، تیر را به نزدیکی هدف بزند، پس تلاش می‌کند چند شلیک به سوی هدف داشته باشد تا در میان این شلیک‌ها تعدادی نیز نزدیک هدف باشند. بنابراین، پژوهشگر به جای آنکه بگوید برآورد میانگین توزیع متغیری مثلاً برابر با ۱۱ است، در حالی که ممکن است مقدار ۱۱ از میانگین واقعی جامعه بسیار کمتر یا بسیار بیشتر باشد، فاصله‌ای از برآوردها را برای میانگین جامعه بیان می‌کند و می‌گوید میانگین جامعه بین ۹ تا ۱۳ است. به این ترتیب، امکان بیشتری وجود دارد که این فاصله شامل میانگین نامعلوم جامعه باشد.

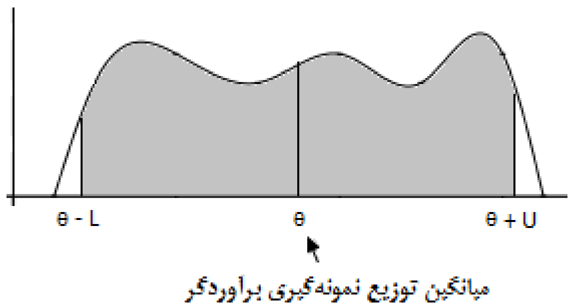
جدول ۲: خطای برآوردهای میانگین و واریانس جامعه مثال جدول ۲ در فصل قبل

پارامتر		شماره نمونه									
		۱	۲	۳	۴	۵	۶	۷	۸	۹	۱۰
میانگین	برآورد	۱۱	۹/۶۷	۸	۱۰/۳۳	۱۱/۳۳	۱۲	۱۰/۶۷	۸/۳۳	۹/۶۷	۹
	خطا	۱	−۰/۳۳	−۲	۰/۳۳	۱/۳۳	۲	۰/۶۷	−۱/۶۷	۰/۳۳	۱
واریانس	برآورد	۱۳	۲/۳۳	۴	۲۰/۳۳	۱۲/۳۳	۷	۲۰/۳۳	۶/۳۳	۲۲/۳۳	۷
	خطا	۱/۵	−۹/۱۷	−۷/۵	۸/۸۳	۰/۸۳	−۴/۵	۹/۳۳	−۵/۱۷	۱۰/۳۸	−۴/۵

منطق برآوردگر فاصله‌ای بر توزیع نمونه‌گیری برآوردگر استوار است. فرض کنید $\hat{\theta}$ برآوردگر نااریبی برای پارامتر نامعلوم θ و توزیع نمونه‌گیری آن مانند منحنی شکل ۲ باشد (چون این برآوردگر نااریب است، پس میانگین توزیع نمونه‌گیری $\hat{\theta}$ برابر با θ است). مساحت ناحیهٔ خاکستری منحنی که بین دو نقطهٔ $\theta+U$ و $\theta-L$ قرار دارد، برابر با ۰/۹۵ است؛

بنابراین ۹۵ درصد از برآوردهای حاصل از این برآوردگر در این فاصله قرار دارند، پس «۹۵ درصد اطمینان» داریم که نامساوی $\theta - L < \hat{\theta} < \theta + U$ برقرار باشد. با یک عملیات سادهٔ ریاضی می‌توان نتیجه گرفت که در ۹۵ درصد از موارد نیز نامساوی $\hat{\theta} - U < \theta < \hat{\theta} + L$ برقرار است. به این نامساوی، برآوردگرِ فاصله‌ای ۹۵ درصدی پارامتر θ می‌گویند و از آنجا که در تفسیر آن گفته می‌شود، ۹۵ درصد «اطمینان» داریم که پارامتر نامعلوم در این «فاصله» قرار داشته باشد، به آن «فاصلهٔ اطمینانِ ۹۵ درصدیِ» پارامتر θ نیز می‌گویند.

شکل ۲: منحنی توزیع نمونه‌گیری برآوردگر پارامتر θ



معمولاً مقادیر L و U بر حسب مضاربی از خطای معیار برآوردگر بیان می‌شوند. بنابراین، برآوردگر فاصله‌ای به صورت $\hat{\theta} - c_1 SE < \theta < \hat{\theta} + c_2 SE$ نوشته می‌شود که در آن SE خطای معیار یعنی جذرِ واریانس برآوردگر و c_1 و c_2 «ضرایب اطمینان» نامیده می‌شوند. همچنین، $\hat{\theta} + c_2 SE$ را حدِ بالای فاصله و $\hat{\theta} - c_1 SE$ را حدِ پایین[1] آن می‌نامند. به این ترتیب، شکل کلی یک برآوردگرِ فاصله‌ای عموماً به صورت زیر است:

(خطای معیار برآوردگر) × ضریب اطمینان + برآورد = حد بالای فاصله

(خطای معیار برآوردگر) × ضریب اطمینان − برآورد = حد پایین فاصله

۱ـ۳ـ برآوردگر فاصله‌ای میانگین

اکنون فاصلهٔ اطمینانی را برای پارامتری مانند میانگین به دست می‌آوریم. به یاد بیاورید که بر اساس قضیهٔ حد مرکزی، توزیع نمونه‌گیری میانگین‌های نمونه‌های حاصل از جامعه‌ای با انحراف معیار σ و میانگین μ، نرمال با انحراف معیار $\frac{\sigma}{\sqrt{n}}$ و میانگین μ است، پس میانگین نمونه، برآوردگر نقطه‌ای نااریبی برای μ قلمداد می‌شود. از آنجا که توزیع نمونه‌گیری این

1. Upper Limit (LL) and Lower Limit (LM)

برآوردگر نرمال است، پس چندک‌های مختلف آن را می‌توان به کمک جدول نرمال استاندارد به دست آورد. اگر $Z_{\frac{\alpha}{۲}}$ مقداری از این جدول باشد که مساحت پس از آن برابر با $\frac{\alpha}{۲}$است، آنگاه قرینهٔ آن یعنی $-Z_{\frac{\alpha}{۲}}$ نیز مقداری از منحنی توزیع نرمال استاندارد است که مساحت پیش از آن $\frac{\alpha}{۲}$ است. بنابراین مساحت بین این دو نقطه $(۱-\alpha)$ است. به این ترتیب مانند شکل ۳، با توجه به آنچه درباره توزیع نمونه‌گیری میانگین و استاندارد کردن بیان شد، مساحت بین دو نقطهٔ $\mu - Z_{\frac{\alpha}{۲}} \frac{\sigma}{\sqrt{n}}$ و $\mu + Z_{\frac{\alpha}{۲}} \frac{\sigma}{\sqrt{n}}$ نیز در منحنی توزیع نمونه‌گیری میانگین برابر با $۱-\alpha$ خواهد بود که به زبان آماری می‌توان نوشت[۱]:

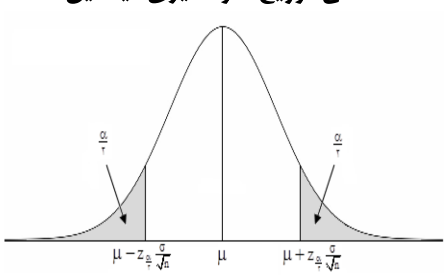
$$P(\mu - z_{\frac{\alpha}{۲}} \frac{\sigma}{\sqrt{n}} < \overline{X} < \mu + z_{\frac{\alpha}{۲}} \frac{\sigma}{\sqrt{n}}) = ۱-\alpha$$

به سادگی می‌توان رابطهٔ فوق را به صورت زیر بازنویسی و فاصلهٔ اطمینان $(۱-\alpha)\times ۱۰۰$ درصدی برای μ ارائه کرد.

$$P(\overline{X} - z_{\frac{\alpha}{۲}} \frac{\sigma}{\sqrt{n}} < \mu < \overline{X} + z_{\frac{\alpha}{۲}} \frac{\sigma}{\sqrt{n}}) = ۱-\alpha$$

برای مثال، اگر از جامعه‌ای با انحراف معیار $\sigma = ۵۰$، نمونه‌ای $n = ۵۰۰$ تایی انتخاب شده باشد که میانگین حاصل از آن $\overline{x} = ۲۰۰$ است، فاصلهٔ اطمینان ۹۰ درصدی برای میانگین جامعه را این گونه به دست می‌آوریم. از آنجا که $(۱-۰/۱)\times ۱۰۰ = ۹۰\%$ است، پس $\alpha = ۰/۱$ و نصف آن برابر با ۰/۰۵ است، بنابراین باید $z_{۰/۰۵}$ را از جدول نرمال استاندارد پیدا کنیم (مقداری که مساحت سمت راست آن ۰/۰۵ باشد). خانهٔ ۰/۰۵۰۵ ما را به سطر ۱/۶ و ستون ۰/۰۴ می‌رساند، پس $z_{۰/۰۵} = ۱/۶۴$ است. مقدار $\frac{\sigma}{\sqrt{n}}$ نیز برابر با ۲/۲۳۶ است و به این ترتیب حدهای بالا و پایین فاصلهٔ اطمینان به شکل زیر به دست می‌آید:

شکل ۳: ناحیه‌ای با مساحت $(۱-\alpha)$ از منحنی توزیع نمونه‌گیری میانگین



حدپایین =۲۰۰−۱/۶۴×۲/۲۳۶=۱۹۶/۳۳   حد بالا=۲۰۰+۱/۶۴×۲/۲۳۶=۲۰۳/۶۷

---

۱. این نماد را این گونه بخوانید: مساحتی از منحنی توزیع نمونه‌گیری $\overline{X}$ که بین دو مقدار $\mu - Z_{\frac{\alpha}{۲}} \frac{\sigma}{\sqrt{n}}$ و $\mu + Z_{\frac{\alpha}{۲}} \frac{\sigma}{\sqrt{n}}$ قرار دارد.

بنابراین، میانگین نامعلوم جامعه بر اساس فاصلهٔ اطمینان ۹۰ درصدی، بین ۱۹۶/۳۳ تا ۲۰۳/۶۷ خواهد بود.

اختلاف حد بالا (UL) و حد پایین (LL) فاصلهٔ اطمینان را پهنای فاصله می‌نامند و هر چه این پهنا کمتر باشد، فاصلهٔ اطمینان دقیق‌تر است. پهنای فاصلهٔ اطمینان میانگین به صورت زیر است. به راحتی می‌توان دید که این پهنا تحت تأثیر سه عامل انحراف معیار جامعه σ، اندازه نمونه n و ضریب اطمینان $Z_{\frac{\alpha}{۲}}$ قرار دارد، به طوری که: ۱. هر چه پراکندگی جامعه بیشتر باشد یعنی σ بزرگ‌تر باشد، فاصله نیز پهن‌تر می‌شود، ۲. هر چه اندازه نمونه بزرگ‌تر باشد، از پهنای فاصله کاسته می‌شود و ۳. هر چه اطمینان به فاصله بیشتر باشد، فاصله پهن‌تر می‌شود.

$$UL - LL = (\overline{X} + z_{\frac{\alpha}{۲}} \frac{\sigma}{\sqrt{n}}) - (\overline{X} - z_{\frac{\alpha}{۲}} \frac{\sigma}{\sqrt{n}}) = ۲\, z_{\frac{\alpha}{۲}} \frac{\sigma}{\sqrt{n}}$$

فاصله‌ای که تاکنون برای میانگینِ نامعلومِ جامعه ارائه شده است، به مقدار σ یعنی انحراف معیار «جامعه» وابسته است. در عمل، همان طور که میانگین جامعه نامعلوم است و تنها از طریق سرشماری می‌توان آن را محاسبه کرد، انحراف معیار جامعه نیز نامعلوم است، یعنی نمی‌توان فاصله اطمینانی را از طریق حدهای بالا و پایین پیشگفته برای میانگین جامعه به دست آورد، بلکه از رابطه‌های زیر که تمام اجزای آن با داده‌های نمونه قابل محاسبه است، استفاده می‌شود:

$$LL = \overline{X} - t_{df,\frac{\alpha}{۲}} \frac{S}{\sqrt{n}} \qquad UL = \overline{X} + t_{df,\frac{\alpha}{۲}} \frac{S}{\sqrt{n}}$$

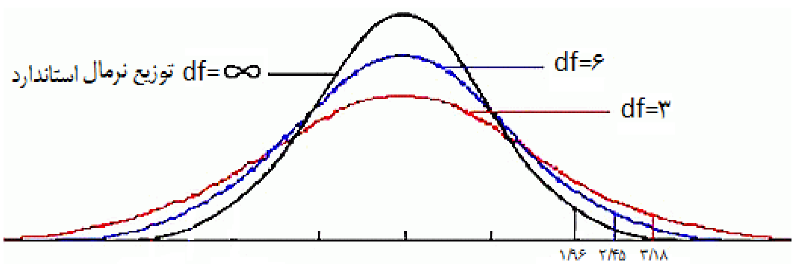
در رابطه‌های جدید، انحراف معیار نمونه جانشین انحراف معیار جامعه شده است و به جای استفاده از ضریب اطمینان z که از جدول نرمال استاندارد به دست می‌آید، از ضریب اطمینان t استفاده می‌شود که در جدول ۳ بر حسب مقدار α و مقدار $df = n - ۱$ ارائه شده است. برای مثال، اگر از جامعه‌ای که هم میانگین و هم انحراف معیار آن نامعلوم است، نمونه‌ای ۱۰۰ تایی انتخاب شده باشد که میانگین آن $\overline{X} = ۲۰۰$ و انحراف معیار آن $S = ۱۸۰$ باشد، فاصلهٔ اطمینان ۹۰ درصدی برای میانگین جامعه را این گونه به دست می‌آوریم: از آنجا که $(۱ - ۰/۱) \times ۱۰۰ = ۹۰\%$ است، پس $\alpha = ۰/۱$ و نصف آن برابر با ۰/۰۵ و از سوی دیگر، $df = ۱۰۰ - ۱ = ۹۹$ است. از همین رو باید $t_{۹۹,۰/۰۵}$ را از جدول ۳ پیدا کنیم که مقداری برابر با ۱/۶۶ دارد (جدول ۳ دارای $df = ۹۹$ نیست، بنابراین از

نزدیک‌ترین مقدار به آن یعنی $df = ۱۰۰$ استفاده می‌کنیم، ولی در انتهای فصل به روش درون‌یابی خطی برای یافتن دقیق آن اشاره می‌شود). مقدار $\frac{S}{\sqrt{n}}$ برابر با ۱۸ است و به این ترتیب حدهای بالا و پایین فاصلهٔ اطمینان به شکل زیر به دست می‌آیند. بنابراین، میانگین نامعلوم جامعه بر اساس فاصلهٔ اطمینان ۹۰ درصدی، مقداری بین ۱۷۰/۱۲ و ۲۲۹/۸۸ دارد.

حدپایین=۲۰۰−۱/۶۶×۱۸=۱۷۰/۱۲   حد بالا=۲۰۰+۱/۶۶×۱۸=۲۲۹/۸۸

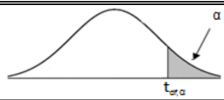
جدول ۳ جدولی مانند جدول نرمال استاندارد (جدول ۳ در فصل قبل) است که چندک‌های منحنی توزیعی را به نام توزیع t به دست می‌دهد. این توزیع نیز مانند توزیع نرمال استاندارد، کاربردهای فراوانی در آمار استنباطی دارد. توزیع t نیز مانند توزیع نرمال، متقارن است و میانگین، میانه و نمای آن برابر هستند، ولی برجستگی آن از توزیع نرمال کمتر است (شکل ۴ را ببینید). جدول ۳ مقادیر متغیری با توزیع t را بر حسب مقدار α و مقدار $df = n - ۱$ ارائه می‌کند.

شکل ۴: مقایسه توزیع t و توزیع نرمال استاندارد



در واقع، همان طور که توزیع‌های نرمال متعددی بر حسب مقدارهای متفاوت میانگین و انحراف معیار وجود دارد، توزیع‌های t متعددی نیز بر حسب مقدارهای مختلف df وجود دارد که برخلاف توزیع نرمال به یکدیگر قابل تبدیل نیستند. شکل ۴ منحنی دو توزیع t با $df = ۳$ و $df = ۶$ را با منحنی توزیع نرمال استاندارد مقایسه می‌کند. همان طور که دیده می‌شود، منحنی‌های توزیع t خوابیده‌تر از توزیع نرمال هستند، پس می‌توان نتیجه گرفت که دُم‌های توزیع t کشیده‌تر از توزیع نرمال است. به همین دلیل است که چندک‌های معادل در توزیع نرمال استاندارد کوچک‌تر از توزیع t است، مثلاً نقطه‌ای از منحنی با مساحت سمت راست ۰/۰۲۵ در توزیع نرمال استاندارد ۱/۹۶، در توزیع t با $df = ۶$ و $df = ۳$ به ترتیب ۲/۴۵ و ۳/۱۸ است. گفتنی است هر چه مقدار df افزایش می‌یابد، منحنی توزیع t به منحنی توزیع نرمال شبیه‌تر می‌شود.

جدول ۳: مساحت سمت راست نقطهٔ $t_{df,\alpha}$ در زیر منحنی توزیع t



α / df	۰/۱۰	۰/۰۵	۰/۰۲۵	۰/۰۱	۰/۰۰۵	۰/۰۰۱	۰/۰۰۰۵
۱	۳/۰۷۸	۶/۳۱۴	۱۲/۷۱	۳۱/۸۲	۶۳/۶۶	۳۱۸/۳۱	۶۳۶/۶۲
۲	۱/۸۸۶	۲/۹۲۰	۴/۳۰۳	۶/۹۶۵	۹/۹۲۵	۲۲/۳۲۷	۳۱/۵۹۹
۳	۱/۶۳۸	۲/۳۵۶	۳/۱۸۲	۴٫۵۴۱	۵/۸۴۱	۱۰/۲۱۵	۱۲/۹۲۴
۴	۱/۵۳۳	۲/۱۳۲	۲/۷۷۶	۳/۷۴۷	۴/۶۰۴	۷/۱۷۳	۸/۶۱۰
۵	۱/۴۷۶	۲/۰۱۵	۲/۵۷۱	۳/۳۶۵	۴/۰۳۲	۵/۸۹۳	۶/۸۶۹
۶	۱/۴۴۰	۱/۹۴۳	۲/۴۴۷	۳/۱۴۳	۳/۷۰۷	۵/۲۰۸	۵/۹۵۹
۷	۱/۴۱۵	۱/۸۹۵	۲/۳۶۵	۲/۹۹۸	۳/۴۹۹	۴/۷۸۵	۵/۴۰۸
۸	۱/۳۹۷	۱/۸۶۰	۲/۳۰۶	۲/۸۹۶	۳/۳۵۵	۴/۵۰۱	۵/۰۴۱
۹	۱/۳۸۳	۱/۸۳۳	۲/۲۶۲	۲/۸۲۱	۳/۲۵۰	۴/۲۹۷	۴/۷۸۱
۱۰	۱/۳۷۲	۱/۸۱۲	۲/۲۲۸	۲/۷۶۴	۳/۱۶۹	۴/۱۴۴	۴/۵۸۷
۱۱	۱/۳۶۳	۱/۷۹۶	۲/۲۰۱	۲/۷۱۸	۳/۱۰۶	۴/۰۲۵	۴/۴۳۷
۱۲	۱/۳۵۶	۱/۷۸۲	۲/۱۷۹	۲/۶۸۱	۳/۰۵۵	۳/۹۳۰	۴/۳۱۸
۱۳	۱/۳۵۰	۱/۷۷۱	۲/۱۶۰	۲/۶۵۰	۳/۰۱۲	۳/۸۵۲	۴/۲۲۱
۱۴	۱/۳۴۵	۱/۷۶۱	۲/۱۴۵	۲/۶۲۴	۲/۹۷۷	۳/۷۸۷	۴/۱۴۰
۱۵	۱/۳۴۱	۱/۷۵۳	۲/۱۳۱	۲/۶۰۲	۲/۹۴۷	۳/۷۳۳	۴/۰۷۳
۱۶	۱/۳۳۷	۱/۷۴۶	۲/۱۲۰	۲/۵۸۳	۲/۹۲۱	۳/۶۸۶	۴/۰۱۵
۱۷	۱/۳۳۳	۱/۷۴۰	۲/۱۱۰	۲/۵۶۷	۲/۸۹۸	۳/۶۴۶	۳/۹۶۵
۱۸	۱/۳۳۰	۱/۷۳۴	۲/۱۰۱	۲/۵۵۲	۲/۸۷۸	۳/۶۱۰	۳/۹۲۲
۱۹	۱/۳۲۸	۱/۷۲۹	۲/۰۹۳	۲/۵۳۹	۲/۸۶۱	۳/۵۷۹	۳/۸۸۳
۲۰	۱/۳۲۵	۱/۷۲۵	۲/۰۸۶	۲/۵۲۸	۲/۸۴۵	۳/۵۵۲	۳/۸۵۰
۲۱	۱/۳۲۳	۱/۷۲۱	۲/۰۸۰	۲/۵۱۸	۲/۸۳۱	۳/۵۲۷	۳/۸۱۹
۲۲	۱/۳۲۱	۱/۷۱۷	۲/۰۷۴	۲/۵۰۸	۲/۸۱۹	۳/۵۰۵	۳/۷۹۲
۲۳	۱/۳۱۹	۱/۷۱۴	۲/۰۶۹	۲/۵۰۰	۲/۸۰۷	۳/۴۸۵	۳/۷۶۸
۲۴	۱/۳۱۸	۱/۷۱۱	۲/۰۶۴	۲/۴۹۲	۲/۷۹۷	۳/۴۶۷	۳/۷۴۵
۲۵	۱/۳۱۶	۱/۷۰۸	۲/۰۶۰	۲/۴۸۵	۲/۷۸۷	۳/۴۵۰	۳/۷۲۵
۲۶	۱/۳۱۵	۱/۷۰۶	۲/۰۵۶	۲/۴۷۹	۲/۷۷۹	۳/۴۳۵	۳/۷۰۷
۲۷	۱/۳۱۴	۱/۷۰۳	۲/۰۵۲	۲/۴۷۳	۲/۷۷۱	۳/۴۲۱	۳/۶۹۰
۲۸	۱/۳۱۳	۱/۷۰۱	۲/۰۴۸	۲/۴۶۷	۲/۷۶۳	۳/۴۰۸	۳/۶۷۴
۲۹	۱/۳۱۱	۱/۶۹۹	۲/۰۴۵	۲/۴۶۲	۲/۷۵۶	۳/۳۹۶	۳/۶۵۹
۳۰	۱/۳۱۰	۱/۶۹۷	۲/۰۴۲	۲/۴۵۷	۲/۷۵۰	۳/۳۸۵	۳/۶۴۶
۴۰	۱/۳۰۳	۱/۶۸۴	۲/۰۲۱	۲/۴۲۳	۲/۷۰۴	۳/۳۰۷	۳/۵۵۱
۶۰	۱/۲۹۶	۱/۶۷۱	۲/۰۰۰	۲/۳۹۰	۲/۶۶۰	۳/۲۳۲	۳/۴۶۰
۸۰	۱/۲۹۲	۱/۶۶۴	۱/۹۹۰	۲/۳۷۴	۲/۶۳۹	۳/۱۹۵	۳/۴۱۶
۱۰۰	۱/۲۹۰	۱/۶۶۰	۱/۹۸۴	۲/۳۶۴	۲/۶۲۶	۳/۱۷۴	۳/۳۹۰
۱۰۰۰	۱/۲۸۲	۱/۶۴۶	۱/۹۶۲	۲/۳۳۰	۲/۵۸۱	۳/۰۹۸	۳/۳۰۰
∞	۱/۲۸۲	۱/۶۴۵	۱/۹۶۰	۲/۳۲۶	۲/۵۷۶	۳/۰۹۰	۳/۲۹۱

منطق فاصلهٔ اطمینان میانگین در شرایطی که انحراف معیار جامعه، نامعلوم است و به جای آن از انحراف معیار نمونه استفاده می‌شود، بر ویژگی خاصی استوار است که با توزیع نرمال و توزیع t ارتباط دارد. این ویژگی بیان می‌کند که اگر متغیر X در جامعه‌ای دارای توزیع نرمال با میانگین μ باشد و نمونه‌ای n تایی از این جامعه انتخاب شده باشد که دارای انحراف معیار S است آنگاه متغیر (آمارهٔ) t که از تبدیل X به صورت زیر به دست می‌آید، دارای توزیع t با $df = n - ۱$ است.

$$t = \frac{X - \mu}{S}$$

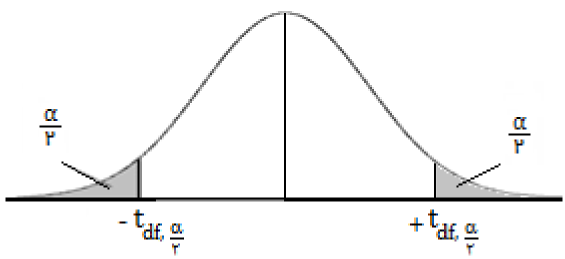
برای مثال اگر ۲۰، یکی از داده‌های نمونه‌ای $n = ۵۰$ نفری با انحراف معیار $S = ۵$ از جامعه‌ای دارای توزیع نرمال با میانگین $\mu = ۱۰$ باشد، آنگاه مقدار t معادل با مقدار ۲۰ عبارت است از $\frac{۲۰-۱۰}{۵} = ۲$.

اکنون این ویژگی را بر روی $\bar{X}$ و S پیاده می‌کنیم که میانگین و انحراف معیار نمونه‌ای n تایی از جامعه‌ای دلخواه با میانگین μ هستند. پیشتر اشاره شد که اگر n به اندازه کافی بزرگ باشد، $\bar{X}$ دارای توزیع نرمال است، پس می‌توان تبدیلِ t برای $\bar{X}$ و $\frac{S}{\sqrt{n}}$ را به صورت زیر نوشت که دارای توزیع t با $df = n - ۱$ است (به یاد بیاورید که طبق قضیهٔ حد مرکزی، انحراف معیار $\bar{X}$ برابر با $\frac{\sigma}{\sqrt{n}}$ است که معادل نمونه‌ای آن $\frac{S}{\sqrt{n}}$ می‌شود).

$$t = \frac{\bar{X} - \mu}{\frac{S}{\sqrt{n}}}$$

برای ساختن فاصلهٔ اطمینان $(۱-\alpha) \times ۱۰۰$ درصدی بر اساس این آماره، باید ناحیه‌ای از منحنی توزیع t تعیین شود که مانند شکل ۵ دارای مساحت $(۱-\alpha)$ باشد.

شکل ۵: ناحیه‌ای با مساحت $(۱-\alpha)$ از منحنی توزیع نمونه‌گیری آمارهٔ t



دو نقطه‌ای که در شکل ۵ دیده می‌شود، همان مرزهایی هستند که مساحت $(1-\alpha)$ را در میان خود جای داده‌اند (به دلیل تقارن منحنی توزیع t، این دو نقطه قرینهٔ یکدیگرند). به عبارت دیگر، $100\times(1-\alpha)$ درصد از مقادیر آمارهٔ t بین این دو نقطه است، یعنی:

$$P(-t_{df,\frac{\alpha}{2}} < t < +t_{df,\frac{\alpha}{2}}) = 1-\alpha$$

با قرار دادن مقدار آمارهٔ t در رابطهٔ فوق به نامساوی زیر می‌رسیم:

$$P(-t_{df,\frac{\alpha}{2}} < \frac{\bar{X}-\mu}{\frac{S}{\sqrt{n}}} < +t_{df,\frac{\alpha}{2}}) = 1-\alpha$$

و با یک عملیات سادهٔ ریاضی به فاصلهٔ اطمینان زیر برای میانگین دست می‌یابیم:

$$P(\bar{X} - t_{df,\frac{\alpha}{2}} \frac{S}{\sqrt{n}} < \mu < \bar{X} + t_{df,\frac{\alpha}{2}} \frac{S}{\sqrt{n}}) = 1-\alpha$$

این همان حد بالا و پایینی است که پیشتر برای فاصلهٔ اطمینان $100\times(1-\alpha)$ درصدی میانگین معرفی شده بود.

۲ـ۳ـ برآوردگر فاصله‌ای نسبت

فرض کنید برآورد نسبتِ وضعیتی از یک متغیر مقوله‌ای در جامعه مورد نیاز است[1]. برای مثال، وضعیت شرکت در انتخابات، متغیری مقوله‌ای با دو وضعیت «شرکت می‌کنم/شرکت نمی‌کنم» است و پژوهشگر مایل است نسبت شرکت‌کنندگان در انتخابات را در جامعه‌ای خاص برآورد کند. برای این منظور نمونه‌ای به اندازه n از این جامعه استخراج می‌شود تا بر اساس داده‌های آن، برآوردی از نسبت ارائه گردد. از آنجا که نسبت، ماهیتی مانند میانگین دارد، می‌توان نسبت حاصل از نمونه را که با p نمایش می‌دهیم، به عنوان برآوردگر نقطه‌ای نااریب برای نسبتِ جامعه قلمداد کنیم. همچنین، ثابت می‌شود که برآوردی از خطای معیار p به صورت $SE=\sqrt{\frac{p(1-p)}{n-1}}$ است. اکنون می‌توان از همان ساختار برآوردگر فاصله‌ای میانگین به شکل زیر برای برآوردگر فاصله‌ای نسبت نیز استفاده کرد:

$$LL = P - t_{df,\frac{\alpha}{2}} \sqrt{\frac{p(1-p)}{n-1}} \qquad UL = P + t_{df,\frac{\alpha}{2}} \sqrt{\frac{p(1-p)}{n-1}}$$

---

۱. پیشتر به تفاوت «نسبت» با «درصد» اشاره شد. در واقع، درصد از ضرب عدد ۱۰۰ در نسبت به دست می‌آید. برای مثال، اگر نسبتِ شرکت‌کنندگان انتخابات ۰/۶۸ باشد، درصدِ شرکت‌کنندگان ۶۸ خواهد بود. بنابراین، با ضرب عدد ۱۰۰ در حد بالا و پایین فاصلهٔ اطمینان مربوط به نسبت، به فاصلهٔ اطمینان مربوط به درصد خواهید رسید.

برای مثال، فرض کنید برای برآورد نسبت مشارکت‌کنندگان در انتخابات نمونه‌ای به اندازه $n=۸۰$ از جامعه انتخاب شده است و از میان آنها ۵۵ نفر در انتخابات شرکت می‌کنند. برای ساختن یک فاصلهٔ اطمینان ۹۰ درصدی برای نسبت شرکت‌کنندگان جامعه، ابتدا نسبت مشارکت‌کنندگان نمونه را به صورت $p=\frac{۵۵}{۸۰}=۰/۶۸۷۵$ برآورد کنید. بر این اساس، خطای معیار عبارت است از:

$$SE=\sqrt{\frac{۰/۶۸۷۵\ (۱-۰/۶۸۷۵)}{۸۰-۱}}=۰/۰۵۲$$

از جدول ۳، مقدار ضریب اطمینان $t_{۷۹,۰/۰۵}=۱/۶۶۴$ به دست می‌آید (در واقع، $n=۸۰$ نزدیک‌ترین مقدار به $n=۷۹$ در جدول ۳ است)، بنابراین حد بالا و پایین فاصلهٔ اطمینان ۹۰ درصدی برای نسبت جامعه عبارت است از:

$$LL=۰/۶۸۷۵-۱/۶۶۴\times ۰/۰۵۲=۰/۶۰۰۹$$
$$UL=۰/۶۸۷۵+۱/۶۶۴\times ۰/۰۵۲=۰/۷۷۴۰$$

براین‌اساس، نسبت نامعلوم شرکت‌کنندگان جامعه بر اساس فاصلهٔ اطمینان ۹۰ درصدی مقداری بین ۰/۶۰ تا ۰/۷۷۶ دارد یعنی، درصد نامعلوم جامعه مقداری بین ۶۰ تا ۷۷ دارد.

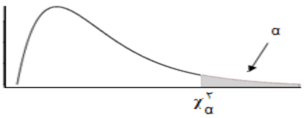
۳ـ۳ـ برآوردگر فاصله‌ای واریانس

برآوردگر فاصله‌ای واریانس اندکی با ساختار برآوردگرهای فاصله‌ای میانگین و درصد تفاوت دارد. فرض کنید نمونه‌ای به اندازه n از جامعه‌ای در دست است که قصد داریم واریانس آن را برآورد کنیم. پیشتر اشاره شد که واریانس نمونه یعنی $S^۲=\frac{۱}{n-۱}\sum_{i=۱}^{n}(X_i-\bar{X})^۲$ برآوردگر نقطه‌ای نااریب برای واریانس جامعه است. بنابراین، از این برآوردگر به شکل زیر برای ساختن یک فاصلهٔ اطمینان $(۱-\alpha)\times ۱۰۰$ درصدی برای واریانس جامعه استفاده می‌کنیم.

$$LL=\frac{(n-۱)S^۲}{\chi^۲_{df,\frac{\alpha}{۲}}} \qquad UL=\frac{(n-۱)S^۲}{\chi^۲_{df,۱-\frac{\alpha}{۲}}}$$

در رابطه‌های فوق $\chi^۲_{df,\frac{\alpha}{۲}}$ (آن را خی‌دو بخوانید) همان ضریب اطمینان است که مقادیر آن در جدول ۴ ارائه شده است؛ مقداری از توزیع که مساحت سمت راست آن $\frac{\alpha}{۲}$ است. توجه

کنید که برخلاف فاصلهٔ اطمینان میانگین و درصد، از دو ضریب اطمینان برای محاسبه حد بالا و حد پایین فاصلهٔ اطمینان واریانس استفاده می‌شود.



جدول ۴: مساحت سمت راست نقطهٔ $\chi^{۲}_{df,\alpha}$ در زیر منحنی توزیع $\chi^{۲}$

df \ α	۰/۹۹۵	۰/۹۹۰	۰/۹۷۵	۰/۹۵۰	۰/۹۰۰	۰/۱۰۰	۰/۰۵۰	۰/۰۲۵	۰/۰۱۰	۰/۰۰۵
۱	۰/۰۰۰	۰/۰۰۰	۰/۰۰۱	۰/۰۰۴	۰/۰۱۶	۲/۷۰۶	۳/۸۴۱	۵/۰۲۴	۶/۶۳۵	۷/۸۷۹
۲	۰/۰۱۰	۰/۰۲۰	۰/۰۵۱	۰/۱۰۳	۰/۲۱۱	۴/۶۰۵	۵/۹۹۱	۳/۳۷۸	۹/۲۱۰	۱۰/۵۹۷
۳	۰/۰۷۲	۰/۱۱۵	۰/۲۱۶	۰/۳۵۲	۰/۵۸۴	۶/۲۵۱	۷/۸۱۵	۹/۳۴۸	۱۱/۳۴۵	۱۱/۸۳۸
۴	۰/۲۰۷	۰/۲۹۷	۰/۴۸۴	۰/۷۱۱	۱/۰۶۴	۷/۷۷۹	۹/۴۸۸	۱۱/۱۴۳	۱۳/۲۷۷	۱۴/۸۶۰
۵	۰/۴۱۲	۰/۵۵۴	۰/۸۳۱	۱/۱۴۵	۱/۶۱۰	۹/۲۳۶	۱۱/۰۷۰	۱۲/۸۳۳	۱۵/۰۸۶	۱۶/۷۵۰
۶	۰/۶۷۶	۰/۸۷۲	۱/۲۳۷	۱/۶۳۵	۲/۲۰۴	۱۰/۶۴۵	۱۲/۵۹۲	۱۴/۴۴۹	۱۶/۸۱۲	۱۸/۵۴۸
۷	۰/۹۸۹	۱/۲۳۹	۱/۶۹۰	۲/۱۶۷	۲/۸۳۳	۱۲/۰۱۷	۱۴/۰۶۷	۱۶/۰۱۳	۱۸/۴۷۵	۲۰/۲۷۸
۸	۱/۳۴۴	۱/۶۴۶	۲/۱۸۰	۲/۷۳۳	۳/۴۹۰	۱۳/۳۶۲	۱۵/۵۰۷	۱۷/۵۳۵	۲۰/۰۹۰	۲۱/۹۵۵
۹	۱/۷۳۵	۲/۰۸۸	۲/۷۰۰	۳/۳۲۵	۴/۱۶۸	۱۴/۶۸۶	۱۶/۹۱۹	۱۹/۰۲۳	۲۱/۶۶۶	۲۳/۵۸۹
۱۰	۲/۱۵۶	۲/۵۵۸	۳/۲۴۷	۳/۹۴۰	۴/۸۶۵	۱۵/۹۸۷	۱۸/۳۰۷	۲۰/۴۸۳	۲۳/۲۰۹	۲۵/۱۸۸
۱۱	۲/۶۰۳	۳/۰۵۳	۳/۸۱۶	۴/۵۷۵	۵/۵۷۸	۱۷/۲۷۵	۱۹/۶۷۵	۲۱/۹۲۰	۲۴/۷۲۵	۲۶/۷۵۷
۱۲	۳/۰۷۴	۳/۵۷۱	۴/۴۰۴	۵/۲۶۶	۶/۳۰۴	۱۸/۵۴۹	۲۱/۰۲۶	۲۳/۳۳۷	۲۶/۲۱۷	۲۸/۳۰۰
۱۳	۳/۵۶۵	۴/۱۰۷	۵/۰۰۹	۵/۸۹۲	۷/۰۴۲	۱۹/۸۱۲	۲۲/۳۶۲	۲۴/۷۳۶	۲۷/۶۸۸	۲۹/۸۱۹
۱۴	۴/۰۷۵	۴/۶۶۰	۵/۶۲۹	۶/۵۷۱	۷/۷۹۰	۲۱/۰۶۴	۲۳/۶۸۵	۲۶/۱۱۹	۲۹/۱۴۱	۳۱/۳۱۹
۱۵	۴/۶۰۱	۵/۲۲۹	۶/۲۶۲	۷/۲۶۱	۸/۵۴۷	۲۲/۳۰۷	۲۴/۹۹۶	۲۷/۴۸۸	۳۰/۵۷۸	۳۲/۸۰۱
۱۶	۵/۱۴۲	۵/۸۱۲	۶/۹۰۸	۷/۹۶۲	۹/۳۱۲	۲۳/۵۴۲	۲۶/۲۹۶	۲۸/۸۴۵	۳۲/۰۰۰	۳۴/۲۶۷
۱۷	۵/۶۹۷	۶/۴۰۸	۷/۵۶۴	۸/۶۷۲	۱۰/۰۸۵	۲۴/۷۶۹	۲۷/۵۸۷	۳۰/۱۹۱	۳۳/۴۰۹	۳۵/۷۱۸
۱۸	۶/۲۶۵	۷/۰۱۵	۸/۲۳۱	۹/۳۹۰	۱۰/۸۶۵	۲۵/۹۸۹	۲۸/۸۶۹	۳۱/۵۲۶	۳۴/۸۰۵	۳۷/۱۵۶
۱۹	۶/۸۴۴	۷/۶۳۳	۸/۹۰۷	۱۰/۱۱۷	۱۱/۶۵۱	۲۷/۲۰۴	۳۰/۱۴۴	۳۲/۸۵۲	۳۶/۱۹۱	۳۸/۵۸۲
۲۰	۷/۴۳۴	۸/۲۶۰	۹/۵۹۱	۱۰/۸۵۱	۱۲/۴۴۳	۲۸/۴۱۲	۳۱/۴۱۰	۳۴/۱۷۰	۳۷/۵۶۶	۳۹/۹۹۷
۲۱	۸/۰۳۴	۸/۸۹۷	۱۰/۲۸۳	۱۱/۵۹۱	۱۳/۲۴۰	۲۹/۶۱۵	۳۲/۶۷۱	۳۵/۴۷۹	۳۸/۹۳۲	۴۱/۴۰۱
۲۲	۸/۶۴۳	۹/۵۴۲	۱۰/۹۸۲	۱۲/۳۳۸	۱۴/۰۴۱	۳۰/۸۱۳	۳۳/۹۲۴	۳۶/۷۸۱	۴۰/۲۸۹	۴۲/۷۹۶
۲۳	۹/۲۶۰	۱۰/۱۹۶	۱۱/۶۸۹	۱۳/۰۹۱	۱۴/۸۴۸	۳۲/۰۰۷	۳۵/۱۷۲	۳۸/۰۷۶	۴۱/۶۳۸	۴۴/۱۸۱
۲۴	۹/۸۸۶	۱۰/۸۵۶	۱۲/۴۰۱	۱۳/۸۴۸	۱۵/۶۵۹	۳۳/۱۹۶	۳۶/۴۱۵	۳۹/۳۶۴	۴۲/۹۸۰	۴۵/۵۵۹
۲۵	۱۰/۵۲۰	۱۱/۵۲۴	۱۳/۱۲۰	۱۴/۶۱۱	۱۶/۴۷۳	۳۴/۳۸۲	۳۷/۶۵۲	۴۰/۶۴۶	۴۴/۳۱۴	۴۶/۹۲۸
۲۶	۱۱/۱۶۰	۱۲/۱۹۸	۱۳/۸۴۴	۱۵/۳۷۹	۱۷/۲۹۲	۳۵/۵۶۳	۳۸/۸۸۵	۴۱/۹۲۳	۴۵/۶۴۲	۴۸/۲۹۰
۲۷	۱۱/۸۰۸	۱۲/۸۷۹	۱۴/۵۷۳	۱۶/۱۵۱	۱۸/۱۱۴	۳۶/۷۴۱	۴۰/۱۱۳	۴۳/۱۹۵	۴۶/۹۶۳	۴۹/۶۴۵
۲۸	۱۲/۴۶۱	۱۳/۵۶۵	۱۵/۳۰۸	۱۶/۹۲۸	۱۸/۹۳۹	۳۷/۹۱۶	۴۱/۳۳۷	۴۴/۴۶۱	۴۸/۲۷۸	۵۰/۹۹۳
۲۹	۱۳/۱۲۱	۱۴/۲۵۶	۱۶/۰۴۷	۱۷/۷۰۸	۱۹/۷۶۸	۳۹/۰۸۷	۴۲/۵۵۷	۴۵/۷۲۲	۴۹/۵۸۸	۵۲/۳۳۶
۳۰	۱۳/۷۸۷	۱۴/۹۵۳	۱۶/۷۹۱	۱۸/۴۹۳	۲۰/۵۹۹	۴۰/۲۵۶	۴۳/۷۷۳	۴۶/۹۷۹	۵۰/۸۹۲	۵۳/۶۷۲
۴۰	۲۰/۷۰۷	۲۲/۱۶۴	۲۴/۴۳۳	۲۶/۵۰۹	۲۹/۰۵۱	۵۱/۸۰۵	۵۵/۷۵۸	۵۹/۳۴۲	۶۳/۶۹۱	۶۶/۷۶۶
۵۰	۲۷/۹۹۱	۲۹/۷۰۷	۳۲/۳۵۷	۳۴/۷۶۴	۳۷/۶۸۹	۶۳/۱۶۷	۶۷/۵۰۵	۷۱/۴۲۰	۷۶/۱۵۴	۷۹,۴۹۰
۶۰	۳۵/۵۳۴	۳۷/۴۸۵	۴۰/۴۸۲	۴۳/۱۸۸	۴۶/۴۵۹	۷۴/۳۹۷	۷۹/۰۸۲	۸۳/۲۹۸	۸۸/۳۷۹	۹۱/۹۵۲
۷۰	۴۳/۲۷۵	۴۵/۴۴۲	۴۸/۷۵۸	۵۱/۷۳۹	۵۵/۳۲۹	۸۵/۵۲۷	۹۰/۵۳۱	۹۵/۰۲۳	۱۰۰/۴۲۵	۱۰۴/۲۱۵
۸۰	۵۱/۱۷۲	۵۳/۵۴۰	۵۷/۱۵۳	۶۰/۳۹۱	۶۴/۲۷۸	۹۶/۵۷۸	۱۰۱/۸۷۹	۱۰۶/۶۲۹	۱۱۲/۳۲۹	۱۱۶/۳۲۱
۹۰	۵۹/۱۹۶	۶۱/۷۵۴	۶۵/۶۴۷	۶۹/۱۲۶	۷۳/۲۹۱	۱۰۷/۵۶۵	۱۱۳/۱۴۵	۱۱۸/۱۳۶	۱۲۴/۱۱۶	۱۲۸/۲۹۹
۱۰۰	۶۷/۳۲۸	۷۰/۰۶۵	۷۴/۲۲۲	۷۷/۹۲۹	۸۲/۳۵۸	۱۱۸/۴۹۸	۱۲۴/۳۴۲	۱۲۹/۵۶۱	۱۳۵/۸۰۷	۱۴۰/۱۶۹

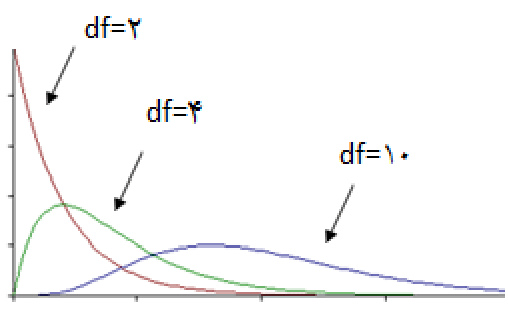
برای مثال، فرض کنید نمونه‌ای ۱۰۰ تایی با واریانس ۵۰ از جامعه‌ای انتخاب شده است تا بر اساس داده‌های آن فاصلهٔ اطمینان ۹۰ درصدی برای واریانس نامعلوم آن جامعه ساخته شود. از آنجا که $(۱-۰/۱)\times۱۰۰=۹۰\%$ است، پس $\alpha=۰/۱$ و نصف آن برابر با ۰/۰۵ است، بنابراین باید $t_{۹۹,۰/۰۵}$ و $t_{۹۹,۰/۹۵}$ را از جدول ۴ پیدا کنیم که به ترتیب برابر با ۱۲۴/۳۴۲ و ۷۷/۹۲۹ هستند (اگرچه $df=۹۹$ است، ولی چون جدول ۴ دارای این df نیست، از نزدیک‌ترین مقدار به آن یعنی $df=۱۰۰$ استفاده می‌کنیم). به این ترتیب حدهای بالا و پایین فاصلهٔ اطمینان به شکل زیر به دست می‌آید:

$$LL=\frac{(۱۰۰-۱)\,۵۰}{۱۲۴/۳۴۲}=۳۹/۸۱ \qquad UL=\frac{(۱۰۰-۱)\,۵۰}{۷۷/۹۲۹}=۶۳/۵۲$$

بنابراین واریانس نامعلوم جامعه بر اساس فاصلهٔ اطمینان ۹۰ درصدی، مقداری بین ۳۹/۸۱ تا ۶۳/۵۲ خواهد داشت.

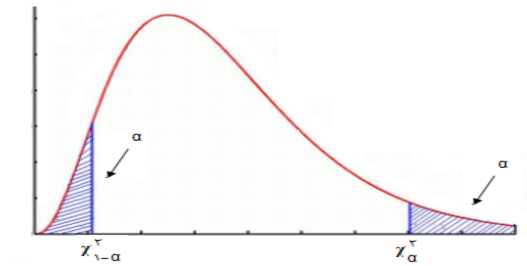
توزیع χ^2 نیز از توزیع‌های پرکاربرد در آمار استنباطی است که چندک‌های مختلف آن به ازای dfها و αهای متفاوت در جدول ۴ ارائه شده است. مانند توزیع t، به ازای dfهای مختلف، توزیع‌های χ^2 متفاوتی نیز وجود دارد که سه منحنی از آن در شکل ۶ نمایش داده شده است. همان‌طور که مشاهده می‌شود، با افزایش df، منحنی توزیع χ^2 به حالت متقارن نزدیک‌تر می‌شود.

شکل ۶: منحنی توزیع χ^2 به ازایِ سه مقدار متفاوت df



برخلاف منحنی توزیع نرمال یا t، منحنی توزیع χ^2 چوله به راست است. همین امر باعث شده است چندک‌های مکمل در توزیع χ^2 قرینهٔ یکدیگر نباشند (شکل ۷ را ببینید). به همین دلیل است که در ساختن فاصلهٔ اطمینان برای واریانس، به یافتن دو مقدار از جدول توزیع χ^2 نیاز است.

شکل ۷: محل قرار گرفتن دو چندکِ مکمل در توزیع $\chi^{۲}$



۴ـ درون‌یابی خطی

در این فصل، به دلیل سهولت، برای یافتن ضریب اطمینان از جدولی مانند جدول ۳ یا جدول ۴، ابتدا به دنبال مقدار مورد نظر می‌گشتیم و در صورت نیافتن آن در جدول، از نزدیک‌ترین df به جای آن استفاده می‌کردیم. برای مثال، از مقدار $t_{۱۰۰,۰/۰۵}=۱/۶۶۰$ به جای $t_{۹۹,۰/۰۵}$ به عنوان ضریب اطمینان استفاده شد، زیرا جدول ۳ شامل $df=۹۹$ نیست و نزدیک‌ترین df به آن، $df=۱۰۰$ است.

راه دقیق‌تر، درون‌یابی خطیِ مقداری است که در جدول وجود ندارد. فرض کنید به دنبال $t_{df,\alpha}$ هستید که در جدول وجود ندارد ولی $t_{L,\alpha}$ و $t_{H,\alpha}$ دو مقدار قبل و بعد از آن هستند که در جدول وجود دارند ($L<df<H$). برای یافتن مقدار مجهولِ $t_{df,\alpha}$، از رابطه زیر استفاده می‌شود (به تناسبِ مثال زیر توجه کنید):

$$t_{df,\alpha}\cong t_{H,\alpha}+\tfrac{H-df}{H-L}\left(t_{L,\alpha}-t_{H,\alpha}\right)$$

برای مثال، جدول ۳ شامل $t_{۳۷,۰/۰۲۵}$ نیست ولی $t_{۳۰,۰/۰۲۵}=۲/۰۴۲$ و $t_{۴۰,۰/۰۲۵}=۲/۰۲۱$ به عنوان دو مقدار قبل و بعد از آن در دسترس هستند، پس $H=۴۰$ و $L=۳۰$. بنابراین:

$$t_{۳۷,۰/۰۲۵}\cong ۲/۰۲۱+\tfrac{۴۰-۳۷}{۴۰-۳۰}\left(۲/۰۴۲-۲/۰۲۱\right)=۲/۰۲۷۳$$

df	t
۳۰	۲/۰۴۲
۳۷	x
۴۰	۲/۰۲۱

۱۰=۴۰−۳۰

۳=۴۰−۳۷

−۰/۰۲۱=۲/۰۲۱−۲/۰۴۲

۲/۰۲۱−x

$$\frac{۱۰}{۳}=\frac{-۰/۰۲۱}{۲/۰۲۱-x}$$

اصطلاحات فصل چهارم

Bias	اریبی	Mean Squared Error (MSE)	میانگین مربع خطا
Chi-square Distribution	توزیع خی‌دو	Minimum Variance	کمترین واریانس
Confidence Interval	فاصله اطمینان	Point Estimator	برآوردگر نقطه‌ای
Estimator	برآوردگر	Standard Error (SE)	خطای معیار
Interval Estimator	برآوردگر فاصله‌ای	t Distribution	توزیع t
Linear Intrapolation	درون‌یابی خطی	Unbias	ناریب
Lower Limit (LL)	حد پایین فاصله	Upper Limit (UL)	حد بالای فاصله

تمرین‌های فصل چهارم

۱. نشان دهید که اگر میانگین خطای برآوردها برابر با ۰ باشد، برآوردگر ناریب است.

۲. بر اساس رابطه MSE، نشان دهید اگر اریبی صفر باشد MSE همان واریانس است.

۳. فرض کنید جامعه‌ای با ۵ عضو مطابق جدول زیر وجود دارد که مقدار متغیر X نیز برای آنها داده شده است.

فرد	جمشید	کامبیز	فریدون	ساسان	شهرام
مقدار X	۵	۹	۳	۴	۸

آ) فضای نمونه تمام نمونه‌های ۳ تایی ممکن از این جامعه را بنویسید (۱۰ نمونه ۳ تایی از این جامعه قابل انتخاب است).

ب) توزیع نمونه‌گیری دامنه را بر اساس این فضای نمونه به دست آورید (جدول فراوانی و نمودار نقطه‌ای).

پ) میانگین توزیع فوق را محاسبه کنید. آیا می‌توان دامنه نمونه را برآوردگر نااریبی برای دامنه جامعه قلمداد کرد؟

۴. فرض کنید دو روش «الف» و «ب» برای برآورد میانه جامعه وجود دارد. در هر یک از حالات زیر، کدام روش بهتر است؟

آ) توزیع نمونه‌گیری روش «الف» دارای اریبی ۱۰ و واریانس ۳ و توزیع نمونه‌گیری روش «ب» دارای اریبی ۵ و واریانس ۷ است.

ب) توزیع نمونه‌گیری روش «الف» نااریب و دارای واریانس ۳ و توزیع نمونه‌گیری روش «ب» دارای اریبی ۱ و واریانس ۲ است.

پ) هر دو روش ناریب هستند، ولی واریانس‌های روش «الف» و «ب» به ترتیب برابر با ۸ و ۵ است.

۵. برای هر یک از حالت‌های زیر فاصله اطمینان محاسبه کنید:

آ) در یک نظرسنجی، به طور تصادفی از ۲۵۶ پاسخگو خواسته شد تا از ۰ تا ۲۰ به کیفیت نوعی خودرو نمره بدهند. اگر واریانس نمرات ۱۰۰ و میانگین آنها ۱۲/۲ باشد، فاصله اطمینانی ۹۵ درصدی برای میانگین ارائه دهید.

ب) از میان ۱۹۶ خانم خانه‌دار که به طور تصادفی انتخاب شده‌اند، تعدا ۱۰۸ نفر از نوعی ماده شوینده استفاده می‌کنند. یک فاصله اطمینان ۹۹ درصدی برای جامعه خانم‌های خانه‌دار استفاده‌کننده از این ماده ارائه دهید.

۶. از نظر کیفی، ابعاد قطعات صنعتی تولیدی باید نزدیک به مقادیر اسمی خود باشد (مقادیر اسمی همان اندازه‌های ایده‌ال برای ابعاد آن قطعه هستند). در کارخانه‌ای، بررسی واحد کنترل کیفیت نشان داده است واریانس قطر نوعی قطعه از فاصله مجاز ۶/۵ تا ۱۸/۱ خارج شده است. از این رو، دستگاه‌های خط تولید این کارخانه تنظیم شدند. برای بررسی اثر این تنظیم، نمونه‌ای ۵۰ تایی به تصادف از قطعات تولیدی این دستگاه‌ها انتخاب شد. واریانس قطر این قطعات، ۱۰/۱ است. به نظر شما تنظیم دستگاه‌ها توانسته است واریانس قطر قطعات را به فاصله مجاز باز گرداند؟

۷. کدام عبارت درباره برآوردگر فاصله‌ای اشتباه است؟

آ) کوتاه‌تر بودن طول فاصله نشان‌دهنده بالاتر بودن دقت آن است.

ب) هر چه اطمینان به فاصله بیشتر باشد، طول آن کوتاه‌تر می‌شود.

پ) خطای فاصله با طول فاصله رابطه مستقیم دارد.

ت) افزایش خطای معیار برآوردگر نقطه‌ای می‌تواند طول برآوردگر فاصله‌ای را افزایش دهد.

ث) اندازه نمونه و طول برآوردگر فاصله‌ای رابطه مستقیم دارند.

۸. آیا می‌توان برای پارامتری برآوردگر فاصله‌ای ۱۰۰ درصدی ارائه داد؟ چرا؟

۹. با درون‌یابی خطی، ضرایب اطمینان فاصله‌های این فصل را بیابید.

فصل پنجم

آمار استنباطی آزمون فرضیه

۵

مبانی آزمون فرضیه و معرفی تعدادی از آزمون‌های آماری (پارامتری) موضوع این فصل است. از آنجا که آزمون‌های آماری بسیار متنوع هستند، تنها به تعداد محدود ولی پرکاربردی از آنها پرداخته می‌شود. با وجود این، خواننده در این فصل هم دربارهٔ آزمون‌های مربوط به میانگین توزیع مانند آزمون t یا تجزیه واریانس مطالبی خواهد یافت و هم دربارهٔ استنباط‌های پیچیده‌تری مانند تحلیل همبستگی و رگرسیون اطلاعاتی به دست خواهد آورد.

۱ـ مفاهیم و اصطلاحات آزمون فرضیه

این بخش سرآغاز یکی از اساسی‌ترین مباحث آمار استنباطی یعنی آزمون فرضیه آماری است. اصطلاحات و مفاهیمی نظیر فرضیه و انواع آن، فرایند عمومی آزمون‌های آماری و خطاهای مرتبط با این آزمون‌ها در این بخش معرفی می‌شوند. بخش‌های بعد به آزمون‌های مختلفی خواهند پرداخت که در استنباط‌های آماری رواج دارند.

۱ـ۱ـ فرضیه آماری

فرضیه به هر ادعایی درباره توزیع متغیرها یا مشخصه‌های آن (پارامترها) در «جامعه آماری» گفته می‌شود که در قالب جمله‌ای خبری (گزاره) بیان شود. بنابراین، جملاتی نظیر «انحراف معیار نمرات هوشبَهر (IQ) حداکثر ۳۰ امتیاز است»، «توزیع شاخص اعتماد به نفس جوانان نرمال است»، «بین معدل دبیرستان و معدل ترم اول دانشجویان رابطهٔ مستقیمی وجود دارد»، «میانگین شاخص سلامت روان متأهلان بیش از مجردان است» یا «توزیع شاخص امید به آینده چوله به راست است»، نمونه‌هایی از فرضیه‌های آماری هستند؛ زیرا ادعایی را درباره خودِ توزیع یا مشخصه‌هایی نظیر انحراف معیار، میانگین یا چولگی مطرح می‌کنند.

در مقابل هر ادعایی، ادعای (فرضیه) دیگری نیز مطرح می‌شود که نقیض آن است. برای مثال، هنگامی که گفته می‌شود «انحراف معیار نمرات هوشبهر حداکثر ۳۰ امتیاز است»، این ادعا (فرضیه) نیز در مقابل آن قابل طرح است که «انحراف معیار نمرات هوشبهر حداکثر ۳۰ امتیاز نیست»، یعنی «انحراف معیار نمرات هوشبهر بیش از ۳۰ امتیاز است». پس در استنباط آماری همواره با دو فرضیه روبه‌رو هستیم که یکی در مقابل دیگری قرار دارد، یعنی نقیض آن است. بنابراین در صورت رد شدن یکی از فرضیه‌ها، فرضیه دیگر جانشین آن خواهد شد.

در آمار استنباطی، فرضیه «ساده‌تر» را «فرضیه صفر» یا «فرضیه خنثی» و فرضیه نقیض آن را «فرضیه جانشین»، «فرضیه یک»، «فرضیه خلاف» یا «فرضیه مقابل» می‌نامند. فرضیه‌ای ساده قلمداد می‌شود که در آن به برابربودن، رابطه‌نداشتن یا بی‌اختلاف‌بودن اشاره شده باشد. بنابراین، «انحراف معیار نمرات هوشبهر حداکثر ۳۰ امتیاز است»، می‌تواند فرضیه صفر باشد، زیرا «حداکثر ۳۰ امتیاز» معادل انحراف معیاری برابر با ۳۰ امتیاز یا کمتر از ۳۰ است، پس حالت تساوی را نیز شامل می‌شود. ولی «میانگین شاخص سلامت روان متأهلان

بیش از مجردان است» نمی‌تواند فرضیه صفر قلمداد شود، زیرا در آن هیچ اشاره‌ای به برابری میانگین شاخص سلامت روان بین متأهلان و مجردان نشده است. این فرضیه درصورتی می‌توانست یک فرضیه صفر باشد که به یکی از صورت‌های «میانگین شاخص سلامت روان متأهلان برابر با مجردان یا بیش از آنان است»، «میانگین شاخص سلامت روان متأهلان برابر با مجردان است» یا «میانگین شاخص سلامت روان متأهلان برابر با مجردان یا کمتر از آنان است» بیان می‌شد.

فرضیه‌های «آماری» با نمادها بیان می‌شوند تا روشن‌تر و دقیق‌تر از جملات نوشتاری باشند. فرضیه‌های یادشده در بند قبل را می‌توان در قالب نمادها به صورت زیر بازنویسی کرد:

انحراف معیار نمرات هوشبهر حداکثر ۳۰ امتیاز است.	$\sigma \leq ۳۰$
میانگین شاخص سلامت روان متأهلان (M) بیش از مجردان (S) است.	$\mu_M > \mu_F$
میانگین شاخص سلامت روان متأهلان برابر با مجردان یا بیش از آنان است.	$\mu_M \geq \mu_F$
میانگین شاخص سلامت روان متأهلان برابر با مجردان است.	$\mu_M = \mu_F$
میانگین شاخص سلامت روان متأهلان برابر با مجردان یا کمتر از آنان است.	$\mu_M \leq \mu_F$

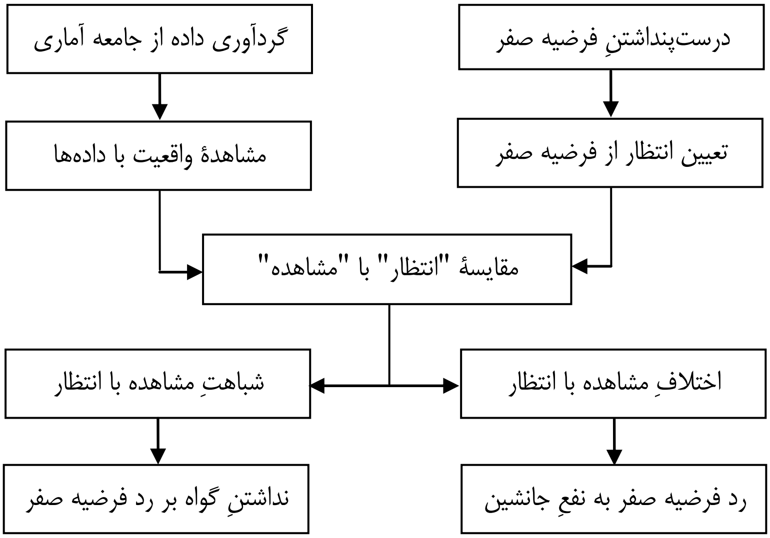
از آنجا که فرضیه‌ها درباره جامعه آماری و نه نمونه هستند، از نمادهای مربوط به پارامترها برای بیان آماری آنها استفاده شده است. توجه کنید که در میان فرضیه‌های فوق، آنهایی فرضیه صفر قلمداد می‌شوند که دارای علامت «تساوی» باشند. پس جز فرضیه دوم، سایر فرضیه‌ها را می‌توان فرضیه صفر به حساب آورد.

۲ـ۱ـ فرایند آزمون‌های آماری

فرایند آزمون‌های آماری مطابق شکل ۱ با درست دانستن فرضیه صفر آغاز می‌شود. پژوهشگر در گام نخست از خود می‌پرسد که اگر فرضیه صفر درست باشد، از جامعه آماری چه انتظاری باید داشته باشم؟ سپس به جامعه آماری مراجعه و در گام دوم، داده‌هایی را در قالب نمونه گردآوری می‌کند. در گام سوم، آنچه را که به کمک داده‌های نمونه از جامعه مشاهده کرده است، با آنچه که در صورتِ درستی فرضیه صفر انتظار داشت، مقایسه می‌کند. اگر واقعیت مشاهده‌شده با واقعیت مورد انتظار اختلاف چشمگیری داشته باشد، به طور منطقی نتیجه می‌گیرد که درست‌پنداشتن فرضیه صفر از ابتدا اشتباه

بوده است، پس فرضیه صفر را به نفعِ فرضیه جانشین رد می‌کند، زیرا آنچه مشاهده کرده نشانه‌ای بر ناسازگاری جامعه آماری با فرضیه صفر است.

شکل ۱: فرایند عمومی آزمون فرضیه آماری



فرایند آزمون‌های آماری به دو معیار نیاز دارد تا قابل اجرا باشد. معیار نخست، چگونگی اندازه‌گیری اختلاف بین «انتظارِ» از درستی فرضیه صفر و «واقعیتِ مشاهده‌شده» را به دست می‌دهد و معیار دوم، «مرزی» را در اختیار می‌گذارد تا بتوان درباره میزان بزرگی این اختلاف قضاوت کرد تا اگر این اختلاف از این مرز فراتر رفت، به رد فرضیه صفر رأی داده شود. معیار اندازه‌گیری اختلاف، آمارهٔ آزمون و مرز بزرگی اختلاف، مقدار بحرانی آزمون نامیده می‌شود.

برای روشن‌تر شدن نقش این دو معیار، این فرضیه صفر را در نظر بگیرید که «میانگین هوشبهر دانشجویان ۱۱۰ است»، یعنی $\mu = ۱۱۰$. پژوهشگری، نمونه‌ای ۴۰۰ نفری از دانشجویان انتخاب می‌کند و میانگین هوشبهر این ۴۰۰ دانشجو $\overline{X} = ۱۱۵_F$ می‌شود. اکنون این پرسش پیش می‌آید که آیا ۱۱۵ به اندازه کافی به ۱۱۰ نزدیک است یا

آمارهٔ آزمون، کمّیتی است که بر اساس داده‌های حاصل از نمونه محاسبه می‌شود. به همین دلیل «آماره» نامیده شده است، پس مانند هر آمارهٔ دیگری دارای توزیع نمونه‌گیری است. در آزمون‌های آماری، همواره توزیع نمونه‌گیری آمارهٔ آزمون در صورتِ درستی فرضیه صفر به دست می‌آید. بنابراین، آمارهٔ آزمون کمّیتی است که از یک سو به داده‌های نمونه یعنی «مشاهده» و از سوی دیگر به فرضیه صفر یعنی «انتظار» بستگی دارد.

چنان با آن اختلاف دارد که باید فرضیه صفر را رد کرد؟ برای پاسخ به این پرسش باید بررسی کنیم که تا چه حد ممکن است میانگینِ نمونه، $۵ = ۱۱۵ - ۱۱۰_F$ واحد از جامعه‌ای با میانگین ۱۱۰ اختلاف داشته باشد. اگر مشاهدهٔ اختلاف ۵ واحدی در چنین جامعه‌ای متداول باشد، پس نمی‌توان دلیلی بر رد فرضیه صفر ارائه کرد، ولی اگر مشاهدهٔ چنین اختلافی در جامعه‌ای با میانگین ۱۱۰ به ندرت اتفاق بیفتد، در این صورت احتمال رد فرضیه صفر نیز وجود دارد. بررسی این واقعیت که اختلاف ۵ واحدی از مواردِ نادر است یا از موارد متداول، تنها با داشتن توزیع نمونه‌گیریِ میانگین امکان‌پذیر است.

اکنون فرض کنید توزیع نمونه‌گیری میانگین بیان می‌کند که ۹۵ درصد از میانگین‌های نمونه‌های انتخاب‌شده از چنین جامعه‌ای بین ۱۰۰ تا ۱۲۲ قرار دارند، پس می‌توان نتیجه گرفت که در ۹۵ درصد از موارد، اختلاف میانگین نمونه از ۱۱۰ نیز بین ۱۰- تا ۱۲+ است و تنها ۵ درصد امکان دارد که این اختلاف کمتر از ۱۰- یا بیشتر از ۱۲+ باشد. بنابراین، مشاهدهٔ اختلافی خارج از فاصلهٔ ۱۰- تا ۱۲+ می‌تواند دلیلی بر رد فرضیه صفر باشد.

در این مثال ساده، آمارهٔ آزمون $\overline{X} - \mu$ یعنی $\overline{X} - ۱۱۰$ و بر اساس توزیع نمونه‌گیری آن، مقادیر بحرانی آزمون نیز ۸- و ۸+ به دست می‌آید. به این ترتیب، مقدار آمارهٔ آزمون بر اساس داده‌های نمونه محاسبه می‌شود و مقدار آن با ۱۰- و ۱۲+ مقایسه می‌گردد تا تعیین شود از آنها فراتر است یا نه؛ در صورت فراتربودن، فرضیه صفر یعنی $\mu = ۱۱۰$ رد می‌شود.

۳ـ۱ـ آزمون‌های یک دامنه و دو دامنه

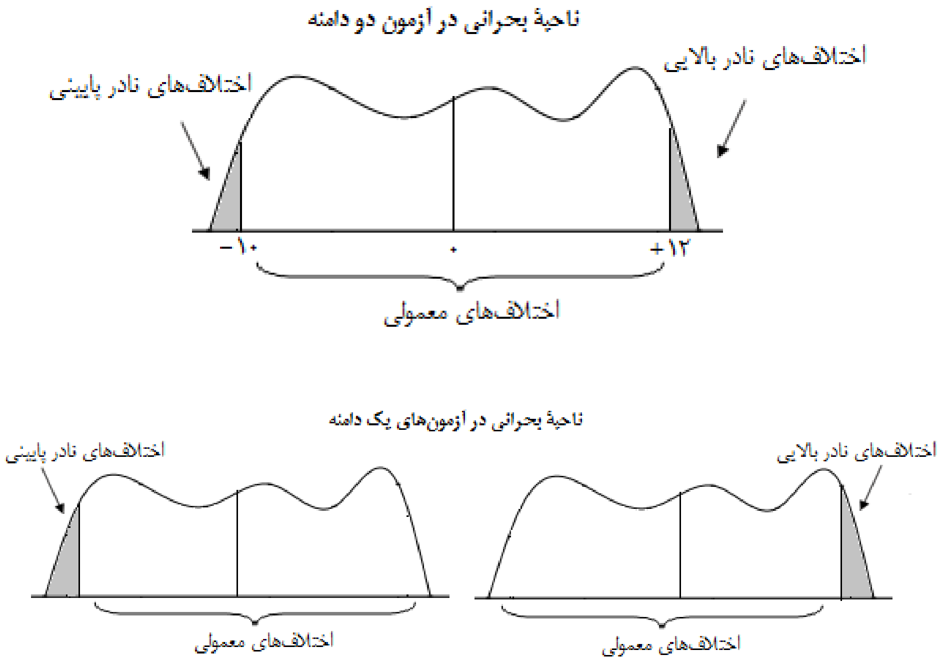
در مثالی که ارائه شد، فرضیه صفر به گونه‌ای بود که هم میانگین‌های نمونه‌ای خیلی بزرگ‌تر از ۱۱۰ و هم میانگین‌های نمونه‌ای خیلی کوچک‌تر از ۱۱۰ می‌توانستند دلیلی بر رد فرضیه صفر $\mu = ۱۱۰$ باشند. از همین رو، دو مقدار بحرانی برای این آزمون مشخص شد تا هم ملاکی برای تشخیص اختلاف‌های بالایی (بیش از ۱۲+) و هم ملاکی برای تشخیص اختلاف‌های پایینی (کمتر از ۱۰-) در دست باشد. آزمون‌هایی را که هم مقادیر بزرگ آمارهٔ آزمون و هم مقادیر کوچک آمارهٔ آزمون می‌توانند موجب رد فرضیه صفر آن شوند، آزمون‌های دو دامنه می‌گویند.

وضعیت برای فرضیه صفری مانند «میانگین هوشبهر دانشجویان حداقل برابر با ۱۱۰ است» یعنی $\mu \geq ۱۱۰$ متفاوت است؛ زیرا تنها میانگین‌های نمونه‌ای خیلی کوچک‌تر از ۱۱۰ می‌توانند به رد این فرضیه صفر منجر شوند و میانگین‌های معمولی یا خیلی بزرگ با

این فرضیه صفر سازگارند. بنابراین، تنها به ارائهٔ یک مقدار بحرانی نیاز است تا ملاکی برای تشخیص اختلاف‌های پایینی باشد. آزمونی را که تنها مقادیر خیلی بزرگ یا مقادیر خیلی کوچک آمارهٔ آزمون دلیلی بر رد فرضیه صفر آن است، آزمون یک دامنه می‌نامند.

آزمون‌های دو دامنه مانند شکل ۲ دارای دو ناحیهٔ بحرانی در دو سوی منحنی توزیع نمونه‌گیری آمارهٔ آزمون هستند، در حالی که آزمون‌های یک دامنه بر حسب نوع فرضیه صفر تنها یک ناحیهٔ بحرانی در یک سوی منحنی توزیع دارند.

شکل ۲: ناحیهٔ بحرانی آزمون‌های یک دامنه و دو دامنه



۴ـ۱ـ خطای نوع اول و دوم، توان آزمون

همان طور که در شکل ۲ نشان داده شد، مقدار بحرانی، منحنی توزیع نمونه‌گیری آمارهٔ آزمون را به دو بخش تقسیم می‌کند: ناحیه‌ای که شامل مقادیر معمولی آماره آزمون است؛ یعنی مقادیری که انتظار داریم در صورتِ درستی فرضیه صفر به طور معمول مشاهده کنیم و ناحیهٔ بحرانی که شامل مقادیری از آماره آزمون است که در صورتِ درستی فرضیه صفر، به ندرت مشاهده می‌شوند.

آزمونی را در نظر بگیرید که ناحیهٔ بحرانی آن شامل ۵ درصد منحنی توزیع نمونه‌گیری آمارهٔ آزمون است و ۹۵ درصد باقیمانده از منحنی، به مقادیر معمولی آماره اختصاص دارد. حال اگر به دلیل قرارگرفتن مقدار آمارهٔ آزمون در ناحیهٔ بحرانی، فرضیه صفر این آزمون رد شده باشد، تصمیمی منطقی درباره فرضیه صفر گرفته‌ایم؛ زیرا مقداری از آمارهٔ آزمون مشاهده شده است که در صورتِ درستی فرضیه صفر دور از انتظار است. با وجود این، نمی‌توان این تصمیم را کاملاً درست دانست، چرا که مقادیر متعلق به ناحیهٔ بحرانی نیز بخشی از منحنی توزیع نمونه‌گیری در صورتِ درستی فرضیه صفر قلمداد می‌شوند و به هر حال امکان مشاهده شدن دارند، حتی اگر این امکان دور از انتظار و بسیار ناچیز یعنی تنها ۵ درصد باشد. بنابراین، ممکن است پژوهشگر به خطا فرضیه صفری را رد کند که واقعاً درست بوده است؛ زیرا مقدار آمارهٔ آزمون مشاهده‌شده، در محدودهٔ دور از انتظار قرار داشته است.

به طور کلی دو نوع خطا نتیجهٔ آزمون فرضیه‌ها را تهدید می‌کند (جدول ۱): رد کردن فرضیه صفری که درست است و رد نکردن فرضیه صفری که نادرست است. این دو نوع خطا را به ترتیب خطای نوع اول و خطای نوع دوم می‌نامند. احتمال ارتکاب خطای نوع اول را با α و احتمال ارتکاب خطای نوع دوم را با β نشان می‌دهند. از آنجا که در مثال فوق، ۵ درصد از مقادیر آمارهٔ آزمون در ناحیهٔ بحرانی قرار دارند، پس $\alpha = ۰/۰۵$ احتمال دارد مرتکب خطای نوع اول شویم؛ یعنی فرضیه صفرِ درستی را به خطا رد کنیم.

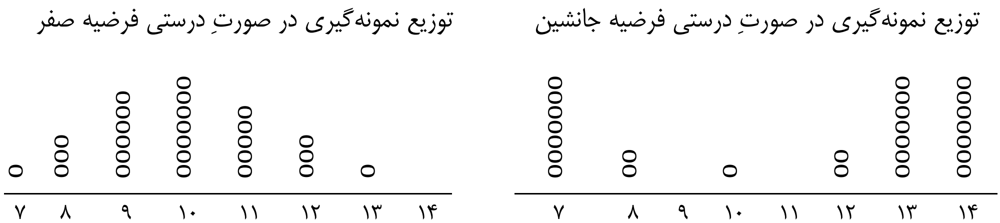
جدول ۱: خطای نوع اول و دوم در آزمون‌های آماری

نتیجهٔ آزمون فرضیه / واقعیت فرضیه صفر	رد کردن فرضیه صفر	رد نکردن فرضیه صفر
فرضیه صفر درست است	خطای نوع اول	نتیجهٔ درست
فرضیه صفر نادرست است	نتیجهٔ درست	خطای نوع دوم

برای روشن شدن مفهوم α و β از مثال سادهٔ شکل ۳ کمک می‌گیریم. فرض کنید توزیع‌های این شکل به آزمون دو دامنه‌ای اختصاص دارند که با مشاهدهٔ مقادیر ۷، ۱۳ و ۱۴ برای آمارهٔ آزمون، فرضیه صفر را رد می‌کند. اگر دایره‌های کوچک بالای هر عدد بیانگر فراوانی آن عدد باشد، مقادیر ۷، ۱۳ و ۱۴ در صورتِ درستی فرضیه صفر به ندرت مشاهده

می‌شوند، زیرا تنها ۲ فراوانی از میان ۲۶ فراوانی به این سه مقدار اختصاص دارد. پس در (۱۰۰×۲/۲۶=) ۷/۸ درصد موارد فرضیه صفر را به اشتباه رد می‌کنیم یعنی α=۰/۰۷۸.

شکل ۳: توزیع نمونه‌گیری آمارهٔ آزمون در صورتِ درستی فرضیه صفر و فرضیه جانشین



برای محاسبه β یعنی احتمال رد نکردن فرضیه صفر درحالی که فرضیه جانشین درست است، باید امکان مشاهدهٔ مقادیری از آمارهٔ آزمون را در صورتِ درستی فرضیه جانشین به دست آوریم که با مشاهدهٔ آنها فرضیه صفر را رد نمی‌کنیم. با توجه به مقادیر ناحیهٔ رد (۷، ۱۳ و ۱۴)، ۸، ۹، ۱۰، ۱۱ و ۱۲ مقادیری از آماره هستند که با مشاهدهٔ آنها فرضیه صفر رد نمی‌شود. فراوانی این مقادیر در توزیع نمونه‌گیری آماره در صورتِ درستی فرضیه «جانشین»، برابر با ۵ فراوانی از ۲۶ فراوانی است. پس در (۱۰۰×۵/۲۶=) ۱۹ درصد از موارد، فرضیه صفرِ نادرست را به نفع فرضیه جانشین رد نمی‌کنیم، یعنی β=۰/۱۹.

> بین احتمال ارتکاب خطای نوع اول و دوم، رابطهٔ معکوس وجود دارد؛ به طوری که اگر احتمال وقوع یکی افزایش یابد، از احتمال وقوع دیگری کاسته می‌شود.

در مباحث آزمون فرضیه، مقدار β−۱ را توان آزمون می‌نامند. در مثالی که گذشت، توان آزمون برابر با (۰/۱۹−۱=) ۰/۸۱ است که بیان می‌کند این آزمون در ۸۱ درصد موارد قادر است فرضیه صفر نادرست را تشخیص دهد و به نفع فرضیه جانشین رد کند.

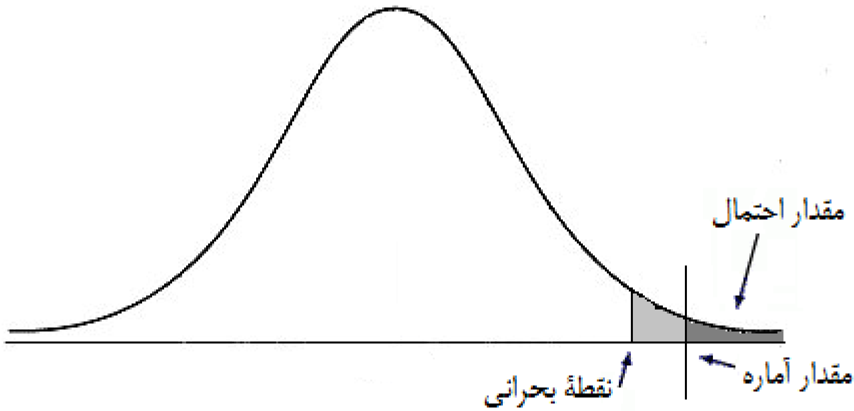
۵ـ۱ـ سطح معناداری و مقدارِ احتمال

شیوهٔ سنتی تصمیم‌گیری درباره رد فرضیه صفر مبتنی بر مقایسهٔ مقدار آمارهٔ آزمون با نقطهٔ بحرانی آزمون است؛ مطابق شکل ۲، اگر مقدار آماره از نقطهٔ بحرانی بزرگ‌تر (دامنهٔ سمت راست) یا کوچک‌تر (دامنهٔ سمت چپ) باشد، فرصیهٔ صفر رد می‌شود. در این شیوه، ابتدا پژوهشگر تعیین می‌کند که چند درصد از مساحت زیر منحنی توزیع نمونه‌گیری آمارهٔ آزمون را به عنوان موارد نادر در نظر بگیرد. سپس ناحیه متناظر با این درصد را در زیر منحنی مشخص می‌کند تا محل نقطهٔ بحرانی آزمون را بیابد (جدول‌هایی مانند جدول ۳ یا

۴ در فصل قبل برای این منظور به کار گرفته می‌شوند). در مرحلهٔ پایانی، مقدار آمارهٔ آزمون را بر اساس داده‌های نمونه محاسبه و آن را با نقطهٔ بحرانی مقایسه می‌کند.

شیوهٔ دیگری که نرم‌افزارهای آماری نیز از آن پیروی می‌کنند، استفاده از مساحت ناحیه‌ها به جای مقدارهای متناظر با آنهاست. برای آشنایی با این شیوه، آزمون یک دامنه‌ای را در نظر بگیرید که مقادیر بزرگ آمارهٔ آزمون می‌توانند به رد فرضیه صفر بینجامند، پس برای تعیین ناحیهٔ رد، با دُم سمت راست منحنی توزیع نمونه‌گیری آمارهٔ آزمون سروکار داریم (شکل ۴).

شکل ۴: مقدار احتمال آزمون یک دامنه در منحنی توزیع نمونه‌گیری آمارهٔ آزمون



فرض کنید ۵ درصدِ انتهای سمت راست نشان‌دهنده مقادیر نادر برای آمارهٔ آزمون باشند، بنابراین نقطهٔ بحرانی محلی است که مساحت سمت راست آن برابر با ۰/۰۵ است. اکنون اگر مقداری که برای آمارهٔ آزمون به دست می‌آید، بزرگ‌تر از نقطهٔ بحرانی باشد، مطابق شکل ۴، در سمت راست نقطهٔ بحرانی قرار می‌گیرد. به این ترتیب، مساحت سمت راست آمارهٔ آزمون کمتر از مساحت سمت راست نقطهٔ بحرانی خواهد بود. پس می‌توان از مقایسهٔ مساحت سمت راست آمارهٔ آزمون با مساحت سمت راست نقطهٔ بحرانی به محل قرار گرفتن آن پی برد و تشخیص داد که مقدار آماره از نقطهٔ بحرانی فراتر رفته است تا فرضیه صفر رد شود. مساحت سمت راست آمارهٔ آزمون را مقدار احتمال و مساحت سمت راست نقطهٔ بحرانی را سطح معناداری آزمون می‌نامند. توجه کنید که سطح معناداری همواره برابر با احتمال وقوع خطای نوع اول آزمون یعنی α است.

۲ـ آزمون‌های مربوط به میانگین‌ها

مشخصهٔ میانگین از پارامترهایی است که در تحلیل‌های آماری متعددی استفاده می‌شود. گاه به برآورد نقطه‌ای یا فاصله‌ای آن و گاه به آزمودن فرضیه‌ای درباره آن نیاز است. آزمون‌های مربوط به میانگین به سه دستهٔ کلی تقسیم می‌شوند: آزمونِ تک میانگین، مقایسهٔ دو میانگین و مقایسهٔ بیش از دو میانگین. هر یک از این آزمون‌ها نیز می‌توانند بر حسب نوع گردآوری داده‌ها و اطلاعات موجود از واریانس جامعه‌های تحت بررسی به انواعی تقسیم شوند که در این بخش به برخی از آنها پرداخته می‌شود.

۱ـ۲ـ آزمونِ تک میانگین

سه نوع فرضیه‌بندی درباره میانگین یک جامعه یعنی μ مطرح می‌شود که یکی آزمونی دو دامنه و بقیه آزمونی یک دامنه هستند. فرض کنید μ_o مقداری مشخص باشد، در این صورت، سه فرضیه‌بندی زیر را خواهیم داشت:

۱. آزمون دو دامنه: فرضیه صفر $\mu = \mu_o$ در مقابل فرضیه جانشین $\mu \neq \mu_o$
۲. آزمون یک دامنه: فرضیه صفر $\mu \geq \mu_o$ در مقابل فرضیه جانشین $\mu < \mu_o$
۳. آزمون یک دامنه: فرضیه صفر $\mu \leq \mu_o$ در مقابل فرضیه جانشین $\mu > \mu_o$

برای مثال، آزمون فرضیه صفرِ «میانگین مدت زمان انتظار برای اتوبوس‌های عمومی حداکثر ۵ دقیقه است» در مقابل فرضیه جانشینِ «میانگین مدت زمان انتظار برای اتوبوس‌های عمومی بیش از ۵ دقیقه است» مانند سومین فرضیه‌بندی آزمونی یک دامنه است، زیرا به صورت فرضیه صفر $\mu \leq ۵$ در مقابل فرضیه جانشین $\mu > ۵$ قابل بیان است، یعنی عدد ۵ به جای μ_o قرار گرفته است.

برای اجرای این آزمون، نمونه‌ای n تایی از جامعه آماری تحت بررسی انتخاب می‌شود. اگر میانگین و انحراف معیار این نمونه به ترتیب $\bar{X}$ و S باشد، آمارهٔ آزمون به شکل زیر محاسبه می‌شود:

$$t = \frac{\bar{X} - \mu_o}{\frac{S}{\sqrt{n}}}$$

به یاد بیاورید که اگر n به اندازه کافی بزرگ باشد، متغیر t دارای توزیع نمونه‌گیری t با $df = n - ۱$ است. پس می‌توان از جدول چندک‌های توزیع t به نقطهٔ بحرانی مناسب برای این آزمون دست یافت (گاهی لازم است نقطه بحرانی از درون‌یابی خطی تعیین شود).

برحسب اینکه $\overline{X}$ بزرگ‌تر یا کوچک‌تر از μ_o باشد، مقدار t می‌تواند مثبت، منفی یا حتی صفر باشد. اگر فرضیه‌بندی اول یعنی آزمون دو دامنه مورد توجه باشد، هم مقادیر خیلی بزرگِ مثبت و هم مقادیر خیلی کوچکِ منفی می‌توانند به رد فرضیه صفر منجر شوند؛ زیرا هم $\overline{X}$های خیلی بزرگ‌تر از μ_o و هم $\overline{X}$های خیلی کوچک‌تر از μ_o با فرضیه صفرِ $\mu = \mu_o$ ناسازگارند. بنابراین، دو نقطهٔ بحرانی برای رد این فرضیه صفر در نظر گرفته می‌شود؛ یکی در سمت راست دُم توزیع به عنوان مرز مقادیر بزرگ آماره و دیگری در سمت چپ دُم توزیع به عنوان مرز مقادیر کوچک آمارهٔ آزمون (شکل ۲ را ببینید).

برای تعیین محل‌های دو نقطهٔ بحرانی در آزمون‌های دو دامنه، سطح معناداری آزمون نصف می‌شود و مانند شکل ۲، نیمی به سمت راست و نیمی به سمت چپ توزیع اختصاص می‌یابد و نقطهٔ متناظر با هر مساحت از جدول توزیع t به دست می‌آید. توجه کنید که به دلیل تقارن منحنی توزیع t، این دو نقطهٔ بحرانی قرینهٔ یکدیگرند و کافی است نقطهٔ بحرانی سمت راست تعیین شود و قرینهٔ آن به عنوان نقطهٔ بحرانی سمت چپ در نظر گرفته شود.

برای مثال می‌توان فرضیه صفر $\mu = ۵$ را در مقابل فرضیه جانشین $\mu \neq ۵$ درباره میانگین مدت زمان انتظار برای اتوبوس‌های عمومی مطرح کرد. فرض کنید نمونه‌ای ۵۰۰ نفری از مسافران درون‌شهری انتخاب و از هر یک، مدت زمان انتظارشان پرسیده شده است. اگر میانگین مدت زمان انتظار این ۵۰۰ نفر ۶/۵ دقیقه و انحراف معیار آنها ۱۵ دقیقه باشد، آمارهٔ آزمون به صورت زیر محاسبه می‌شود:

$$t = \frac{۶/۵ - ۵}{\frac{۱۵}{\sqrt{۵۰۰}}} = ۲/۲۴$$

اگر پژوهشگر تصمیم گرفته باشد آزمونی در سطح معناداری $\alpha = ۰/۰۵$ اجرا کند، یعنی ۵ درصد از مقادیر خیلی بزرگ یا خیلی کوچک آماره را از مقادیر نادر و نامعمول به حساب آورد، باید نیمی از این مقدار یعنی $\frac{\alpha}{۲} = ۰/۰۲۵$ را به عنوان ناحیهٔ بحرانی سمت راست و نیم دیگر را به عنوان ناحیهٔ بحرانی سمت چپ منحنی توزیع آماره در نظر بگیرد

و نقاط بحرانی متناظر با آنها را به ازای $df = ۴۹۹$ از جدول توزیع t به دست آورد. نقطهٔ بحرانی سمت راستی یعنی $t_{۴۹۹,۰/۰۲۵}$ در جدول توزیع t از تقاطع ستون ۰/۰۲۵ و سطر ۱۰۰۰ برابر با ۱/۹۶۲ به دست می‌آید (نزدیک‌ترین df به ۴۹۹، $df = ۱۰۰۰$ است)، پس نقطهٔ بحرانی سمت چپی ۱/۹۶۲- است. بر این اساس، فرضیه صفر را هنگامی رد می‌کنیم که مقدار آمارهٔ آزمون بزرگ‌تر از ۱/۹۶۲ یا کوچک‌تر از ۱/۹۶۲- باشد. از آنجا که ۲/۲۴ از ۱/۹۶۲ بزرگ‌تر است، پس فرضیه صفر $\mu = ۵$ در سطح معناداری ۵ درصد به نفع فرضیه جانشین $\mu \neq ۵$ رد می‌شود؛ یعنی می‌توان ادعا کرد که میانگین مدت زمان انتظار برای اتوبوس‌های عمومی برابر با ۵ دقیقه نیست.

اکنون همین اطلاعات آماری را برای آزمون یک دامنهٔ فرضیه صفر $\mu \leq ۵$ در مقابل فرضیه جانشین $\mu > ۵$ به کار می‌گیریم. محاسبه آمارهٔ آزمون مانند آزمون دو دامنهٔ پیشین است و تنها تفاوت به تعیین نقطهٔ بحرانی باز می‌گردد. در این آزمون یک دامنه، $\overline{X}$های خیلی بزرگ‌تر از ۵ می‌توانند دلیلی بر رد فرضیه صفر باشند؛ پس یک نقطهٔ بحرانی در سمت راست توزیع کافی است تا مرز تشخیص بزرگ بودن مقدار آماره آزمون باشد. از این رو مقدار α را نصف نمی‌کنیم و در جدول توزیع t برای سطح معناداری $\alpha = ۰/۰۵$ به دنبال $t_{۴۹۹,۰/۰۵}$ به عنوان نقطهٔ بحرانی خواهیم بود. از تقاطع ستون ۰/۰۵ و سطر ۱۰۰۰ به رقم ۱/۶۴۶ خواهید رسید. از آنجا که مقدار آمارهٔ ۲/۲۴ از ۱/۶۴۶ بزرگ‌تر است، پس فرضیه صفر $\mu \leq ۵$ را به نفع فرضیه جانشین $\mu > ۵$ در سطح ۵ درصد رد می‌کنیم و نتیجه می‌گیریم که میانگین مدت زمان انتظار برای اتوبوس‌های عمومی بیش از ۵ دقیقه است.

آزمون یک دامنه بعدی، عکس آزمون فوق است؛ آزمون فرضیه صفر $\mu \geq ۵$ در مقابل فرضیه جانشین $\mu < ۵$. آمارهٔ آزمون مانند قبل محاسبه می‌شود، ولی نقطهٔ بحرانی آن قرینهٔ نقطهٔ بحرانی آزمون یک دامنهٔ پیشین است؛ زیرا در این آزمون مقادیر کوچک آمارهٔ آزمون با فرضیه صفر ناسازگار هستند. بر این اساس، نقطهٔ بحرانی، ۱/۶۴۶- است و مقدار آماره (۲/۲۴) از آن کوچک‌تر نیست، پس نمی‌توان فرضیه صفر $\mu \geq ۵$ را به نفع فرضیه جانشین $\mu < ۵$ رد کرد؛ زیرا داده‌ها اختلافی را بین انتظار ما از درستی فرضیه صفر با میانگین مشاهده‌شده از نمونه نشان نمی‌دهند (فرضیه صفر می‌گوید میانگین حداقل برابر با ۵ است و میانگین مشاهده شده از نمونه نیز ۶/۵ است که با فرضیه صفر سازگاری دارد).

این سه آزمون با روش سنتی یعنی مقایسهٔ مقدار آمارهٔ آزمون با نقطهٔ بحرانی به دست‌آمده از جدول توزیع t صورت گرفتند. اکنون آنها را با شیوهٔ مقدار احتمال اجرا می‌کنیم. به کمک نرم‌افزارهای آماری، مقدار احتمال یعنی مساحت سمت راست آمارهٔ آزمون ۲/۲۴ برابر با ۰/۰۱۲۸ است. این رقم برای دو فرضیه‌بندی زیر مفید است؛ زیرا ناحیهٔ رد آنها در دُم سمت راست توزیع قرار دارد:

۱. آزمون دو دامنه: فرضیه صفر $\mu = ۵$ در مقابل فرضیه جانشین $\mu \neq ۵$
۲. آزمون یک دامنه: فرضیه صفر $\mu \leq ۵$ در مقابل فرضیه جانشین $\mu > ۵$

برای آزمون دو دامنه، مقدار احتمال را با نصف سطح معناداری آزمون مقایسه کنید؛ یعنی برای سطح معناداری $\alpha = ۰/۰۵$، ۰/۰۱۲۸ را با $\frac{\alpha}{۲} = ۰/۰۲۵$ تا بتوانید درباره رد فرضیه صفر تصمیم بگیرید. از آنجا که در این آزمون مقدار احتمال از نصف سطح‌معناداری کوچک‌تر است، پس مقدار آماره از نقطهٔ بحرانی بزرگ‌تر بوده است و فرضیه صفر رد می‌شود. برای آزمون یک دامنهٔ فوق نیز، مقدار احتمال را با سطح معناداری آزمون مقایسه کنید؛ یعنی ۰/۰۱۲۸ را با $\alpha = ۰/۰۵$. بار دیگر فرضیه صفر این آزمونِ یک دامنه، به دلیل کوچک‌تر بودنِ ۰/۰۱۲۸ نسبت به ۰/۰۵ رد می‌شود.

مقدار احتمال برای آزمون فرضیه صفر $\mu \geq ۵$ در مقابل فرضیه جانشین $\mu < ۵$ فرق می‌کند، زیرا ناحیهٔ رد این آزمون یک دامنه در دُم سمت چپ توزیع قرار دارد. بنابراین مقدار احتمال، برابر با مساحت سمت چپ مقدار آماره است و باید با سطح معناداری آزمون مقایسه شود. مساحت سمت چپ آماره (۲/۲۴) برابر با ۰/۹۸۷۲ بوده که از $\alpha = ۰/۰۵$ بزرگ‌تر است، پس فرضیه صفر رد نمی‌شود.

همواره به تفاوت تفسیر نتیجهٔ آزمون آماری در دو وضعیت رد کردن فرضیه صفر و رد نکردن آن توجه کنید. اگر فرضیه صفر رد شود، درستی فرضیه جانشین را از آن نتیجه می‌گیریم، زیرا داده‌ها گواهی بر نادرستی فرضیه صفر هستند. اگر داده‌ها قادر نباشند فرضیه صفر را رد کنند، نباید نتیجه گرفت که فرضیه صفر درست است و تنها می‌توان گفت که داده‌های «موجود» دلیلی بر رد فرضیه صفر ارائه نکرده‌اند. به یاد بیاورید که فرضیه صفر در فرایند آزمون‌های آماری بدون هیچ دلیلِ مبتنی بر واقعیتی درست «پنداشته می‌شود» تا مبنایی برای آغاز فرایند اجرای آزمون باشد و این درست پنداشتن می‌تواند کاملاً اشتباه باشد. بنابراین، وقتی فرضیه صفر رد نمی‌شود، نباید با بازگشت به نقطهٔ ابتدایی فرایند

آزمون مدعی شد که این فرضیه درست است، بلکه تنها می‌توان گفت داده‌هایی ناسازگار با این فرضیه به دست نیامده است.

۲ـ۲ـ آزمون مقایسهٔ دو میانگین

پژوهشگری در پی آن است که میانگین هوشبهر دانشجویان مقطع کارشناسی را با دانشجویان مقاطع تحصیلات تکمیلی مقایسه کند، زیرا مدعی است میانگین هوشبهر دانشجویان مقاطع تحصیلات تکمیلی بیش از میانگین هوشبهر دانشجویان مقطع کارشناسی است. این فرضیه با دو میانگین مرتبط است که به دو جامعه مختلف تعلق دارند: جامعه دانشجویان مقطع کارشناسی و جامعه دانشجویان مقاطع تحصیلات تکمیلی. فرضیه‌هایی از این دست، آزمون‌های مقایسهٔ دو میانگین را پیش می‌کشند.

گاه مانند مثال فوق، دو میانگین به دو جامعه جدا از هم و به اصطلاح مستقل از یکدیگر تعلق دارند و گاه دو جامعه دارای چنین تمایزی نیستند. مثل زمانی که پژوهشگری می‌خواهد میانگین معدل سال اول دانشجویان مقطع کارشناسی را با میانگین معدل آنان در سال آخر دبیرستان مقایسه کند، زیرا مدعی است دانشجویان پس از ورود به دانشگاه در مقایسه با دوران دبیرستان دچار افت تحصیلی می‌شوند. او نمونه‌ای از دانشجویان سال اول مقطع کارشناسی را انتخاب و میانگین درسی آنان را در دانشگاه و دبیرستان گردآوری می‌کند. در این بررسی نیز دو میانگین وجود دارد که برخلاف مثال قبل، به دو جامعه با افراد متفاوت تعلق ندارند، بلکه دو دسته داده از گروهی «ثابت» گردآوری شده است: معدل دانشجویان در دانشگاه و معدل «همان» دانشجویان در دبیرستان. این فرضیه نیز به آزمون مقایسهٔ دو میانگین باز می‌گردد؛ با این تفاوت که دو جامعه وابسته به هم و مرتبط تحت بررسی است.

همواره در آزمون مقایسهٔ دو میانگین، از دو متغیر صحبت به میان می‌آید: متغیری مقوله‌ای با تنها دو مقوله و متغیری پیوسته (فاصله‌ای یا نسبتی) که قرار است استنباط آماری درباره میانگین آن صورت بگیرد. در مثال نخست، متغیر مقوله‌ای، مقطع تحصیلی با دو مقولهٔ «مقطع کارشناسی» و «مقطع دوره‌های تحصیلات تکمیلی» و متغیر پیوسته، هوشبهر بود که مقایسهٔ میانگین آن مورد توجه قرار داشت. در مثال دوم نیز یکی از متغیرها مقطع تحصیلی با دو مقولهٔ «دبیرستان» و «کارشناسی» و متغیر دیگر، معدل بود که میانگین آن، در دو مقطع مورد مقایسه قرار می‌گرفت.

اگرچه مقایسهٔ دو میانگین در دو حالت فوق موضوع آزمون آماری است، ولی آزمون، با دو آمارهٔ آزمون متفاوت اجرا می‌شود. با وجود این، فرضیه‌بندی مشابهی برای هر دو حالت قابل بیان است (μ_1 و μ_2 میانگین‌های دو جامعه هستند):

۱. آزمون دو دامنه: فرضیه صفر $\mu_1 = \mu_2$ در مقابل فرضیه جانشین $\mu_1 \neq \mu_2$
۲. آزمون یک دامنه: فرضیه صفر $\mu_1 \geq \mu_2$ در مقابل فرضیه جانشین $\mu_1 < \mu_2$
۳. آزمون یک دامنه: فرضیه صفر $\mu_1 \leq \mu_2$ در مقابل فرضیه جانشین $\mu_1 > \mu_2$

توجه کنید که این سه فرضیه‌بندی را می‌توان به صورت زیر نیز بازنویسی کرد:

۱. آزمون دو دامنه: فرضیه صفر $\mu_1 - \mu_2 = 0$ در مقابل فرضیه جانشین $\mu_1 - \mu_2 \neq 0$
۲. آزمون یک دامنه: فرضیه صفر $\mu_1 - \mu_2 \geq 0$ در مقابل فرضیه جانشین $\mu_1 - \mu_2 < 0$
۳. آزمون یک دامنه: فرضیه صفر $\mu_1 - \mu_2 \leq 0$ در مقابل فرضیه جانشین $\mu_1 - \mu_2 > 0$

۱ـ۲ـ۲ـ مقایسهٔ میانگین‌های دو جامعه مستقل

برای مقایسهٔ میانگین دو جامعه، ابتدا باید از هر جامعه نمونه‌ای «مستقل» از جامعه دیگر انتخاب کرد. فرض کنید آماره‌های زیر از این دو نمونه به دست آمده باشد:

نمونه	اندازه نمونه	میانگین نمونه	واریانس نمونه
اول	n_1	$\bar{X}_1$	S_1^2
دوم	n_2	$\bar{X}_2$	S_2^2

آمارهٔ آزمونِ مقایسهٔ دو میانگین نیز مانند آزمون تک میانگین به اطلاعات آماری جدول فوق نیاز دارد که بر اساس داده‌های دو نمونه انتخاب‌شده از دو جامعه تحت بررسی محاسبه می‌شوند. بر حسب اینکه واریانس‌های دو جامعه برابر باشند یا خیر، دو آمارهٔ آزمون مختلف مطرح می‌شود. اگر واریانس دو جامعه نابرابر باشند، آمارهٔ آزمون مقایسهٔ دو میانگین به صورت است:

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

این آماره نیز دارای توزیع نمونه‌گیری t است، مشروط بر آنکه نمونه انتخاب‌شده به اندازه کافی بزرگ باشد. برای یافتن df آن از تقریب سَتِرزوِیت[1] به صورت مقابل استفاده می‌شود:

1. Satterthwaite

$$df = \frac{\frac{S_۱^۲}{n_۱} + \frac{S_۲^۲}{n_۲}}{\frac{\left(\frac{S_۱^۲}{n_۱}\right)^۲}{n_۱ - ۱} + \frac{\left(\frac{S_۲^۲}{n_۲}\right)^۲}{n_۲ - ۱}}$$

فرایند تصمیم‌گیری درباره رد فرضیه صفر بر حسب یک دامنه یا دو دامنه بودنِ آزمون دقیقاً مانند آزمون تک میانگین است. بنابراین، ابتدا سطح معناداری آزمون، تعیین و سپس با مراجعه به جدول توزیع t، نقطهٔ بحرانیِ متناظر با سطح معناداری و df به دست می‌آید.

اگر واریانس دو جامعه برابر باشد، این امکان وجود دارد که داده‌های دو نمونه با هم ترکیب شوند و واریانس واحدی به نام واریانس آمیخته به صورت زیر محاسبه شود تا جانشین $S_۱^۲$ و $S_۲^۲$ در مخرج کسر آمارهٔ آزمون گردد.

$$S_p^۲ = \frac{(n_۱ - ۱)\, S_۱^۲ + (n_۲ - ۱)\, S_۲^۲}{n_۱ + n_۲ - ۲}$$

بنابراین، در صورتِ برابری واریانس‌های دو جامعه، آمارهٔ آزمون مقایسهٔ دو میانگین به صورت زیر خواهد بود:

$$t = \frac{\bar{X}_۱ - \bar{X}_۲}{\sqrt{\frac{S_p^۲}{n_۱} + \frac{S_p^۲}{n_۲}}}$$

این آماره نیز دارای توزیع نمونه‌گیری t ولی با $df = n_۱ + n_۲ - ۲$ است. از آنجا که مخرج کسر این آمارهٔ آزمون، متفاوت با آمارهٔ پیشین است، df آن نیز متفاوت با df توزیع نمونه‌گیری آن آماره است. بنابراین، نه تنها مقدار آمارهٔ آزمون در صورتِ برابری واریانس‌ها با آمارهٔ آزمون در صورتِ نابرابری واریانس‌ها تفاوت دارد، بلکه نقطهٔ بحرانی آزمون نیز در این دو وضعیت یکسان نخواهد بود، زیرا دارای dfهای مختلفی هستند.

برای مثال، مسئله مقایسهٔ میانگین هوشبهر دانشجویان مقطع کارشناسی و دانشجویان مقاطع تحصیلات تکمیلی را در نظر بگیرید. پژوهشگری می‌خواهد فرضیه صفر «میانگین هوشبهر دانشجویان مقاطع تحصیلات تکمیلی حداکثر برابر با میانگین هوشبهر دانشجویان مقطع کارشناسی است» را در مقابل فرضیه جانشین «میانگین هوشبهر دانشجویان مقاطع تحصیلات تکمیلی بیش از میانگین هوشبهر دانشجویان مقطع کارشناسی است» بیازماید. اگر جامعه اول را دانشجویان مقطع کارشناسی و جامعه دوم را دانشجویان مقطع تحصیلات

تکمیلی در نظر بگیریم، او با آزمونی یک دامنه، به صورت فرضیه صفر $\mu_۱ \geq \mu_۲$ در مقابل فرضیه جانشین $\mu_۱ < \mu_۲$ روبه‌روست.

این پژوهشگر، نمونه‌ای ۳۰۰ تایی از دانشجویان مقطع کارشناسی و نمونه‌ای ۱۱۰ تایی از دانشجویان مقاطع تحصیلات تکمیلی انتخاب می‌کند تا به مقایسهٔ دو میانگین بپردازد. فرض کنید میانگین هوشبهر ۳۰۰ دانشجوی مقطع کارشناسی ۱۱۵ با واریانس ۱۲۱ و میانگین هوشبهر ۱۱۰ دانشجوی مقاطع تحصیلات تکمیلی ۱۱۷ با واریانس ۱۰۰ است. پس اطلاعات زیر از این دو نمونه در دست است:

$$S_۱^۲ = ۱۲۱ \qquad \overline{X}_۱ = ۱۱۵ \qquad n_۱ = ۳۰۰$$
$$S_۲^۲ = ۱۰۰ \qquad \overline{X}_۲ = ۱۱۷ \qquad n_۲ = ۱۱۰$$

بر این اساس، آماره آزمون و df برای وضعیت واریانس‌های نابرابر به صورت زیر محاسبه می‌شود:

$$t = \frac{۱۱۵ - ۱۱۷}{\sqrt{\frac{۱۲۱}{۳۰۰} + \frac{۱۰۰}{۱۱۰}}} = -۱/۷۵ \qquad df = ۱۶۱/۵۱$$

اگر قرار باشد آزمون در سطح معناداری $\alpha = ۰/۰۵$ اجرا شود، باید در جدول توزیع t به دنبال $t_{۱۶۱/۵۱۹۹,۰/۰۵}$ به عنوان نقطهٔ بحرانی آزمون بود (توجه کنید که ناحیهٔ بحرانی این آزمونِ یک دامنه در دُم سمت چپ منحنی توزیع نمونه‌گیری قرار دارد، بنابراین از قرینهٔ چندک‌های t استفاده می‌شود). این نقطه برابر با ۱/۶۴۶- است که مقدار آمارهٔ ۱/۷۵- از آن کوچک‌تر است، پس فرضیه صفر به نفع فرضیه جانشین رد می‌شود. بر اساس شیوهٔ مقدار احتمال نیز همین نتیجه به دست می‌آید، زیرا مساحت سمت چپ مقدار ۱/۷۵- در توزیع t با $df = ۱۶۱/۵۱$ برابر با ۰/۰۴۱۰ است که کوچک‌تر از سطح معناداری ۰/۰۵ است (این مساحت به کمک نرم‌افزارهای آماری به دست می‌آید).

آمارهٔ آزمون، df و واریانس آمیخته در شرایطی که واریانس‌های دو جامعه برابرند، به صورت زیر محاسبه می‌شوند:

$$s_p^۲ = \frac{(۳۰۰-۱)\times ۱۲۱ + (۱۱۰-۱)\times ۱۰۰}{۳۰۰+۱۱۰-۲} = ۱۱۵,\ t = \frac{۱۱۵-۱۱۷}{\sqrt{\frac{۱۱۵}{۳۰۰}+\frac{۱۱۵}{۱۱۰}}} = -۱/۶۵,\ df = ۴۰۸$$

برای اجرای این آزمون نیز باید در جدول توزیع t به دنبال $-t_{۴۰۸,۰/۰۵}$ به عنوان نقطهٔ بحرانی باشیم که مقدار آن برابر با ۱/۶۴۶- است. در این وضعیت نیز بار دیگر فرضیه صفر رد می‌شود، زیرا مقدار آمارهٔ ۱/۶۵- از نقطهٔ بحرانی کوچک‌تر است. بر اساس مقدار

احتمال نیز همین نتیجه به دست می‌آید، زیرا مساحت سمت چپ مقدار ۱/۶۵- در توزیع t با $df = ۴۰۸$ برابر با ۰/۰۴۷۸ است که از ۰/۰۵ کوچک‌تر است.

همان طور که در این مثال دیده شد، مقدار آماره برای وضعیت واریانس‌های نابرابر (۱/۷۵-) متفاوت با وضعیت واریانس‌های برابر (۱/۶۵-) است. همچنین، مقدار احتمال این دو آماره (۰/۰۴۱۰ و ۰/۰۴۷۸) با یکدیگر متفاوت هستند. در این مثال، فرضیه صفر در هر دو حالت واریانس‌های برابر و نابرابر رد شد، ولی لزوماً در همهٔ موارد چنین نیست.

۲ـ۲ـ۲ـ مقایسهٔ میانگین‌های دو جامعه وابسته

یکی از معمول‌ترین وضعیت‌هایی که به مقایسهٔ میانگین‌های دو جامعهٔ وابسته می‌انجامد، مقایسهٔ میانگین‌های دو دسته از داده‌هاست که به افراد یکسان در دو مقطع زمانی مختلف تعلق دارند. مثالِ مقایسهٔ میانگین معدل دانشجویان در سال اول دانشگاه با میانگین معدل سال آخر دبیرستان از همین نوع است، زیرا در زمان ثبت معدل سال اولِ دانشجو (مقطع زمانی دوم)، معدل سال آخر دبیرستان او (مقطع زمانی اول) نیز ثبت می‌شود، پس دو داده از هر دانشجو در دست است که به دو مقطع زمانی مختلف (سال اول دانشگاه و سال آخر دبیرستان) تعلق دارند.

در مثالی دیگر، پژوهشگری را در نظر بگیرید که می‌خواهد میزان اثربخشی یک برنامهٔ آموزشی را بسنجد. او ابتدا نمونه‌ای از افراد را انتخاب و اطلاعات آنان را درباره محتویات این برنامهٔ آموزشی اندازه‌گیری می‌کند. سپس افراد نمونه را مطابق این برنامه، آموزش می‌دهد و در پایان آموزش، بار دیگر اطلاعات آنان را می‌سنجد. اگر میانگین اطلاعات شرکت‌کنندگان این آموزش در اندازه‌گیری دوم بیش از اندازه‌گیری اول باشد، او نتیجه خواهد گرفت که برنامهٔ آموزشی اثربخش بوده است. در این مثال نیز از هر فرد نمونه دو داده در دست است: دادهٔ مربوط به اطلاعات فرد پیش از ارائه آموزش و دادهٔ مربوط به اطلاعات فرد پس از آموزش. بنابراین، یک متغیر، دو بار در گروه ثابتی اندازه‌گیری شده است.

از آنچه گذشت می‌توان دریافت که از هر فرد، زوجی از داده‌ها وجود دارد، پس تعداد داده‌های دو مقطع برابر است، در حالی که در مقایسهٔ میانگین‌های دو جامعهٔ مستقل لزوماً چنین نبود؛ دو نمونه از افراد «مختلف» و معمولاً با تعدادی نابرابر در دست بود. بنابراین، در مقایسهٔ میانگین دو جامعهٔ وابسته، اگر سنجش متغیر در مقطع اول بر روی مثلاً ۵۰۰ نفر

صورت می‌گیرد، در مقطع دوم نیز این متغیر از همان ۵۰۰ نفر سنجیده می‌شود، از همین رو دو دسته دادهٔ ۵۰۰ تایی برای انجام آزمون خواهیم داشت.

در مقایسهٔ میانگین‌های دو جامعهٔ وابسته، برای آنکه نشان داده شود متغیر واحدی مانند X در دو مقطع اندازه‌گیری شده است، از دو اندیس ۱ و ۲ برای هر یک از مقاطع استفاده می‌شود. پس، $X_۱$ بیانگر اندازه متغیر X در مقطع اول و $X_۲$ بیانگر اندازه متغیر X در مقطع دوم است. بنابراین، زوج $X_{۱i}$ و $X_{۲i}$ نیز اندازه‌های متغیر X را از فرد iاُم نمونه در دو مقطع نشان می‌دهد. آزمون آماری با محاسبه تفاضل اندازه‌های مقطع دوم از مقطع اول مانند جدول ۲ آغاز می‌شود. با محاسبه این تفاضل‌ها، دو دسته داده تبدیل به یک دسته داده یعنی $D_۱$، $D_۲$،...، D_n می‌شود. توجه کنید که بین میانگین $X_{۱i}$ها و میانگین $X_{۲i}$ها با میانگین D_i ها ارتباط وجود دارد (سطر آخر جدول ۲ را ببینید) که در اجرای این آزمون مفید است.

جدول ۲: محاسبه تفاضل اندازه‌های دو مقطع برای آزمون مقایسهٔ میانگین‌های دو جامعه وابسته

فرد نمونه	نمونه اول	نمونه دوم	تفاضل
۱	$X_{۱۱}$	$X_{۲۱}$	$D_۱ = X_{۱۱} - X_{۲۱}$
۲	$X_{۱۲}$	$X_{۲۲}$	$D_۲ = X_{۱۲} - X_{۲۲}$
.	.	.	.
.	.	.	.
.	.	.	.
n	$X_{۱n}$	$X_{۲n}$	$D_n = X_{۱n} - X_{۲n}$
میانگین	$\bar{X}_۱$	$\bar{X}_۲$	$\bar{D} = \bar{X}_۱ - \bar{X}_۲$

بر اساس این رابطه، اگر اختلاف بزرگ $\bar{X}_۱$ از $\bar{X}_۲$ با فرضیه صفرِ $\mu_۱ = \mu_۲$ ناسازگار است و می‌تواند به رد آن منجر شود، آنگاه اختلاف زیاد $\bar{D} = \bar{X}_۱ - \bar{X}_۲$ از صفر نیز همان نتیجه را خواهد داشت. به همین ترتیب، اگر خیلی بزرگ‌تربودن $\bar{X}_۱$ از $\bar{X}_۲$ با فرضیه صفرِ $\mu_۱ \leq \mu_۲$ ناسازگار است، آنگاه خیلی بزرگ‌بودنِ $\bar{D} = \bar{X}_۱ - \bar{X}_۲$ از صفر نیز با این فرضیه ناسازگار خواهد بود. با استفاده از این رابطه، می‌توان سه فرضیه‌بندی مقایسهٔ میانگین‌های دو جامعه را بازنویسی کرد (اندیسِ D نشان می‌دهد که μ_D میانگین تفاضل‌های دو جامعه است):

۱. آزمون دو دامنه: فرضیه صفر $\mu_D = ۰$ در مقابل فرضیه جانشین $\mu_D \neq ۰$
۲. آزمون یک دامنه: فرضیه صفر $\mu_D \geq ۰$ در مقابل فرضیه جانشین $\mu_D < ۰$
۳. آزمون یک دامنه: فرضیه صفر $\mu_D \leq ۰$ در مقابل فرضیه جانشین $\mu_D > ۰$

برای اجرای آنها نیز می‌توان از همان آزمون مربوط به تک میانگین استفاده کرد، مشروط بر آنکه به جای μ_o عدد صفر به کار رود. به این ترتیب آمارهٔ آزمون مقایسهٔ میانگین‌های دو جامعه وابسته عبارت است از:

$$t = \frac{\bar{D}}{\frac{S_D}{\sqrt{n}}}$$

توجه کنید که در آمارهٔ آزمون فوق $\bar{D}$ و S_D به ترتیب میانگین و انحراف معیار تفاضل‌های جدول ۲ یعنی $D_۱$، $D_۲$،...، D_n هستند. این آمارهٔ آزمون دارای توزیع نمونه‌گیری t با $df = n - ۱$ است. پس می‌توان از جدول چندک‌های توزیع t، به نقطهٔ بحرانی مناسب برای آزمون دست یافت.

برای مثال، پژوهشگری به دنبال مقایسهٔ میانگین معدل سال اول دانشجویان با میانگین معدل سال آخر دبیرستان آنان است. او قصد دارد فرضیه صفر «میانگین معدل سال آخر دبیرستان دانشجویان حداکثر برابر با میانگین معدل سال اول دانشگاه آنان است» را در مقابل فرضیه جانشین «میانگین معدل سال آخر دبیرستان دانشجویان بیش از میانگین معدل سال اول دانشگاه آنان است» بیازماید؛ یعنی آزمونِ یک دامنهٔ فرضیه صفرِ $\mu_۱ \leq \mu_۲$ در مقابل فرضیه جانشین $\mu_۱ > \mu_۲$ (جامعه اول معدل‌های سال آخر دبیرستان و جامعه دوم معدل‌های سال اول دانشگاه است).

او نمونه‌ای ۳۰۰ نفری از دانشجویان سال اول انتخاب و معدل سال آخر دبیرستان و سال اول دانشگاه آنان را ثبت می‌کند؛ یعنی دو معدل از هر دانشجوی نمونه در اختیار است. اگر میانگین معدل سال آخر دبیرستان این ۳۰۰ دانشجو ۱۷/۵ و میانگین معدل سال اول دانشگاه آنان ۱۶ باشد، آنگاه $\bar{D} = ۱۷/۵ - ۱۶ = ۱/۵$. اگر انحراف معیار تفاضل‌های معدل‌های سال اول دانشگاه از معدل‌های سال آخر دبیرستان برابر با ۵ باشد، در این صورت آمارهٔ آزمون عبارت است از:

$$t = \frac{۱۷/۵ - ۱۶}{\frac{۵}{\sqrt{۳۰۰}}} = ۵/۱۹۶ \qquad df = ۳۰۰ - ۱ = ۲۹۹$$

اگر آزمون در سطح معناداری ۰/۰۵ اجرا شود، با توجه به اینکه ناحیهٔ رد آزمون در دُم سمت راست توزیع نمونه‌گیری قرار دارد، باید به دنبال $t_{۲۹۹,۰/۰۵}$ که برابر با ۱/۶۴۶ است به عنوان نقطهٔ بحرانی آزمون باشیم. از آنجا که مقدارآماره (۵/۱۹۶) از نقطهٔ بحرانی بزرگ‌تر است، پس فرضیه صفر به نفع فرضیه جانشین رد می‌شود؛ یعنی میانگین معدل سال آخر دبیرستان دانشجویان بیش از میانگین معدل سال اول دانشگاه آنان است. بر اساس مقدار احتمال نیز مساحت سمت راست آمارهٔ آزمون ۰/۰۰۶۹ است که از سطح معناداری ۰/۰۵ کوچک‌تر است، پس فرضیه صفر رد می‌شود.

آزمون مقایسهٔ میانگین‌های دو جامعه وابسته تنها به اندازه‌گیری یک متغیر از افرادی ثابت در دو مقطع زمانی محدود نمی‌شود. ممکن است دو متغیر به طور همزمان از افراد ثابتی اندازه‌گیری شوند. برای مثال، پرسشنامه‌ای طراحی می‌شود که شامل پرسش‌هایی برای اندازه‌گیری امید به آیندهٔ اقتصادی و پرسش‌هایی برای اندازه‌گیری امید به آیندهٔ سیاسی افراد است. پس این پرسشنامه دو نمره برای هر فرد به دست می‌دهد که نشان‌دهنده امید به آیندهٔ اقتصادی و امید به آیندهٔ سیاسی اوست. اگر پژوهشگری علاقه‌مند باشد که بداند امید به آیندهٔ سیاسی در بین مردم بیشتر است یا امید به آیندهٔ اقتصادی، در این صورت باید میانگین نمرات امید به آیندهٔ سیاسی را با میانگین نمرات امید به آیندهٔ اقتصادی مقایسه کند. از آنجا که از افراد ثابتی دو نمره (زوجی از داده‌ها) در دست است، مقایسهٔ او بین دو جامعه وابسته صورت می‌گیرد[1].

وضعیت دیگری که می‌تواند به مقایسهٔ میانگین‌های دو جامعهٔ وابسته بینجامد اندازه‌گیری یک متغیر از دو فردِ به شدت مرتبط مانند «زن و شوهر» یا دو فرد نسبتاً مشابه مانند «دوقلوها» است. اگرچه در این وضعیت با دو جامعه از افراد روبه‌رو هستیم، ولی به دلیل ارتباط یا شباهت زیاد آنان به یکدیگر معمولاً این دو جامعه وابسته در نظر گرفته می‌شوند.

۱. از همین رو، مقایسهٔ میانگین‌های دو جامعه وابسته را مقایسهٔ زوجی (Paired Comparison) نیز می‌نامند.

۳ـ۲ـ آزمون مقایسهٔ بیش از دو میانگین (تجزیهٔ واریانس)

پژوهشگری می‌خواهد بداند آیا تفاوتی بین میانگین معدل دانش‌آموختگان مقطع کارشناسی سه دانشگاه در رشتهٔ آمار وجود دارد. او در پایان سال تحصیلی جاری، ۹ دانش‌آموخته را از دانشگاه «الف»، ۹ دانش‌آموخته را از دانشگاه «ب» و ۱۰ دانش‌آموخته را از دانشگاه «پ» انتخاب و معدل فارغ‌التحصیلی آنها را مطابق جدول ۳ ثبت می‌کند.

جدول ۳: معدل‌های دانش‌آموختگان سه دانشگاه

<table dir="rtl">
<tr><th>دانشگاه</th><th>الف (۱)</th><th>ب (۲)</th><th>پ (۳)</th><th>کل</th></tr>
<tr><td>معدل فارغ‌التحصیلی</td><td>۱۸، ۱۹/۲۵، ۱۹
۱۹/۵، ۱۸/۲۵
۱۷/۷۵، ۱۷/۵
۱۹، ۱۸</td><td>۱۶، ۱۷/۵
۱۶/۲۵، ۱۵
۱۵/۵، ۱۶
۱۴، ۱۳/۵، ۱۷</td><td>۱۷، ۱۵، ۱۹
۱۳، ۱۴، ۱۹/۵
۱۸/۵، ۱۹
۱۷/۵، ۱۸</td><td></td></tr>
<tr><td>تعداد</td><td>۹</td><td>۹</td><td>۱۰</td><td>۲۸</td></tr>
<tr><td>میانگین</td><td>۱۸/۴۷۲</td><td>۱۵/۶۳۸</td><td>۱۷/۰۵</td><td>۱۷/۰۵۴</td></tr>
<tr><td>واریانس</td><td>۰/۵۲۳</td><td>۱/۷۰۵</td><td>۵/۱۹۲</td><td>۳/۷۲۹</td></tr>
<tr><td rowspan="2">تغییرپذیری</td><td>۴/۱۸۴</td><td>۱۳/۶۴</td><td>۴۶/۷۲۸</td><td rowspan="2">۱۰۰/۶۷۷</td></tr>
<tr><td colspan="3">۶۴/۵۵۲</td></tr>
</table>

این جدول میانگین کل ۲۸ داده (۱۷/۰۵۴) و میانگین داده‌های تک‌تک دانشگاه‌ها را نیز ارائه کرده است. تغییرپذیری موجود در ۲۸ دادهٔ جدول ۳ بر تفاوت معدل‌های دانشجویان دلالت دارد. این تغییرپذیری را می‌توان با محاسبه مجموع مربعات اختلاف داده‌ها از میانگین کل به صورت زیر اندازه‌گیری کرد:

$$(۱۸-۱۷/۰۵۴)^{۲}+(۱۹/۲۵-۱۷/۰۵۴)^{۲}+\cdots+(۱۸-۱۷/۰۵۴)^{۲}=۱۰۰/۶۸۳$$

ممکن است بخشی از این تغییرپذیری به دانشگاهی برگردد که دانشجو در آن تحصیل می‌کند، زیرا برنامهٔ آموزشی دانشگاه‌ها و میزان سختگیری آنها می‌تواند متفاوت باشد و به این ترتیب دانش‌آموختگان یک دانشگاه دارای معدلی بالاتر یا پایین‌تر از دانش‌آموختگان دانشگاه دیگر باشند. میانگین معدل‌های دانش‌آموختگان دانشگاه‌ها در جدول ۳ یعنی ۱۸/۴۷۲، ۱۵/۶۳۸ و ۱۷/۰۵ بر همین تفاوت اشاره دارد. این تغییرپذیری یعنی تغییرپذیری «بینِ» دانشگاه‌ها به صورت زیر قابل محاسبه است:

$$۹\times(۱۸/۴۷۲-۱۷/۰۵۴)^{۲}+۹\times(۱۵/۶۳۸-۱۷/۰۵۴)^{۲}+۱۰\times(۱۷/۰۵-۱۷/۰۵۴)^{۲}=۳۶/۱۲۵$$

اگرچه بخشی از تغییرپذیری برگرفته از تفاوت دانشگاه‌هاست، ولی نمی‌توان تمام آن را به دانشگاه‌ها نسبت داد، زیرا دانش‌آموختگانی هستند که به یک دانشگاه تعلق دارند، ولی معدل یکسانی کسب نکرده‌اند.

برای مثال، در میان معدل دانش‌آموختگان دانشگاه «الف»، هم معدل ۱۹/۲۵ و هم معدل ۱۷/۵ و در میان معدل دانش‌آموختگان دانشگاه «ب»، هم معدل ۱۳/۵ و هم معدل ۱۷/۵ به چشم می‌خورد. به عبارت دیگر، میان دانش‌آموختگان یک دانشگاه نیز تغییرپذیری مشاهده می‌شود؛ هرچند همه آنان تحت یک برنامهٔ آموزشی و یک میزان سختگیری قرار داشته‌اند. این تغییرپذیری را باید ناشی از طبیعت افراد، یعنی ویژگی‌های فردی نظیر استعداد و تلاش آنان دانست که حتی در شرایط کاملاً یکسان، رفتاری متفاوت از خود نشان می‌دهند. تغییرپذیری معدل‌های هر کدام از دانشگاه‌ها به صورت زیر محاسبه می‌شود:

$$۴/۱۸۴=(۱۸-۱۸/۴۷)^۲+(۱۹/۲۵-۱۸/۴۷)^۲+\cdots+(۱۸-۱۸/۴۷)^۲$$

$$۱۳/۶۴=(۱۶-۱۵/۶۳)^۲+(۱۷/۵-۱۵/۶۳)^۲+\cdots+(۱۷-۱۵/۶۳)^۲$$

$$۴۶/۷۲۸=(۱۷-۱۷/۰۵)^۲+(۱۵-۱۷/۰۵)^۲+\cdots+(۱۸-۱۷/۰۵)^۲$$

از جمع این سه تغییرپذیری، تغییرپذیری «درونِ» دانشگاه‌ها به صورت زیر به دست می‌آید:

$$۶۴/۵۵۲=۴/۱۸۱+۱۳/۶۴+۴۶/۷۲۸$$

پس می‌توان گفت:

تغییرپذیری ناشی از تفاوت‌های فردی+تغییرپذیری ناشی از تفاوت دانشگاه‌ها=تغییرپذیری کل

این رابطه در بین ارقامی که پیشتر به دست آوردیم نیز دیده می‌شود، یعنی:

$$۱۰۰/۶۷۷=۳۶/۱۲۵+۶۴/۵۵۲$$

در واقع، تغییرپذیری کل داده‌ها به دو بخش تجزیه شده است: تغییرپذیری مربوط به تفاوت دانشگاه‌ها (تغییرپذیری بین دانشگاه‌ها) و تغییرپذیری ناشی از تفاوت‌های فردی که در داخل دانشگاه‌ها مشاهده می‌شود (تغییرپذیری درون دانشگاه‌ها).

تجزیهٔ تغییرپذیریِ کل، به دو مؤلفه، مبنایی برای اجرای آزمون مقایسهٔ میانگین معدل‌هایِ دانش‌آموختگان سه دانشگاه فراهم می‌آورد. اگر تغییرپذیری بین دانشگاه‌ها از تغییرپذیریِ درون دانشگاه‌ها بیشتر باشد می‌توان نتیجه گرفت که تغییرپذیری کل، بیشتر از

تفاوتِ دانشگاه‌ها ناشی می‌شود تا ویژگی‌های فردی دانش‌آموختگان. بنابراین، دانشگاه‌ها باعث این تغییرپذیری شده‌اند و می‌توان پذیرفت که میانگین‌های آنها با یکدیگر تفاوت دارند.

اکنون این پرسش مطرح است که آیا مقایسهٔ این دو تغییرپذیری یعنی ۳۶/۱۲۵ با ۶۴/۵۵۲ برای این منظور مناسب است؟ پاسخ منفی است؛ زیرا ۳۶/۱۲۵ از ۳ داده (میانگین‌های ۳ دانشگاه) ولی ۶۴/۵۵۲ از ۲۸ داده به دست آمده است. پس دور از انتظار نیست که مقدار تغییرپذیری درون دانشگاه‌ها بیش از تغییرپذیری بین دانشگاه‌ها باشد، زیرا امکان مشاهدهٔ داده‌های متنوع‌تر و به تبع آن تغییرپذیری بیشتر در میان ۲۸ داده بیش از ۳ داده است. از این رو به جای استفاده از مجموع مربعات (تغییرپذیری)، از میانگین مربعات استفاده می‌شود (میانگین مربعات کمّیتی همانند واریانس است). به این ترتیب، کمّیتی به دست می‌آید که کمتر تحت تأثیر تعداد داده‌هاست. این محاسبات برای داده‌های جدول ۳، در جدول ۴ ارائه شده است که به آن تجزیهٔ واریانس گفته می‌شود.

جدول ۴: اطلاعات تجزیهٔ واریانس برای آزمون مقایسهٔ میانگین‌ها

منابع تغییرات	مجموع مربعات	df	میانگین مربعات
بین دانشگاه‌ها (میانگین‌ها)	۳۶/۱۲۵	۲	۱۸/۰۶۳
درون دانشگاه‌ها (افراد)	۶۴/۵۵۲	۲۵	۲/۵۸۲
کل	۱۰۰/۶۷۷	۲۷	

ارقامی که در ستون df در جدول ۴ آمده است، اعدادی هستند که مجموعِ مربعات بر آن تقسیم می‌شود تا میانگینِ مربعات (یا همان واریانس) به دست آید؛ یعنی ۱۸/۰۶۳ حاصلِ تقسیم ۳۶/۱۲۵ بر عدد ۲ است. حاصل تقسیم میانگین مربعاتِ تفاوت‌های بین دانشگاه‌ها بر میانگین مربعاتِ داخل دانشگاه‌ها (افراد) یعنی ۶/۹۹=۱۸/۰۶۳÷۲/۵۸۲ همان مقدار آمارهٔ آزمون برای مقایسهٔ این سه میانگین است. بزرگی آن نشان می‌دهد که واریانس میانگین‌ها بزرگ‌تر از واریانس افراد است، پس میانگین‌ها اختلاف چشمگیری با یکدیگر دارند که باعث این واریانس بزرگ شده است. نقطهٔ بحرانی این آزمون از جدول‌های مربوط به توزیع F به دست می‌آید، زیرا توزیع نمونه‌گیری آمارهٔ این آزمون توزیع F است.

اکنون آزمون مقایسهٔ میانگین‌های k جامعه «مستقل» را در حالت کلی بررسی می‌کنیم. این آزمون دارای یک فرضیه‌بندی است: فرضیه صفرِ $\mu_1 = \mu_2 = \cdots = \mu_k$ در مقابل فرضیه

جانشینِ «حداقل در میانگین‌های دو جامعه اختلاف وجود دارد». برای اجرای آزمون، ابتدا از هر جامعه نمونه‌ای انتخاب می‌شود. مانند آزمون مقایسهٔ میانگین‌های دو جامعه مستقل، تعداد نمونه‌های انتخاب‌شده از هر جامعه می‌تواند متفاوت با جامعه دیگر باشد. برای اجرای آزمون، ابتدا باید اطلاعات جدول ۵ بر اساس داده‌های این نمونه‌ها استخراج شود (جدولی شبیه به جدول ۳).

جدول ۵: اطلاعات لازم از نمونه‌ها برای اجرای آزمون مقایسهٔ K میانگین

نمونه	اول	دوم	...	K اُم	کل
تعداد نمونه	$n_۱$	$n_۲$	...	n_k	n
میانگین نمونه	$\bar{X}_۱$	$\bar{X}_۲$	...	$\bar{X}_k$	$\bar{X}$
واریانس نمونه	$S_۱^۲$	$S_۲^۲$	...	$S_k^۲$	$S^۲$

با توجه به توضیحات گذشته، تغییرپذیری کل، تغییرپذیری بین جامعه‌ها و تغییرپذیری داخل جامعه‌ها به صورت زیر محاسبه می‌شود. توجه کنید که تغییرپذیری با SS به معنای مجموع مربعات[1] و حروف t، b و w نیز به عنوان زیرنویسِ SSها برای اشاره به تغییرپذیری کل (total)، تغییرپذیری بین جامعه‌ها (between) و تغییرپذیری درون جامعه‌ها (within) به کار می‌روند.

$$SS_w = (n_۱ - ۱)S_۱^۲ + (n_۲ - ۱)S_۱^۲ + \cdots (n_k - ۱)S_k^۲ = \sum_{j=۱}^{k}(n_j - ۱)S_j^۲$$

$$SS_b = n_۱(\bar{X}_۱ - \bar{X})^۲ + n_۲(\bar{X}_۲ - \bar{X})^۲ + \cdots + n_k(\bar{X}_k - \bar{X})^۲ = \sum_{j=۱}^{k} n_j(\bar{X}_j - \bar{X})^۲$$

$$SS_t = (n - ۱)S^۲ = SS_b + SS_w \qquad n = \sum_{j=۱}^{k} n_j$$

اما به یاد بیاورید که برای محاسبه آمارهٔ آزمون (F)، به جای مجموع مربعات یعنی SSها، از میانگین مربعات[2] یعنی MSها به صورت زیر استفاده می‌شود:

$$MS_b = \frac{SS_b}{k-۱} \qquad MS_w = \frac{SS_w}{n-k} \qquad F = \frac{MS_b}{MS_w} = \frac{n-k}{k-۱}\,\frac{SS_b}{SS_w}$$

مانند جدول ۴، این اطلاعات در قالب جدولی به نام «تجزیهٔ واریانس» نمایش داده می‌شوند که حالت کلی آن را در ادامه مشاهده می‌کنید.

1. Sum of Squares (SS)
2. Mean of Squares (MS)

جدول تجزیهٔ واریانس برای مقایسهٔ k میانگین

منابع تغییرپذیری	SSها	dfها	MSها	آمارهٔ آزمون
تغییرپذیری بین جامعه‌ها	SS_b	$df_۱ = k - ۱$	MS_b	$F = \frac{MS_b}{MS_w}$
تغییرپذیری درون جامعه‌ها	SS_w	$df_۲ = n - k$	MS_w	
تغییرپذیری کل داده‌ها	SS_t	$n - ۱$		

اگر مقدار آمارهٔ آزمون بزرگ باشد، نتیجه خواهیم گرفت که MS_b از MS_w بزرگ‌تر است، پس سهم تفاوت بین جامعه‌ها در تغییرپذیری کل بیشتر از سهم تفاوت‌های فردی (درون جامعه‌ها) است و به این ترتیب می‌توان استنباط کرد که میانگین‌های جامعه‌ها با یکدیگر اختلاف دارند. بنابراین، آزمون مقایسهٔ میانگین‌ها آزمونی یک دامنه است، زیرا تنها مقادیر بزرگ آمارهٔ آزمون باعث رد فرضیه صفر می‌شوند. همان طور که پیشتر اشاره شد، آمارهٔ F دارای توزیع نمونه‌گیری به نام F با $df_۱ = k - ۱$ و $df_۲ = n - k$ است که با توجه به سطح معناداری برای یافتن نقطهٔ بحرانی آزمون به کار می‌رود (از همین رو به آزمون تجزیهٔ واریانس، آزمون F نیز گفته می‌شود). جدول مربوط به چندک‌های منحنی توزیع F به ازای $\alpha = ۰/۱$، $\alpha = ۰/۰۵$ و $\alpha = ۰/۰۱$ برای $df_۱$ها و $df_۲$های مختلف در پیوست ارائه شده است و از آن می‌توان مقدار $F_{df_۱, df_۲, \alpha}$ را به عنوان نقطهٔ بحرانی آزمون به دست آورد.

در مثالی که بحث با آن آغاز شد، اگر آزمون در سطح معناداری $\alpha = ۰/۰۵$ اجرا شود، نقطهٔ بحرانی $F_{۲,۲۵,۰/۰۵}$ خواهد بود، زیرا $k = ۳$ و $n = ۲۸$ است. این نقطه از جدولِ توزیع F در پیوست، برابر با ۳/۳۹ به دست می‌آید. آمارهٔ آزمون با توجه به محاسبات پیشین ۶/۹۹ و بزرگ‌تر از نقطهٔ بحرانی است، پس فرضیه صفرِ $\mu_۱ = \mu_۲ = \mu_۳$ رد می‌شود؛ یعنی «حداقل» میانگین معدل‌های دو دانشگاه با هم اختلاف دارند. بر اساس مقدار احتمال نیز مساحت سمت راست ۶/۹۹ برابر با ۰/۰۰۴ است که از $\alpha = ۰/۰۵$ کوچک‌تر است.

اکنون که آزمون نشان داد بین میانگین معدل‌های دانشگاه‌ها اختلاف وجود دارد، آیا باید نتیجه گرفت که میانگین‌های هر سه دانشگاه متفاوت هستند؟ ممکن است تفاوت تنها بین دو دانشگاه باشد. مثلاً میانگین‌های دانشگاه‌های «الف» و «پ» در جدول ۳ برابر باشند و تنها میانگین دانشگاه «ب» با این دو اختلاف داشته باشد. آزمون مقایسهٔ چند میانگین تنها به

وجود اختلاف بین میانگین‌ها اشاره دارد، ولی مشخص نمی‌کند این اختلاف بین کدامیک از میانگین‌هاست. برای بررسی این موضوع باید از آزمون‌های تعقیبی استفاده شود.

۱ـ۳ـ۲ـ آزمون‌های تعقیبی (مقایسه‌های دوبه‌دو)

آزمون‌های تعقیبی برای مقایسهٔ دوبه‌دوی میانگین‌ها به کار می‌روند تا مشخص شود کدام میانگین‌ها با هم اختلاف دارند. از همین رو به این آزمون‌ها مقایسه‌های دوبه‌دو یا مقایسه‌های چندگانه نیز گفته می‌شود. برای مثال، اگر آزمون مقایسهٔ میانگین‌های $k = ۴$ جامعه معنادار باشد، یعنی بر وجود تفاوت بین این میانگین‌ها دلالت کند، اکنون باید ۶ مقایسهٔ جدول ۶ را بین دوبه‌دوی میانگین‌ها صورت دهیم تا میانگین‌های دارای اختلاف مشخص شوند (به طور کلی، k جامعه به $c = \frac{k(k-۱)}{۲}$ مقایسهٔ دوبه‌دو می‌انجامد).

جدول ۶: شش مقایسهٔ دوبه‌دو بین میانگین‌های چهار جامعه

جامعه	۱	۲	۳	۴
۱	–	۱ با ۲	۱ با ۳	۱ با ۴
۲	تکراری	–	۲ با ۳	۲ با ۴
۳	تکراری	تکراری	–	۳ با ۴
۴	تکراری	تکراری	تکراری	–

واژهٔ «تکراری» در این جدول نشان می‌دهد که پیشتر، به آن مقایسه به شکلی دیگر اشاره شده است. برای مثال، مقایسهٔ جامعه دوم با جامعه اول تکراری است، زیرا پیشتر مقایسهٔ جامعه اول با جامعه دوم ذکر شده و تنها شماره جامعه‌ها جابه‌جا شده است (مقایسهٔ جامعه‌های ۱ با ۲ فرقی با مقایسهٔ جامعه‌های ۲ با ۱ ندارد).

نخستین راهکاری که برای مقایسه‌های دوبه‌دو به نظر می‌رسد، استفاده از آزمون مقایسهٔ میانگین‌های دو جامعه مستقل با واریانس برابر است که به آن «کمترین اختلاف معنادار» می‌گویند. به این ترتیب، به تعداد $c = \frac{k(k-۱)}{۲}$، آزمون‌هایِ t جداگانه بر اساس اطلاعات جدول ۵ و MS_w محاسبه‌شده به صورت زیر اجرا می‌کنیم تا مشخص شود کدام میانگین‌ها با هم اختلاف دارند (در واقع، MS_w نوعی واریانس آمیخته است).

$$t = \frac{\bar{X}_r - \bar{X}_s}{\sqrt{\frac{MS_w}{n_r} + \frac{MS_w}{n_s}}} \qquad df = n - k$$

در آمارۀ آزمون فوق، r و s شماره‌های جامعه‌هایی هستند که با هم مقایسه می‌شوند. برای مثال، اگر جامعه اول با جامعه سوم مقایسه می‌شود $r=۱$ و $s=۳$ است. این مقایسه‌ها برای مثالِ جدول‌های ۳ و ۴ به شکل جدول ۷ صورت می‌گیرد. توجه کنید که ۳ مقایسۀ دوبه‌دو برای این مثال لازم است، زیرا $k=۳$ است. جدول ۷ نشان می‌دهد که در سطح معناداری ۵ درصد، تنها بین میانگین معدل‌های دانش‌آموختگان دانشگاه «الف» و «ب» اختلاف وجود دارد.

جدول ۷: نتایج مقایسه‌های دوبه‌دو با روش کمترین اختلاف معنادار

r	s	مقایسه	مقدار آمارۀ t	مقدار احتمال	نتیجۀ آزمون
۱	۲	دانشگاه «الف» با «ب»	$۳/۷۴=\frac{۱۸/۴۷۲-۱۵/۶۳۸}{\sqrt{\frac{۲/۵۸۲}{۹}+\frac{۲/۵۸۲}{۹}}}$	۰/۰۰۱	اختلاف وجود دارد
	۳	دانشگاه «الف» با «پ»	$۱/۹۳=\frac{۱۸/۴۷۲-۱۷/۰۵}{\sqrt{\frac{۲/۵۸۲}{۹}+\frac{۲/۵۸۲}{۱۰}}}$	۰/۰۶۵	اختلافی مشاهده نشد
۲	۳	دانشگاه «ب» با «پ»	$۱/۹۱=\frac{۱۵/۶۳۸-۱۷/۰۵}{\sqrt{\frac{۲/۵۸۲}{۹}+\frac{۲/۵۸۲}{۱۰}}}$	۰/۰۶۷	اختلافی مشاهده نشد

استفاده از چندین آزمونِ جداگانه می‌تواند به چندین بار ارتکاب خطای نوع اول منجر شود. به عبارت دیگر، نتیجه‌گیری «کلی» درباره اختلاف میانگین‌ها که از برآیند آزمون‌های مقایسه‌های دوبه‌دو به دست می‌آید با خطای نوع اولی روبه‌روست که بیش از خطای نوع اول تک‌تک آزمون‌هاست. در واقع، اگر تعداد c آزمون مقایسه‌ای دوبه‌دو صورت گیرد و سطح معناداری تک‌تک این آزمون‌ها α باشد، آنگاه احتمال خطای نوع اول نتیجه‌گیری «کلی» برابر با $۱-(۱-\alpha)^c$ است. برای مثال، مقایسۀ دوبه‌دوی میانگین‌های ۴ جامعه به ۶ آزمون t نیاز دارد که اگر تک‌تک آنها در سطح معناداری $\alpha=۰/۰۵$ اجرا شوند، احتمال خطای نوع اول نتیجه‌گیری «کلی» برابر با ۰/۲۶۵ است که خیلی بزرگ‌تر از سطح معناداری تک‌تک این آزمون‌هاست. به این ترتیب، نتیجه‌گیری «کلی» از این مقایسه‌ها چندان قابل اعتماد نخواهد بود.

برای آنکه سطح معناداری نتیجه‌گیری کلی همان α باقی بماند، از روش بانفرونی[1] برای اصلاح روش کمترین اختلاف معنادار استفاده می‌شود. این روش، خطای نوع اول مقایسه‌های دوبه‌دو را $\frac{\alpha}{c}$ در نظر می‌گیرد. برای مثال، اگر قرار باشد ۴ جامعه با یکدیگر مقایسه شوند $c=۶$ مقایسهٔ دوبه‌دو خواهیم داشت. بنابراین برای حفظ سطح معناداری کلی در همان سطح $\alpha=۰/۰۵$، باید سطح معناداری این ۶ مقایسهٔ دوبه‌دو برابر با $\frac{۰/۰۵}{۶}=۰/۰۰۸۳$ باشد. به این ترتیب، نتیجه‌گیری کلی حاصل از این ۶ مقایسه دارای سطح معناداری ۰/۰۵ خواهد بود. توجه کنید که مقدار آمارهٔ آزمون در روش بانفرونی مانند روش کمترین اختلاف معنادار است و تنها در تعیین نقطهٔ بحرانی تفاوت وجود دارد. نقطهٔ بحرانی در روش کمترین اختلاف معنادار $t_{n-k,\frac{\alpha}{۲}}$، ولی در روش بانفرونی $t_{n-k,\frac{\alpha}{۲c}}$ است (به یاد بیاورید که آزمون‌های مقایسهٔ دوبه‌دو آزمون‌هایی دو دامنه هستند و در آزمون‌های دو دامنه از $\frac{\alpha}{۲}$ به جای α برای یافتن نقطهٔ بحرانی استفاده می‌شود. بنابراین $\frac{\alpha}{۲}$ بر c تقسیم شده است).

اکنون آزمون‌های جدول ۷ را با روش بانفرونی اجرا می‌کنیم. برای $\alpha=۰/۰۵$ و $c=۳$ باید مقدار احتمال‌ها را با $\frac{۰/۰۲۵}{۳}=۰/۰۰۸۳$ به عنوان سطح معناداری مقایسه کنیم. با این روش نیز تنها اختلاف دانشگاه «الف» با «ب» معنادار است. اگرچه برحسب اتفاق روش کمترین اختلاف معنادار و روش بانفرونی به نتیجهٔ یکسانی رسیدند، ولی این نتیجه همواره رخ نمی‌دهد. همچنین، سطح معناداری روش بانفرونی کاملاً متفاوت با روش کمترین اختلاف معنادار است.

با وجود اینکه روش بانفرونی به حفظ سطح معناداری در حدی قابل قبول می‌انجامد، ولی کوچک کردن سطح معناداری تک‌تک مقایسه‌های دوبه‌دو باعث می‌شود فرضیه برابری دو میانگین به ندرت رد شود و اختلافی بین دو میانگین نشان داده نشود؛ زیرا سطح معناداری مقایسه‌های زوجی آنقدر کوچک است که امکان فراتر رفتن آمارهٔ آزمون از نقطهٔ بحرانی را بسیار کم می‌کند. از این رو، روش بانفرونی آزمونی محافظه‌کارانه[2] برای مقایسه‌های دوبه‌دو قلمداد می‌شود (هر چه تعداد مقایسه‌ها بیشتر باشد، روش بانفرونی محافظه‌کارانه‌تر می‌شود).

---

1. Bonferroni

۲. رد فرضیه صفر در آزمون‌های محافظه‌کارانه (Conservative) به سختی امکان‌پذیر است.

روش‌های دیگری برای آزمون مقایسه‌های دوبه‌دو ارائه شده‌اند که هم به آزادی روش کمترین اختلاف معنادار و هم به محافظه‌کاری روش بانفرونی نباشند. این روش‌ها کمتر در محاسبه آمارهٔ آزمون با یکدیگر اختلاف دارند و اختلاف اساسی آنها در یافتن نقاط بحرانی است. روش توکی[1] که به آن «اختلاف معنادار واقعی» نیز گفته می‌شود، دارای آمارهٔ آزمونی به صورت زیر است که تنها تفاوت آن با آمارهٔ آزمون روش کمترین اختلاف معنادار و روش بانفرونی، در مخرج کسر است.

$$w = \frac{\overline{X}_r - \overline{X}_s}{\sqrt{\frac{۱}{۲}\left[\frac{MS_w}{n_r} + \frac{MS_w}{n_s}\right]}} \qquad df = n - k$$

نقطهٔ بحرانی آزمون توکی بر اساس مقدار df، k و α از جدول مربوط به آن در پیوست به دست می‌آید. این روش برای حالتی مناسب‌تر است که تعداد نمونه برابری از هر جامعه انتخاب شده باشد، زیرا در صورت وجود اختلاف بین تعداد نمونه‌ها به محافظه‌کارانه بودن تمایل دارد.

روش شِفه[2] روش دیگری برای مقایسه‌های دوبه‌دو، با آمارهٔ آزمونی دقیقاً مانند روش کمترین اختلاف معنادار و روش بانفرونی است، ولی نقطهٔ بحرانی آن با استفاده از نقطهٔ بحرانی آزمون تجزیهٔ واریانس یعنی F به صورت $\sqrt{df_۱ F_{df_۱, df_۲, \alpha}}$ به دست می‌آید. روش شفه کاملاً با نتیجهٔ آزمون تجزیهٔ واریانس سازگار است؛ یعنی اگر آزمون تجزیهٔ واریانس معنادار نباشد، نتایج هیچ‌یک از آزمون‌های شفه برای مقایسه‌های دوبه‌دو معنادار نمی‌شود و اگر آزمون تجزیهٔ واریانس معنادار باشد، به یقین یکی از آزمون‌های مقایسه‌های دوبه‌دو با روش شفه معنادار خواهد بود.

در برخی موارد، یکی از جامعه‌ها از اهمیت خاصی برخوردار است و مقایسه‌های دوبه‌دو تنها به مقایسهٔ میانگین‌های جامعه‌های دیگر با میانگین آن جامعه محدود می‌شود. به این ترتیب، به جای $c = \frac{k(k-۱)}{۲}$ مقایسه، تنها به $k - ۱$ مقایسهٔ دوبه‌دو نیاز است. برای مثال، اگر دانشگاه «الف» در جدول ۳ بهترین دانشگاه کشور باشد، تنها مقایسهٔ دو دانشگاه دیگر با این دانشگاه برای پژوهشگر جالب است. بنابراین، او به جای اجرای سه مقایسهٔ جدول ۷

1. Tukey
2. Scheffe

به اجرای دو مقایسه نیاز دارد: دانشگاه «ب» با «الف» و دانشگاه «پ» با «الف». روش دانِت[1] اصلاحی بر روش کمترین اختلاف معنادار است، در شرایطی که تنها مقایسهٔ یکی از جامعه‌ها با بقیهٔ آنها هدف مقایسه‌های دوبه‌دو باشد. آمارهٔ آزمون در روش دانت مانند روش‌های کمترین اختلاف معنادار، بانفرونی و شفه است، ولی نقطهٔ بحرانی آن بر اساس مقدار df، k و α از جدول مربوط به این آزمون در پیوست به دست می‌آید.

۲ـ۳ـ۲ـ پیش‌فرض‌های آزمون تجزیهٔ واریانس

مقایسهٔ میانگین‌های چند جامعه در شرایطی از طریق آزمون تجزیهٔ واریانس یا همان آزمون F امکان‌پذیر است که داده‌های گردآوری‌شده از این جامعه‌ها دارای سه ویژگی باشند: مستقل بودن نمونه‌ها، برابری واریانس‌ها و نرمال بودن توزیع متغیر. در واقع، این ویژگی‌ها پیش‌فرض‌های آزمون تجزیهٔ واریانس هستند.

استقلال نمونه‌ها

درست مانند آزمون مقایسهٔ میانگین‌های دو جامعه مستقل، آزمون F نیز با این فرض صورت می‌گیرد که نمونه‌های انتخاب‌شده از جامعه‌ها مستقل از یکدیگرند. این بدان معناست که افراد مختلفی از هر جامعه بدون ارتباط با جامعه دیگر به عنوان نمونه انتخاب می‌شوند تا داده‌ها از آنان گردآوری شود.

برای مثال، در مقایسهٔ میانگین‌های سه دانشگاه در جدول ۳، نمونه‌ای ۹ تایی از جامعه دانش‌آموختگان دانشگاه «الف»، نمونه‌ای ۹ تایی از جامعه دانش‌آموختگان دانشگاه «ب» و نمونه‌ای ۱۰ تایی از جامعه دانش‌آموختگان دانشگاه «پ» انتخاب شده بود. این نمونه‌ها از سه جامعه متفاوت و بی‌ارتباط با یکدیگر انتخاب شده بودند، به طوری که مثلاً ۹ دانش‌آموختهٔ انتخابی از دانشگاه «الف» افرادی کاملاً بی‌ارتباط با ۱۰ دانش‌آموختهٔ انتخابی از دانشگاه «پ» بودند. از این رو، این سه نمونه مستقل از یکدیگر قلمداد می‌شوند.

در مقابل، درنظر بگیرید که پژوهشگری در پی مقایسهٔ میانگین معدل‌های مقطع کارشناسی ارشد، کارشناسی و دیپلم دانش‌آموختگان مقطع کارشناسی ارشد دانشگاه‌هاست. برای این منظور، او نمونه‌ای از دانش‌آموختگان مقطع کارشناسی ارشد را انتخاب و معدل کارشناسی ارشد، کارشناسی و دیپلم آنان را ثبت می‌کند. به این ترتیب، او سه دسته داده در

1. Dunnett

اختیار دارد که شامل معدل‌های کارشناسی ارشد، کارشناسی و دیپلم دانش‌آموختگانی است که به عنوان نمونه انتخاب شده‌اند. هدف پژوهشگر نیز مقایسهٔ میانگین‌های این سه دسته داده است، ولی این داده‌ها به یکدیگر وابسته‌اند، زیرا از سه گروه مستقل از دانش‌آموختگان گردآوری نشده‌اند، بلکه سه دسته داده از گروه ثابتی از دانش‌آموختگان مقطع کارشناسی ارشد هستند. بنابراین، نمی‌توان از آزمون F (تجزیهٔ واریانس) برای مقایسهٔ این سه میانگین کمک گرفت (این مثال مشابه مقایسهٔ میانگین دو جامعه وابسته است).

برابری یاهمگنی واریانس‌ها (آزمون وِلچ)

ضروری است در آزمون F، متغیر در جامعه‌های تحت بررسی دارای واریانس یکسانی باشد. در مثال جدول ۳ با سه واریانس روبه‌رو هستیم: واریانس معدل‌های دانش‌آموختگان دانشگاه «الف»، واریانس معدل‌های دانش‌آموختگان دانشگاه «ب» و واریانس معدل‌های دانش‌آموختگان دانشگاه «پ». استفاده از آزمون F (تجزیهٔ واریانس) برای مقایسهٔ میانگین‌های معدل‌های دانش‌آموختگان سه دانشگاه تنها در شرایطی امکان‌پذیر است که این سه واریانس برابر باشند، یعنی پراکندگی معدل‌ها در سه دانشگاه به یک میزان باشد.

در واقع، آزمون F (تجزیهٔ واریانس) معادل آزمون مقایسهٔ میانگین‌های دو جامعه مستقل با واریانس‌های برابر است. از همین رو، MS_w در آزمون F نیز به ترتیب زیر مانند واریانس آمیخته در آزمون t محاسبه می‌شود. به عبارت دیگر، MS_w برآوردی از واریانس جامعه‌ها تحت فرضیه برابری آنها به دست می‌دهد که از ترکیب داده‌های نمونه‌های مختلف محاسبه شده است.

$$MS_w = \frac{SS_w}{n-k} = \frac{(n_۱-۱)S_۱^۲ + (n_۲-۱)S_۱^۲ + \cdots (n_k-۱)S_k^۲}{(n_۱-۱)+(n_۲-۱)+\cdots+(n_k-۱)}$$

اگر واریانس‌ها برابر نباشند، آزمون وِلچ[1] به جای آزمون تجزیهٔ واریانس برای مقایسهٔ میانگین‌های چند جامعه با واریانس نابرابر به کار می‌رود. آزمون ولچ از منطقی مانند آزمون مقایسهٔ میانگین‌های دو جامعه مستقل با واریانس نابرابر پیروی می‌کند. آمارهٔ آزمون آن نیز دارای توزیع F ولی با dfهای متفاوت با آزمون تجزیهٔ واریانس است (به یاد بیاورید که در آزمون t نیز آمارهٔ آزمون دارای توزیع t بود و df آن از تقریب سترزویت محاسبه می‌شد).

1. Welch

گفتنی است آنچه آزمون تجزیهٔ واریانس را در شرایطی که واریانس‌ها برابرند، برتر از آزمون ولچ می‌سازد، توان بیشتر آزمون تجزیهٔ واریانس است (در واقع، آماره F در آزمون ولچ به طور تقریبی دارای توزیع F ولی آماره F در تجزیه واریانس تحت برقراری شرایط آن، دقیقاً دارای توزیع F است).

نرمال‌بودن توزیع

باید توزیع متغیر در تک‌تک جامعه‌ها نرمال باشد. در واقع، نرمال‌بودن توزیع متغیر است که استفاده از جدول چندک‌های توزیع F را برای یافتن نقاط بحرانی آزمون مجاز می‌سازد. بنابراین، بدون برقراری این پیش‌فرض، مرزی در اختیار نخواهید داشت تا آمارهٔ محاسبه‌شده را با آن مقایسه کنید و درباره رد فرضیه صفر تصمیم بگیرید.

برای روشن شدن موضوع، مثال جدول ۳ را به یاد بیاورید که میانگین‌های معدل‌های دانش‌آموختگان سه دانشگاه را مقایسه می‌کرد. معدل‌های دانش‌آموختگان هر دانشگاه متغیری پیوسته و دارای منحنی توزیع خاص خود است. پس با سه منحنی توزیع روبه‌رو هستیم که هر یک باید مانند منحنی توزیع نرمال باشد تا بتوانیم از آزمون تجزیهٔ واریانس برای مقایسهٔ میانگین‌های این سه دانشگاه استفاده کنیم. خوشبختانه اگر تعداد نمونه انتخابی از هر جامعه به اندازه کافی بزرگ باشد، نرمال‌نبودن توزیع متغیر در جامعه‌ها نگران‌کننده نخواهد بود. در غیر این صورت، اطمینان از نرمال‌بودن توزیع متغیر در تک‌تک جامعه‌ها ضروری است[1].

۳ـ۳ـ۲ـ شدت ارتباط (ضریب اِتا دو)

پیشتر اشاره شد که در آزمون مقایسهٔ چند میانگین، از دو متغیر صحبت به میان می‌آید: متغیری مقوله‌ای که مقوله‌های آن، همان جامعه‌ها هستند و متغیری پیوسته (فاصله‌ای یا نسبتی) که میانگین آن بین جامعه‌ها (مقوله‌ها) مقایسه می‌شود. در مثال جدول ۳، متغیر مقوله‌ای «دانشگاه» با سه وضعیت دانشگاه «الف»، «ب» و «پ» است. معدل دانش‌آموختگان نیز متغیر پیوسته‌ای است که استنباط آماری درباره آن صورت می‌گیرد؛ یعنی میانگین آن بین سه دانشگاه مقایسه می‌شود.

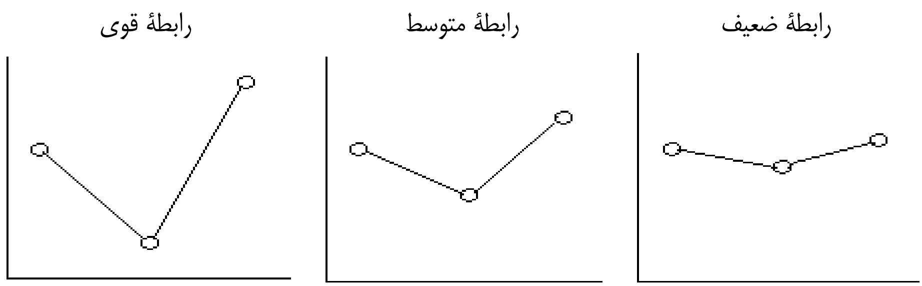
---

۱. پیش‌فرض نرمال‌بودن تنها به آزمون تجزیهٔ واریانس اختصاص ندارد، بلکه آزمون مربوط به یک میانگین و آزمون‌های مقایسهٔ دو میانگین نیز به این پیش‌فرض وابسته‌اند، به طوری که تحت برقراری آن می‌توان از چندک‌های توزیع t برای یافتن نقاط بحرانی این آزمون‌ها استفاده کرد. از همین رو، هنگام بحث از این آزمون‌ها بر کافی‌بودن تعداد نمونه تأکید می‌شد.

اگر فرضیه صفر در آزمون مقایسهٔ میانگین‌ها رد شود، نتیجه خواهیم گرفت که حداقل میانگین‌های متغیر پیوسته در دو مقوله (جامعه) با یکدیگر اختلاف دارند. این بدان معناست که پیوندی بین متغیر مقوله‌ای و متغیر پیوسته وجود دارد، به طوری که میانگین متغیر پیوسته با تغییر وضعیت متغیر مقوله‌ای تغییر می‌کند. برای مثال، میانگین معدل‌های دانش‌آموختگان دانشگاه «الف» متفاوت با دانشگاه «ب» است، پس اگر وضعیت دانشگاه از «الف» به «ب» تغییر کند، با تغییر میانگین معدل‌های دانش‌آموختگان روبه‌رو می‌شویم.

در آمار، هنگامی که تغییر در وضعیت متغیری با تغییر در وضعیت‌های متغیر دیگری همراه باشد، بین آن دو متغیر رابطه وجود دارد. رابطه می‌تواند از شدت‌های گوناگونی برخوردار باشد. رابطهٔ قوی بین دو متغیر بیان می‌کند که تغییر در وضعیت یک متغیر با تغییری چشمگیر در وضعیت متغیر دیگر همراه است. شکل ۵ در قالب سه نمودار، چند حالت فرضی را برای میانگین‌های معدل‌های دانش‌آموختگان سه دانشگاه مثال جدول ۳ نشان می‌دهد. هر دایره، میانگین معدل‌های یک دانشگاه را نشان می‌دهد (فرض کنید از راست به چپ دایره‌ها به ترتیب به دانشگاه «الف»، «ب» و «پ» اختصاص دارند). این دایره‌ها با خط‌هایی به یکدیگر متصل شده‌اند تا نوسان‌های میانگین‌ها بهتر نمایان شود. مقایسهٔ میانگین‌ها در قالب چنین نموداری را تحلیل نیمرخی آنها می‌نامند.

شکل ۵: تحلیل نیمرخی برای نمایش شدت رابطه (اختلاف میانگین‌ها)



اگرچه در هر سه نمودار، میانگین برای دانشگاه «الف» بیش از دو دانشگاه دیگر و میانگین دانشگاه «ب» کمتر از دو دانشگاه دیگر است، ولی این اختلاف میانگین‌ها در نمودارها از شدت یکسانی برخوردار نیست. در رابطهٔ قوی، اختلاف میانگین‌ها زیاد است، از همین رو

شاهد افت‌وخیز قابل توجهی در نمودار هستیم، در حالی که این نوسان‌ها در رابطهٔ متوسط کم می‌شود و در رابطهٔ ضعیف به حداقل می‌رسد.

تحلیل نیمرخی، ابزاری تصویری است که از طریق نمودار و به عبارتی به صورت چشمی، به درک شدت ارتباط کمک می‌کند، ولی معیاری کمّی به دست نمی‌دهد تا مبنای قضاوت باشد. معیار «اِتا دو[1]» که با $\eta^۲$ نشان داده می‌شود، شدت رابطهٔ متغیر مقوله‌ای و متغیر پیوسته را در آزمون تجزیهٔ واریانس اندازه می‌گیرد و بر اساس آن می‌توان قضاوت کرد که تا چه اندازه رابطهٔ دو متغیر (اختلاف میانگین‌ها) قوی است (به آن، معیار اندازه اثر نیز می‌گویند). این ضریب به شکل زیر از تقسیم تغییرپذیری بین جامعه‌ها بر تغییرپذیری کل محاسبه می‌شود:

$$\eta^۲ = \frac{SS_b}{SS_t}$$

بنابراین، مقدار $\eta^۲$ برابر با نسبتی از تغییرپذیری کل است که ناشی از تفاوت بین جامعه‌هاست. مقدار آن نیز همواره از ۰ تا ۱ است، به طوری که هر چه به ۱ نزدیک‌تر باشد، شدت رابطه قوی‌تر است. مشکل اساسی $\eta^۲$ بزرگ جلوه‌دادن شدت رابطه است. از این رو به جای آن از معیار دقیق‌تری به نام «اُمِگا دو[2]» برای اندازه‌گیری شدت رابطه استفاده می‌شود. این معیار نیز همواره مقداری از ۰ تا ۱ دارد که مانند مقادیر $\eta^۲$ تفسیر می‌شود.

$$\omega^۲ = \frac{SS_b - (k-۱)MS_w}{SS_t + MS_w}$$

این دو معیار با توجه به اطلاعات جدول ۴ به صورت زیر محاسبه می‌شوند. بر اساس مقدار $\eta^۲$ باید گفت حدود ۳۶ درصد از تغییرات معدل‌های دانش‌آموختگان ناشی از تفاوت دانشگاه‌هاست؛ این رقم شدت ارتباط متغیر معدل با متغیر نوع دانشگاه را در پایین‌ترین حد متوسط نشان می‌دهد[3]. از سوی دیگر، معیار $\omega^۲$ بیان می‌کند که حدود ۳۰ درصد از این تغییرات ناشی از تفاوت دانشگاه‌هاست و شدت رابطه تقریباً در حد ضعیف است.

$$\eta^۲ = \frac{۳۶/۱۲۵}{۱۰۰/۶۷۷} = ۰/۳۵۸ \qquad \omega^۲ = \frac{۳۶/۱۲۵ - (۳-۱)۲/۵۸۲}{۱۰۰/۶۷۷ + ۲/۵۸۲} = ۰/۲۹۹$$

1. Eta-square

2. Omega-square

۳. معمولاً مقدار ۰ تا ۰/۳ به عنوان رابطهٔ ضعیف، ۰/۳ تا ۰/۷ به عنوان رابطهٔ متوسط و بیش از ۰/۷ به عنوان رابطهٔ قوی تفسیر می‌شود. بنابراین، ۰/۳۵۸ را پایین‌ترین حد رابطهٔ متوسط در نظر می‌گیریم.

۳ـ آزمون برابری واریانس‌ها

برابری (همگنی یا همسانی) واریانس‌ها یکی از پیش‌فرض‌های آزمون تجزیهٔ واریانس است که در بخش قبل به آن اشاره شد. همچنین، دو آزمون برای مقایسهٔ میانگین‌های دو جامعه مستقل معرفی شد که تفاوت آنها در برابری یا نابرابری واریانس‌های دو جامعه بود. بنابراین، پژوهشگر پیش از به‌کارگیری این آزمون‌ها باید بداند که آیا اختلافی بین واریانس‌های جامعه‌ها وجود دارد؟ آزمون‌های برابری واریانس‌ها قادرند به این پرسش پاسخ دهند.

فرض کنید k جامعه وجود دارد که واریانس متغیر تحت بررسی را در این جامعه‌ها با $\sigma_۱^۲$، $\sigma_۲^۲$، ... و $\sigma_k^۲$ نشان می‌دهیم. فرضیه صفرِ آزمون برابری واریانس‌ها عبارت است از $\sigma_۱^۲ = \sigma_۲^۲ = \cdots = \sigma_k^۲$. در این بخش سه آزمون برای مقایسهٔ واریانس‌ها معرفی می‌شود: آزمون بارتلِت، آزمون لِوین و آزمون براون‌ـفورسیت[1].

آزمون بارتلت تحت شرایطی مناسب است که توزیع متغیرِ تحت بررسی در تک‌تک جامعه‌ها نرمال و اندازه نمونه، بزرگ باشد. آمارهٔ این آزمون به صورت زیر است:

$$Q = \frac{(n-k)\mathrm{Ln}(MS_w) - \sum_{j=۱}^{k}(n_j - ۱)\mathrm{Ln}(S_j^۲)}{۱ + \frac{۱}{۳(k-۱)}\left[\sum_{j=۱}^{k}\frac{۱}{n_j - ۱} - \left(\frac{۱}{n-k}\right)\right]}$$

در رابطهٔ فوق، Ln نشان‌دهنده لگاریتم طبیعی است و نمادهای دیگر نیز در بحث از آزمون تجزیهٔ واریانس معرفی شده‌اند. نقطهٔ بحرانی این آزمون از جدول چندک‌های توزیع $\chi^۲$ به ازای $df = k-۱$ و α مورد نظر تعیین می‌شود.

آمارهٔ آزمون بارتلت بر اساس اطلاعات جدول ۳ و ۴ به صورت زیر محاسبه می‌شود:

$$Q = \frac{۲۵\mathrm{Ln}(۲/۵۸۲) - \left[۸\mathrm{Ln}(۰/۵۲۳) + ۸\mathrm{Ln}(۱/۷۰۵) + ۹\mathrm{Ln}(۵/۱۹۲)\right]}{۱ + \frac{۱}{۳\times۲}\left[\left(\frac{۱}{۸} + \frac{۱}{۸} + \frac{۱}{۹}\right) - \frac{۱}{۲۵}\right]} = ۱۱/۴۲۹$$

اگر برای اجرای آزمون $\alpha = ۰/۰۵$ در نظر بگیریم، باید مقدار آمارهٔ فوق را با $\chi^۲_{۲,۰/۰۵}$ به

1. Bartlett, Levene, Brown-Forsythe

عنوان نقطهٔ بحرانی آزمون مقایسه کنیم. با مراجعه به جدول توزیع $\chi^۲$ درمی‌یابیم که مقدار آماره (۱۱/۴۲۹) از نقطهٔ بحرانی $\chi^۲_{۲,۰/۰۵} = ۵/۹۹۱$ بزرگ‌تر است، پس فرضیه صفر برابری واریانس‌های معدل‌های دانش‌آموختگان سه دانشگاه یعنی $\sigma_۱^۲ = \sigma_۲^۲ = \sigma_۳^۲$ رد می‌شود. این بدان معناست که ما مجاز نبودیم از آزمون تجزیهٔ واریانس برای مقایسهٔ میانگین‌های این سه دانشگاه استفاده کنیم!

آزمون لوین کمتر از آزمون بارتلت به نرمال‌بودن توزیع متغیر وابسته است. برای اجرای این آزمون مطابق جدول ۸، ابتدا با محاسبه قدر مطلق تفاضل هر داده از میانگین آن تبدیلی بر داده‌ها اعمال می‌شود و داده‌های تبدیل‌شده برای اجرای آزمون به کار گرفته می‌شوند.

جدول ۸: تبدیل داده‌ها برای اجرای آزمون لوین

	نمونه اول	نمونه دوم	…	نمونه kام
داده‌های اصلی	$X_{۱۱}, X_{۱۲}, \ldots, X_{۱n_۱}$	$X_{۲۱}, X_{۲۲}, \ldots, X_{۲n_۲}$	…	$X_{k۱}, X_{k۲}, \ldots, X_{kn_k}$
تعداد	$n_۱$	$n_۲$	…	n_k
میانگین	$\bar{X}_۱$	$\bar{X}_۲$	…	$\bar{X}_k$
داده‌های تبدیل‌شده	$A_{۱i} = \left\lvert X_{۱i} - \bar{X}_۱ \right\rvert$ $A_{۱۱}, A_{۱۲}, \ldots, A_{۱n_۱}$	$A_{۲i} = \left\lvert X_{۲i} - \bar{X}_۲ \right\rvert$ $A_{۲۱}, A_{۲۲}, \ldots, A_{۲n_۲}$	…	$A_{ki} = \left\lvert X_{ki} - \bar{X}_k \right\rvert$ $A_{k۱}, A_{k۲}, \ldots, A_{kn_k}$
میانگین	$\bar{A}_۱ = \frac{\sum_{i=۱}^{n_۱} A_{۱i}}{n_۱}$	$\bar{A}_۲ = \frac{\sum_{i=۱}^{n_۲} A_{۲i}}{n_۲}$	…	$\bar{A}_k = \frac{\sum_{i=۱}^{n_k} A_{ki}}{n_k}$
واریانس	$S_{۱A}^۲ = \frac{\sum_{i=۱}^{n_۱} (A_{۱i} - \bar{A}_۱)^۲}{n_۱ - ۱}$	$S_{۲A}^۲ = \frac{\sum_{i=۱}^{n_۲} (A_{۲i} - \bar{A}_۲)^۲}{n_۲ - ۱}$	…	$S_{kA}^۲ = \frac{\sum_{i=۱}^{n_k} (A_{ki} - \bar{A}_k)^۲}{n_k - ۱}$

در مرحلهٔ بعد، آمارهٔ آزمون، درست مانند محاسبات آزمون تجزیهٔ واریانس، بر روی داده‌های تبدیل‌شده به صورت زیر محاسبه می‌شود:

$$W = \frac{MS_b}{MS_w} = \frac{n-k}{k-۱} \frac{\sum_{j=۱}^{k} n_j (\bar{A}_j - \bar{A})}{\sum_{j=۱}^{k} (n_j - ۱) S_{jA}^۲} = \frac{n-k}{k-۱} \frac{\sum_{j=۱}^{k} n_j (\bar{A}_j - \bar{A})}{\sum_{j=۱}^{k} \sum_{i=۱}^{n_j} (A_{ji} - \bar{A}_j)^۲}$$

نقطهٔ بحرانی این آزمون نیز مانند آزمون تجزیهٔ واریانس $F_{k-1,n-k,\alpha}$ است، پس فرضیه برابری واریانس‌ها هنگامی رد می‌شود که مقدار W از $F_{k-1,n-k,\alpha}$ بزرگ‌تر باشد.

اکنون این آزمون را برای داده‌های جدول ۳ به کار می‌گیریم. ابتدا باید قدرمطلق تفاضل داده‌های هر نمونه از «میانگین» آن نمونه را محاسبه کنیم تا بتوان میانگین‌ها ($\bar{A}_j$ ها) و واریانس‌های (S^2_{Aj}ها) داده‌های تبدیل‌شده را محاسبه کرد. جدول زیر مقادیر میانگین و واریانس را پس از تبدیل داده‌ها ارائه می‌کند.

دانشگاه	الف (۱)	ب (۲)	پ (۳)	کل
تعداد	$n_1 = ۹$	$n_2 = ۹$	$n_3 = ۱۰$	$n = ۲۸$
میانگین	$\bar{A}_1 = ۰/۶۳۶۹$	$\bar{A}_2 = ۱/۰۱۲۴$	$\bar{A}_3 = ۱/۸۴$	$\bar{A} = ۱/۱۸۷۳$
واریانس	$S^2_{A1} = ۰/۰۶۷$	$S^2_{A2} = ۰/۵۵۲$	$S^2_{A3} = ۱/۴۳$	

با توجه به اطلاعات جدول فوق، آمارهٔ آزمون به صورت زیر محاسبه می‌شود. از آنجا که مقدار آماره (۵/۰۹۳) از نقطهٔ بحرانی $F_{2,25,0/05} = ۳/۳۹$ بزرگ‌تر است، فرضیه برابری واریانس‌های معدل‌های دانش‌آموختگان سه دانشگاه یعنی $\sigma_1^2 = \sigma_2^2 = \sigma_3^2$ رد می‌شود.

$$w = \frac{۲۵}{۲}\ \frac{۹\times(۰/۶۳۶۹-۱/۱۸۷۳)^2 + ۹\times(۱/۰۱۲۴-۱/۱۸۷۳)^2 + ۱۰\times(۱/۸۴-۱/۱۸۷۳)^2}{۸\times۰/۰۶۷ + ۸\times۰/۵۵۲ + ۹\times۱/۴۳} = ۵/۰۹۳$$

آزمون براون‌ـفورسیت دارای فرایندی مشابه آزمون لوین است، با این تفاوت که در تبدیل داده‌ها از قدرمطلق تفاضل داده‌های هر نمونه از «میانهٔ» آن نمونه استفاده می‌شود، یعنی ابتدا میانهٔ نمونه jام یعنی M_j محاسبه می‌شود و سپس داده‌های این نمونه به صورت $A_{ij} = \left|X_{ji-M_j}\right|$ تبدیل می‌شود و آمارهٔ آزمون مانند آزمون لوین بر اساس این داده‌ها به دست می‌آید. آزمون براون‌ـفورسیت برای داده‌هایی که دارای توزیع نرمال نیستند، مناسب‌تر از آزمون لوین است. اطلاعات لازم برای اجرای آزمون براون‌ـفورسیت برای داده‌های جدول ۳ در زیر ارائه شده‌اند.

دانشگاه	الف (۱)	ب (۲)	پ (۳)	کل
تعداد	$n_1 = ۹$	$n_2 = ۹$	$n_3 = ۱۰$	$n = ۲۸$
میانگین	$\bar{A}_1 = ۰/۶۱۱۱$	$\bar{A}_2 = ۰/۹۷۲۲$	$\bar{A}_3 = ۱/۷۵$	$\bar{A} = ۱/۱۳۳۹$
واریانس	$S^2_{A1} = ۰/۱۵۸$	$S^2_{A2} = ۰/۷۸۸$	$S^2_{A3} = ۲/۰۱۴$	

اکنون می‌توان آمارهٔ آزمون را به صورت زیر به دست آورد:

$$w = \frac{۲۵}{۲}\ \frac{۹\times(۰/۶۱۱۱-۱/۱۳۳۹)^۲+۹\times(۰/۹۷۲۲-۱/۱۳۳۹)^۲+۱۰\times(۱/۷۵-۱/۱۳۳۹)^۲}{۸\times۰/۱۵۸+۸\times۰/۷۸۸+۹\times۲/۰۱۴} = ۳/۱۵۸$$

برخلاف دو آزمون بارتلت و آزمون لوین، مقدار آمارهٔ آزمون براون‌ـفورسیت یعنی ۳/۱۵۸، کمتر از نقطهٔ بحرانی $F_{۲,۲۵,۰/۰۵} = ۳/۳۹$ است، پس فرضیه صفر برابری واریانس‌های معدل‌های دانش‌آموختگان سه دانشگاه یعنی $\sigma_۱^۲ = \sigma_۲^۲ = \sigma_۳^۲$ رد نمی‌شود. بنابراین، استفاده از تجزیهٔ واریانس با استناد به نتیجهٔ این آزمون مجاز خواهد بود.

۴ـ ضریب همبستگی

پیشتر اشاره شد که هر گاه تغییر مقادیر متغیری با تغییر مقادیر متغیر دیگری همراه باشد، بین آن دو متغیر رابطه وجود دارد. همچنین، دو معیار $\eta^۲$ و $\omega^۲$ در پایان بحث آزمون تجزیهٔ واریانس برای اندازه‌گیری شدت رابطهٔ دو متغیر مقوله‌ای و پیوسته معرفی شد. در این بخش درباره معیاری به نام ضریب همبستگی پیرسون[1] صحبت خواهیم کرد که رابطهٔ دو متغیر پیوسته را اندازه می‌گیرد. بحث را با معرفی نمودار پراکَنش آغاز می‌کنیم که به‌خوبی می‌تواند چگونگی رابطهٔ دو متغیر را به تصویر بکشد. سپس معیار کوواریانس، ضریب همبستگی و آزمون آماری مربوط به آن معرفی می‌شود. بحث همبستگی با معرفی ضریب همبستگی دو رشته‌ای و دو رشته‌ای نقطه‌ای پایان می‌یابد.

۱ـ۴ـ نمودار پراکنش

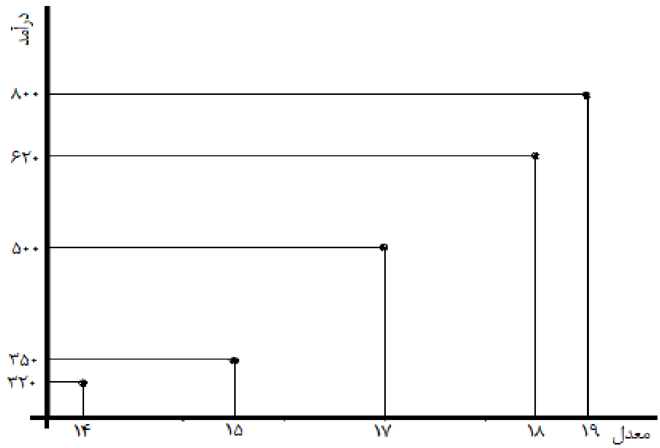
ابتدا توجه کنید که در اندازه‌گیری رابطهٔ دو متغیر از هر فرد نمونه، دو داده در دست است. برای مثال برای بررسی رابطهٔ متغیر معدل دانش‌آموختگی از دانشگاه و میزان درآمد، لازم است نمره معدل و میزان درآمد هر فرد نمونه ثبت شود. فرض کنید جدول زیر داده‌های مربوط به این دو متغیر را برای ۵ نفر ارائه می‌دهد.

فرد نمونه	اول	دوم	سوم	چهارم	پنجم
معدل	۱۸	۱۵	۱۹	۱۴	۱۷
درآمد	۶۲۰	۳۵۰	۸۰۰	۳۲۰	۵۰۰

1. Pearson

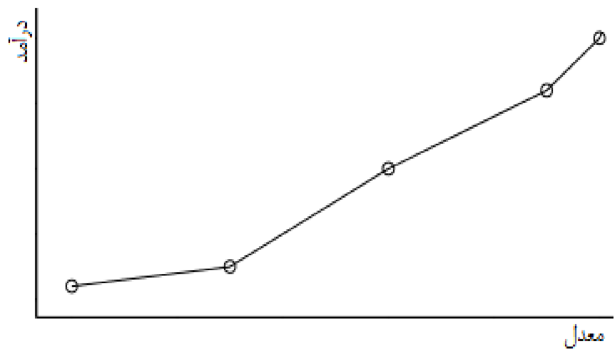
برای نمایش رابطهٔ این دو متغیر در قالب نمودار پراکنش، ابتدا باید دو محور مختصات عمود بر هم مطابق شکل ۶ رسم کرد. هر محور برای نمایش داده‌های مربوط به یکی از دو متغیر به کار می‌رود. در شکل ۶ از محور افقی برای متغیر معدل و از محور عمودی برای متغیر درآمد استفاده شده است. هر فرد با یک نقطه در این نمودار نشان داده می‌شود و محل این نقطه از تقاطع مقدار معدل او با میزان درآمد او مشخص می‌گردد. برای مثال، محل نقطهٔ مربوط به فرد سوم از تقاطع مقدار معدل ۱۹ در محور افقی با میزان درآمد ۸۰۰ هزار تومان در محور عمودی به دست آمده است. هنگامی که نقاط تمام افراد نمونه بر روی نمودار رسم شوند، به نمودار پراکنشِ متغیر معدل در برابرِ متغیر درآمد می‌رسیم.

شکل ۶: نمودار پراکنش متغیر درآمد در برابرِ متغیر معدل



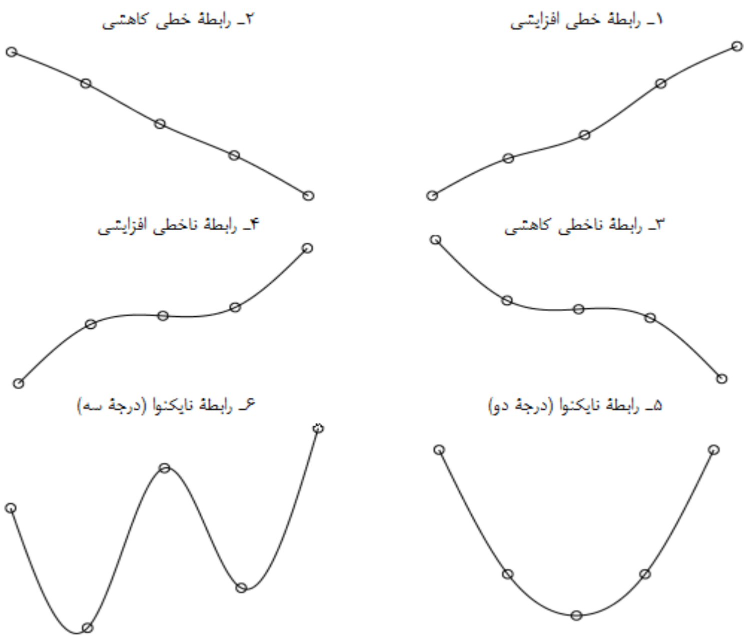
اگر دو متغیر با یکدیگر در ارتباط باشند، الگویی از نقاط نمودار پراکنش آشکار می‌شود که بر این رابطه دلالت دارد. اغلب اتصال نقاط نمودار پراکنش می‌توانند به شناسایی این الگو کمک کند. برای مثال، این نقاط می‌توانند به گونه‌ای در نمودار قرار گرفته باشند که شکلی شبیه به یک منحنی یا خط مستقیم تشکیل شود. شکل ۷ با اتصال نقاط شکل ۶ به الگویی رسیده است که بیانگر رابطه‌ای افزایشی بین دو متغیر معدل و میزان درآمد است، به طوری که هر چه معدل بیشتر باشد، درآمد نیز بیشتر می‌شود. با توجه به الگویی که در نمودار پراکنش مشاهده می‌شود، رابطه‌ها می‌توانند بسیار متنوع باشند. با وجود این، رابطه‌ها بر اساس برخی از الگوهای بااهمیت دسته‌بندی می‌شوند.

شکل ۷: الگوی نمودار پراکنش در شکل ۶



در یک دسته‌بندی، رابطه به یکنوا و نـایکنوا و رابطـهٔ یکنـوا نیـز بـه رابطـهٔ افزایشـی و کاهشی تقسیم می‌شود. در الگوی رابطه‌های یکنوا از نوسان و افت‌وخیزهای جـدی خبـری نیست، در حالی که الگوی رابطهٔ نایکنوا بالا و پایین مـی‌رود. الگوهـای ۵ و ۶ در شـکل ۸ مثال‌هایی از رابطهٔ نایکنوا هستند. به عبارت دیگر، برخلاف رابطهٔ نایکنوا، جهـت ثـابتی در الگوی رابطه‌های یکنوا مشاهده می‌شود.

شکل ۸: برخی از انواع الگوهای ارتباط دو متغیر

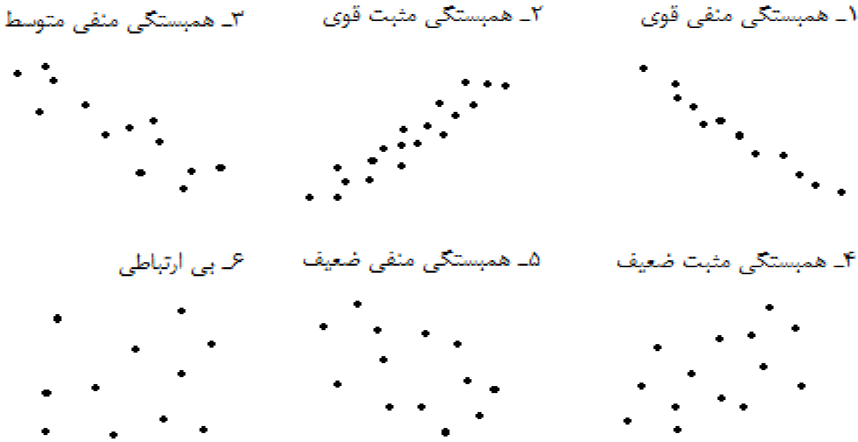


رابطهٔ یکنوا بر اساس جهت به دو دستهٔ افزایشی و کاهشی تقسیم می‌شود. اگر در رابطهٔ یکنوا با افزایش مقدار یکی از متغیرها، مقدار متغیر دیگر نیز افزایش یابد، رابطهٔ آن دو متغیر یکنوایِ افزایشی است (رابطهٔ صعودی، مستقیم یا مثبت نام‌های دیگری برای رابطهٔ یکنوایِ افزایشی هستند). در مقابل، اگر با افزایش مقدار یک متغیر، مقدار متغیر دیگر کاهش یابد، رابطهٔ آن دو متغیر یکنوایِ کاهشی است (رابطهٔ نزولی، معکوس یا منفی نیز نام‌های دیگر رابطهٔ یکنوایِ کاهشی هستند). الگوهای ۱، ۲، ۳ و ۴ در شکل ۸ مثال‌هایی از رابطه‌های یکنوای افزایشی و کاهشی ارائه می‌کنند. در مقابل، توجه کنید که رابطهٔ نایکنوای الگوی ۵ با حالتی کاهشی آغاز و سپس به حالتی افزایشی تبدیل شده است و رابطهٔ نایکنوای الگوی ۶ با نوسان روبه‌روست، یعنی ابتدا کاهش، بعد افزایش، بار دیگر کاهش و سپس افزایش یافته است.

در دسته‌بندی دیگری، رابطه به خطی و ناخطی تقسیم می‌شود. نقاط نمودار پراکنش رابطهٔ خطی، الگویی نزدیک به یک خط راست را از خود نشان می‌دهد، در حالی که رابطه‌های ناخطی می‌توانند دارای هر الگویی مانند الگوی درجهٔ دو، درجهٔ سه، سینوسی، نمایی و ... باشند. الگوهای ۱ و ۲ در شکل ۸ نمایانگر دو رابطهٔ خطی هستند. توجه کنید که اگرچه نقاط در این دو الگو دقیقاً روی یک خط راست نیستند، ولی الگو بسیار به یک الگوی خطی نزدیک است.

اغلب برای بیان رابطهٔ خطی از اصطلاح «همبستگی» استفاده می‌شود. پس اگر گفته شود بین دو متغیر همبستگی وجود دارد، این بدان معناست که بین آن دو، رابطهٔ خطی (افزایشی یا کاهشی) وجود دارد. همبستگی بین دو متغیر را می‌توان با معیار ضریب همبستگی پیرسون اندازه‌گیری کرد. این معیار هم «شدت» همبستگی و هم «جهت» آن را به دست می‌دهد. شدت همبستگی به میزان قوی‌بودن رابطهٔ «خطی» بین دو متغیر اشاره دارد و جهت همبستگی، افزایشی (مثبت) یا کاهشی (منفی) بودن این رابطه را نشان می‌دهد. نمودارهای پراکنش در شکل ۹ به درک این موضوع کمک می‌کنند. نقاط نمودارهای ۱ و ۲ در شکل ۹ بیش از سایر نمودارها در امتداد یک خط راست قرار دارند. بنابراین، شدت همبستگی دو متغیر در آنها قوی‌تر است. نقاط نمودار ۳ تاحدی به خط راست نزدیک است، ولی نقاط نمودارهای ۴ و ۵ نسبت به نمودارهای دیگر از الگوی خط راست دور هستند. نقاط در نمودار ۶ به شکلی کاملاً تصادفی پراکنده شده‌اند و بر نبودِ ارتباط بین دو متغیر اشاره دارند (الگوی خاصی در نقاط مشاهده نمی‌شود).

شکل ۹: شدت همبستگی دو متغیر در حالت‌های مختلف



۲ـ۴ـ کوواریانس و ضریب همبستگی

نمودارهای شکل ۹ به ایده‌ای برای تشخیص جهت همبستگی می‌انجامند. برای بیان این ایده، فرض کنید به دنبال یافتن جهت همبستگیِ دو متغیر X و Y هستید و برای این منظور نمونه‌ای n تایی انتخاب و این دو متغیر را از تک‌تک افراد نمونه اندازه‌گیری کرده‌اید. این اندازه‌ها را مانند سه ستون اول جدول ۹ ثبت کنید.

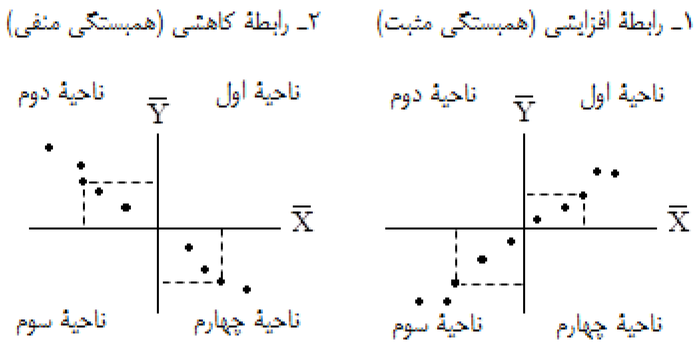
جدول ۹: اطلاعات لازم برای محاسبه ضریب همبستگی پیرسون

فرد نمونه (i)	X	Y	$X_i-\bar{X}$	$Y_i-\bar{Y}$	$(X_i-\bar{X})\times(Y_i-\bar{Y})$
۱	X_1	Y_1	$X_1-\bar{X}$	$Y_1-\bar{Y}$	$(X_1-\bar{X})\times(Y_1-\bar{Y})$
۲	X_2	Y_2	$X_2-\bar{X}$	$Y_2-\bar{Y}$	$(X_2-\bar{X})\times(Y_2-\bar{Y})$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
n	X_n	Y_n	$X_n-\bar{X}$	$Y_n-\bar{Y}$	$(X_n-\bar{X})\times(Y_n-\bar{Y})$
میانگین	$\bar{X}$	$\bar{Y}$	۰	۰	
انحراف معیار	S_X	S_Y	S_X	S_Y	

دو خط عمود بر هم مطابق شکل ۱۰ رسم کنید، به طوری که خط افقی نشان‌دهنده میانگین متغیر X و خط عمودی نشان‌دهنده میانگین متغیر Y باشد. این دو خط عمود بر هم، چهار ناحیه ایجاد می‌کنند. اکنون نمودار پراکنش دو متغیر را بر اساس دستورالعمل

شکل ۶ رسم کنید. اگر نقطه‌ای سمت راست محور عمودی قرار بگیرد، مقدار X آن بیشتر از $\overline{X}$ و اگر در سمت چپ آن باشد، مقدار X آن کمتر از $\overline{X}$ است. همچنین، اگر نقطه‌ای بالای محور افقی باشد، مقدار Y آن از $\overline{Y}$ بیشتر و اگر پایین آن باشد، مقدار Y آن کمتر از $\overline{Y}$ است.

شکل ۱۰: وضعیت قرار گرفتن نقاط در ناحیه‌های ایجادشده از طریق خط‌های عمود بر هم



بر این اساس، ویژگی نقاط ناحیهٔ اول آن است که هم مقدار X آنها از $\overline{X}$ و هم مقدار Y آنها از $\overline{Y}$ بیشتر است. ویژگی نقاط ناحیهٔ دوم آن است که مقدار X آنها از $\overline{X}$ کمتر، ولی مقدار Y آنها از $\overline{Y}$ بیشتر است. ویژگی نقاط ناحیهٔ سوم آن است که هم مقدار X آنها از $\overline{X}$ و هم مقدار Y آنها از $\overline{Y}$ کمتر است و سرانجام ویژگی نقاط ناحیهٔ چهارم آن است که مقدار X آنها از $\overline{X}$ بیشتر، ولی مقدار Y آنها از $\overline{Y}$ کمتر است.

اکنون ارتباط ناحیه‌ها را با جهت همبستگی بررسی می‌کنیم. همان‌طور که در نمودار ۱ از شکل ۱۰ دیده می‌شود، نقاط در رابطهٔ افزایشی (همبستگی مثبت) در ناحیهٔ اول و سوم قرار دارند، در حالی که نقاط در رابطهٔ کاهشی (همبستگی منفی) مانند نمودار ۲ از شکل ۱۰، در ناحیهٔ دوم و چهارم قرار دارند. به عبارت دیگر، در همبستگی مثبت، هر دو مقدار X و Y مربوط به یک نقطه، از میانگین‌های متناظرشان یا بیشتر یا کمتر هستند، در حالی که در همبستگی منفی، باید یکی از دو مقدار X یا Y مربوط به یک نقطه از میانگین آن بیشتر و مقدار دیگر از میانگین آن کمتر باشد.

بار دیگر به جدول ۹ باز می‌گردیم. سه ستون آخر این جدول شامل تفاضل مقدار متغیر X هر نقطه از میانگین آن یعنی $X_i - \overline{X}$، تفاضل مقدار متغیر Y هر نقطه از میانگین آن یعنی $Y_i - \overline{Y}$ و حاصل‌ضرب هر تفاضل یعنی $(X_i - \overline{X}) \times (Y_i - \overline{Y})$ است.

حاصل‌ضرب برای نقاطی که در ناحیهٔ اول قرار دارند، همواره مثبت است؛ زیرا هم مقدار X آنها از $\overline{X}$ و هم مقدار Y آنها از $\overline{Y}$ بیشتر و به این ترتیب $X_i - \overline{X}$ و $Y_i - \overline{Y}$ هر دو مثبت هستند. با استدلال مشابهی، حاصل‌ضرب‌های نقاط ناحیهٔ سوم نیز مثبت هستند. بنابراین، اگر نقاط در ناحیهٔ اول و سوم قرار داشته باشند، با حاصل‌ضرب‌های مثبتی روبه‌رو هستیم. در مقابل، اگر نقاط در ناحیهٔ دوم و چهارم قرار داشته باشند با حاصل‌ضرب‌های منفی روبه‌رو هستیم، زیرا علامت‌های $X_i - \overline{X}$ و $Y_i - \overline{Y}$ متفاوت و در نتیجه حاصل‌ضرب آنها منفی است.

محاسبات جدول ۹ برای داده‌های شکل ۶ در جدول ۱۰ ارائه شده است. تفاضل‌های مثبت نشان می‌دهند که داده از میانگین خود بیشتر و تفاضل‌های منفی نشان می‌دهند که داده از میانگین خود کمتر است. برای مثال، مقدار $X_i - \overline{X}$ برای فرد پنجم مثبت است، زیرا معدل ۱۷ از میانگینِ ۱۶/۶ بیشتر، ولی مقدار $Y_i - \overline{Y}$ منفی است، زیرا درآمد ۵۰۰ هزارتومان از میانگینِ ۵۱۸ هزارتومان کمتر است. از همین رو، حاصل‌ضرب، منفی و نقطهٔ مربوط به فرد پنجم در ناحیهٔ دوم قرار دارد. از سوی دیگر، نقطهٔ مربوط به فرد سوم در ناحیهٔ اول است، زیرا هم مقدار $X_i - \overline{X}$ و هم مقدار $Y_i - \overline{Y}$ برای او مثبت و به این ترتیب حاصل‌ضرب آنها نیز مثبت است.

جدول ۱۰: اطلاعات لازم برای محاسبه کوواریانس و ضریب همبستگی شکل ۶

فرد نمونه	معدل (X)	درآمد (Y)	$X_i - \overline{X}$	$Y_i - \overline{Y}$	حاصل‌ضرب
۱	۱۸	۶۲۰	۱/۴	۱۰۲	۱۴۲/۸
۲	۱۵	۳۵۰	۱/۶−	۱۶۸−	۲۶۸/۸
۳	۱۹	۸۰۰	۲/۴	۲۸۲	۶۷۶/۸
۴	۱۴	۳۲۰	۲/۶−	۱۹۸−	۵۱۴/۸
۵	۱۷	۵۰۰	۰/۴	۱۸−	۷/۲−
میانگین	۱۶/۶	۵۱۸	۰	۰	
انحراف معیار	۲/۰۷۴	۱۹۸/۵۴۳	۲/۰۷۴	۱۹۸/۵۴۳	

از آنچه گذشت می‌توان نتیجه گرفت که حاصل‌ضرب‌های مثبت بر قرارگرفتن نقاط در ناحیهٔ اول و سوم و در نتیجه همبستگی مثبت و حاصل‌ضرب‌های منفی بر قرارگرفتن نقاط در ناحیهٔ دوم و چهارم و در نتیجه همبستگی منفی دلالت دارند. در عمل، این

حاصل‌ضرب‌ها با یکدیگر جمع و میانگین آنها محاسبه می‌شود تا معیاری به نام کوواریانس[1] به صورت زیر برای تشخیص جهت همبستگی به دست آید (مانند واریانس، از $n-1$ به جای n در محاسبه این میانگین استفاده می‌شود).

$$\mathrm{Cov}(X,Y)=\frac{۱}{n-۱}\sum_{i=۱}^{n}(X_i-\bar{X})\times(Y_i-\bar{Y})$$

مقدار مثبت و منفیِ کوواریانس به ترتیب، همبستگی مثبت (رابطهٔ افزایشی) و همبستگی منفی (رابطهٔ کاهشی) را نشان می‌دهد. کوواریانس متغیر معدل و متغیر درآمد در شکل ۶، با توجه به اطلاعات جدول ۱۰ به صورت زیر است:

$$\mathrm{Cov}(X,Y)=\frac{۱}{۵-۱}(۱۴۲/۸+۲۶۸/۸+۶۷۶/۸+۵۱۴/۸-۷/۲)=۳۹۹$$

الگوی افزایشی شکل ۷ به‌خوبی مقدار مثبت به دست‌آمده را تأیید می‌کند.

اگرچه کوواریانس قابلیت تعیین جهت رابطهٔ دو متغیر را دارد، ولی نمی‌توان از مقدار آن به شدت رابطه پی برد، زیرا معیاری با مقادیر نامحدود است؛ یعنی می‌تواند هر مقدار مثبت یا منفی را اختیار کند. برای مثال، آیا باید رقم ۳۹۹ را به عنوان کوواریانس معدل و درآمد بزرگ قلمداد کنیم، در حالی که کوواریانس می‌تواند ۵۰۰، ۱۰۰۰ یا مقداری کمتر یا بیشتر از آنها نیز داشته باشد. ضریب همبستگی پیرسون معیاری است که نه تنها مانند کوواریانس، جهت همبستگی دو متغیر را نشان می‌دهد، بلکه شدت آن را نیز ارائه کند؛ زیرا برخلاف کوواریانس، مقادیر آن در محدودهٔ مشخصی قرار دارد. این ضریب به صورت زیر از تقسیم کوواریانسِ دو متغیر بر حاصل‌ضرب انحراف معیارهای آنها به دست می‌آید:

$$r_{XY}=\frac{\mathrm{Cov}(X,Y)}{S_X\,S_Y}$$

تقسیم کردن، هم باعث می‌شود که ضریب همبستگی به واحد اندازه‌گیری دو متغیر بستگی نداشته باشد و هم مقدار آن را در محدوده‌ای از ۱- تا ۱+ نگه‌می‌دارد؛ یعنی قوی‌ترین همبستگی‌ها نیز نمی‌توانند مقداری فراتر از ۱- یا ۱+ داشته باشند (درحالی که چنین محدوده‌ای برای کوواریانس وجود ندارد).

---

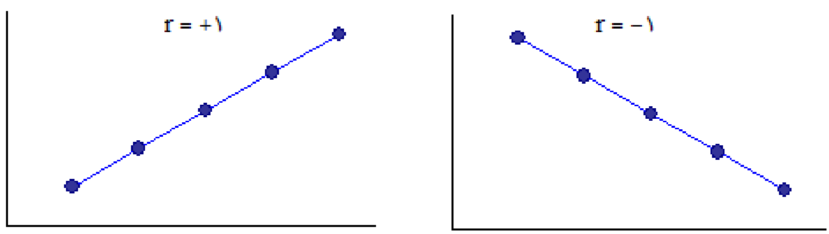
۱. کوواریانس (Covariance) به معنای هم‌تغییری است. همان‌طور که واریانس به عنوان میانگین تفاضل‌های مقادیر یک متغیر از میانگین آنها، بیانگر تغییرات آن متغیر است، کوواریانس نیز به عنوان میانگینِ حاصل‌ضرب تفاضل‌ها بیانگر تغییرات توأم دو متغیر است.

ضریب همبستگی نیز جهت رابطه را مانند کوواریانس با علامت خـود نشـان مـی‌دهـد؛ مقدار منفی ضریب همبستگی بر رابطهٔ خطی کاهشی و مقدار مثبت ضـریب همبسـتگی بـر رابطهٔ خطی افزایشی دلالت دارد. هر چه مقدار این ضریب به ۱- یـا ۱+ نزدیـک‌تـر باشـد، الگوی نقاط در نمودار پراکنش به خط راست نزدیک‌تر است؛ به طوری که مقدار ۱- یا ۱+ بیانگر رابطهٔ خطی کاملی است که در آن، نقاط درست روی یک خط راسـت ماننـد شـکل ۱۱ قرار دارند. مقدار ضریب همبستگی برای اطلاعات جدول ۱۰ که پیشتر کوواریـانس آن محاسبه شد، به قرار زیر است:

$$r = \frac{۳۹۹}{۲/۰۷۴ \times ۱۹۸/۵۴۳} = ۰/۹۶۹$$

این رقم به ۱+ بسیار نزدیک است، بنابراین همبستگی افزایشی قوی بین دو متغیـر معـدل و درآمد وجود دارد.

شکل ۱۱: همبستگی (رابطهٔ خطی) کامل بین دو متغیر



نکته قابل توجه در ضریب همبستگی، تفسیر مقدار صفر است. مقدار صـفر یـا مقـداری نزدیک به آن بیانگر نبود یا بسیار ضعیف بودن رابطهٔ «خطی» بین دو متغیر اسـت، ولـی بـه معنای آن نیست که دو متغیر اصلاً با یکدیگر رابطـه ندارنـد؛ زیـرا ممکـن اسـت رابطـه‌ای ناخطی (مانند نمودارهای ۳ و ۴ در شکل ۸) یا رابطه‌ای نایکنوا (مانند نمودارهای ۵ و ۶ در شکل ۸) بین آنها وجود داشته باشد. ضریب همبستگی پیرسون تنها قابلیت شناسایی جهـت و شدت رابطهٔ «خطی» بین دو متغیر را دارد، نه انواع دیگر رابطه‌ها را.

همان‌طور که در بحث آزمون‌های مربوط به میانگین‌ها این پرسش مطـرح بـود کـه اگـر بین میانگین متغیری در دو نمونه اختلاف مشاهده شود آیا این اختلاف را مـی‌تـوان بـه دو جامعه آماری مربوط به آن دو نمونه نیز تعمیم داد و مدعی شد میانگین‌های ایـن متغیـر در

دو جامعه نیز تفاوت دارند، اکنون این پرسش را می‌توان درباره ضریب همبستگی حاصل از نمونه نیز مطرح کرد که آیا مقدار همبستگی دو متغیر که از داده‌های نمونه به دست آمده، به جامعه آماری آن نمونه قابل تعمیم است؟ به عبارت دیگر، آیا می‌توان ادعا کرد که این همبستگی در جامعه نیز بین دو متغیر وجود دارد؟

همان‌طور که در بحث میانگین، از دو نماد $\overline{X}$ و μ استفاده شد تا بین آماره و پارامتر نظیر آن تمایز ایجاد کنیم، در بحث همبستگی نیز بین ضریب همبستگی حاصل از داده‌های نمونه که آماره قلمداد و با r نشان داده می‌شود با همبستگی واقعی بین دو متغیر در جامعه آماری تمایز قایل می‌شویم که پارامتری نامعلوم است و با حرف یونانی ρ نشان داده می‌شود. با این نمادگذاری پرسش پیشین را چنین بیان می‌کنیم: آیا از $r \neq ۰$ می‌توان $\rho \neq ۰$ را نتیجه گرفت؟ آزمونی وجود دارد که فرضیه صفر $\rho = ۰$ را در مقابل فرضیه جانشین $\rho \neq ۰$ می‌آزماید. به این ترتیب با رد فرضیه صفر نتیجه خواهیم گرفت که همبستگی مشاهده‌شده در نمونه در جامعه آماری نیز وجود دارد. این آزمون دو دامنه است؛ زیرا هم مقادیر بزرگ r و هم مقادیر کوچک r می‌توانند به رد فرضیه صفر بینجامند. آمارهٔ آزمون به صورت زیر محاسبه می‌شود:

$$t = \frac{r\sqrt{n-۲}}{\sqrt{۱-r^{۲}}}$$

نقطهٔ بحرانی آزمون نیز از جدول چندک‌های توزیع t به ازای $df = n-۲$ و $\frac{\alpha}{۲}$ تعیین می‌شود. اگر مقدار آماره بزرگ‌تر از $t_{n-۲,\frac{\alpha}{۲}}$ یا کوچک‌تر از $-t_{n-۲,\frac{\alpha}{۲}}$ باشد، فرضیه صفر رد می‌شود و نتیجه خواهیم گرفت که بین دو متغیر همبستگی وجود دارد. همواره در ذهن داشته باشید که رد نشدن فرضیه صفر به معنای وجود نداشتن ارتباط بین دو متغیر نیست.

اکنون می‌توان این محاسبات را برای اطلاعات جدول ۱۰ صورت داد. هدف، آزمون فرضیه صفر «بین معدل و درآمد همبستگی وجود ندارد» در مقابل فرضیه جانشین «بین معدل و درآمد همبستگی وجود دارد» است. آمارهٔ آزمون با توجه به $r = ۰/۹۶۹$ و $n = ۵$ عبارت است از:

$$t = \frac{۰/۹۶۹ \times \sqrt{۵-۲}}{\sqrt{۱-(۰/۹۶۹)^{۲}}} = ۶/۷۹$$

از طرفی نقطهٔ بحرانی آزمون از جدول توزیع t برابر با $t_{۳,۰/۰۲۵} = ۳/۱۸۲$ است، پس فرضیه صفر در سطح معناداری $\alpha = ۰/۰۵$ رد می‌شود؛ یعنی می‌توان نتیجه گرفت که بین معدل و

درآمد همبستگی مثبت وجود دارد. بر اساس مقدار احتمال نیز نتیجه‌ای مشابه به دست می‌آید؛ زیرا مساحت سمت راست ۶/۷۹ در توزیع t با $df=3$ با ۰/۰۰۳۲ برابر و از $\frac{\alpha}{2}=0/025$ کوچک‌تر است.

مانند آزمون تجزیهٔ واریانس، آزمون مربوط به ضریب همبستگی نیز دارای پیش‌فرضی است که در صورتِ برقراری آن، مجاز به استفاده از این آزمون هستیم. آزمون مربوط به ضریب همبستگی پیرسون برای دو متغیر پیوسته با توزیع نرمال قابل اجراست، هر چند با نگاه توصیفی می‌توان این ضریب را برای دو متغیر فاصله‌ای یا نسبتی دلخواه (بدون توزیع نرمال) نیز محاسبه کرد، مشروط بر آنکه از به‌کارگیری آزمون آن صرفنظر کنید. در فصل بعد با ضریب‌های دیگری آشنا خواهید شد که به برقراری این پیش‌فرض نیازی ندارند.

۳ـ۴ـ ضریب همبستگی دورشته‌ای و دو رشته‌ای نقطه‌ای

این دو ضریب، حالت‌های خاصی از ضریب همبستگی پیرسون هستند. فرض کنید در پی بررسی رابطه متغیر پیوسته Y با متغیر مقوله‌ای دوحالتی X با مقوله‌های ۰ و ۱ هستید؛ برای مثال، رابطه متغیر سن با متغیر وضعیت اشتغال با دو مقولهٔ بیکار=۰ و شاغل=۱. نمونه‌ای nتایی انتخاب می‌شود، به طوری که n_1 نفر از آن را شاغلان و n_0 نفر را بیکاران تشکیل می‌دهند؛ توجه کنید که $n_0+n_1=n$ و $p=\frac{n_1}{n}$ نسبتِ شاغلان نمونه است. اگر انحراف معیار متغیر سن را با s_Y، میانگین سنی شاغلان را با $\bar{y}_1$ و میانگین سنی بیکاران را با $\bar{y}_0$ نشان دهیم، همبستگی دو رشته‌ای نقطه‌ای به صورت زیر محاسبه می‌شود:

$$r_{pb}=\frac{\bar{y}_1-\bar{y}_0}{s_Y}\sqrt{\frac{n_1 n_0}{n^2}}=\frac{\bar{y}_1-\bar{y}_0}{s_Y}\sqrt{p(1-p)}$$

آزمون آماری مربوط به این ضریب مانند آزمون ضریب همبستگی پیرسون است، با این تفاوت که r_{pb} به جای r در آماره آزمون قرار می‌گیرد.

شرایطی وجود دارد که متغیر مقوله‌ای دوحالتی از دسته‌بندی یک متغیر پیوسته به دو وضعیت ایجاد شده است. برای مثال، دانش‌آموزان بر حسب نمره امتحانیِ کمتر از ۱۰ یا بیشتر از ۱۰ به دو وضعیتِ «مردود» یا «قبول» در یک درس دسته‌بندی می‌شوند. اگر ساعاتی را که دانش‌آموز به طور معمول در هفته صرف مطالعه آن درس می‌کند، متغیری

پیوسته با توزیع «نرمال» در نظر بگیریم، رابطه آن با متغیر دوحالتی وضعیت دانش‌آموز (مردود/قبول) از طریق ضریب همبستگی دورشته‌ای به صورت زیر تعیین می‌شود:

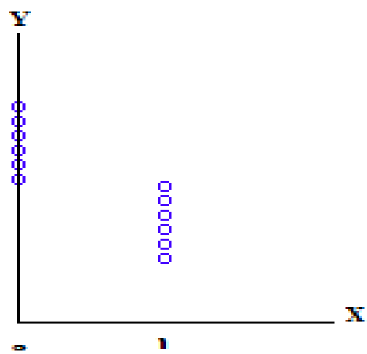
$$r_b = \frac{\bar{y}_1 - \bar{y}_0}{s_Y} \frac{p(1-p)}{f(z_p)}$$

مقدار z_p، چندکی از توزیع نرمال استاندارد است که مساحت سمت راست آن برابر با p است و $f(z_p)$، چگالی منحنی توزیع نرمال استاندارد در نقطه z_p است (رسم بافت‌نگار را از فصل دوم به یاد بیاورید). آزمون آماری مربوط به این ضریب نیز مانند آزمون ضریب همبستگی پیرسون است، ولی r_b به جای r در آماره آزمون قرار می‌گیرد.

همان طور که اشاره شد، این دو ضریب مبتنی بر ضریب همبستگی پیرسون هستند. در واقع، اگر اعداد ۰ و ۱ به وضعیت‌های متغیر مقوله‌ای X اختصاص یابد، مقدار ضریب همبستگی پیرسون X و Y با ضریب همبستگی دو رشته‌ای نقطه‌ای این دو متغیر برابر خواهد بود. با این وصف، چگونه می‌توان رابطه یک متغیر پیوسته را با متغیری مقوله‌ای از طریق ضریب همبستگی پیرسون که به رابطه‌های خطی اختصاص دارد، تفسیر کرد؟

برای پاسخ، ابتدا نمودار پراکنش متغیر پیوسته Y را در مقابل متغیر مقوله‌ای X رسم می‌کنیم. اگر اعداد ۰ و ۱ به مقوله‌ها اختصاص یابند، این نمودار شبیه به شکل ۱۲ خواهد بود، یعنی مقادیر متغیر Y در هر مقوله از متغیر X در امتداد یک خط عمودی قرار می‌گیرند. بنابراین، هرگز نمی‌توان الگویی خطی از نمودار پراکنش یک متغیر پیوسته در مقابل متغیری مقوله‌ای انتظار داشت.

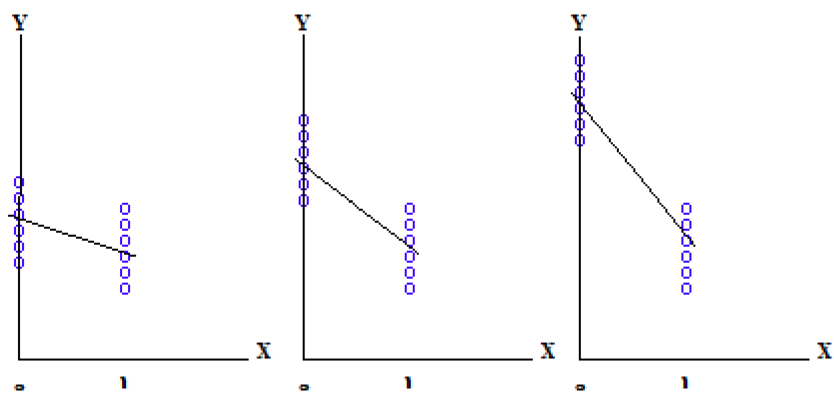
شکل ۱۲: نمودار پراکنش متغیر پیوسته در مقابل متغیر مقوله‌ای دوحالتی



با وجود این، نگاهی دقیق‌تر به شکل ۱۲ نشان می‌دهد که مقادیر Y در مقوله ۰ بالاتر از مقادیر Y در مقوله ۱ قرار دارند؛ گو اینکه ارتباطی کاهشی بین Y و X وجود دارد، به طوری که با حرکت از مقوله ۰ به مقوله ۱، «برآیند» مقادیر Y کاهش می‌یابد. بنابراین، اگر همبستگی دو رشته‌ای نقطه‌ای، منفی باشد، می‌توان نتیجه گرفت حرکت از مقوله دارای رقم کمتر به مقوله دارای رقم بیشتر با کاهش برآیند مقادیر متغیر پیوسته همراه است. از این رو، تفسیر ضریب همبستگی دو رشته‌ای نقطه‌ای کاملاً به چگونگی اختصاص اعداد به دو مقوله بستگی دارد، به طوری که اگر این اعداد بین دو مقوله جابه‌جا شوند، علامت این ضریب نیز تغییر می‌کند.

یکی از عوامل اثرگذار بر مقدار ضریب همبستگی دورشته‌ای نقطه‌ای، اختلاف میانگین متغیر پیوسته در دو مقوله یعنی اختلاف $\bar{y}_1$ با $\bar{y}_0$ است. شکل ۱۳ سه حالت مختلف را نشان می‌دهد. خط‌های مستقیم، میانگین‌های دو گروه را به هم وصل کرده‌اند. هرچه شیب این خط بیشتر باشد، رابطه دو متغیر قوی‌تر است. بنابراین، هر چه اختلاف میانگین‌های دو مقوله بیشتر باشد، خطی با شیب بیشتر و رابطه‌ای قوی‌تر خواهیم داشت (فرض کنید تحلیل نیمرخی را برای دو میانگین صورت می‌دهید؛ شکل ۵ را به یاد بیاورید). برای مثال، اگر میانگین سنی شاغلان اختلاف زیادی با میانگین سنی بیکاران داشته باشد، انتظار می‌رود همبستگی دو رشته‌ای نقطه‌ای، رابطه سن و وضعیت اشتغال را قوی نشان دهد.

شکل ۱۳: تأثیر اختلاف میانگین بر مقدار ضریب همبستگی دو رشته‌ای نقطه‌ای



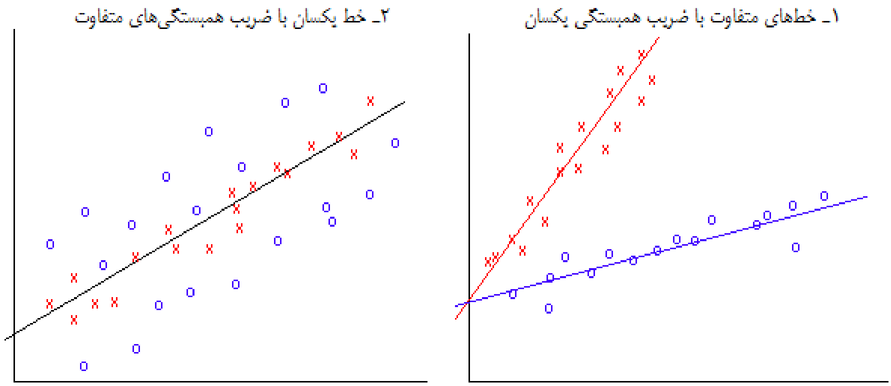
۵ـ رگرسیون

ضریب همبستگی معیاری مناسب برای تشخیص خطی‌بودن الگوی رابطهٔ دو متغیر است، ولی اطلاعی درباره معادلهٔ خطی که الگو به آن نزدیک است، نمی‌دهد. همچنین با دانستن ضریب همبستگی به شدت و جهت رابطهٔ خطی پی می‌بریم، ولی از چگونگی این رابطه اطلاعی نداریم. برای مثال، مقدار ۰/۹۶۹ برای ضریب همبستگی دو متغیر معدل و درآمد از اطلاعات جدول ۱۰ به دست آمد؛ آزمون آماری نیز همبستگی معناداری را بین این دو متغیر نشان داد و بر این اساس می‌توان دریافت که متغیر معدل با متغیر درآمد همبستگی «قوی افزایشی» دارد. با وجود این، هنوز می‌توان سؤال کرد که تغییر در مقدار معدل تا چه میزان با تغییر در مقدار درآمد همراه است؟

برای روشن شدن اهمیت این پرسش، فرض کنید معادلهٔ خطی که الگوی نمودار پراکنش دو متغیر X (معدل) و Y (درآمد) به آن نزدیک است، به صورت $Y = ۹۳X - ۱۰۲۲$ باشد. این معادله بیان می‌کند که اگر معدل فردی مثلاً $X = ۱۵$ باشد، درآمدی برابر با $Y = ۹۳ \times ۱۵ - ۱۰۲۲ = ۳۷۳$ هزار تومان خواهد داشت، در حالی که اگر معدل فردی مثلاً $X = ۱۸$ باشد، درآمد او برابر با $Y = ۹۳ \times ۱۸ - ۱۰۲۲ = ۶۵۲$ هزار تومان خواهد بود. پس می‌توان نتیجه گرفت که با تغییر مقدار متغیر X از ۱۵ به ۱۸ مقدار متغیر Y از ۳۷۳ به ۶۵۲ افزایش می‌یابد؛ یعنی ۳ نمره افزایش در معدل با ۲۷۹ هزار تومان افزایش در درآمد همراه است. به دست آوردن این اطلاعات تنها با داشتن معادله‌ای امکان‌پذیر است که رابطهٔ خطی این دو متغیر را ارائه می‌دهد.

بحث رگرسیون، راهکاری برای یافتن معادلهٔ خطی فراهم می‌کند که الگوی رابطهٔ بین دو متغیر به آن نزدیک است. با وجود این، همواره به یاد داشته باشید که اگرچه این معادله اطلاعاتی را درباره رابطهٔ دو متغیر به دست می‌دهد که ضریب همبستگی قادر به ارائه آنها نیست، ولی معادلهٔ خط تنها در کنار ضریب همبستگی می‌تواند تصویری روشن از رابطهٔ خطی دو متغیر را ترسیم کند. شکل ۱۴ دو نمودار را نشان می‌دهد که هر یک دو مجموعه از نقاط (با علامت × و o) را شامل می‌شوند. این دو مجموعه در نمودار ۱ دارای ضریب همبستگی برابری هستند، ولی الگوهای آنها به دو خط متفاوت نزدیک‌اند، در حالی که الگوهای مجموعه نمودار ۲ به یک خط نزدیک‌اند، ولی ضرایب همبستگی متفاوتی دارند (مجموعه نقاطی که به خط نزدیک‌ترند همبستگی قوی‌تری نسبت به مجموعه نقاط دورتر دارند).

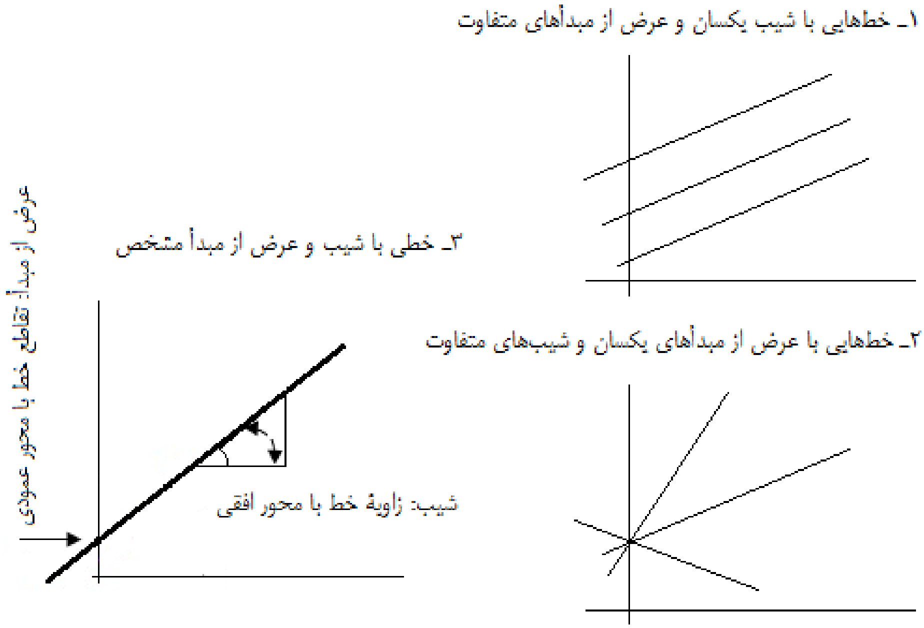
شکل ۱۴: نقش‌های تکمیلی ضریب همبستگی و معادلهٔ خط رگرسیونی



نزدیک‌ترین خط به الگوی آشکارشده از نقاط نمودار پراکنش را خط رگرسیونی می‌نامند و یافتن معادلهٔ این خط اصلی‌ترین هدف در بحث رگرسیون است. از ریاضیات به یاد دارید که محل هر خط با دو مؤلفه شیب و عرض از مبدأ مشخص می‌شود. شیبِ خط، زاویه‌ای را نشان می‌دهد که خط با محور افقی می‌سازد و عرض (فاصله) از مبدأ نقطه‌ای را نشان می‌دهد که خط با محور عمودی برخورد می‌کند. این دو مؤلفه در نمودار ۳ از شکل ۱۵ مشاهده می‌شوند.

مؤلفه‌های شیب و عرض از مبدأ با هم می‌توانند محل یک و تنها یک خط را مشخص کنند، در حالی که هر یک به تنهایی بر محل تعداد بی‌شماری خط اشاره دارند. در نمودار ۱ از شکل ۱۵ خط‌هایی با شیب یکسان مشاهده می‌شوند، زیرا همگی با هم موازی هستند. خط‌های موازی دیگری را نیز می‌توان به این نمودار اضافه کرد که دارای همان شیب باشند. بنابراین، تعیین مقدار شیب به تنهایی نمی‌تواند محل خط منحصربه‌فردی را مشخص کند. از سوی دیگر، خط‌های نمودار ۲ همگی عرض از مبدأ یکسانی دارند؛ زیرا در یک نقطه محور عمودی را قطع کرده‌اند. مانند نمودار ۱، خط‌های دیگری را نیز می‌توان در نمودار ۲ رسم کرد که در همان نقطه محور عمودی را قطع کنند. بنابراین، عرض از مبدأ نیز به تنهایی نمی‌تواند محل خط منحصر به‌فردی را مشخص کند.

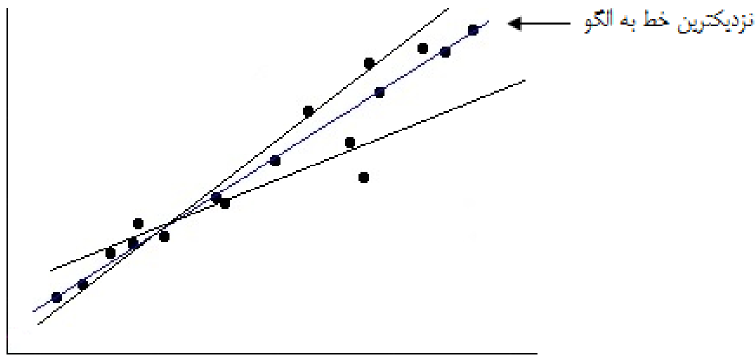
شکل ۱۵: نقش شیب و عرض از مبدأ در تعیین محل خط



اگرچه شیب یا عرض از مبدأ به تنهایی برای تعیین محل خط کافی نیستند، ولی همان‌طور که اشاره شد، هر دو با هم محل خط منحصربه‌فردی را نشان می‌دهند. پس باید معادلهٔ خطی را که به الگوی آشکارشده در نمودار پراکنش از هر خط دیگری نزدیک‌تر است، با تعیین شیب و عرض از مبدأ آن مشخص کنیم. به این ترتیب، محل خط به شکلی یکتا تعیین می‌شود و با هیچ خط دیگری اشتباه نخواهد شد. یافتن خطی نزدیک به الگوی نقاط نمودار پراکنش را برازش مدل رگرسیون یا خط رگرسیون بر داده‌ها می‌نامند.

برای یافتن معادلهٔ خط رگرسیون، معادلهٔ کلی خط را به صورت $Y = bX + a$ در نظر بگیرید که b مقدار شیب خط و a عرض از مبدأ آن است. در بحث رگرسیون به b و a ضرایب رگرسیون گفته می‌شود. با دانستن مقادیر b و a می‌توان به‌راحتی محل خط را در نمودار به دست آورد. مسئله اصلی در یافتن معادلهٔ خط رگرسیون، تعیین مقادیر b و a، یعنی ضرایب رگرسیون به صورتی است که مانند شکل ۱۶، از میان خط‌های مختلفی که از الگوی نقاط نمودار پراکنش رد می‌شوند، «نزدیک‌ترین» خط به این الگو به عنوان خط رگرسیون انتخاب شود.

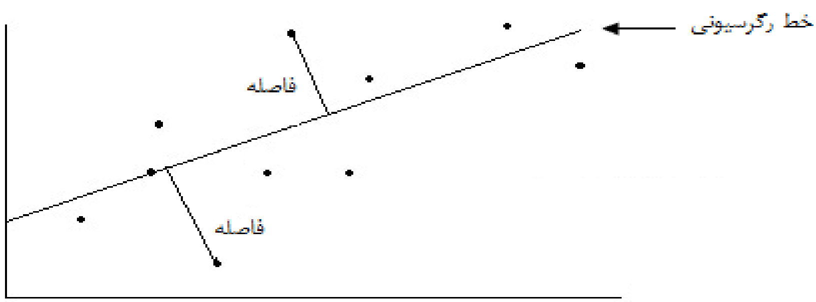
شکل ۱۶: خط رگرسیون، نزدیک‌ترین خط به الگوی نقاط نمودار پراکنش

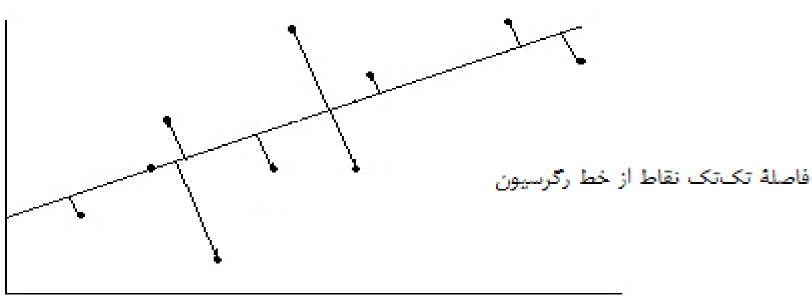


۱ـ۵ـ برآورد ضرایب رگرسیون

روشی به نام حداقل مربعات معمولی (یا کمترین توان‌های دوم معمولی) وجود دارد که مقادیر b و a را در معادلهٔ خط رگرسیون طوری تعیین می‌کند تا مجموع مربعات فاصلهٔ تک‌تک نقاط از آن خط، به کمترین مقدار در مقایسه با سایر خط‌ها برسد. همان طور که در شکل ۱۷ مشاهده می‌شود، فاصلهٔ تک‌تک نقاط از خط تعیین می‌شوند، سپس به توان دو می‌رسند و با هم جمع می‌شوند تا مجموع مربعات فاصله‌ها به دست آید.

شکل ۱۷: فاصلهٔ نقاط نمودار پراکنش از خط رگرسیونی





پس از آن، با حل معادلاتی، مقادیر a و b طوری تعیین می‌شوند که این مجموع مربعات دارای کمترین مقدار در مقایسه با مجموع مربعات مربوط بـه خـط‌هـای دیگـر باشـد. ایـن مقادیر با روش کمترین مربعات به صورت زیر محاسبه می‌شوند:

$$b = \frac{Cov(X,Y)}{S_X^۲} = \frac{S_Y}{S_X} r_{XY} \qquad a = \overline{Y} - b\overline{X}$$

اکنون این محاسبات را برای داده‌های شکل ۶ با توجه به اطلاعات جدول ۱۰ صورت می‌دهیم.

$$\overline{X} = ۱۶/۶, \quad \overline{Y} = ۵۱۸, \quad S_X = ۲/۰۷۴, \quad S_Y = ۱۹۸/۵۴۳, \quad r = ۰/۹۶۹$$

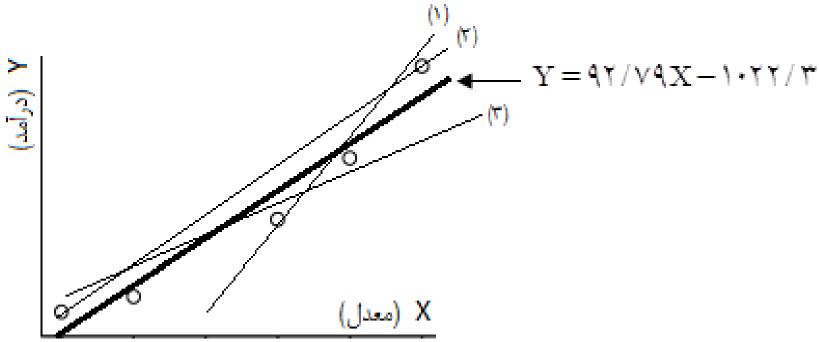
$$b = \frac{۱۹۸/۵۴۳}{۲/۰۷۴} \times ۰/۹۶۹ = ۹۲/۷۹ \quad a = ۵۱۸ - ۹۲/۷۹ \times ۱۶/۶ = -۱۰۲۲/۳$$

پس معادلهٔ خط رگرسیون برای دو متغیر معدل و درآمد عبارت است از:

$$Y = ۹۲/۷۹X - ۱۰۲۲/۳$$

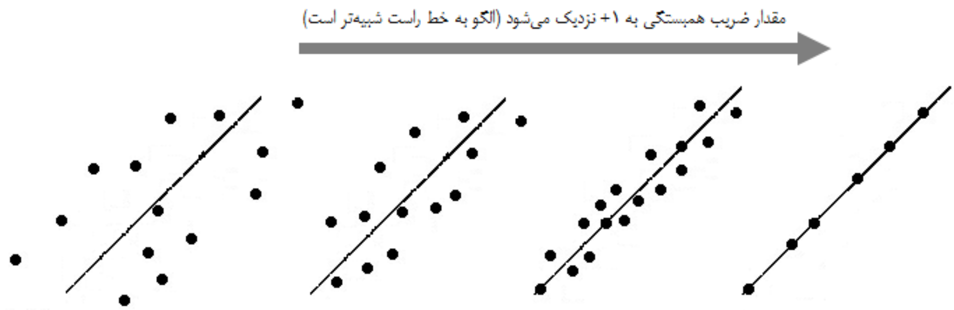
این خط در شکل ۱۸ همراه با خطوط دیگری مشاهده می‌شود که آنها نیز از میان نقـاط نمودار پراکنش می‌گذرند. با وجود این، فاصلهٔ نقاط از خطی که با استفاده از روش حـداقل مربعات معمولی به دست می‌آید، کمتر از هر خط دیگری است کـه از میـان نقـاط نمـودار پراکنش می‌گذرد. برای مثال، خط (۱) از نزدیکی سه نقطه می‌گذرد ولی از دو نقطـهٔ دیگـر به شدت دور است. خط (۲) دقیقاً از دو نقطه عبور می‌کنـد، ولـی از سـه نقطـهٔ دیگـر دور است. همچنین، خط (۳) به سه نقطه نسبتاً نزدیک، ولی از دو نقطهٔ دیگر دور است. امـا در خط رگرسیون نوعی تعادل در نزدیکی وجود دارد؛ به طوری که تمام نقاط کـم و بـیش بـه آن نزدیک‌اند با اینکه هیچ نقطه‌ای درست روی آن قرار ندارد.

شکل ۱۸: خط رگرسیونی نزدیک به الگوی نمودار پراکنش داده‌های جدول ۱۰



به یاد داشته باشید که هرگز نمی‌توان خطی یافت که تمام نقاط الگو، روی آن قرار داشته باشند، مگر آنکه مقدار ضریب همبستگی درست برابر با ۱- یا ۱+ باشد (شکل ۱۱ را ببینید). همچنین، میزان شباهت الگوی نقاط نمودار پراکنش به یک خط راست کاملاً به مقدار ضریب همبستگی آنها بستگی دارد. الگوی نقاطی که دارای ضریبی نزدیک به ۱- یا ۱+ هستند، بیشتر شبیه به یک خط راست است، از همین رو فاصلهٔ این نقاط از خط رگرسیونی به دست‌آمده کم است؛ یعنی نقاط در اطراف این خط قرار دارند. در مقابل، اگر ضریب همبستگی از ۱- یا ۱+ دور و به صفر نزدیک باشد، الگوی نقاط نمودار پراکنش شباهت کمتری به خط راست خواهد داشت و این نقاط از خط رگرسیون به دست‌آمده دور خواهند بود. شکل ۱۹، این ویژگی را برای ضرایب همبستگی مثبت (رابطهٔ افزایشی) به تصویر کشیده است.

شکل ۱۹: رابطهٔ مقدار ضریب همبستگی و نزدیکی نقاط به خط رگرسیون



خط رگرسیون نیز مانند ضریب همبستگی می‌تواند جهت رابطهٔ دو متغیر X و Y را از طریق علامت b نشان دهد؛ زیرا مقدار b با توجه به $b=\frac{S_Y}{S_X}r_{XY}$ با ضریب همبستگی هم‌علامت است. بنابراین اگر r مثبت باشد، b نیز مثبت و اگر r منفی باشد، b نیز منفی است. اما تفاوتی بین معادلهٔ خط رگرسیون و ضریب همبستگی وجود دارد و آن تفاوت نقش X و Y در معادلهٔ خط رگرسیون است. در بحث رگرسیون، اینکه کدام متغیر در سمت راست معادلهٔ $Y=bX+a$ قرار دارد و کدام متغیر در سمت چپ آن، مهم است. معادلهٔ $Y=bX+a$، رگرسیون متغیر Y بر X را نشان می‌دهد، در حالی که معادلهٔ $X=bY+a$، رگرسیون متغیر X بر Y را به نمایش می‌گذارد. توجه کنید که اغلب،

مقادیر b و a متفاوتی برای این دو معادله به دست می‌آید که تفسیر متفاوتی را نیز از رابطهٔ خطی دو متغیر ارائه می‌دهند.

تصمیم درباره اینکه کدام متغیر در سمت چپ و کدام در سمت راست معادله باشد، به منطقی باز می‌گردد که در رابطهٔ دو متغیر وجود دارد. برای مثال در رابطهٔ بین متغیر معدل و متغیر درآمد، انتظار داریم معدل بالاتر با شرایط کاری بهتر و به تبع آن درآمد بالاتر همراه باشد، پس منطقی است تغییرات درآمدی را در پاسخ به تغییرات معدل بدانیم نه بر عکس و به دنبال رگرسیون متغیر درآمد بر متغیر معدل باشیم. همچنین، اگر در پی بررسی رابطهٔ متغیر هوشبهر و متغیر اعتماد به نفس باشیم، منطقی است رابطه را از سوی هوشبهر به اعتماد به نفس بدانیم، زیرا انتظار داریم فردی با هوشبهر بالا در موقعیت‌های خاص مسلط‌تر و تواناتر بوده و اعتماد به نفس بالایی داشته باشد، یعنی تغییرات متغیر اعتماد به نفس به تغییرات متغیر هوشبهر باز می‌گردد، بنابراین باید به دنبال رگرسیون متغیر اعتماد به نفس بر متغیر هوشبهر باشیم.

برای نشان‌دادن نقش متغیرهای X و Y در معادلهٔ خط رگرسیون $Y = bX + a$ از عنوان «متغیر مستقل» برای X و از عنوان «متغیر پاسخ» یا «متغیر وابسته» برای Y استفاده می‌شود. واژهٔ «وابسته» بیان می‌کند که این تغییرات متغیر Y است که به تغییرات متغیر X باز می‌گردد و به آن بستگی دارد یا واژهٔ «پاسخ» بر این واقعیت تأکید دارد که تغییرات متغیر Y در پاسخ به تغییراتی است که در متغیر X رخ‌داده است. بنابراین، پیش از تعیین خط رگرسیون باید تشخیص داد که کدام متغیر را مستقل و کدام متغیر را وابسته قلمداد کنیم تا متغیر مستقل در سمت راست و متغیر وابسته در سمت چپ معادله رگرسیون قرار گیرد.

۲ـ۵ـ آزمون معادلهٔ خط رگرسیون

به یاد بیاورید که هدف از یافتن معادلهٔ خط رگرسیون، بیان رابطهٔ خطی دو متغیر از طریق یک معادلهٔ ریاضی است که به آن مدل رگرسیونی یا مُدل آماری نیز گفته می‌شود[۱]. از سوی دیگر، منظور از رابطه، همراه‌بودنِ تغییرات یک متغیر با تغییرات متغیر دیگر است. بنابراین، معادلهٔ خط رگرسیون چگونگی تغییراتِ یک متغیر را بر اساس تغییرات متغیر دیگر بیان می‌کند. به عبارت

۱. به طور معمول، مدل آماری (Statistical Model)، بیانی ریاضی از رابطهٔ دو یا چند متغیر در قالب یک معادله است. از همین رو، به معادلهٔ خط رگرسیون نیز مدل آماری یا مدل رگرسیونی می‌گویند.

دیگر، مدل رگرسیون توضیح می‌دهد که چگونه تغییرات متغیر وابسته به تغییرات متغیر مستقلِ معادله باز می‌گردد. برای مثال اگر معادلهٔ رگرسیونِ $Y = ۹۲/۷۹X - ۱۰۲۲/۳$ بین متغیر معدل (X) به عنوان متغیر مستقل و متغیر درآمد (Y) به عنوان متغیر وابسته به شکلی «مناسب» برقرار باشد، می‌توان ادعا کرد که بخش قابل توجهی از تفاوت (تغییرات) بین درآمدهای افراد به تفاوت (تغییرات) بین معدل‌های آنان باز می‌گردد. چگونگی این رابطه نیز به کمک ضریب متغیر مستقل X یعنی $b = ۹۲/۷۹$ قابل بیان است (۱ نمره تغییر در مقدار معدل، با ۹۲/۷۹ هزارتومان تغییر در مقدار درآمد همراه است). از همین رو گفته می‌شود متغیر X بخشی از تغییرات متغیر Y را از طریق مدل رگرسیون توضیح می‌دهد. بنابراین، معادلهٔ خطی (مدلی) را باید مناسب ارزیابی کرد که تغییرات قابل توجهی از متغیر وابسته را به کمک تغییرات متغیر مستقل معادله توضیح می‌دهد.

ارزیابی مناسب‌بودن معادلهٔ خط رگرسیون از طریق آزمونی صورت می‌گیرد که منطقی مانند آزمون تجزیهٔ واریانس (آزمون F) دارد. در این آزمون، تغییرات مقادیر مشاهده‌شده از متغیر وابسته اندازه‌گیری می‌شود و سپس به دو بخش تجزیه می‌گردد: تغییراتی که با متغیر مستقل از طریق معادله توضیح داده می‌شود و تغییراتی که بدون توضیح باقی می‌ماند. اگر تغییرات توضیح‌داده شده در مقایسه با تغییرات باقیمانده چشمگیر باشد، نتیجه خواهیم گرفت که معادله به کمک متغیر مستقل به‌خوبی می‌تواند تغییرات متغیر وابسته را به متغیر مستقل نسبت دهد.

برای روشن شدن فرایند ارزیابی مناسب‌بودن معادلهٔ خط رگرسیون از اطلاعات جدول ۱۰ و معادلهٔ $\hat{Y} = ۹۲/۷۹X - ۱۰۲۲/۳$ کمک می‌گیریم. فرد اول در این جدول دارای معدلِ $X = ۱۸$ و درآمد $Y = ۶۲۰$ هزار تومان است. معادلهٔ خط به ازای معدلِ ۱۸، درآمدِ $\hat{Y} = ۶۴۷/۹$ هزارتومان را به دست می‌دهد، زیرا $۶۴۷/۹ = ۹۲/۷۹ \times ۱۸ - ۱۰۲۲/۳$، پس اختلافِ ۶۴۷/۹ با ۶۲۰ یعنی ۲۷/۹- باقیمانده‌ای است که معادلهٔ رگرسیون برای معدل ۱۸ قادر به توضیح آن نیست. جدول ۱۱، مقدار درآمد را که از طریق معادله برای Xهای مختلف جدول ۱۰ به دست می‌آید، همراه با باقیمانده‌های معادلهٔ خط رگرسیون نشان می‌دهد. در این جدول از نماد $\hat{Y}$ (وای هَت) برای مقدار درآمدِ حاصل از معادله استفاده شده است تا از درآمد مشاهده‌شده یعنی Y تمایز داشته باشد. مقدار $\hat{Y}$ که به ازای X خاصی از معادله به دست می‌آید، پیش‌بینی یا برآورد معادلهٔ خط رگرسیون برای Y نیز نامیده می‌شود. بنابراین، $\hat{Y} = ۶۴۷/۹$ پیش‌بینی یا برآورد معادلهٔ خط رگرسیونی $\hat{Y} = ۹۲/۷۹X - ۱۰۲۲/۳$ برای Y به ازای $X = ۱۸$ است.

جدول ۱۱: اطلاعات لازم برای ارزیابی معادلهٔ خط رگرسیون مربوط به جدول ۱۰

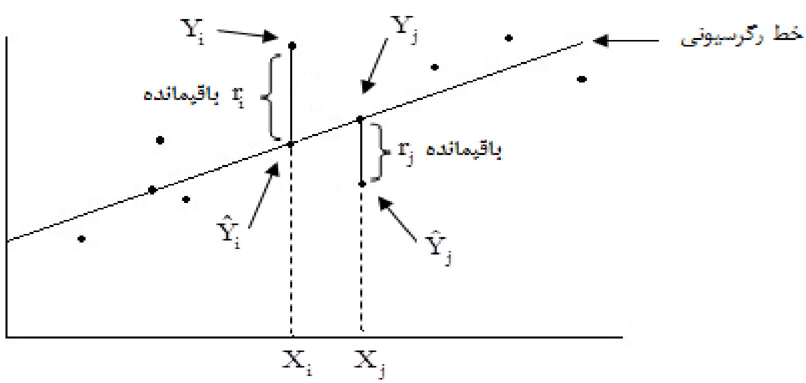
فرد نمونه (i)	معدل (X)	درآمد (Y)	درآمد حاصل از معادله $(\hat{Y})$	باقیمانده $(r = Y - \hat{Y})$
۱	۱۸	۶۲۰	۶۴۷/۹	-۲۷/۹
۲	۱۵	۳۵۰	۳۶۹/۵	-۱۹/۵
۳	۱۹	۸۰۰	۷۴۰/۷	۵۹/۳
۴	۱۴	۳۲۰	۲۷۶/۷	۴۳/۳
۵	۱۷	۵۰۰	۵۵۵/۱	-۵۵/۱
میانگین	۱۶/۶	۵۱۸	۵۱۸	۰
تغییرپذیری (SS)	۱۷/۱۹۹	۳۹۴۲۰	۳۷۰۲۳/۴۸۸	۹۵۸۶/۰۴۷

بر این اساس، هر مقدار مشاهده‌شده از متغیر Y را می‌توان به صورت مجموع مقدار حاصل از رگرسیون و مقدار باقیمانده به شکل زیر در نظر گرفت:

$$Y_i = \hat{Y}_i + r_i = \hat{Y}_i + (Y_i - \hat{Y}_i)$$

توجه کنید که اگر معادله کاملاً مناسب باشد، هیچ خطایی رخ نمی‌دهد، یعنی $r_i = ۰$ یا $Y_i = \hat{Y}_i$ است. در چنین حالتی گفته می‌شود معادلهٔ خط رگرسیون با استفاده از متغیر X به طور کامل متغیر Y را توضیح می‌دهد، زیرا این معادله به ازای هر مقدار از متغیر X دقیقاً همان مقداری را به عنوان $\hat{Y}$ ارائه می‌دهد که پیشتر از متغیر Y مشاهده شده است و به این ترتیب تغییرات Y کاملاً به تغییرات X باز می‌گردد.

شکل ۲۰: باقیمانده‌های خط رگرسیون



از سوی دیگر، معادلهٔ ناکامل نمی‌تواند مقداری درست برابر با مقدار مشاهده‌شده به دست دهد، بلکه همواره باقیمانده‌ای برای مقدار حاصل از معادلهٔ رگرسیون وجود دارد که این معادله قادر نیست آن را توضیح دهد. برای مثال، مقدار ۲۷/۹- باقیماندهٔ فرد اول از معادله و مقدار ۵۹/۳ باقیماندهٔ فرد سوم از این معادله است.

بزرگی مقادیر باقیمانده‌ها در مقایسه با مقادیر حاصل از معادلهٔ رگرسیون می‌تواند ناتوانی این معادله را در توضیح تغییرات متغیر وابسته از طریق متغیر مستقل نشان دهد. برای این منظور از تجزیهٔ واریانس به صورت زیر برای تغییرپذیری این سه مؤلفه (مشاهده Y، مقدار حاصل از رگرسیون $\hat{Y}$ و باقیمانده r) استفاده می‌کنیم:

$$SS_Y = SS_{\hat{Y}} + SS_r$$

$$SS_Y = \sum_{i=۱}^{n}(Y_i - \bar{Y})^۲,\ \ SS_{\hat{Y}} = \sum_{i=۱}^{n}(\hat{Y}_i - \bar{Y})^۲,\ \ SS_r = \sum_{i=۱}^{n} r_i^۲$$

این محاسبات برای اطلاعات جدول ۱۱ به شرح زیر است. توجه کنید که همواره میانگین مقادیر متغیر Y با میانگین مقادیر $\hat{Y}$ برابر و میانگین باقیمانده‌ها صفر است.

$$SS_Y = (۶۲۰-۵۱۸)^۲ + (۳۵۰-۵۱۸)^۲ + (۸۰۰-۵۱۸)^۲ + (۳۲۰-۵۱۸)^۲ + (۵۰۰-۵۱۸)^۲ = ۱۵۷۶۸۰$$

$$SS_{\hat{Y}} = (۶۴۷/۹-۵۱۸)^۲ + (۳۶۹/۵-۵۱۸)^۲ + (۷۴۰/۷-۵۱۸)^۲ + (۲۷۶/۷-۵۱۸)^۲ + (۵۵۵/۱-۵۱۸)^۲ = ۱۴۸۰۹۳$$

$$SS_r = (-۲۷/۹)^۲ + (-۱۹/۵)^۲ + (۵۹/۳)^۲ + (۴۳/۳)^۲ + (-۵۵/۱)^۲ = ۹۵۸۶$$

در عمل مانند آزمون تجزیهٔ واریانس به جای استفاده از مجموع مربعات (SSها) از میانگین مربعات (MSها) با تقسیم SS بر df متناظر آن در قالب جدول زیر استفاده می‌شود (k برابر با تعداد ضرایب معادله رگرسیون یعنی ۲ است).

جدول ۱۲: جدول تجزیهٔ واریانس برای ارزیابی مدل رگرسیون

تغییرات	SS	df	MS	آمارهٔ آزمون
رگرسیون	$SS_{\hat{Y}}$	$df_۱ = k-۱$	$MS_{\hat{Y}} = \frac{SS_{\hat{Y}}}{k-۱}$	$F = \frac{MS_{\hat{Y}}}{MS_r}$
باقیمانده	SS_r	$df_۲ = n-k$	$MS_r = \frac{SS_r}{n-k}$	
مشاهدات	SS_Y	$n-۱$		

نقطهٔ بحرانی این آزمون یک دامنه از جدول چندک‌های توزیع F در پیوست به ازای $df_۱ = k-۱$، $df_۲ = n-k$ و α به دست می‌آید. اگر مقدار آمارهٔ آزمون از $F_{df_۱, df_۲, \alpha}$

بزرگ‌تر باشد، فرضیه صفرِ «معادلهٔ خط رگرسیون تغییرات متغیر وابسته را توضیح نمی‌دهد» رد می‌شود و می‌توان نتیجه گرفت مدل رگرسیونی کم و بیش می‌تواند تغییرات متغیر وابسته را از طریق متغیر (یا متغیرهای) مستقل خود توضیح دهد. با توجه به محاسبات جدول ۱۱، جدول تجزیهٔ واریانس برای مدل رگرسیونی برازش‌داده‌شده به صورت زیر خواهد بود:

جدول ۱۳: جدول تجزیهٔ واریانس برای اطلاعات جدول ۱۱

تغییرات	SS	df	MS	آمارهٔ آزمون
رگرسیون	۱۴۸۰۹۳/۹۵	$df_۱ = ۱$	۱۴۸۰۹۳/۹۵	$F = ۴۶/۳۴۷$
باقیمانده	۹۵۸۶/۰۴۷	$df_۲ = ۳$	۳۱۹۵/۳۴۹	
مشاهدات	۱۵۷۶۸۰	۴		

از آنجا که مقدار آماره ۴۶/۳۴۷ از مقدار نقطهٔ بحرانی $F_{۱,۳,۰/۰۵} = ۱۰/۱۳$ بزرگ‌تر است، نتیجه می‌گیریم متغیر معدل از طریق مدل رگرسیونی $\hat{Y} = ۹۲/۷۹X - ۱۰۲۲/۳$ به‌خوبی تغییرات متغیر درآمد را توضیح می‌دهد.

آزمون تجزیهٔ واریانس برای مدل رگرسیونی، ارتباط نزدیکی با ضریب همبستگی دارد، به طوری که می‌توان نشان داد:

$$SS_{\hat{Y}} = (n-۱)S_Y^۲ r^۲,\ SS_r = (n-۱)S_Y^۲(۱-r^۲),\ F = \frac{MS_{\hat{Y}}}{MS_r} = \frac{n-k}{k-۱}\frac{r^۲}{۱-r^۲}$$

همچنین $r^۲$ که ضریب تعیین مدل رگرسیونی نامیده می‌شود دارای تفسیری مانند $\eta^۲$ است؛ نسبتی از تغییرپذیری مشاهدات که از طریق مدل رگرسیونی توضیح داده می‌شود، پس

$$r^۲ = \frac{SS_{\hat{Y}}}{SS_Y}$$

به‌راحتی می‌توان بررسی کرد که $r^۲ = \frac{S_{\hat{Y}}^۲}{S_Y^۲} = b^۲\frac{S_X^۲}{S_Y^۲} = r_{Y\hat{Y}}^۲$. مقدار $r^۲$ بر اساس اطلاعات جدول ۱۳ برابر با $r^۲ = \frac{۱۴۸۰۹۳/۹۵}{۱۵۷۶۸۰} = ۰/۹۳۹$ است، پس ۹۴ درصد از تغییرات متغیر درآمد به متغیر معدل باز می‌گردد. توجه کنید که از مربع ضریب همبستگی متغیر معدل و متغیر درآمد نیز به رقم $۰/۹۶۹^۲ = ۰/۹۳۹$ می‌رسیدیم.

ارتباط بین آزمون F و ضریب همبستگی کاملاً منطقی است؛ زیرا آمارهٔ F مناسب‌بودن مدل رگرسیون را در توضیح تغییرات متغیر وابسته می‌آزماید. از سوی دیگر، اگر خط

برازش‌داده‌شده بر داده‌ها به الگوی نقاط نمودار پراکنش نزدیک باشد و نقاط نمودار فاصلهٔ چندانی از آن نداشته باشند، آنگاه این خط به شکل مناسبی بر داده‌ها برازش‌یافته است و متغیر مستقل به‌خوبی می‌تواند از طریق معادلهٔ خط، تغییرات متغیر وابسته را توضیح دهد. به یاد بیاورید که هر چه ضریب همبستگی دو متغیر به ۱+ یا ۱− نزدیک‌تر باشد، الگوی نقاط نمودار پراکنش نیز به یک خط راست شبیه‌تر است (شکل ۱۹ را ببینید) و به این ترتیب امکان بیشتری وجود دارد که رابطهٔ دو متغیر با معادلهٔ «خط» قابل بیان باشد. بنابراین، نزدیکی r به ۱+ یا ۱− به معنی کارآمدی معادلهٔ خط در به تصویرکشیدن الگوی نمودار پراکنش و در نتیجه توانایی معادلهٔ خط در توضیح تغییرات متغیر وابسته است. در چنین حالتی، مقدار آمارهٔ F نیز بزرگ می‌شود و فرضیه صفرِ «معادلهٔ خط رگرسیون تغییرات متغیر وابسته را توضیح نمی‌دهد» رد می‌گردد.

۳ـ۵ـ آزمون ضرایب خط رگرسیون

در بحث ضریب همبستگی بین آمارهٔ r و پارامتر ρ تمایز قایل شدیم تا نشان دهیم r مقداری است که بر اساس داده‌های نمونه به عنوان همبستگی بین دو متغیر محاسبه می‌شود و ρ مقدار واقعی همبستگی دو متغیر است که محاسبه آن تنها بر اساس داده‌های تمام اعضای جامعه امکان‌پذیر است و به همین دلیل، آزمونی را برای قابل تعمیم‌بودن مقدار r به جامعه معرفی کردیم. در بحث رگرسیون نیز باید بین ضرایب معادلهٔ خط رگرسیون نمونه (آماره) و ضرایب معادلهٔ خط رگرسیون جامعه (پارامتر) تفاوت قایل شویم.

مقادیر ضرایب رگرسیون b و a که از داده‌های متعلق به نمونه‌ای از جامعه به دست می‌آیند، به معادلهٔ خط رگرسیون نمونه می‌انجامند. این معادله برآوردی از معادلهٔ خطی است که رابطهٔ بین دو متغیر را در جامعه نشان می‌دهد. اگر $Y = BX + A$ معادلهٔ خط رگرسیون جامعه باشد، b برآوردی از B و a برآوردی از A است. در واقع، ضرایب B و A پارامترهای نامعلومی هستند که تنها با داشتن داده‌های کل جامعه قابل محاسبه‌اند، پس در عمل با مقادیر آماره‌های b و a برآورد می‌شوند. بنابراین با این پرسش روبه‌رو هستیم که اگر مقادیری برای b و a از داده‌های نمونه به دست آمد و معادلهٔ خطی با استفاده از آنها برای رابطهٔ خطی دو متغیر تعیین شد، آیا مجاز هستیم این خط را به عنوان تقریبی قابل قبول از رابطهٔ خطی دو متغیر در «جامعه» نیز درنظر بگیریم؟ برای پاسخگویی به این پرسش از آزمون‌های آماری مربوط به ضرایب رگرسیون استفاده می‌شود.

نخستین آزمون، فرضیه صفرِ $B=۰$ را در مقابل فرضیه جانشین $B\neq۰$ می‌آزماید. در واقع، فرضیه صفر این آزمونِ دو دامنه بیانگر آن است که بین دو متغیر X و Y رابطهٔ خطی وجود ندارد، زیرا اگر مقدار B در معادلهٔ $Y=BX+A$ برابر با صفر باشد، متغیر X از معادله حذف و ارتباط خطی متغیر Y با X قطع می‌شود. این آزمون با استفاده از آمارهٔ t به صورت زیر به اجرا در می‌آید:

$$t=\frac{b}{SE_b} \qquad SE_b=\sqrt{\frac{MS_r}{SS_X}}=\sqrt{\frac{MS_r}{\sum_{i=۱}^{n}(X_i-\overline{X})^۲}}$$

نقطهٔ بحرانی این آزمون دو دامنه از جدول چندک‌های توزیع t به ازای $df=n-۲$ و $\frac{\alpha}{۲}$ تعیین می‌شود. اگر مقدار آماره بزرگ‌تر از $t_{n-۲,\frac{\alpha}{۲}}$ یا کوچک‌تر از $-t_{n-۲,\frac{\alpha}{۲}}$ باشد، فرضیه صفرِ $B=۰$ رد می‌شود و نتیجه خواهیم گرفت که متغیر X از طریق معادله با متغیر Y رابطهٔ خطی دارد.

آزمونِ دو دامنهٔ مربوط به عرض از مبدأ معادلهٔ خط رگرسیون نیز فرضیه صفرِ $A=۰$ را در مقابل فرضیه جانشین $A\neq۰$ می‌آزماید و از آمارهٔ t زیر برای اجرای این آزمون استفاده می‌کند. نقطهٔ بحرانی این آزمون نیز مانند آزمون ضریبِ B است.

$$t=\frac{a}{SE_a} \qquad SE_a=\sqrt{MS_r\left(\frac{۱}{n}+\frac{\overline{X}^۲}{SS_X}\right)}=\sqrt{MS_r\left(\frac{۱}{n}+\frac{\overline{X}^۲}{\sum_{i=۱}^{n}(X_i-\overline{X})^۲}\right)}$$

نتایج این دو آزمون برای ضرایب معادلهٔ $\hat{Y}=۹۲/۷۹X-۱۰۲۲/۳$ با توجه به اطلاعات جدول‌های ۱۱ و ۱۳ به صورت زیر است:

$$MS_r=۳۱۹۵/۳۴۹,\ SS_X=۱۷/۱۹۹,\ \overline{X}=۱۶/۶$$

$$SE_a=\sqrt{\frac{۳۱۹۵/۳۴۹}{۱۷/۱۹۹}\times\left(\frac{۱}{۵}+\frac{۱۶/۶^۲}{۱۷/۱۹۹}\right)}=۲۲۷/۶۶۵\Rightarrow t=\frac{-۱۰۲۲/۳}{۲۲۷/۶۶۵}=-۴/۴۹$$

$$SE_b=\sqrt{\frac{۳۱۹۵/۳۴۹}{۱۷/۱۹۹}}=۱۳/۶۳\Rightarrow t=\frac{۹۲/۷۹}{۱۳/۶۳}=۶/۸۰۸$$

از آنجا که مقدار احتمال برای آمارهٔ ۶/۸۰۸ برابر با ۰/۰۰۶ و از $\frac{\alpha}{۲}=۰/۰۲۵$ کوچک‌تر است، فرضیه صفر $B=۰$ رد می‌شود، پس می‌توان گفت متغیر درآمد (Y) با متغیر معدل (X) از طریق مدل فوق رابطهٔ خطی دارد. همچنین، مقدار احتمال آمارهٔ ۴/۴۹- برابر با ۰/۰۲۱ و از $\frac{\alpha}{۲}=۰/۰۲۵$ کوچک‌تر است، پس فرضیه صفر $A=۰$ نیز رد می‌شود؛ یعنی

خط رگرسیون از مبدأ مختصات رد نمی‌شود (این بدان معناست که اگر معدل فردی صفر باشد، یعنی $X=۰$، بدون درآمد نخواهد بود).

۴ـ۵ـ پیش‌فرض‌های مدل رگرسیونی

پیش از بحث درباره پیش‌فرض‌های مدل رگرسیونی باید به این نکته توجه کرد که در مثال جدول ۱۱ به ازای هر مقدار از معدل تنها یک درآمد ثبت شده بود، ولی در عمل افراد مختلفی با مثلاً معدل دانش‌آموختگی ۱۸ مانند جدول ۱۴، وجود دارند که لزوماً درآمد آنان یکسان نیست. بنابراین، ممکن است فردی با معدل ۱۸ دارای درآمد ۶۰۰ هزار تومان و فرد دیگری با همان معدل دارای درآمد ۶۹۰ هزار تومان و... باشد. به عبارت دیگر، به ازای معدل ۱۸ مجموعه‌ای از درآمدها برای افراد مختلف به چشم می‌خورد که دارای توزیعی با میانگین و واریانس خاص خود است. به همین ترتیب، برای معدل ۱۷ نیز مجموعه‌ای از درآمدهای متفاوت قابل تصور است و به ازای معدل‌های ۱۹، ۱۴ و... نیز شرایط مشابهی وجود دارد.

جدول ۱۴: تکرار به ازای مقادیر مختلف متغیر مستقل (معدل)

معدل (X)	درآمدها (Y)	تعداد	میانگین ($\bar{Y}$)	واریانس (S^2)
۱۸	۶۰۰، ۶۲۰، ۶۵۰، ۵۸۰، ۶۹۰	۵	۶۲۸	۱۸۷۰
۱۵	۳۳۰، ۳۴۵، ۳۵۰، ۴۲۰، ۵۰۰	۵	۳۸۹	۵۰۵۵
۱۹	۷۳۰، ۸۰۰، ۸۷۰	۳	۸۰۰	۴۹۰۰
۱۴	۲۵۰، ۲۹۰، ۳۲۰، ۳۷۰	۴	۳۰۷/۵	۲۵۵۸
۱۷	۴۳۰، ۵۰۰، ۵۵۰	۳	۴۹۳/۳	۳۶۳۳

بر این اساس، مقادیر متغیر وابسته برای هر مقدار از متغیر مستقل دارای توزیعی با میانگین و واریانس خاصی است و خط رگرسیون مانند شکل ۲۱ از میانگین این توزیع‌ها می‌گذرد؛ زیرا میانگین، نمایندهٔ عمومی مقادیر متغیر وابسته به ازای هر مقدار از متغیر مستقل است. پس وقتی در معادلهٔ $\hat{Y}=۹۲/۷۹X-۱۰۲۲/۳$ به جای معدل، مقدار $X=۱۸$ را قرار می‌دهید و به ازای آن مقدار $\hat{Y}=۶۴۷$ هزار تومان را به دست می‌آورید، عدد ۶۴۷ می‌تواند چنین تفسیر شود: متوسط درآمد افرادی با معدل ۱۸، ۶۴۷ هزار تومان پیش‌بینی می‌شود[1].

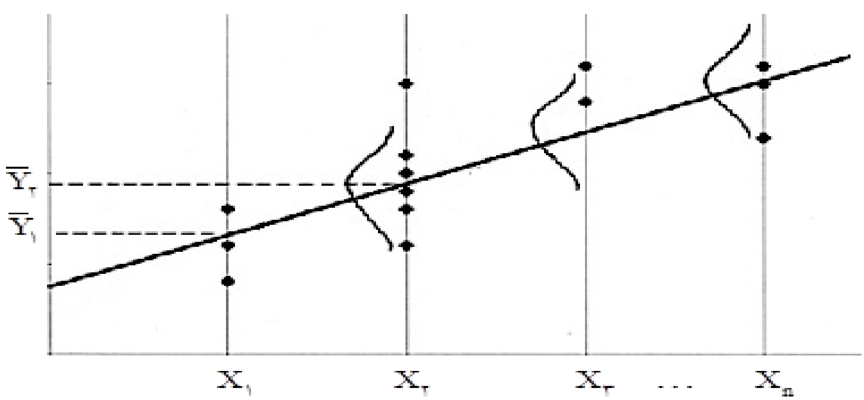
۱. پیشتر عدد ۶۴۷ هزارتومان چنین تفسیر می‌شد: درآمد «فردی» با معدل ۱۸ از طریق مدل رگرسیون، ۶۴۷ هزارتومان پیش‌بینی می‌شود. تفسیر ۶۴۷ هزارتومان به متوسط درآمد افراد یا درآمد فرد، هر دو صحیح است، هر چند پیش‌بینی یکسانی برای هر دو مورد به کار رفته است. با وجود این، استفاده از این پیش‌بینی برای میانگین درآمد افراد از خطای کمتری در مقایسه با درآمد فرد برخوردار است. در اینجا به شیوهٔ محاسبه این خطا و تفاوت آنهَا اشاره نمی‌شود.

استقلال مشاهدات از هم: مشاهدات یا همان Yها باید از یکدیگر مستقل باشند، به این معنا که مقدار آنها از یکدیگر تأثیر نپذیرند. برای مثال اگر داده‌های مربوط به دو متغیر درآمد و معدل در جدول ۱۱ این گونه به دست آمده بود: فردی با معدل ۱۷ دیپلم می‌گیرد و کاری با درآمد ۵۰۰ هزارتومان به دست می‌آورد. او ادامه تحصیل می‌دهد و با معدل ۱۸ فوق دیپلم می‌گیرد و شغلی با درآمد ۶۲۰ هزار تومان به دست می‌آورد و به همین ترتیب، لیسانس، فوق لیسانس و دکترا می‌گیرد و شغل‌هایی با درآمدهای متفاوت کسب می‌کند. داده‌هایی که در اینجا به دست آمده‌اند، داده‌هایی متعلق به یک فرد هستند، نه افراد متفاوتِ مستقل از هم، از همین رو در چنین حالتی نمی‌توان از مدل‌های رگرسیونی برای بیان رابطهٔ معدل و درآمد استفاده کرد.

برابری واریانس‌های متغیر وابسته به ازای مقادیر مختلف متغیر مستقل: روش حداقل مربعات معمولی فرض می‌کند که واریانس‌های توزیع‌های متغیر وابسته به ازای مقادیر مختلف متغیر مستقل با هم برابرند (پیش‌فرضی مانند تجزیهٔ واریانس). اگر این پیش‌فرض برقرار نباشد، باید از روش حداقل مربعات موزون برای برآورد ضرایب معادلهٔ خط رگرسیون استفاده کرد.

نرمال‌بودن توزیع متغیر پاسخ: برازش مدل رگرسیون بر نقاط نمودار پراکنش یعنی یافتن ضرایب معادلهٔ خط با روش حداقل مربعات معمولی (یا موزون)، تا زمانی که آزمونی درباره ضرایب این معادله یا مناسب‌بودن آن اجرا نشود، به هیچ پیش‌فرضی درباره نرمال‌بودن توزیع متغیر وابسته نیاز ندارد. اگر به دنبال بررسی مناسب‌بودن مدل با آزمون F یا بررسی ضرایب مدل با آزمون‌های t هستید، ضروری است توزیع متغیر وابسته به ازای هر مقدار از متغیر مستقل توزیع نرمال باشد (شکل ۲۱ را ببینید).

شکل ۲۱: خط رگرسیونی گذرنده از میانگین توزیع متغیر وابسته به ازای مقادیر مختلف متغیر مستقل



۵ـ۵ـ مدل رگرسیون با دو متغیر مستقل

تاکنون تنها از یک متغیر مستقل در معادلهٔ رگرسیون برای توضیح تغییرات متغیر وابسته استفاده می‌شد، از این رو، آن را مدل رگرسیون ساده می‌نامند. در اینجا به معرفی مدلی می‌پردازیم که از دو متغیر مستقل X و Z برای توضیح تغییرات متغیر وابستهٔ Y استفاده می‌کند، بنابراین به آن مدل رگرسیون چندگانه گفته می‌شود. به این ترتیب این امکان وجود دارد که تغییرات متغیر وابسته را با کمک این دو متغیر، بیشتر از حالت تک متغیر توضیح دهیم.

شکل مدل رگرسیون با دو متغیر مستقل به صورت $\hat{Y} = cZ + bX + a$ است که مانند قبل، a، b و c ضرایب مدل رگرسیون نامیده می‌شوند. برازش مدل رگرسیون با دو متغیر مستقل با رسم نمودار پراکنش آغاز می‌شود؛ پژوهشگر باید نمودار پراکنش متغیر وابسته (Y) را با تک‌تک متغیرهای مستقل (X و Z) بررسی کند. اگر الگوی این نمودارهای پراکنش نمایانگر رابطه‌ای خطی بود، او به دنبال برازش مدل $\hat{Y} = cZ + bX + a$ می‌رود. محاسبه ضریب همبستگی پیرسون متغیر وابسته با متغیرهای مستقل یعنی r_{XY} و r_{ZY} نیز به تصمیم‌گیری او کمک می‌کند.

در گام بعد، برازش مدل با محاسبه ضرایب آن از طریق روش حداقل مربعات معمولی از رابطه‌های زیر صورت می‌گیرد:

$$c = \frac{S_Y}{S_Z}\frac{r_{ZY} - r_{XZ}r_{XY}}{۱ - r_{XZ}^۲},\ b = \frac{S_Y}{S_X}\frac{r_{XY} - r_{XZ}r_{ZY}}{۱ - r_{XZ}^۲},\ a = \overline{Y} - b\overline{X} - c\overline{Z}$$

نمادهای S_X، S_Z و S_Y به ترتیب انحراف معیار متغیرهای X، Z و Y را نشان می‌دهند. نماد r نیز مانند قبل بیانگر ضریب همبستگی پیرسون دو متغیر است. توجه کنید که به دلیل وجود سه متغیر، سه ضریب همبستگی نیز می‌توان بین دوبه‌دوی آنها محاسبه کرد.

مانند مدل رگرسیون ساده، مدل رگرسیون $\hat{Y} = cZ + bX + a$ با دو متغیر مستقل نیز تقریبی از مدل رگرسیون جامعه یعنی $Y = CZ + BX + A$ است؛ زیرا ضرایب a، b و c با استفاده از داده‌های نمونه محاسبه شده‌اند. بنابراین، باید مناسب‌بودن مدل با آزمون F (تجزیهٔ واریانس) و ضرایب آن با آزمون‌های t بررسی شوند. فرایند اجرای آزمون F مانند جدول ۱۲ است. این آزمون به طور همزمان فرضیه صفر $C = B = ۰$، یعنی وجود نداشتن هیچ رابطهٔ خطی بین متغیر Y با دو متغیر X و Z را از طریق مدل رگرسیونی $Y = CZ + BX + A$ می‌آزماید. پس با رد این فرضیه نتیجه خواهیم گرفت مدل $\hat{Y} = cZ + bX + a$ می‌تواند «بخشی» از تغییرات متغیر وابستهٔ Y را به کمک متغیرهای مستقل خود توضیح دهد.

معناداری آزمون F یعنی رد فرضیه صفر به این معنا نیست که تک‌تک متغیرهای مستقل در توضیح تغییرات متغیر وابسته نقشی جدی دارند. بنابراین، باید برای هر یک از ضرایب مدل نیز آزمون t جداگانه‌ای اجرا شود. آمارهٔ آزمون فرضیه صفر $C=۰$ در مقابل فرضیه جانشین $C\neq ۰$ عبارت است از:

$$t=\frac{c}{SE_c} \qquad SE_c=\sqrt{\frac{MS_r}{(۱-r_{XZ}^{۲})SS_Z}}=\sqrt{\frac{MS_r}{(۱-r_{XZ}^{۲})\sum_{i=۱}^{n}(Z_i-\bar{Z})^{۲}}}$$

آمارهٔ آزمون فرضیه صفر $B=۰$ در مقابل فرضیه جانشین $B\neq ۰$ نیز عبارت است از:

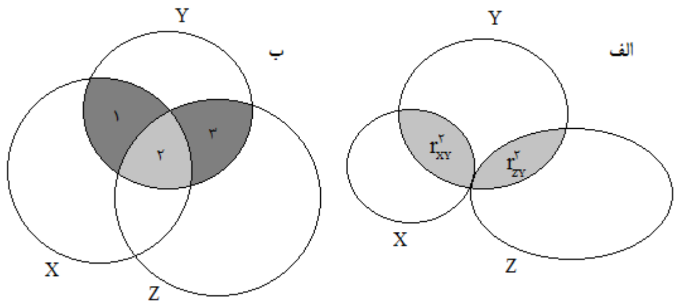
$$t=\frac{b}{SE_b} \qquad SE_b=\sqrt{\frac{MS_r}{(۱-r_{XZ}^{۲})SS_X}}=\sqrt{\frac{MS_r}{(۱-r_{XZ}^{۲})\sum_{i=۱}^{n}(X_i-\bar{X})^{۲}}}$$

آمارهٔ آزمون فرضیه صفر $A=۰$ در مقابل فرضیه جانشین $A\neq ۰$ نیز به صورت $t=\frac{a}{\sqrt{SE_a}}$ است که به دلیل پیچیدگی، شیوهٔ محاسبه SE_a آورده نمی‌شود. گفتنی است هر سه آزمون، دو دامنه هستند و از $t_{n-۳,\frac{\alpha}{۲}}$ به عنوان نقطهٔ بحرانی استفاده می‌کنند.

در مدلی با دو متغیر مستقل نیز از ضریب تعیین $r^{۲}=r_{Y\hat{Y}}^{۲}=\frac{SS_{\hat{Y}}}{SS_Y}$ برای پاسخگویی به این پرسش استفاده می‌شود که متغیرهای مستقل مدل «با هم» چه نسبتی از تغییرات متغیر وابسته را توضیح می‌دهند؟ علاوه بر این، دو پرسش دیگر نیز درباره متغیرهای مستقل قابل طرح است: هر یک از متغیرهای مستقل «به تنهایی» چه نسبتی از تغییرات متغیر وابسته را توضیح می‌دهند؟ پاسخ به این پرسش در شرایطی که دو متغیر مستقل X و Z دارای همبستگی نباشند، یعنی $r_{XZ}=۰$ یا این همبستگی ضعیف و قابل چشم‌پوشی باشد، از $r^{۲}=r_{XY}^{۲}+r_{ZY}^{۲}$ به سادگی به دست می‌آید. بر این اساس، سهم هر متغیر مستقل در توضیح متغیر وابسته، برابر با مربع همبستگی آن متغیر با متغیر وابسته است.

اگر دو متغیر مستقل دارای همبستگی باشند، یعنی $r_{XZ}\neq ۰$، یافتن سهم هر متغیر در پیش‌بینی تغییرات متغیر وابسته به سازوکار خاصی نیاز دارد. برای روشن شدن تفاوت این دو وضعیت به شکل ۲۲ توجه کنید. هر دایره در این شکل نشان‌دهنده تغییرات یک متغیر است. در حالت «الف»، بین متغیرهای مستقل X و Z همبستگی وجود ندارد؛ زیرا دایره‌های مربوط به این دو متغیر دارای هیچ اشتراکی نیستند. قسمت‌های خاکستری نیز سهمی را نشان می‌دهند که هر متغیر از تغییرات متغیر وابسته توضیح می‌دهد. پس جمع این دو برابر با سهمی است که هر دو متغیر با هم توضیح می‌دهند.

شکل ۲۲: تفاوت وضعیت متغیرهای مستقلِ همبسته و ناهمبسته در توضیح تغییرات متغیر وابسته



دو متغیر X و Z در حالت «ب» دارای همبستگی هستند، زیرا دایره‌های آنها در قسمت خاکستریِ کمرنگ‌تر با یکدیگر اشتراک دارند. مانند حالت «الف»، سهمی از تغییرات متغیر وابستهٔ Y که از طریق متغیرهای مستقل X و Z توضیح داده می‌شود، برابر با اشتراک دایره‌های این دو متغیر با دایرهٔ متغیر Y است؛ یعنی ناحیه‌های ۱، ۲ و ۳. برخلاف حالت «الف» که این ناحیه به‌راحتی به ناحیهٔ مربوط به X و ناحیهٔ مربوط به Z تفکیک می‌شد، در حالت «ب» نمی‌توان این ناحیه را به دو قسمت مجزا تفکیک کرد، زیرا سهم متغیر X شامل ناحیه‌های ۱ و ۲ و سهم متغیر Z شامل ناحیه‌های ۲ و ۳ است؛ یعنی ناحیهٔ ۲ بین هر دو متغیر مستقل مشترک است. اکنون چگونه می‌توان تعیین کرد سهم اختصاصی هر متغیر دقیقاً کدام است؟ اگر ناحیهٔ ۲ را متعلق به هر دو متغیر بدانیم، در این صورت این ناحیه دوبار به حساب آمده است. از سوی دیگر، ناحیهٔ ۲، سهم «درهم‌تنیدهٔ» هر دو متغیر است که نمی‌توان آن را به دو تکه تقسیم کرد. این چالش باعث مطرح شدن معیار همبستگی دیگری به نام ضریب همبستگی جزئی می‌شود که در این کتاب بررسی نخواهد شد.

اکنون بحث مدل رگرسیون را در قالب مثالی مرور می‌کنیم. فرض کنید معلمی می‌خواهد تغییرات نمرات دانش‌آموزان خود را در انشانویسی بررسی کند. او یکی از متغیرهای مهم را نمرهٔ دانش‌آموز در درس دستور زبان فارسی می‌داند. از آنجا که واژه‌های عربی نیز در زبان فارسی زیاد استفاده می‌شوند، او انتظار دارد نمرهٔ درس عربی نیز در توضیح تغییرات نمرهٔ درس انشا مؤثر باشد. بنابراین به دنبال برازش مدلی است که در آن نمرهٔ دستور زبان فارسی (متغیر X) و نمرهٔ عربی (متغیر Z) متغیرهای مستقل و نمرهٔ انشا

(متغیر Y) متغیر وابستهٔ آن است. این معلم داده‌های مربوط به این سه متغیر را مطابق جدول ۱۵ از ۱۰ دانش‌آموز گردآوری می‌کند.

جدول ۱۵: اطلاعات سه متغیر نمرهٔ عربی، نمرهٔ دستورزبان و نمرهٔ انشانویسی

دانش‌آموز (i)	عربی (Z)	دستورزبان (X)	انشانویسی (Y)	پیش‌بینی ($\hat{Y}$)	باقیمانده (r)
۱	۱۸	۱۶	۱۶	۱۶/۷۶	−۰/۷۶
۲	۱۵	۲۰	۱۸	۱۸/۲۸	−۰/۲۸
۳	۱۷	۱۶	۱۵	۱۶/۱۸	−۱/۱۸
۴	۱۲	۱۴	۱۱	۱۱/۶۶	−۰/۶۶
۵	۲۰	۱۸	۱۹	۱۹/۵۴	−۰/۱۴
۶	۱۹	۱۸	۱۹	۱۸/۹۷	−۰/۱۷
۷	۱۴	۱۳	۱۲	۱۲/۰۱	−/۰۱
۸	۱۸	۱۹	۲۰	۱۹/۲	۰/۸
۹	۱۶	۱۵	۱۷	۱۴/۷۹	۲/۲۱
۱۰	۱۲	۱۷	۱۴	۱۴/۱۱	−۰/۱۹
میانگین	۱۶/۱	۱۶/۶	۱۶/۱۵	۱۶/۱۵	۰
انحراف معیار	۲/۸۰۷	۲/۲۲۱	۳/۰۷۴		
تغییرپذیری (SS)	۷۰/۹۱	۴۴/۴	۸۵/۰۶۵	۷۶/۹۶۳	۸/۱۰۲

ضرایب همبستگی دوبه‌دویِ متغیرها به صورت زیر است:

$$r_{XZ} = ۰/۴۵۳,\quad r_{YX} = ۰/۸۲۷,\quad r_{YZ} = ۰/۷۹۴$$

پس ضرایب رگرسیونی به صورت زیر محاسبه می‌شوند:

$$c = \frac{۳/۰۷۴}{۲/۸۰۷} \times \frac{۰/۷۹۴ - ۰/۴۵۳ \times ۰/۸۲۷}{۱ - ۰/۴۵۳^{۲}} = ۰/۵۷۸$$

$$b = \frac{۳/۰۷۴}{۲/۲۲۱} \times \frac{۰/۸۲۷ - ۰/۴۵۳ \times ۰/۷۹۴}{۱ - ۰/۴۵۳^{۲}} = ۰/۸۱۳$$

$$a = ۱۶/۱۵ - ۰/۸۱۳ \times ۱۶/۶ - ۰/۵۷۸ \times ۱۶/۱ = -۶/۶۶۱$$

پس مدل رگرسیونی برازش‌داده‌شده به صورت $\hat{Y} = ۰/۵۷۸\,Z + ۰/۸۱۳\,X - ۶/۶۶۱$ است. دو ستون آخر جدول ۱۵، شامل پیش‌بینی‌ها و باقیمانده‌های حاصل از این مدل است. با این اطلاعات می‌توان فرضیه صفر $C = B = ۰$ را با آزمون تجزیهٔ واریانس از طریق جدول ۱۶ بررسی کرد (توجه کنید که $k = ۳$ است).

جدول ۱۶: جدول تجزیهٔ واریانس برای اطلاعات جدول ۱۵

تغییرات	SS	df	MS	آمارهٔ آزمون
رگرسیون	۷۶/۹۶۳	$df_1=۲$	۳۸/۴۸۲	$F=۳۳/۲۴۷$
باقیمانده	۸/۱۰۲	$df_2=۷$	۱/۱۵۷	
مشاهدات	۸۵/۰۶۵	۹		

نقطهٔ بحرانی آزمون از جدول چندک‌های توزیع F در پیوست، $F_{۲,۷,۰/۰۵}=۴/۷۴$ به دست می‌آید و بر رد فرضیه صفر دلالت می‌کند (مقدار احتمال آزمون نیز ۰/۰۰۰۳ از $\alpha=۰/۰۵$ کوچک‌تر است). بنابراین، نتیجه خواهیم گرفت حداقل یکی از دو ضریب B یا C صفر نیست.

برای بررسی این امر باید آزمون ضرایب رگرسیون را صورت دهیم. آزمون مربوط به فرضیه صفر $C=۰$ در مقابل فرضیه جانشین $C\neq۰$ با استفاده از اطلاعات جدول‌های ۱۵ و ۱۶ عبارت است از:

$$SE_c=\sqrt{\frac{۱/۱۵۷}{(۱-۰/۴۵۳^۲)\times۷۰/۹۱}}=۰/۱۴۳\Rightarrow t=\frac{۰/۵۷۸}{۰/۱۴۳}=۴/۰۳۵$$

مقدار احتمال این آزمون دودامنه ۰/۰۰۵ از $\frac{\alpha}{۲}=۰/۰۲۵$ کوچک‌تر است، پس ضریب نمرهٔ عربی در مدل صفر نیست و این متغیر از مدل حذف نمی‌شود. آزمون مربوط به فرضیه صفر $B=۰$ در مقابل فرضیه جانشین $B\neq۰$ عبارت است از:

$$SE_b=\sqrt{\frac{۱/۱۵۷}{(۱-۰/۴۵۳^۲)\times۴۴/۴}}=۰/۱۸۱\Rightarrow t=\frac{۰/۸۱۳}{۰/۱۸۱}=۴/۴۹۲$$

مقدار احتمال این آزمونِ دو دامنه ۰/۰۰۳ از $\frac{\alpha}{۲}=۰/۰۲۵$ کوچک‌تر است و فرضیه صفر این آزمون نیز رد می‌شود، پس ضریب نمرهٔ دستورزبان در مدل صفر نیست و این متغیر نیز در مدل باقی می‌ماند. بنابراین هر دو متغیر نمرهٔ عربی و نمرهٔ دستورزبان از طریق مدل $\hat{Y}=۰/۵۷۸Z+۰/۸۱۳X-۶/۶۶۱$ با متغیر نمرهٔ انشانویسی مرتبط هستند. حال به سراغ آزمون فرضیه صفر $A=۰$ در مقابل فرضیه جانشین $A\neq۰$ می‌رویم (به دلیل طولانی‌بودن رابطه، شیوه محاسبه خطای معیار a، ارائه نمی‌شود).

$$SE_a=۲/۸۶۳\Rightarrow t=\frac{a}{SE_a}=\frac{-۶/۶۶۱}{۲/۸۶۳}=-۲/۳۲۷$$

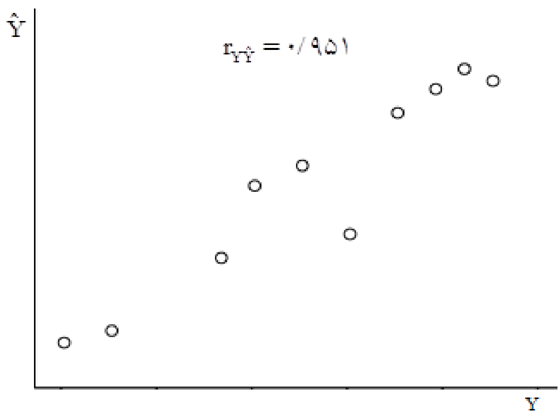
مقدار احتمال این آزمون ۰/۰۵۳ از $\frac{\alpha}{۲}=۰/۰۲۵$ بزرگ‌تر است، پس نمی‌توان فرضیه صفر $A=۰$ را رد کرد. این فرضیه صفر مدعی است که مدل باید از مبدأ مختصات بگذرد؛ یعنی اگر نمرهٔ دانش‌آموزی در درس عربی صفر و در درس دستورزبان نیز صفر باشد، آنگاه نمرهٔ او در درس انشانویسی نیز باید صفر باشد (این نتیجه‌ای منطقی است، زیرا اگر فردی در دو درس مرتبط با انشانویسی نمره‌ای کسب نکرده است انتظار نمی‌رود نمره‌ای در درس انشانویسی به دست آورد).

علامت ضرایب در مدلی با دو متغیر مستقل نیز نشان‌دهنده جهت رابطهٔ متغیرهای مستقل با متغیر وابسته است. بنابراین، هر دو متغیر X و Z با متغیر Y رابطهٔ مثبت (افزایشی) دارند، زیرا هر دوی آنها در مدل $\hat{Y}=۰/۵۷۸\,Z+۰/۸۱۳\,X-۶/۶۶۱$ دارای ضرایب مثبتی هستند. ضریب تعیین این مدل که با توجه به اطلاعات جدول ۱۶ به صورت زیر محاسبه می‌شود، بیانگر آن است که دو متغیر مستقل X و Z از طریق این مدل قادرند ۹۰ درصد از تغییرات متغیر وابستهٔ Y را توضیح دهند.

$$r^{۲}=\frac{SS_{\hat{Y}}}{SS_{Y}}=\frac{۷۶/۹۶۳}{۸۵/۰۶۵}=۰/۹۰۵$$

این بدان معناست که رابطهٔ خطی نزدیکی بین مشاهدات (Yها) و پیش‌بینی‌های مدل ($\hat{Y}$ها) وجود دارد، پس مدل به شکل مناسبی قادر است رابطهٔ متغیر وابسته را با متغیرهای مستقل بیان کند. کارآمدی مدل از طریق نمودار پراکنش $\hat{Y}$ دربرابر Y در شکل ۲۳ به‌خوبی دیده می‌شود (به یاد بیاورید که ضریب تعیین از مربع همبستگی $\hat{Y}$ و Y نیز قابل محاسبه است).

شکل ۲۳: نمودار پراکنش $\hat{Y}$ در برابر Y (داده‌های جدول ۱۵)



اصطلاحات فصل پنجم

Alternative Hypothesis	فرضیه جانشین (مقابل، یک)	Ordinary Least Squares	حداقل مربعات معمولی
Analysis of Variance	تجزیه واریانس	Pairwise Comparisons	مقایسه‌های دوبه‌دو
Assumption	پیش‌فرض	Partial Correlation Coefficient	ضریب همبستگی جزئی
Biserial Correlation	همبستگی دو رشته‌ای	Pattern	الگو
Coefficient of Determination	ضریب تعیین	Coefficient of Correlation	ضریب همبستگی
Correlation	همبستگی	Point biserial Correlation	همبستگی دورشته‌ای نقطه‌ای
Critical Value	مقدار بحرانی	Pooled variance	واریانس آمیخته
Decreasing (descending)	کاهشی، نزولی	Post-hoc Tests	آزمون‌های تعقیبی
Dependency	وابستگی	Predict or Estimate	پیش‌بینی یا برآورد
Dependent Variable	متغیر وابسته	Profile Analysis	تحلیل نیمرخی
Equation	معادله	P-value	مقدار احتمال
Eta-square	اتا دو (مربع اتا)	Regression Coefficients	ضرایب رگرسیونی
Homogeneity of Variances	همسانی (همگنی) واریانس‌ها	Regression Line	خط رگرسیون
Honestly Significant Difference	اختلاف معنادار واقعی	Regression Model Fit	برازش مدل رگرسیون
Increasing (ascending)	افزایشی، صعودی	Relationship	رابطه
Independent	مستقل	Residual	باقیمانده
Independent Variable	متغیر مستقل	Response Variable	متغیر پاسخ
Intercept	عرض از مبدأ	Satterthwaite Approximation	تقریب سترزویت
Least Significant Difference	کمترین اختلاف معنادار	Scatter Plot	نمودار پراکنش
Linear and Nonlinear	خطی و ناخطی	Significance Level	سطح معناداری
Monotonic and Nonmonotonic	یکنوا و نایکنوا	Slope	شیب (خط)
Multiple Comparisons	مقایسه‌های چندگانه	Test Statistic	آماره آزمون
Multiple Regression Model	مدل رگرسیونی چندگانه	Two-tailed Test	آزمون دو دامنه
Null Hypothesis	فرضیه صفر	Type I Error	خطای نوع اول
Omega-square	امگا دو (مربع امگا)	Type II Error	خطای نوع دوم
One-tailed Test	آزمون یک دامنه	Weighted Least Squares	حداقل مربعات وزنی

تمرین‌های فصل پنجم

۱. چه فرضیه‌ای می‌تواند جانشین هر یک از فرضیه‌های زیر باشد؟

آ) متوسط درآمد مردم تهران حداقل برابر با ۳۰۰۰۰۰ تومان است.

ب) درصد بیننده تلویزیون در شهرهای اصفهان، شیراز و مشهد برابر است.

پ) با افزایش هوش بر میزان اعتماد به نفس افراد افزوده می‌شود.

ت) بین رضایت از تلویزیون و میزان تماشای آن رابطه مستقیمی وجود دارد.

ث) رضایت از زندگی افراد جوان برابر با افراد میانسال است.

ج) توزیع امید به آینده چوله به چپ است (چولگی منفی دارد).

چ) در بین تهرانی‌ها، میانه مدت زمان تماشای تلویزیون فصل گذشته کمتر از میانه این فصل است.

ح) رابطه سن و مدت زمان تماشای تلویزیون در میان مردان ضعیف‌تر از زنان است.

۲. کدام‌یک از فرضیه‌های تمرین ۱ را می‌توان به عنوان فرضیه صفر انتخاب کرد؟

۳. کدام‌یک از فرضیه‌های تمرین ۱ منجر به آزمون یک دامنه و کدام منجر به آزمون دو دامنه می‌شود؟

۴. خطای نوع اول خطرناک‌تر است یا خطای نوع دوم؟

۵. فرض کنید بر اساس نمونه‌های جدول ۲ از فصل سوم در پی آزمون فرضیهٔ صفر μ=۱۰ هستید.

آ) اگر اختلاف میانگین نمونه از ۱۰، بیش از ۱ واحد باشد، این فرضیهٔ صفر را رد می‌کنید. احتمال وقوع خطای نوع اول در این آزمون چقدر است؟

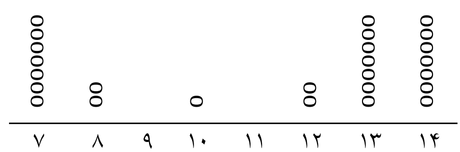
ب) اگر اختلاف میانگین نمونه از ۱۰، بیش از ۱/۵ واحد باشد، این فرضیهٔ صفر را رد می‌کنید. احتمال وقوع خطای نوع اول در این آزمون چقدر است؟

۶. دو فرضیه زیر را در نظر بگیرید:

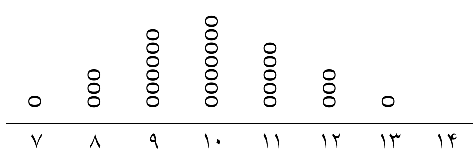
H_0: میانه توزیع متغیر X برابر با ۱۰ است.

H_1: میانه توزیع متغیر X برابر با ۱۰ نیست.

توزیع نمونه‌گیری تحت درستی H_1



توزیع نمونه‌گیری تحت درستی H_0



نمودارهای نقطه‌ای فوق توزیع نمونه‌گیری آماره آزمون را تحت درستی هر یک از فرضیه‌ها نشان می‌دهند. اکنون به پرسش‌های زیر پاسخ دهید:

آ) فرضیه‌بندی فوق، یک دامنه است یا دو دامنه؟

ب) فرض کنید H_0 را هنگامی رد می‌کنیم که مقدار آماره آزمون برابر با ۷، ۱۳ یا ۱۴ باشد. احتمال خطای نوع اول، احتمال خطای نوع دوم و توان آزمون چقدر است؟

۷. کدام تعریف درباره «مقدار احتمال» آماره آزمون درست است؟

آ) احتمال اینکه داده‌ها به نفع فرضیه جانشین باشند، تحت درستی فرضیه صفر.

ب) احتمال درستی نتیجه‌گیری در یک آزمون آماری.

پ) کوچک‌ترین سطح معناداری برای رد فرضیه صفر.

ت) احتمال رد فرضیه صفر وقتی واقعاً غلط است.

۸. پژوهشگری در مطالعه مدت زمان مطالعه ساکنان شهر خاصی، فرضیه صفر «میانگین مطالعه مردم ۴۵ دقیقه است» را مطرح می‌کند، ولی میانگین مطالعه این جامعه واقعاً ۴۵ دقیقه نیست. با وجود این، او نمی‌تواند بر اساس داده‌های گردآوری شده، فرضیه صفر را رد کند. آیا خطایی در آزمون فرضیه او رخ داده است؟ اگر خطایی رخ داده است، چگونه می‌توان از بروز آن خودداری کرد؟

۹. یک نامزد انتخاباتی ادعا می‌کند میزان محبوبیت او در بین رأی‌دهندگان از ۰ تا ۲۰، برابر با ۱۶ است. نمونه‌ای ۳۰۰ نفری به تصادف از میان رأی‌دهندگان انتخاب می‌شود و میزان محبوبیت این نامزد ۱۶/۳ با انحراف معیار ۳ به دست می‌آید. آیا می‌توان ادعای این نامزد انتخاباتی را پذیرفت؟

۱۰. دادگاهی با گزارش کم‌وزنی نان‌های یک نانوایی روبه‌روست، زیرا هر نان باید حداقل ۳۰۰ گرم باشد. دادگاه برای بررسی این پرونده، ۴۰۰ نان پخته‌شده از سوی این نانوایی را به تصادف انتخاب می‌کند. میانگین و انحراف معیار وزن این نان‌ها به ترتیب ۳۰۵ گرم و ۴۰ گرم است.

آ) دادگاه باید چه فرضیه‌های صفر و جانشینی را برای تصمیم‌گیری درباره این پرونده در نظر بگیرد؟

ب) خطاهای نوع اول و دوم را با توجه به فرضیه‌ها بیان کنید.

پ) اگر آزمون در سطح معناداری ۰/۰۱ اجرا شود، آیا باید این نانوایی محکوم شود؟

ت) اگر سطح معناداری ۰/۰۰۱ باشد، نتیجه آزمون چیست؟

۱۱. نمونه‌ای ۲۵ نفری به تصادف از جامعه‌ای انتخاب شده است. میانگین تعداد سال‌های تحصیل این نمونه ۱۲/۵ سال با انحراف معیار ۲/۲ سال است. در سطح ۵ درصد آزمون کنید که آیا

آ) میانگین تعداد سال‌های تحصیل افراد جامعه برابر با ۱۲ سال است؟

ب) میانگین سال‌های تحصیل افراد جامعه حداقل ۱۲ سال است؟

پ) دو آزمون فوق را در سطح معناداری ۰/۰۱ اجرا کنید.

ت) آزمون‌های فوق را با این فرض که تعداد نمونه ۱۰۰ نفر است تکرار کنید.

۱۲. روان‌شناسی معتقد است کسانی که حداقل ۶ ساعت در شبانه روز استراحت می‌کنند در یک آزمون مهارت‌های شناختی موفق‌ترند. او دو نمونه تصادفی به اندازه‌های ۳۵ و ۳۲ نفر به ترتیب از افرادی با حداقل ۶ ساعت استراحت در شبانه روز و کمتر از ۶ ساعت انتخاب و آزمون را در میان آنان اجرا می‌کند. اطلاعات جدول زیر از داده‌های این دو گروه به دست آمده است:

گروه	تعداد	میانگین نمره آزمون	انحراف معیار نمرات
حداقل ۶ ساعت	۳۵	۸	۲
کمتر از ۶ ساعت	۳۲	۶/۵	۲/۵

آ) اعتقاد این روان‌شناس را با فرضیه‌های صفر و جانشین مناسب برای آزمون بیان کنید.

ب) این آزمون را در سطح ۵ درصد، هم با فرض برابری واریانس‌ها و هم با فرض نابرابری اجرا کنید. آیا عقیده این روان‌شناس پذیرفته می‌شود؟

۱۳. کدامیک از حالت‌های زیر به آزمون مقایسه میانگین‌های دو جامعه وابسته مربوط است؟

آ) مقایسه تأثیر دو نوع کود شیمیایی بر رشد یک گونه گیاهی

ب) مقایسه عملکرد دانش‌آموزان در دو درس ریاضی و شیمی

پ) مقایسه شدت علاقه به طبیعت در بین متولدین روستا و شهر

ت) مقایسه مدت زمان مطالعه کتاب‌های الکترونیکی با معمولی (کاغذی) در بین دانشجویان

ث) میزان افزایش اطلاعات آماری خوانندگان این کتاب نسبت به قبل

۱۴. معلم زبان انگلیسی، دانش‌آموزان کلاس ۳۰ نفری خود را به سه دسته تقسیم کرده است: کسانی که در کلاس‌های زبان خارج از مدرسه نیز شرکت می‌کنند (۱۱ نفر)، کسانی

که در چنین کلاس‌هایی شرکت نمی‌کنند، ولی در درس زبان قوی هستند (۷ نفر) و کسانی که در چنین کلاس‌هایی شرکت نمی‌کننـد و در درس زبـان ضـعیف هسـتند (۱۲ نفـر). او نمرات آزمون زبان میان‌ترم این سه گروه را با روش تجزیه واریانس مقایسه می‌کند.

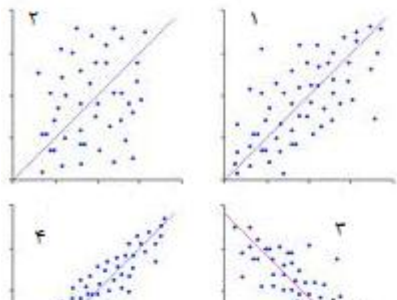
منبع تغییرات	SS	df	MS	F
تفاوت بین سه دسته	۳۰			
خطا	۳۰			
کل	۶۰			

آ) خانه‌های جدول تجزیه واریانس فوق را تکمیل کنید.

ب) فرضیه‌های آزمون این معلم را بنویسید و آزمون را در سطح ۰/۰۵ اجرا و نتیجه خود را بیان کنید.

پ) صرفنظر از نتیجه آزمون، این سه گروه را دو به دو با چهار روش آزمـون‌هـای تعقیبـی مقایسه کنید.

۱۵. چهار نمودار پراکنش زیر را در نظر بگیرید.



آ) هر یک از چهار ضریب همبستگی ۰/۹، ۰/۸، ۰/۸۵– و ۰/۵ را به یکـی از چهـار نمـودار نسبت دهید.

ب) اگر این ضریب‌ها بر اساس داده‌های یک نمونه تصادفی ۵۰ تایی به دست آمده باشـند، کدامیک در سطح ۱ درصد معنادار است؟

۱۶. اگر داده‌ها در عددی ضرب و/یا با عددی جمع شوند، ضریب همبستگی آنهـا چـه تغییری می‌کند؟ این موضوع را در قالب مثالی ساختگی بررسی کنید.

۱۷. مطالعه‌ای نشان داده است بین میزان مصرف بنزین بر حسب گالون و وزن خودرو بر حسب پوند، رابطه رگرسیونی $\hat{Y} = -۰/۰۰۷X + ۴۷$ وجود دارد.

آ) عدد ۰/۰۰۷- بیانگر چیست؟

ب) اگر خودرویی ۲۵۰۰ پوند وزن داشته باشد، پیش‌بینی می‌کنید مصرف سوخت آن چقدر باشد؟

۱۸. رابطه درآمد کارمندان یک شرکت با تعداد سال‌های تحصیل و تعداد سال‌های تجربه کاری آنان از طریق برازش مدل رگرسیون خطی با دو متغیر مستقل بررسی شد. برای این منظور مقدار درآمد، تعداد سال‌های تحصیل و تعداد سال‌های تجربه کاری ۱۰ کارمند به کار رفت که برخی از اطلاعات آنها در جدول زیر آمده است.

متغیر	تعداد	میانگین	واریانس
درآمد (بر حسب ده هزار تومان)	۱۰	۹۲/۷	۶۴۸
تعداد سال‌های تحصیل	۱۰	۱۴/۱	۱۲/۸
تعداد سال‌های تجربه کاری	۱۰	۴/۵	۲/۳

آ) ابتدا جدول تجزیه واریانس مربوط به این مدل را تکمیل و مناسب‌بودن مدل را آزمون کنید.

منبع تغییرات	SS	df	MS	F
رگرسیون	۵۷۸۵/۸۷۷			
باقیمانده				
کل	۵۸۳۴/۱			

ب) مقدار ضریب تعیین مدل چیست؟ آن را چگونه تفسیر می‌کنید؟

پ) همبستگی دو به دوی متغیرها در جدول زیر ارائه شده است. ضرایب مدل را برآورد کنید و مدل را بنویسید. کدام متغیر در پیش‌بینی درآمد بااهمیت‌تر است؟

	درآمد	سال‌های تحصیل	سال‌های تجربه
درآمد	۱	۰/۹۹۶	۰/۸۳۷
سال‌های تحصیل	۰/۹۹۶	۱	۰/۸۳۴
سال‌های تجربه	۰/۸۳۷	۰/۸۳۴	۱

ت) آزمون مربوط به ضریب متغیر سال‌های تحصیل و ضریب متغیر سال‌های تجربه را پس از نوشتن فرضیه‌های آنها، اجرا و نتیجه‌گیری کنید.

ث) فردی که ۸ سال تجربه کاری دارد و دارای مدرک کارشناسی ارشد (۱۸ سال تحصیل) است، چه حقوقی دریافت می‌کند؟

۱۹. چگونه می‌توان با درون‌یابی خطی (به فصل قبل مراجعه کنید)، نقطه بحرانی مربوط به توزیع F را از نزدیک‌ترین نقاط موجود در جدول یافت؟

فصل ششم

۶ آمار استنباطی
آزمون‌های ناپارامتری

تاکنون، آزمون‌هایی معرفی شده‌اند که تحت برقراری پیش‌فرض‌هایی خاص مانند توزیع نرمال یا همسانی واریانس‌ها برای متغیرهای فاصله‌ای یا نسبتی مناسب هستند. در عمل، شرایطی وجود دارد که این پیش‌فرض‌ها برقرار نیستند یا داده‌ها مربوط به متغیرهای اسمی یا ترتیبی هستند. آزمون‌های ناپارامتری می‌توانند به عنوان ابزارهای استنباطی در چنین شرایطی به کار گرفته شوند. این فصل ابتدا به آزمون‌هایی می‌پردازد که معادل ناپارامتری آزمون‌های مربوط به میانگین‌ها هستند؛ نظیر آزمون علامت، ویلکاکسون یا کروسکال-والیس. سپس، دو ضریبِ اسپیرمن و کندال را برای اندازه‌گیری رابطه دو متغیر معرفی می‌کند و در نهایت روش تحلیل جدول‌های توافقی و آزمون مربوط به نسبت‌ها را ارائه می‌دهد.

۱ـ آزمون‌های ناپارامتری (آزاد توزیع)

تاکنون آزمون‌هایی برای استنباط درباره داده‌ها معرفی شده‌اند که در آمار با عنوان آزمون‌های «پارامتری» شناخته می‌شوند. استفاده از آزمون‌های پارامتری تنها با برقراری پیش‌فرض‌هایشان مجاز است. برای مثال، آزمون‌های مربوط به میانگین‌ها، مقایسهٔ واریانس‌ها و رگرسیون به برقراری پیش‌فرض‌هایی نظیر نرمال‌بودن توزیع متغیر تحت بررسی، همسانی واریانس‌ها یا پیوسته‌بودن متغیر نیاز دارند. اما از آنجا که در عمل، پژوهشگران با متغیرهایی سروکار دارند که لزوماً از این پیش‌فرض‌ها پیروی نمی‌کنند، استفاده از آزمون‌های پارامتری با چالش روبه‌رو می‌شود. آزمون‌های ناپارامتری جانشینی برای این آزمون‌ها ارائه می‌دهند، درحالی که برقراری آن پیش‌فرض‌ها ضرورتی ندارد.

ممکن است این پرسش برای خواننده مطرح شود که اگر آزمون‌های ناپارامتری وجود دارند که آزادانه و بدون نیاز به پیش‌فرضی خاص قادرند فرضیه‌های آزمون‌های پارامتری را بیازمایند، چرا از همان ابتدا این آزمون‌ها به جای آزمون‌های پارامتری به کار نمی‌روند؟ پاسخ در توانمندی آزمون‌های پارامتری در مقایسه با آزمون‌های ناپارامتری نهفته است. این توانمندی نیز برگرفته از همان پیش‌فرض‌هایی است که آزمون‌های پارامتری بر اساس آنها شکل گرفته‌اند. برای نمونه، اگر بدانید متغیر تحت بررسی در جامعه آماری دارای توزیع نرمال است، آنگاه اطلاعی از این متغیر در دست دارید که بر اساس آن می‌توانید آزمون دقیق‌تری برای میانگین آن متغیر در جامعه اجرا کنید.

به طور کلی، آزمون‌های پارامتری در صورتِ درستی پیش‌فرض‌هایشان از توان بیشتری نسبت به آزمون‌های ناپارامتری برخوردارند؛ یعنی بیشتر می‌توانند یک فرضیه صفر نادرست را شناسایی کنند. ضعف دیگر آزمون‌های ناپارامتری کاهش سطح سنجش متغیرهاست. خواهیم دید که اغلب آزمون‌های ناپارامتری با متغیرهای فاصله‌ای یا نسبتی مانند یک متغیر ترتیبی عمل می‌کنند. این به معنای نادیده‌گرفتن دقتی است که متغیرهای فاصله‌ای یا نسبتی در اندازه‌گیری دارا هستند.

در این فصل به معرفی تعدادی از آزمون‌های ناپارامتری خواهیم پرداخت. بحث را با آن دسته از آزمون‌های ناپارامتری آغاز می‌کنیم که جانشین آزمون‌های پارامتری مربوط به میانگین‌ها هستند. سپس دو ضریب برای اندازه‌گیری رابطهٔ دو متغیر به عنوان جانشین ضریب همبستگی پیرسون معرفی می‌شوند. پایان‌بخش بحث آزمون‌های ناپارامتری به

آزمون‌هایی اختصاص دارد که رابطهٔ دو متغیر اسمی یا ترتیبی را در جدول‌های توافقی بررسی می‌کنند.

۲ـ آزمون علامت

آزمون علامت، معادلی ناپارامتری برای آزمون درباره نقطهٔ تمرکز توزیع یک متغیر در جامعه است. پس معادلی ناپارامتری برای آزمون پارامتری تک میانگین تلقی می‌شود؛ با این تفاوت که «میانه» به جای میانگین به عنوان نقطهٔ تمرکز توزیع در نظر گرفته می‌شود. مانند آزمون تک میانگین، سه فرضیه‌بندی زیر درباره میانهٔ توزیع مطرح است (حرف M میانهٔ توزیع متغیر X را نشان می‌دهد):

۱. آزمون دو دامنه: فرضیه صفر $M = M_o$ در مقابل فرضیه جانشین $M \neq M_o$
۲. آزمون یک دامنه: فرضیه صفر $M \leq M_o$ در مقابل فرضیه جانشین $M > M_o$
۳. آزمون یک دامنه: فرضیه صفر $M \geq M_o$ در مقابل فرضیه جانشین $M < M_o$

اگر توزیع متغیر متقارن باشد، آنگاه میانگین می‌تواند به جای میانه در فرضیه‌بندی‌های فوق به کار رود، زیرا میانگین و میانه در توزیع متقارن برابرند. به این ترتیب، آزمون علامت می‌تواند فرضیه‌ای را درباره میانگین توزیع بیازماید.

آزمون بر اساس داده‌های نمونه‌ای n تایی از جامعه صورت می‌گیرد. در نخستین گام، داده‌ها به سه دسته تفکیک می‌شوند: داده‌های بزرگ‌تر از M_o (پس $X - M_o$ مثبت است)، داده‌های برابر با M_o (پس $X - M_o = ۰$) و داده‌های کوچک‌تر از M_o (پس $X - M_o$ منفی است). تعداد داده‌های بزرگ‌تر از M_o را با n^+ و تعداد داده‌های کوچک‌تر از M_o را با n^- نشان می‌دهیم و بزرگ‌ترین آنها را به عنوان آمارهٔ آزمون به کار می‌بریم و با T نمایش می‌دهیم. نقطهٔ بحرانی آزمون به ازای α و $n^- + n^+$ از جدول نقاط بحرانی آزمون علامت در پیوست تعیین می‌شود. اگر مقدار آماره برابر با مقدار بحرانی این جدول یا بزرگ‌تر از آن باشد، فرضیه صفر رد می‌شود.

برای مثال، فرض کنید ۷ داور میزان نمایش خشونت در فیلمی را ارزیابی و به آن نمره‌ای از ۰ تا ۱۰ داده‌اند (جدول ۱). می‌خواهیم بدانیم میانهٔ این نمرات می‌تواند بیش از ۸ باشد. بنابراین، فرضیه صفر $M \leq ۸$ در مقابل فرضیه جانشین $M > ۸$ آزموده می‌شود. علامت‌های مثبت در جدول ۱ نشان‌دهنده نمرات بیش از ۸ و علامت‌های منفی نشان‌دهنده

نمرات کمتر از ۸ هستند، پس $n^{+} = ۴$ و $n^{-} = ۲$. بنابراین آمارهٔ آزمون برابر با بزرگ‌ترین آنها یعنی $T = ۴$ است.

جدول ۱: نمرات داوران به میزان خشونت در یک فیلم

داور	۱	۲	۳	۴	۵	۶	۷
نمره	۹	۹/۵	۱۰	۸	۸/۵	۷	۵
علامت	+	+	+	۰	+	−	−

جدول زیر بخشی از جدول نقاط بحرانی آزمون علامت است. اگر آزمون به ازای $\alpha = ۰/۰۵$ صورت گیرد، مقدار نقطهٔ بحرانی برای این آزمون یک دامنه به ازای $n^{-} + n^{+} = ۶$ برابر با ۶ است، پس فرضیه صفر رد نمی‌شود زیرا $T < ۶$ است.

بخشی از نقاط بحرانی آزمون علامت

α دو دامنه	۰/۱	۰/۰۵	۰/۰۲	۰/۰۱	α دو دامنه	۰/۱	۰/۰۵	۰/۰۲	۰/۰۱
α یک دامنه	۰/۰۵	۰/۰۲۵	۰/۰۱	۰/۰۰۵	α یک دامنه	۰/۰۵	۰/۰۲۵	۰/۰۱	۰/۰۰۵
$n^{-}+n^{+}$	نقاط بحرانی				$n^{-}+n^{+}$	نقاط بحرانی			
۱	−	−	−	−	۳۱	۱۱	۱۳	۱۵	۱۷
۲	−	−	−	−	۳۲	۱۲	۱۴	۱۶	۱۶
۳	−	−	−	−	۳۳	۱۱	۱۳	۱۵	۱۷
۴	−	−	−	−	۳۴	۱۲	۱۴	۱۶	۱۶
۵	۵	−	−	−	۳۵	۱۱	۱۳	۱۵	۱۷
۶	۶	۶	−	−	۳۶	۱۲	۱۴	۱۶	۱۸
۷	۷	۷	۷	−	۳۷	۱۱	۱۳	۱۷	۱۷
۸	۶	۸	۸	۸	۳۸	۱۲	۱۴	۱۶	۱۸

هنگامی که $n^{-} + n^{+} > ۱۰$ است، از آمارهٔ آزمون زیر برای اجرای آزمون علامت استفاده می‌شود:

$$z = \frac{T - \frac{n^{-}+n^{+}}{۲}}{\sqrt{\frac{n^{-}+n^{+}}{۴}}}$$

نقاط بحرانی برای آزمون دو دامنه $Z_{\frac{\alpha}{۲}}$ و $-Z_{\frac{\alpha}{۲}}$ از جدول توزیع نرمال استاندارد و برای آزمون‌های یک دامنه بر حسب نوع فرضیه صفر Z_{α} یا $-Z_{\alpha}$ خواهد بود. بر اساس داده‌های

پیشین، مقدار z به صورت زیر محاسبه می‌شود، هر چند با توجه به $n^- + n^+ = ۶ < ۱۰$، همان آمارهٔ T و مراجعه به جدول پیشین برای داده‌های جدول ۱ مناسب است.

$$z = \frac{۴ - \frac{۶}{۲}}{\sqrt{\frac{۶}{۴}}} = ۰/۸۱۶$$

نقطهٔ بحرانی آزمون یک دامنهٔ فرضیه صفر $M \leq ۸$ در مقابل فرضیه جانشین $M > ۸$ به ازای $\alpha = ۰/۰۵$ از جدول توزیع نرمال استاندارد برابر با $Z_{۰/۰۵} = ۱/۶۴$ است. از آنجا که مقدار آماره ۰/۸۱۶ از ۱/۶۴ بزرگ‌تر نیست، پس نمی‌توان فرضیه صفر را به نفع فرضیه جانشین رد کرد.

آمارهٔ آزمون علامت تنها بر اساس تعداد داده‌های بیش از میانه یا کمتر از میانه یعنی تعداد علامت‌های مثبت و منفی تعیین می‌شود، بنابراین به میزان اختلاف داده‌ها از میانه اهمیتی داده نمی‌شود. برای مثال، مقدار آمارهٔ آزمون علامت برای فرضیه صفر $M \leq ۸$ در هر دو مجموعه دادهٔ زیر یکسان است، هرچند اختلاف داده‌ها از میانه در این دو مجموعه کاملاً متفاوت است (داده‌ها در مجموعه دوم برخلاف مجموعه اول، در اطراف عدد ۸ هستند).

مجموعه اول	۹	۹/۵	۱۰	۸	۸/۵	۷	۵
علامت	+	+	+	۰	+	−	−
مجموعه دوم	۸/۵	۸/۵	۸/۷۵	۸	۸/۲۵	۷/۵	۷/۷۵
علامت	+	+	+	۰	+	−	−

آزمون رتبهٔ علامت‌دار ویلکاکسون[1] آزمون ناپارامتری دیگری است که تاحدی اختلاف داده‌ها از میانه را در تعیین آمارهٔ آزمون به حساب می‌آورد. در این آزمون، ابتدا تفاضل هر داده از میانه محاسبه می‌شود و داده‌های برابر با میانه یعنی همان تفاضل‌های صفر حذف می‌گردند. سپس قدرمطلق تفاضل‌ها، یعنی بدون درنظر گرفتن علامت آنها، از کمترین به بیشترین مرتب می‌شوند و از ۱ تا $n^- + n^+$ رتبه می‌گیرند. در نهایت، مجموع رتبه‌های تفاضل‌های مثبت (T^+) و مجموع رتبه‌های تفاضل‌های منفی (T^-) محاسبه و مقدار کوچک‌تر (T) به عنوان آمارهٔ آزمون به کار می‌رود. اکنون این آزمون را برای داده‌های جدول ۱ به کار می‌گیریم.

1. Wilcoxon

ممکن است در حین رتبه‌دادن، با داده‌ای تکراری روبه‌رو شوید پس باید درباره رتبهٔ آنها نیز تصمیم بگیرید. برای مثال، در جدول زیر، دو عددِ ۱ در ردیف قدر مطلق تفاضل‌ها به چشم می‌خورند. برای تعیین رتبهٔ آنها، به داده‌های مرتب‌شده بدون توجه به تکرار برخی از داده‌ها از ۱ تا ۶ رتبه بدهید، سپس میانگینِ رتبه‌های داده‌های یکسان را برای همگی آنها به کار ببرید. در جدول زیر، رتبه‌های ۲ و ۳ به دو تفاضل تکراریِ ۱ نسبت داده شده است. میانگین ۲ و ۳ برابر با ۲/۵ است، پس رتبهٔ این دو عددِ ۱ به طور یکسان ۲/۵ خواهد بود. از آنجا که رتبهٔ یکسانی به دو عدد ۱ اختصاص یافته است، به آنها همرُتبگی در داده‌ها گفته می‌شود.

داور	۱	۲	۳	۴	۵	۶	۷
نمره	۹	۹/۵	۱۰	۸	۸/۵	۷	۵
تفاضل	+۱	+۱/۵	+۲	۰	+۰/۵	−۱	−۳
قدر مطلق تفاضل	۱	۱/۵	۲	حذف	۰/۵	۱	۳
رتبه	۲	۴	۵	حذف	۱	۳	۶
همرتبگی	۲/۵					۲/۵	

بر این اساس، مقدار آمارهٔ آزمون رتبهٔ علامت‌دار ویلکاکسون با توجه به دو کمیت زیر برابر با $T = ۸/۵$ است.

$$T^{+} = ۲/۵ + ۴ + ۵ + ۱ = ۱۲/۵ \qquad T^{-} = ۲/۵ + ۶ = ۸/۵$$

نقطهٔ بحرانی آزمون از جدول نقاط بحرانی آزمون رتبهٔ علامت‌دار ویلکاکسون در پیوست به ازای α و $n^{-}+n^{+}$ تعیین می‌شود. اگر مقدار آماره کوچک‌تر یا برابر با نقطهٔ بحرانی باشد، فرضیه صفر به نفع فرضیه جانشین رد می‌شود. مقدار آماره در مثال پیشین برابر با ۸/۵ است، در حالی که مقدار نقطهٔ بحرانی برای این آزمون یک دامنه به ازای $\alpha = ۰/۰۵$ و $n^{-}+n^{+} = ۱۰$ برابر با ۲ است، پس فرضیه صفر $M \leq ۸$ رد نمی‌شود.

بخشی از جدول نقاط بحرانی آزمون رتبهٔ علامت‌دار ویلکاکسون

α یک دامنه		α دو دامنه			$n^{-}+n^{+}$	α یک دامنه		α دو دامنه			$n^{-}+n^{+}$
۰/۰۱	۰/۰۵	۰/۰۰۱	۰/۰۱	۰/۰۵		۰/۰۱	۰/۰۵	۰/۰۰۱	۰/۰۱	۰/۰۵	
۱۸۵	۲۲۷	۱۳۰	۱۷۱	۲۰۸	۳۶	−	۲	−	−	۰	۶
۱۹۸	۲۴۱	۱۴۰	۱۸۲	۲۲۱	۳۷	۰	۳	−	−	۲	۷
۲۱۱	۲۵۶	۱۵۰	۱۹۴	۲۳۵	۳۸	۱	۵	−	۰	۳	۸
۲۲۴	۲۷۱	۱۶۱	۲۰۷	۲۴۹	۳۹	۳	۸	−	۱	۵	۹
۲۳۸	۲۸۶	۱۷۲	۲۲۰	۲۶۴	۴۰	۵	۱۰	−	۳	۸	۱۰

اگر $n^{-}+n^{+}>۱۰$ باشد، از آمارهٔ آزمون زیر برای اجرای آزمون رتبهٔ علامت‌دار ویلکاکسون استفاده می‌شود:

$$z=\frac{T-\frac{(n^{-}+n^{+})(n^{-}+n^{+}+۱)}{۴}}{\sqrt{\frac{(n^{-}+n^{+})(n^{-}+n^{+}+۱)[۲(n^{-}+n^{+})+۱]}{۲۴}}}$$

نقاط بحرانی برای آزمون دو دامنه $Z_{\frac{\alpha}{۲}}$ و $-Z_{\frac{\alpha}{۲}}$ از جدول توزیع نرمال استاندارد و برای آزمون‌های یک دامنه بر حسب نوع فرضیه صفر Z_{α} یا $-Z_{\alpha}$ خواهد بود.

بر اساس داده‌های پیشین، مقدار z برای آزمون رتبهٔ علامت‌دار ویلکاکسون به صورت زیر محاسبه می‌شود؛ هر چند با توجه به $n^{-}+n^{+}=۶<۱۰$، همان آمارهٔ T و مراجعه به جدول پیشین برای آن مناسب است.

$$z=\frac{۸/۵-\frac{۶\times۷}{۴}}{\sqrt{\frac{۶\times۷\times۱۳}{۲۴}}}=-۰/۴۱۹$$

پس فرضیه صفر $M\leq۸$ رد نمی‌شود، زیرا مقدار آماره به ازای $\alpha=۰/۰۵$ از $Z_{۰/۰۵}=۱/۶۴$ بزرگ‌تر نیست.

۳ـ آزمون رتبهٔ علامت‌دار ویلکاکسون برای دو جامعه وابسته

دو دسته داده از n فردِ نمونه گردآوری شده است تا نقطهٔ تمرکز توزیع متغیر در دو جامعه وابسته مقایسه شود. به یاد بیاورید که آزمون پارامتری مقایسهٔ میانگین‌های دو جامعه وابسته، این دو دسته داده را با محاسبه تفاضل‌ها مطابق جدول ۲ فصل قبل، به یک دسته داده تبدیل و به این ترتیب از همان آزمون تک میانگین برای مقایسهٔ دو میانگین استفاده می‌کرد. آزمون ناپارامتری متناظر با آن نیز از همان منطق پیروی می‌کند؛ یعنی با محاسبه تفاضل‌ها، دو دسته داده به یک دسته تبدیل می‌شود و از آزمون علامت یا آزمون رتبهٔ علامت‌دار ویلکاکسون برای اجرای مقایسهٔ نقطهٔ تمرکز توزیع متغیر در دو جامعه وابسته استفاده می‌گردد. در اینجا، آزمون رتبهٔ علامت‌دار ویلکاکسون برای این مقایسه به کار گرفته می‌شود، هرچند آزمون علامت نیز به‌راحتی قابل استفاده است.

ابتدا تفاضل داده‌ها یعنی D_iها را مشابه جدول ۲ فصل قبل، برای آزمون ناپارامتری مقایسهٔ میانه‌های دو جامعه وابسته در نظر بگیرید. بر این اساس، فرضیه‌های زیر درباره میانهٔ D_iها قابل بیان هستند:

۱. آزمون دو دامنه: فرضیه صفر $M_D = ۰$ در مقابل فرضیه جانشین $M_D \neq ۰$
۲. آزمون یک دامنه: فرضیه صفر $M_D \geq ۰$ در مقابل فرضیه جانشین $M_D < ۰$
۳. آزمون یک دامنه: فرضیه صفر $M_D \leq ۰$ در مقابل فرضیه جانشین $M_D > ۰$

برای اجرای آزمون، قدر مطلق تفاضل‌ها از کوچک به بزرگ مرتب و تفاضل‌های ۰ حذف می‌شوند. اگر تعداد تفاضل‌های مثبت را با n^+ و تعداد تفاضل‌های منفی را با n^- نشان دهیم، به تفاضل‌ها بدون توجه به علامت از ۱ تا $n^- + n^+$ رتبه داده می‌شود. سپس مجموع رتبه‌های تفاضل‌های مثبت (T^+) و مجموع رتبه‌های تفاضل‌های منفی (T^-) جداگانه محاسبه و کوچک‌ترین آنها (T) به عنوان آمارهٔ آزمون انتخاب می‌شود. چگونگی یافتن نقطهٔ بحرانی و تصمیم‌گیری درباره فرضیه صفر مانند آزمون رتبهٔ علامت‌دار ویلکاکسون برای میانهٔ یک جامعه است که پیشتر شرح داده شد. مثال زیر به روشن شدن استفاده از آزمون رتبهٔ علامت‌دار ویلکاکسون برای مقایسهٔ نقطهٔ تمرکز یک متغیر در دو جامعه وابسته کمک می‌کند.

نمرات درس روش تحقیق و آمار ۱۰ دانشجو در جدول زیر آمده است. اگر M_r و M_s به ترتیب میانهٔ نمرات درس روش‌تحقیق و درس آمار باشد، فرضیه $M_r = M_s$ را در مقابل فرضیه صفر جانشین $M_r \neq M_s$، با آزمون رتبهٔ علامت‌دار ویلکاکسون می‌آزماییم.

روش تحقیق	آمار	تفاضل	قدرمطلق تفاضل	رتبه	همرتبگی
۱۸	۱۵	۳	۳	۶	۷
۱۷	۱۷	۰	۰	حذف	حذف
۱۹	۱۶	۳	۳	۷	۷
۱۴	۱۶	−۲	۲	۲	۳/۵
۲۰	۱۸	۲	۲	۳	۳/۵
۱۶	۱۷	−۱	۱	۱	
۱۵	۱۳	۲	۲	۴	۳/۵
۱۸	۱۵	۳	۳	۸	۷
۱۹	۱۷	۲	۲	۵	۳/۵
۱۵	۱۵	۰	حذف	حذف	حذف

آمارهٔ آزمون، $T = ۴/۵$ است، زیرا:

$$T^{+} = ۷+۷+۳/۵+۳/۵+۷+۳/۵ = ۳۲/۵ \quad T^{-} = ۳/۵+۱ = ۴/۵$$

مقدار نقطهٔ بحرانی برای این آزمون دو دامنه به ازای $\alpha = ۰/۰۵$ و $n^{-} + n^{+} = ۹$ از جدول نقاط بحرانی آزمون رتبهٔ علامت‌دار ویلکاکسون، برابر با ۵ است، پس فرضیه صفرِ برابری میانه‌های نمرات دو درس رد می‌شود، یعنی نقطهٔ میانهٔ نمرات درس روش تحقیق بیش از میانهٔ نمرات درس آمار است. مقدار آماره برای نمونه‌های بزرگ نیز به صورت زیر محاسبه می‌شود:

$$z = \frac{۴/۵ - \frac{۹\times۱۰}{۴}}{\sqrt{\frac{۹\times۱۰\times۱۹}{۲۴}}} = -۲/۱۳$$

این مقدار به ازای $\alpha = ۰/۰۵$ از $-Z_{۰/۰۲۵} = -۱/۹۶$ کوچک‌تر است، پس فرضیه صفرِ برابری میانه‌های نمرات دو درس رد می‌شود.

۴ـ آزمون مجموع رتبه‌ای ویلکاکسون و آزمون من‌ـویتنی

آزمون مجموع رتبه‌ای ویلکاکسون برای آزمودن یکسانیِ «توزیع» یک متغیر در دو جامعه مستقل به کار می‌رود، ولی در اینجا به منظور آزمودن برابریِ نقطهٔ تمرکز توزیع یک متغیر در دو جامعه مستقل معرفی می‌شود، پس جانشینی ناپارامتری برای آزمون پارامتری t برای مقایسهٔ میانگین‌های دو جامعه مستقل است؛ خواه واریانس‌های آنها برابر باشد یا نباشد.

برای اجرای این آزمون، نمونه‌ای به اندازه $n_۱$ از جامعه اول و نمونه‌ای به اندازه $n_۲$ از جامعه دوم انتخاب می‌شود. سپس داده‌های هر دو نمونه از کمترین به بیشترین مقدار مرتب و به آنها از ۱ تا $n_۱ + n_۲$ رتبه داده می‌شود. در نهایت، جمع رتبه‌های هر نمونه یعنی $W_۱$ و $W_۲$ محاسبه می‌گردد و بزرگ‌ترین آنها به عنوان آمارهٔ آزمون به کار می‌رود[۱]. توجه داشته باشید که همواره رابطهٔ زیر بین جمع‌های رتبه‌ای دو نمونه برقرار است. بنابراین می‌توان با محاسبه یکی از آنها، دیگری را نیز از این رابطه به دست آورد.

$$W_۱ + W_۲ = \frac{(n_۱ + n_۲)(n_۱ + n_۲ + ۱)}{۲}$$

برای روشن شدن این فرایند، از داده‌های جدول بعد کمک می‌گیریم. فرض کنید گروه

۱. در واقع، هر دو جمع می‌توانند به عنوان آمارهٔ آزمون استفاده شوند، مشروط بر آنکه جدول نقاط بحرانی مناسب در اختیار باشد. از آنجا که ما جدول نقاط بحرانی برای مجموع بزرگ‌تر را به کار می‌گیریم، مجموع بزرگ‌تر را نیز به عنوان آمارهٔ آزمون استفاده می‌کنیم.

اول با ۶ نمونه شامل نمراتی از ۰ تا ۱۰۰ است که انگیزهٔ کاری کارگران یک کارخانه را نشان می‌دهد و گروه دوم با ۵ نمونه شامل نمرات انگیزهٔ کاری کارمندان همان کارخانه است. هدف، بررسی این امر است که آیا انگیزهٔ کاری کارمندان با کارگران برابر است؟

گروه	داده‌ها	اندازه نمونه
کارگران (۱)	۶۰، ۳۰، ۱۵، ۵۰، ۴۵، ۷۵	$n_۱ = ۶$
کارمندان (۲)	۷۵، ۹۰، ۶۰، ۷۸، ۹۵	$n_۲ = ۵$

از جامعه کارگران $n_۱ = ۶$ نمونه و از جامعه کارمندان $n_۲ = ۵$ نمونه در دست است. پس در مجموع ۱۱ داده در اختیار داریم که آنها را از کمترین به بیشترین، به صورت جدول زیر مرتب می‌کنیم و از ۱ تا ۱۱ رتبه می‌دهیم.

گروه	۱	۱	۱	۱	۱	۲	۱	۲	۲	۲	۲
داده‌های مرتب‌شده	۱۵	۳۰	۴۵	۵۰	۶۰	۶۰	۷۵	۷۵	۷۸	۹۰	۹۵
رتبه‌ها	۱	۲	۳	۴	۵	۶	۷	۸	۹	۱۰	۱۱
همرتبگی‌ها					۵/۵	۵/۵	۷/۵	۷/۵			

در جدول فوق دو مورد همرتبگی وجود دارد، زیرا دو عدد ۶۰ و دو عدد ۷۵ در ردیف داده‌های مرتب‌شده به چشم می‌خورند. برای تعیین رتبهٔ آنها، ابتدا به داده‌های مرتب‌شده بدون توجه به تکرار برخی از داده‌ها از ۱ تا ۱۱ رتبه داده‌ایم و سپس میانگین رتبه‌های داده‌های یکسان را برای همگی آنها به کار برده‌ایم. دو رتبهٔ ۵ و ۶ به دو دادهٔ تکراری ۶۰ نسبت داده شده است. میانگینِ ۵ و ۶ برابر با ۵/۵ است، پس رتبهٔ این دو عددِ ۶۰ به طور یکسان ۵/۵ خواهد بود. به همین ترتیب، میانگین رتبه‌های ۷ و ۸ یعنی ۷/۵ را برای دو عددِ ۷۵ به کار می‌گیریم.

جمع رتبه‌های هر یک از نمونه‌ها به صورت زیر محاسبه و بزرگ‌ترین جمع یعنی ۴۳ به عنوان مقدار آمارهٔ آزمون تعیین می‌شود.

$$W_۱ = ۱+۲+۳+۴+۵/۵+۷/۵ = ۲۳ \quad W_۲ = ۵/۵+۷/۵+۹+۱۰+۱۱ = ۴۳$$

توجه کنید که رابطهٔ پیشگفته برای این مثال به صورت زیر برقرار است:

$$۲۳+۴۳ = \frac{(۶+۵)\times(۶+۵+۱)}{۲} = ۶۶$$

نقطهٔ بحرانی این آزمون به ازای مقادیر $n_۱$، $n_۲$ و α از جدول نقاط بحرانی مربوط به

آزمون مجموع رتبه‌ای ویلکاکسون در پیوست تعیین می‌شود. بخشی از این جدول را در زیر مشاهده می‌کنید که دارای دو قسمت با عنوان‌های «دامنهٔ پایینی» و «دامنهٔ بالایی» است. برای آزمون‌های دو دامنه، دو نقطهٔ بحرانی یکی از قسمت «دامنهٔ پایینی» و یکی از قسمت «دامنهٔ بالایی» انتخاب می‌شود، در حالی که برای آزمون‌های یک دامنه تنها یک نقطهٔ بحرانی بر حسب نوع فرضیه از قسمت «دامنهٔ پایینی» یا «دامنهٔ بالایی» انتخاب می‌گردد.

فرض کنید در مثال کنونی با آزمونی دو دامنه روبه‌رو هستید که فرضیه صفر «انگیزهٔ کاری کارمندان با کارگران برابر است» را در مقابل فرضیه جانشین «انگیزهٔ کاری کارمندان با کارگران برابر نیست» می‌آزمایید. اگر آزمون در سطح معناداری $\alpha = ۰/۰۵$ اجرا شود، در جدول نقاط بحرانی باید به دنبال دو مقدار به ازای $n_{min} = n_۲ = ۵$، $n_{max} = n_۱ = ۶$ و $\frac{\alpha}{۲} = ۰/۰۲۵$ باشید (توجه کنید که آزمون دو دامنه است، پس α نصف می‌شود). بنابراین، نقطهٔ بحرانی پایینی برابر با ۱۸ و نقطهٔ بحرانی بالایی برابر با ۴۲ است[1].

بخشی از جدول نقاط بحرانی آزمون مجموع رتبه‌ای ویلکاکسون

n_{min}	n_{max}	دامنهٔ پایین ۰/۰۰۵	۰/۰۱	۰/۰۲۵	۰/۰۵	۰/۱۰	۰/۲۰	دامنهٔ بالا ۰/۲۰	۰/۱۰	۰/۰۵	۰/۰۲۵	۰/۰۱	۰/۰۰۵
۵	۵	۱۵	۱۶	۱۷	۱۹	۲۰	۲۲	۳۳	۳۵	۳۶	۳۸	۳۹	۴۰
	۶	۱۶	۱۷	**۱۸**	۲۰	۲۲	۲۴	۳۶	۳۸	۴۰	**۴۲**	۴۳	۴۴
	۷	۱۶	۱۸	۲۰	۲۱	۲۳	۲۶	۳۹	۴۲	۴۴	۴۵	۴۷	۴۹
	۸	۱۷	۱۹	۲۱	۲۳	۲۵	۲۸	۴۲	۴۵	۴۷	۴۹	۵۱	۵۳
	۹	۱۸	۲۰	۲۲	۲۴	۲۷	۳۰	۴۵	۴۸	۵۱	۵۳	۵۵	۵۷
	۱۰	۱۹	۲۱	۲۳	۲۶	۲۸	۳۲	۴۸	۵۲	۵۴	۵۷	۵۹	۶۱
	۱۱	۲۰	۲۲	۲۴	۲۷	۳۰	۳۴	۵۱	۵۵	۵۸	۶۱	۶۳	۶۵
	۱۲	۲۱	۲۳	۲۶	۲۸	۳۲	۳۶	۵۴	۵۸	۶۲	۶۴	۶۷	۶۹

اگر آمارهٔ آزمون از نقطهٔ بحرانی پایینی کوچک‌تر یا از نقطهٔ بحرانی بالایی بزرگ‌تر باشد، فرضیه صفر به نفع فرضیه جانشین رد می‌شود. در مثال ما، مقدار آماره ۴۳ از نقطهٔ بحرانی بالایی ۴۲ بزرگ‌تر است، پس فرضیه صفر رد می‌شود و نتیجه خواهیم گرفت که انگیزهٔ کاری کارمندان با کارگران برابر نیست.

---

۱. اگر جمع کوچک‌تر به عنوان نقطهٔ بحرانی انتخاب شود، با استفاده از نقاط بحرانی این جدول می‌توان نقاط بحرانی مناسب برای این آمارهٔ آزمون را نیز به دست آورد. برای این منظور کافی است نقاط بحرانی از $W_۱ + W_۲$ کم شوند. پس اگر ۱۸ و ۴۲ نقاط بحرانی برای آمارهٔ $W_۲ = ۴۳$ باشند، آنگاه نقاط بحرانی آمارهٔ کوچک‌تر $W_۱ = ۲۳$ برابر با ۲۴ و ۴۸ خواهد بود. توجه داشته باشید که نتیجهٔ به‌کارگیری هر دو آماره همیشه یکسان است.

هنگامی که از هر جامعه حداقل ۱۰ نمونه انتخاب شده باشد، مقدار آمارهٔ آزمون از رابطهٔ زیر محاسبه و نقطهٔ بحرانی آن از جدول توزیع نرمال استاندارد به ازای α مورد نظر تعیین می‌شود.

$$z = \frac{W - \frac{n_1(n_1+n_2+1)}{2}}{\sqrt{\frac{n_1 n_2 (n_1+n_2+1)}{12}}}$$

مقادیر بحرانی برای آزمون دو دامنه $Z_{\frac{\alpha}{2}}$ و $-Z_{\frac{\alpha}{2}}$ از جدول توزیع نرمال استاندارد و برای آزمون‌های یک دامنه بر حسب نوع فرضیه Z_α یا $-Z_\alpha$ است. این آماره برای مثال پیشین به صورت زیر محاسبه می‌شود (البته با توجه به اندازه نمونه‌ها همان W مناسب است و نباید از z استفاده شود. محاسبات زیر تنها برای ارائه شیوهٔ یافتن مقدار آمارهٔ z صورت می‌گیرد):

$$z = \frac{43 - \frac{5\times(5+6+1)}{2}}{\sqrt{\frac{5\times 6\times(5+6+1)}{12}}} = \frac{43-30}{5/48} = 2/37$$

اگر مقدار آماره در این آزمون دو دامنه از $Z_{0/025} = 1/96$ بزرگ‌تر یا از $-Z_{0/025} = -1/96$ کوچک‌تر باشد، فرضیه صفر رد می‌شود. در این مثال، مقدار ۲/۳۷ از ۱/۹۶ بزرگ‌تر است، پس مانند قبل نتیجه خواهیم گرفت توزیع متغیر در این دو جامعه دارای نقاط تمرکز متفاوتی است.

آزمون من‌ـویتنی[1] آزمون ناپارامتری دیگری برای آزمودن فرضیه صفر «یکسانی نقطهٔ تمرکز توزیع متغیر در دو جامعه مستقل» است. این آزمون نیز فرایندی مانند آزمون مجموع رتبه‌ای ویلکاکسون دارد، با این تفاوت که آمارهٔ آن به شکل دیگری محاسبه می‌شود. برای این منظور پس از محاسبه جمع‌های رتبه‌ای دو نمونه یعنی W_1 و W_2، دو کمیت زیر محاسبه و از کوچک‌ترین آنها به عنوان آمارهٔ آزمون استفاده می‌شود.

$$U_1 = n_1 n_2 + \frac{n_1(n_1+1)}{2} - W_1 \qquad U_1 = n_1 n_2 + \frac{n_2(n_2+1)}{2} - W_2$$

همواره رابطهٔ زیر بین U_1 و U_2 برقرار است، پس می‌توان با محاسبه یکی از آنها دیگری را از این رابطه به دست آورد.

$$U_1 + U_2 = n_1 n_2$$

1. Mann-Whitney

با توجه به اطلاعات مثال پیشین و محاسبات زیر، مقدار ۲ را به عنوان مقدارآمارهٔ آزمون انتخاب می‌کنیم، زیرا کوچک‌تر است.

$$U_1 = ۶ \times ۵ + \frac{۶ \times ۷}{۲} - ۲۳ = ۲۸ \qquad U_2 = ۶ \times ۵ + \frac{۵ \times ۶}{۲} - ۴۳ = ۲$$

اگر مقدار آماره از نقطهٔ بحرانی به ازای مقادیر n_1، n_2 و α کوچک‌تر باشد، فرضیه صفر رد می‌شود. جدول ۲، نقاط بحرانی آزمون من‌ـویتنی را برای حالت دو دامنه به ازای $\alpha = ۰/۰۵$ نشان می‌دهد. مقدار آمارهٔ ۲ از نقطهٔ بحرانی ۳ کوچک‌تر است، پس مانند آزمون مجموع رتبه‌ای ویلکاکسون فرضیه صفر رد می‌شود و نتیجه خواهیم گرفت که انگیزهٔ کاری کارمندان با کارگران برابر نیست[1].

جدول ۲: نقاط بحرانی آزمون دو دامنهٔ من‌ـویتنی برای $\alpha = ۰/۰۵$

n_{min} \ n_{max}	۵	۶	۷	۸	۹	۱۰	۱۱	۱۲
۳	۰	۱	۱	۲	۲	۳	۳	۴
۴	۱	۲	۳	۴	۴	۵	۶	۷
۵	۲	$\boxed{۳}$	۵	۶	۷	۸	۹	۱۱
۶		۵	۶	۸	۱۰	۱۱	۱۳	۱۴
۷			۸	۱۰	۱۲	۱۴	۱۶	۱۸
۸				۱۳	۱۵	۱۷	۱۹	۲۲
۹					۱۷	۲۰	۲۳	۲۶
۱۰						۲۳	۲۶	۲۹
۱۱							۳۰	۳۳
۱۲								۳۷

مانند آزمون مجموع رتبه‌ای ویلکاکسون، اگر از هر جامعه حداقل ۱۰ نمونه انتخاب شده باشد، مقدار آمارهٔ آزمون از رابطهٔ زیر محاسبه و نقطهٔ بحرانی آن از جدول توزیع نرمال استاندارد به ازای α مورد نظر تعیین می‌شود (اهمیتی ندارد کدامیک از Wها در آمارهٔ زیر به کار رود). شیوهٔ رد فرضیه صفر مانند آزمون مجموع رتبه‌ای ویلکاکسون است.

$$z = \frac{U - \frac{n_1 n_2}{۲}}{\sqrt{\frac{n_1 n_2 (n_1 + n_2 + ۱)}{۱۲}}}$$

۱. در آزمون من‌ـویتنی نیز مانند آزمون مجموع رتبه‌ای ویلکاکسون می‌توان هم از مقدار کوچک‌تر و هم از مقدار بزرگ‌تر Uها استفاده کرد. ما از جدولی برای یافتن نقاط بحرانی استفاده می‌کنیم که برای U کوچک‌تر تنظیم شده است.

مقدار این آمارهٔ آزمون برای اطلاعات مثال قبل و به ازای $U_1 = ۲۸$ به صورت زیر است. بر این اساس نیز بار دیگر فرضیه صفر رد می‌شود.

$$z = \frac{۲۸ - \frac{۵\times۶}{۲}}{\sqrt{\frac{۵\times۶\times۱۲}{۱۲}}} = ۲/۳۷$$

همان طور که پیشتر اشاره شد، آزمون من‌ـویتنی معادل آزمون مجموع رتبه‌ای ویلکاکسون است، به این معنی که اگر فرضیه صفر با یکی از آنها رد شود، بی‌تردید با آزمون دیگر نیز رد می‌شود. همچنین، مقدار آمارهٔ z برای هر دو آزمون برابر به دست می‌آید.

۵ـ آزمون کروسکال‌ـوالیس

آزمون کروسکال‌ـوالیس[1] معادل ناپارامتری آزمون تجزیهٔ واریانس برای مقایسهٔ چند میانگین است و فرایندی مانند تجزیهٔ واریانس دارد، با این تفاوت که بر روی رتبه‌های متناظر با داده‌ها اعمال می‌شود. این آزمون را در قالب مثال زیر شرح می‌دهیم.

در سازمانی، رضایت شغلی کارمندان سه واحد بررسی شده است. از واحد اول ۴ کارمند، از واحد دوم ۵ کارمند و از واحد سوم ۳ کارمند انتخاب و رضایت شغلی آنها مطابق جدول ۳ اندازه‌گیری شده است. بنابراین در مجموع $n = ۱۲$ داده در دست است.

جدول ۳: رضایت شغلی کارمندان سه واحد از یک سازمان

واحد	واحد ۱	واحد ۲	واحد ۳
داده‌ها	۸۵، ۹۰،۸۰، ۸۸	۷۰، ۷۵، ۷۰، ۸۰، ۹۰	۵۰، ۸۰، ۶۰
رتبه‌ها	۹، ۱۱/۵، ۷، ۱۰	۳/۵، ۵، ۳/۵، ۷، ۱۱/۵	۱، ۷، ۲
جمع رتبه‌ها (T_i)	$T_1 = ۳۷/۵$	$T_2 = ۳۰/۵$	$T_3 = ۱۰$
تعداد نمونه (n_i)	$n_1 = ۴$	$n_2 = ۵$	$n_3 = ۳$

ابتدا این داده‌ها از کوچک‌ترین به بزرگ‌ترین مرتب می‌شوند و از ۱ تا ۱۲ مطابق جدول زیر رتبه می‌گیرند. آنچه در سطر رتبه‌های جدول ۳ ارائه شده است، بر اساس این رتبه‌دهی است.

داده‌ها	۵۰	۶۰	۷۰	۷۰	۷۵	۸۰	۸۰	۸۰	۸۵	۸۸	۹۰	۹۰
رتبه	۱	۲	۳	۴	۵	۶	۷	۸	۹	۱۰	۱۱	۱۲
همرتبگی‌ها			۳/۵	۳/۵		۷	۷	۷			۱۱/۵	۱۱/۵

1. Kruskal-Wallis

آمارهٔ آزمون با استفاده از مجموع رتبه‌ها و اندازه‌های نمونه به صورت زیر محاسبه می‌شود (k تعداد گروه‌هایی است که مقایسه می‌شوند، پس $k=3$).

$$H=\frac{12}{n(n+1)}\sum_{i=1}^{k}\frac{T_i^2}{n_i}-3(n+1)$$

$$=\frac{12}{12\times 13}\left(\frac{37/5^2}{4}+\frac{30/5^2}{5}+\frac{10^2}{3}\right)-3\times 13=4/92$$

نقطهٔ بحرانی آزمون به ازای مقدار α، k و اندازه‌های نمونه‌های گروه‌ها از جدول نقاط بحرانی آزمون کروسکال‌ـوالیس به دست می‌آید. بخشی از این جدول در زیر ملاحظه می‌شود. مقدار نقطهٔ بحرانی برای آزمون فوق از این جدول به ازای $\alpha=0/05$ و $k=3$ با ۵/۶۵۶ برابر است، پس فرضیه صفر رد نمی‌شود.

بخشی از جدول نقاط بحرانی آزمون کروسکال‌ـوالیس برای $\alpha=0/05$ و $\alpha=0/01$

$k=5$							$k=4$						$k=3$				
۰/۰۱	۰/۰۵	اندازه نمونه					۰/۰۱	۰/۰۵	اندازه نمونه				۰/۰۱	۰/۰۵	اندازه نمونه		
۸/۴۲۴	۷/۵۷۶	۱	۱	۳	۳	۳	–	–	۱	۱	۱	۴	۷/۱۴۴	۵/۵۹۸	۳	۴	۴
۹/۰۵۱	۷/۷۶۹	۱	۲	۳	۳	۳	–	۵/۸۳۳	۱	۱	۲	۴	۷/۶۵۴	۵/۶۹۲	۴	۴	۴
۹/۵۰۵	۸/۰۴۴	۲	۲	۳	۳	۳	۷/۰۰۰	۶/۱۳۳	۱	۲	۲	۴	–	۵/۰۰۰	۱	۲	۵
۹/۴۵۱	۸/۰۰۰	۱	۳	۳	۳	۳	۷/۳۹۱	۶/۵۴۵	۲	۲	۲	۴	۶/۵۳۳	۵/۱۶۰	۲	۲	۵
۹/۸۷۶	۸/۲۰۰	۲	۳	۳	۳	۳	۷/۰۶۷	۶/۱۷۸	۱	۱	۳	۴	–	۴/۹۶۰	۱	۳	۵
۱۰/۲۰	۸/۳۳۳	۳	۳	۳	۳	۳	۷/۴۵۵	۶/۳۰۹	۱	۲	۳	۴	۶/۹۰۹	۵/۲۵۱	۲	۳	۵
							۷/۸۷۱	۶/۶۲۱	۲	۲	۳	۴	۷/۱۷۹	۵/۶۴۸	۳	۳	۵
							۷/۷۵۸	۶/۵۴۵	۱	۳	۳	۴	۶/۹۵۵	۴/۹۸۵	۱	۴	۵
							۸/۳۳۳	۶/۷۹۵	۲	۳	۳	۴	۷/۲۰۵	۵/۲۷۳	۲	۴	۵
							۸/۶۵۹	۶/۹۸۴	۳	۳	۳	۴	۷/۴۴۵	۵/۶۵۶	۳	۴	۵
							۷/۹۰۹	۵/۹۴۵	۱	۱	۴	۴	۷/۷۶۰	۵/۶۵۷	۴	۴	۵

اگر برای هر گروه حداقل ۵ نمونه انتخاب شده باشد، نقطهٔ بحرانی از جدول توزیع χ^2 تعیین می‌شود. در این صورت، کافی است مقدار H از $\chi^2_{df,\alpha}$ به ازای $df=k-1$ بزرگ‌تر باشد تا فرضیه صفرِ برابری نقاط تمرکز توزیع متغیر در k جامعه رد شود. نقطهٔ بحرانی برای آزمون مثال پیشین به ازای $df=2$ و $\alpha=0/05$ برابر با $\chi^2_{2,0/05}=5/991$ به دست می‌آید، ولی آماره H از آن کوچک‌تر است، پس فرضیه صفر رد نمی‌شود[1].

۱. تمرین چهارم در پایان این فصل، آزمون میانه را که بر اساس آزمون نیکویی برازش است، برای مقایسه چند گروه معرفی می‌کند. این آزمون کم‌توان‌تر از آزمون کروسکال‌ـوالیس است.

۶ـ ضریب همبستگی اسپیرمن

معمولاً ضریب همبستگی اسپیرمن[1] به عنوان معادل ناپارامتری ضریب همبستگی پیرسون معرفی می‌شود، ولی این ضریب تنها به اندازه‌گیری رابطهٔ خطی بین دو متغیر اختصاص ندارد، بلکه قادر است هر رابطهٔ «یکنوایی» را اندازه‌گیری کند. مانند ضریب همبستگی پیرسون، دو متغیر X و Y از هر فرد نمونه n تایی اندازه‌گیری شده است. برای محاسبه ضریب همبستگی اسپیرمن، داده‌های هر یک از متغیرها جداگانه از کمترین به بیشترین رتبه می‌گیرند و سپس ضریب همبستگی پیرسونِ رتبه‌های دو متغیر به عنوان ضریب همبستگی اسپیرمن آن دو متغیر به کار می‌رود.

برای مثال، فرض کنید داده‌های جدول ۴ نتایج نمرهٔ آزمون کتبی (X) و مصاحبهٔ علمی (Y) مربوط به $n=۷$ داوطلب باشد. دو ستون نیز در این جدول شامل رتبه‌هایی است که به نمرات آزمون کتبی و نمرات مصاحبهٔ علمی از کمترین به بیشترین داده شده است. از آنجا که داوطلب سوم و ششم در متغیر X همرتبه و داوطلب سوم و هفتم نیز در متغیر Y همرتبه هستند، از میانگین رتبه‌های آنان برای این همرتبگی‌ها استفاده شده است. نمادهای R_i^X و R_i^Y به ترتیب رتبه‌های دو متغیر X و Y را نشان می‌دهند.

جدول ۴: نمرات آزمون کتبی و مصاحبهٔ علمی ۷ داوطلب

داوطلب (i)	آزمون کتبی (X_i)	رتبه آزمون (R_i^X)	مصاحبهٔ علمی (Y_i)	رتبه مصاحبه (R_i^Y)	$R_i^X - R_i^Y$
۱	۱۸	۵	۱۹	۷	−۲
۲	۱۵	۲	۱۶	۳	−۱
۳	۱۹	۶/۵	۱۷	۴/۵	۲
۴	۱۴	۱	۱۱	۱	۰
۵	۱۷	۴	۱۵	۲	۲
۶	۱۹	۶/۵	۱۸	۶	۰/۵
۷	۱۶	۳	۱۷	۴/۵	−۱/۵

ضریب همبستگی پیرسون دو متغیر X و Y به طور مستقیم بر اساس داده‌های این دو متغیر یعنی ستون X_iها و ستون Y_iها محاسبه می‌شود. حال اگر ضریب همبستگی پیرسون را به

1. Spearman

جای محاسبه بر اساس داده‌های دو متغیر، بر اساس رتبه‌های دو متغیر یعنی ستون R_i^Xها و R_i^Yها محاسبه کنید، به ضریب همبستگی اسپیرمن آن دو متغیر خواهید رسید. معمولاً به جای استفاده از رابطهٔ ضریب همبستگی پیرسون از رابطه‌های ساده‌تری برای محاسبه ضریب همبستگی اسپیرمن دو متغیر استفاده می‌شود، بویژه اگر همرتبگی در داده‌ها وجود نداشته باشد (نماد r_s برای نشان‌دان ضریب همبستگی اسپیرمن به کار می‌رود).

$$r_s = ۱ - \frac{۶\sum_{i=۱}^{n}(R_i^X - R_i^Y)^۲}{n(n^۲ - ۱)} = \frac{\sum_{i=۱}^{n} R_i^X R_i^Y - n\left(\frac{n+۱}{۲}\right)^۲}{\frac{n(n^۲-۱)}{۱۲}}$$

جدول ۴ شامل ستونی است که مقادیر $R_i^X - R_i^Y$ را در خود جای داده است. بر این اساس، ضریب همبستگی اسپیرمن دو متغیرِ نمرهٔ آزمون کتبی و نمرهٔ مصاحبهٔ علمی با توجه به رابطه فوق، عبارت است از:

$$r_s = ۱ - \frac{۶\times[(-۲)^۲ + (-۱)^۲ + ۲^۲ + ۰^۲ + ۲^۲ + ۰/۵^۲ + (-۱/۵)^۲]}{۷\times(۷^۲ - ۱)} = ۰/۷۱۸$$

مانند ضریب همبستگی پیرسون، ضریب همبستگی اسپیرمن نیز می‌تواند مقداری از ۱− تا ۱+ داشته باشد، به طوری که مقدار مثبت نشان‌دهنده رابطهٔ «یکنوای» افزایشی و مقدار منفی نشان‌دهنده رابطهٔ «یکنوای» کاهشی است. هر چه مقدار ضریب به ۱− یا ۱+ نزدیک‌تر باشد، شدت رابطهٔ یکنوا قوی‌تر است[۱]. مقدار ۱+ تنها زمانی به دست می‌آید که داده‌های مربوط به متغیرهای X و Y برای هر فرد نمونه، رتبهٔ یکسانی داشته باشند؛ یعنی $R_i^X = R_i^Y$ و مقدار ۱− نیز زمانی به دست می‌آید که رابطهٔ $R_i^X + R_i^Y = n+۱$ بین رتبه‌های داده‌های مربوط به متغیر X و Y برای هر فرد نمونه برقرار باشد؛ یعنی برای مثال اگر متغیر X برای فردی دارای رتبه r است، متغیر Y برای همین فرد دارای رتبه $n+۱-r$ است.

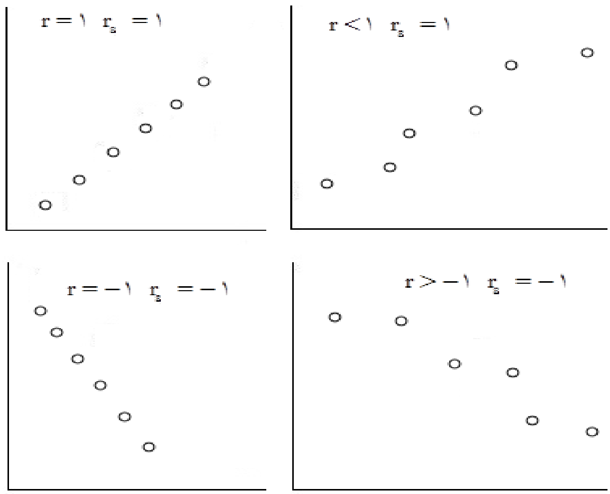
اغلب جهت رابطهٔ دو متغیر از طریق هر دو ضریب همبستگی پیرسون و اسپیرمن یکسان به دست می‌آید، ولی شدت رابطه لزوماً یکسان ارزیابی نمی‌شود، زیرا ضریب

۱. برخلاف ضریب همبستگی پیرسون که به‌خوبی می‌توانست شدت و جهت رابطهٔ «خطی» را ارائه دهد، ضریب همبستگی اسپیرمن به دلیل اتکای آن بر رتبه‌ها و نه داده‌ها، در ارائهٔ جهت رابطه، بهتر از شدت آن عمل می‌کند.

همبستگی اسپیرمن به دنبالِ رابطهٔ یکنواست، در حالی که ضریب همبستگی پیرسون تنها نسبت به رابطهٔ «خطی» حساس است. شکل ۱ به درک این موضوع کمک می‌کند.

مشابه همبستگی پیرسون، r_s نیز بر اساس داده‌های نمونه به دست می‌آید و برآوردی از ضریب همبستگی واقعی جامعه است که با ρ_s نشان داده می‌شود. بنابراین باید بررسی کرد که آیا r_s محاسبه‌شده برای دو متغیر بر اساس داده‌های نمونه بیانگر رابطهٔ آن دو متغیر در جامعه نیز هست؟ این، به معنای آزمون فرضیه صفر $\rho_s = ۰$ در مقابل فرضیه جانشین $\rho_s \neq ۰$ است. اگر $n < ۱۰$ باشد، از ضریب همبستگی اسپیرمن نمونه به عنوان آمارهٔ آزمون استفاده می‌شود. نقطهٔ بحرانی این آزمون نیز به ازای α و n از جدول نقاط بحرانی آزمون ضریب همبستگی اسپیرمن در پیوست محاسبه می‌شود. اگر قدرمطلق ضریب از نقطهٔ بحرانی بزرگ‌تر باشد، فرضیه صفر این آزمونِ دو دامنه رد می‌شود.

شکل ۱: ارزیابی شدت رابطه از طریق ضرایب همبستگی پیرسون و اسپیرمن



مقدار ضریب همبستگی اسپیرمن برای داده‌های جدول ۴ برابر با ۰/۷۱۸ بود. از طرفی مقدار نقطهٔ بحرانی به ازای $\alpha = ۰/۰۵$ و $n = ۷$ از جدول مقابل، برابر با ۰/۷۱۴ است، پس فرضیه صفر آزمون یک دامنهٔ $\rho_s \leq ۰$ رد می‌شود و می‌توان نتیجه گرفت بین نمرهٔ آزمون کتبی داوطلبان و نمرهٔ مصاحبهٔ علمی آنان رابطهٔ مستقیم یکنوا وجود دارد.

بخشی از جدول نقاط بحرانی آزمون ضریب همبستگی اسپیرمن

n	α دو دامنه		α یک دامنه	
	۰/۰۵	۰/۰۱	۰/۰۵	۰/۰۱
۵	۱	–	۰/۹۰۰	۱
۶	۰/۸۸۶	۱	۰/۸۲۹	۰/۹۴۳
۷	۰/۷۸۶	۰/۹۲۹	۰/۷۱۴	۰/۸۹۳
۸	۰/۷۳۸	۰/۸۸۱	۰/۶۴۳	۰/۸۳۳
۹	۰/۷۰۰	۰/۸۳۳	۰/۶۰۰	۰/۷۸۳
۱۰	۰/۶۴۸	۰/۷۹۴	۰/۵۶۴	۰/۷۴۵
۱۱	۰/۶۱۸	۰/۷۵۵	۰/۵۳۶	۰/۷۰۹

اگر $n \geq ۱۰$ باشد، از آمارهٔ زیر برای آزمون فرضیه صفر $\rho_s = ۰$ در مقابل فرضیه جانشین $\rho_s \neq ۰$ استفاده می‌شود. نقطهٔ بحرانی این آزمون دو دامنه از جدول توزیع نرمال استاندارد به ازای $\frac{\alpha}{۲}$ تعیین می‌شود. اگر مقدار آماره بزرگ‌تر از $Z_{\frac{\alpha}{۲}}$ یا کوچک‌تر از $-Z_{\frac{\alpha}{۲}}$ باشد، فرضیه صفر رد می‌شود و نتیجه خواهیم گرفت که بین دو متغیر رابطهٔ یکنوا برقرار است.

$$z = r_s\sqrt{n-۱}$$

۷ـ ضریب کندال

ضریب همبستگی اسپیرمن از منطق ضریب همبستگی پیرسون برای اندازه‌گیری رابطهٔ دو متغیر استفاده می‌کند، ولی به دلیل به‌کارگیری رتبه‌ها به جای داده‌ها، چندان در تعیین شدت رابطه موفق نیست. ضریب کندال[1] معیار ناپارامتری دیگری است که با منطقی متفاوت، جهت و شدت رابطهٔ «یکنوای» دو متغیر را می‌سنجد. در محاسبه این ضریب نیز مانند ضریب همبستگی اسپیرمن، زوجی از داده‌ها (مقدار متغیر X و مقدار متغیر Y) از هر یک از افراد نمونه n تایی در دست است.

محاسبه ضریب کندال مبتنی بر تعداد زوج‌های هماهنگ و ناهماهنگ است که از مقایسهٔ دوبه‌دویِ زوج‌ها مشخص می‌شوند. دو زوج را هماهنگ می‌نامند اگر مقدار مثلاً متغیر X برای یکی، بزرگ‌تر از مقدار متغیر X برای دیگری بود، آنگاه مقدار متغیر Y آن نیز بزرگ‌تر از مقدار متغیر Y دیگری باشد. برای نمونه، مقدار X برای داوطلب اول در جدول ۴ برابر با ۱۸ و بزرگ‌تر از مقدار X برای داوطلب دوم (۱۵) است. از سوی دیگر مقدار Y نیز برای داوطلب اول (۱۹) بزرگ‌تر از مقدار Y برای داوطلب دوم (۱۶) است، پس این دو

1. Kendall

زوج هماهنگ هستند. در مقابل، اگر مقدار X زوجی بزرگ‌تر از مقدار X زوج دیگر، ولی مقدار Y آن کوچک‌تر باشد، آنگاه آن دو زوج ناهماهنگ نامیده می‌شوند؛ مانند داوطلب پنجم و هفتم.

تشخیص تعداد زوج‌های هماهنگ و ناهماهنگ با مقایسهٔ دوبه‌دویِ زوج‌ها با یکدیگر امکان‌پذیر است. تعداد کل این مقایسه‌های دوبه‌دو برابر با $\frac{n(n-۱)}{۲}$ است. پس در مثال جدول ۴ به $\frac{۷\times۶}{۲}=۲۱$ مقایسهٔ دوبه‌دو نیاز است. یک راهکار عملی برای شمارش تعداد زوج‌های هماهنگ و ناهماهنگ، مرتب کردن زوج‌ها بر اساس مقدار یکی از متغیرها مثلاً متغیر X است. این کار برای داده‌های جدول ۴ در جدول زیر صورت گرفته است.

داوطلب	۴	۲	۷	۵	۱	۳	۶
X	۱۴	۱۵	۱۶	۱۷	۱۸	۱۹	۱۹
Y	۱۱	۱۶	۱۷	۱۵	۱۹	۱۷	۱۸

اکنون می‌توان با مقایسهٔ مقدار متغیر Y هر داوطلب با داوطلب‌های پس از آن، به شناسایی زوج‌های هماهنگ و ناهماهنگ دست زد (مثلاً داوطلب ۵ با داوطلب‌های بعد از خودش یعنی داوطلب‌های ۱، ۳ و ۶ مقایسه می‌شود). توجه کنید که اگر مقادیر X دو زوج یا مقادیر متغیر Y دو زوج برابر باشد، همرتبه به حساب می‌آیند و در تعیین تعداد زوج‌های هماهنگ و ناهماهنگ دخالت داده نمی‌شوند؛ مانند داوطلبان سوم و ششم یا داوطلبان هفتم و سوم. جدول زیر از مقایسهٔ دوبه‌دویِ زوج‌ها با این راهکار ارائه شده است.

زوج‌های هماهنگ	تعداد	زوج‌های ناهماهنگ	تعداد
داوطلب ۴ با داوطلب‌های ۲، ۷، ۵، ۱، ۳ و ۶	۶	داوطلب ۴ با هیچ‌کدام	۰
داوطلب ۲ با داوطلب‌های ۷، ۱، ۳ و ۶	۴	داوطلب ۲ با داوطلب ۵	۱
داوطلب ۷ با داوطلب‌های ۱ و ۶	۲	داوطلب ۷ با داوطلب ۵	۱
داوطلب ۵ با داوطلب‌های ۱، ۳ و ۶	۳	داوطلب ۵ با هیچ‌کدام	۰
داوطلب ۱ با هیچ‌کدام	۰	داوطلب ۱ با داوطلب‌های ۳ و ۶	۲
داوطلب ۳ با هیچ‌کدام	۰	داوطلب ۳ با هیچ‌کدام	۰
جمع زوج‌های هماهنگ	۱۵	جمع زوج‌های ناهماهنگ	۴

از آنجا که زوج‌های هماهنگ نشان‌دهنده رابطهٔ افزایشی و زوج‌های ناهماهنگ نشان‌دهنده رابطهٔ کاهشی هستند، پس با مقایسهٔ تعداد زوج‌های هماهنگ (n_c) با تعداد

زوج‌های ناهماهنگ (n_d) به جهت رابطه پی‌خواهیم برد. بنابراین اگر $n_c - n_d$ مثبت باشد، نتیجه خواهیم گرفت که تعداد زوج‌های هماهنگ بیشتر از زوج‌های ناهماهنگ است و رابطه‌ای افزایشی بین دو متغیر برقرار است. برعکس، اگر $n_c - n_d$ منفی باشد، تعداد زوج‌های ناهماهنگ بیشتر از زوج‌های هماهنگ است و رابطه‌ای کاهشی بین دو متغیر برقرار است. همچنین اگر $n_c - n_d$ برابر با صفر باشد، به آن معناست که به تعداد برابر زوج هماهنگ و ناهماهنگ وجود دارد، پس الگوی رابطهٔ دو متغیر «یکنوا» نیست و نوسان دارد. در عمل مقدار $n_c - n_d$ به صورت زیر بر تعداد کل مقایسه‌های دوبه‌دو یعنی $\frac{n(n-۱)}{۲}$ تقسیم می‌شود تا ضریبی به دست آید که در محدودهٔ ۱- تا ۱+ تغییر می‌کند.

$$r_{ken} = \frac{n_c - n_d}{\frac{n(n-۱)}{۲}}$$

هنگامی $r_{ken} = +۱$ است که هیچ زوج ناهماهنگ و هیچ همرتبگی وجود نداشته باشد؛ یعنی همه $\frac{n(n-۱)}{۲}$ مقایسهٔ دوبه‌دویِ زوج‌ها هماهنگ باشند. به این ترتیب، دو متغیر دارای رابطهٔ افزایشی کاملی هستند. از سوی دیگر، هنگامی $r_{ken} = -۱$ است که هیچ زوج هماهنگ و هیچ همرتبگی وجود نداشته باشد؛ یعنی همه $\frac{n(n-۱)}{۲}$ مقایسهٔ دوبه‌دویِ زوج‌ها ناهماهنگ باشند. در این صورت، دو متغیر دارای رابطهٔ کاهشی کاملی هستند.

در مثال جدول ۴ که تعداد زوج‌های هماهنگ و ناهماهنگ آن در جدول پیشین برابر با $n_c = ۱۵$ و $n_d = ۴$ به دست آمد، مقدار ضریب کندال به صورت زیر محاسبه می‌شود:

$$r_{ken} = \frac{۱۵ - ۴}{\frac{۷ \times ۸}{۲}} = ۰/۵۲۴$$

مثبت‌بودن این مقدار بر رابطهٔ یکنوای افزایشی بین دو متغیر نمرهٔ آزمون کتبی و نمرهٔ مصاحبهٔ علمی دلالت دارد؛ یعنی هرچه نمره آزمون کتبی بیشتر باشد، نمرهٔ مصاحبهٔ علمی نیز بیشتر می‌شود. همچنین از آنجا که مقدار این ضریب حداکثر می‌تواند ۱+ باشد، شدت رابطه، متوسط ارزیابی می‌شود.

ضریب r_{ken} بیانگر رابطهٔ دو متغیر در نمونه است و برآوردی از ضریبی قلمداد می‌شود که رابطهٔ دو متغیر را در جامعه نشان می‌دهد. برای تأکید بر این تمایز، ضریب کندالِ جامعه را با τ_{ken} نشان می‌دهیم (حرف یونانی « τ »، تاو خوانده می‌شود). پس با این پرسش

روبه‌رو هستیم که آیا ضریب حاصل از نمونه قابل تعمیم به جامعه است؟ پاسخ به این پرسش در آزمون فرضیه صفر $\tau_{ken} = ۰$ در مقابل فرضیه جانشین $\tau_{ken} \neq ۰$ نهفته است. از r_{ken} به عنوان آمارهٔ این آزمون برای $n < ۱۰$ استفاده می‌شود. نقطهٔ بحرانی آزمون از جدول نقاط بحرانی آزمون ضریب کندال در پیوست تعیین می‌گردد. اگر مقدار r_{ken} حداقل برابر با نقطهٔ بحرانی باشد، ضمن رد فرضیه صفر، نتیجه خواهیم گرفت که بین دو متغیر در جامعه رابطهٔ یکنوا برقرار است.

آزمون دو دامنهٔ فرضیه صفر $\tau_{ken} = ۰$ در مقابل فرضیه جانشین $\tau_{ken} \neq ۰$ برای داده‌های جدول ۴ با آمارهٔ $r_{ken} = ۰/۵۲۴$ صورت می‌گیرد. نقطهٔ بحرانی آزمون به ازای $\alpha = ۰/۰۵$ و $n = ۷$ از جدول نقاط بحرانی آزمون ضریب کندال که بخشی از آن در زیر مشاهده می‌شود، برابر با ۰/۷۱۴ است، پس نمی‌توان فرضیه صفر را رد کرد.

بخشی از جدول نقاط بحرانی برای آزمون ضریب کندال

n	یک دامنه		دو دامنه	
	۰/۰۵	۰/۰۱	۰/۰۵	۰/۰۱
۴	۱			
۵	۰/۸۰۰	۱	۱	
۶	۰/۷۳۳	۰/۸۶۷	۰/۸۶۷	۱
۷	۰/۶۱۹	۰/۸۱۰	$\boxed{۰/۷۱۴}$	۰/۹۰۵
۸	۰/۵۷۱	۰/۷۱۴	۰/۶۴۳	۰/۷۸۶

اگر $n \geq ۱۰$ باشد، از آمارهٔ زیر برای آزمون فرضیه صفر $\tau_{ken} = ۰$ در مقابل فرضیه جانشین $\tau_{ken} \neq ۰$ استفاده می‌شود. نقطهٔ بحرانی این آزمون دو دامنه از جدول توزیع نرمال استاندارد به ازای $\frac{\alpha}{۲}$ تعیین می‌گردد. اگر مقدار آماره بزرگ‌تر از $Z_{\frac{\alpha}{۲}}$ یا کوچک‌تر از $-Z_{\frac{\alpha}{۲}}$ باشد، فرضیه صفر رد می‌شود و نتیجه خواهیم گرفت که بین دو متغیر رابطهٔ یکنوا وجود دارد.

$$z = r_{ken}\sqrt{\frac{۹n(n-۱)}{۲(۲n+۵)}}$$

مقدار آماره برای داده‌های جدول ۴ به ازای $r_{ken} = ۰/۵۲۴$ و $n = ۷$ عبارت است از:

$$z = ۰/۵۲۴\sqrt{\frac{۹\times۷\times۶}{۲\times۱۹}} = ۱/۶۵$$

البته این آماره به دلیل کوچک‌بودن اندازه نمونه معتبر نیست و تنها برای آشنایی با شیوهٔ محاسبه ارائه شده است.

۸ـ تحلیل جدول‌های توافقی

رابطهٔ یک متغیر پیوسته با یک متغیر مقوله‌ای دو حالتی یا چند حالتی از طریق آزمون‌های مقایسهٔ میانگین‌ها بررسی می‌شود. همچنین رابطهٔ دو متغیر پیوسته از طریق رگرسیون و ضریب همبستگی پیرسون قابل بررسی است. معادل‌های ناپارامتری این روش‌ها نیز در قسمت‌های قبل معرفی شدند. در این قسمت به چگونگی بررسی رابطهٔ متغیرهای مقوله‌ای (اسمی یا ترتیبی) با یکدیگر خواهیم پرداخت. این بحث را با معرفی جدول توافقی که ابزاری «توصیفی» برای نمایش توزیع همزمان دو متغیر مقوله‌ای است، آغاز می‌کنیم و سپس به معرفی معیارها و آزمون‌های رابطهٔ این گونه متغیرها می‌پردازیم.

در فصل پیش آموختیم که جدول توزیع فراوانی به منظور نمایش توزیع «یک» متغیر مقوله‌ای به کار می‌رود. این جدول اطلاعاتی درباره فراوانی، درصد و درصد تجمعی هر مقوله از متغیر به دست می‌دهد. جدول توافقی تعمیمی از جدول توزیع فراوانی است و به منظور نمایش توزیع همزمان (توأم) دو متغیر استفاده می‌شود. برای آگاهی از چگونگی تشکیل جدول توافقی برای دو متغیر مقوله‌ای از مجموعه دادهٔ زیر کمک می‌گیریم.

از ۱۰ پاسخگو به عنوان یک نمونه ۱۰تایی پرسیده شد که چه نوع موسیقی را بیشتر دوست دارید؟ باکلام (۱)، بی‌کلام (۲) یا هردو (۳). همچنین از آنان پرسیده شد که از چه ضرباهنگ موسیقی بیشتر خوشتان می‌آید؟ تند (۱) یا ملایم (۲). داده‌های این دو پرسش در جدول زیر ارائه شده است.

پاسخ‌های ۱۰ پاسخگو به دو پرسش درباره ضرباهنگ و نوع موسیقی

نمونه	پاسخگو	ضرباهنگ موسیقی (X)	نوع موسیقی (Y)
۱	پژمان	تند (۱)	باکلام (۱)
۲	سامان	ملایم (۲)	باکلام (۱)
۳	کامران	تند (۱)	بی‌کلام (۲)
۴	سیاوش	تند (۱)	باکلام (۱)
۵	مهرزاد	تند (۱)	هردو (۳)
۶	فرهاد	ملایم (۲)	هردو (۳)
۷	پوپک	تند (۱)	باکلام (۱)
۸	پویا	تند (۱)	بی‌کلام (۲)
۹	پریا	ملایم (۲)	هردو (۳)
۱۰	سیروس	ملایم (۲)	هردو (۳)

این دو پرسش، دو متغیر مقوله‌ای اسمی را ایجاد می‌کنند که یکی دارای دو مقوله (تند/ملایم) و دیگری دارای سه مقوله (باکلام/ بی‌کلام/ هردو) است. عدد ۱ را به مقولهٔ «تند» و عدد ۲ را به مقولهٔ «ملایم» و اعداد ۱، ۲ و ۳ را نیز به ترتیب به مقوله‌های «باکلام»، «بی‌کلام» و «هردو» نسبت می‌دهیم. در نخستین مرحله از توصیف داده‌های جدول قبل، جدول توزیع فراوانی هر متغیر را به صورت زیر رسم می‌کنیم. توجه کنید که به دلیل اسمی‌بودن دو متغیر، از آوردن درصد تجمعی خودداری شده است.

جدول ۵: توزیع فراوانی تک‌تک متغیرهای نوع موسیقی و ضرباهنگ موسیقی

جدول توزیع فراوانی ضرباهنگ موسیقی

مقوله	فراوانی	درصد
تند	۶	۶۰
ملایم	۴	۴۰
جمع	۱۰	۱۰۰

جدول توزیع فراوانی نوع موسیقی

مقوله	فراوانی	درصد
باکلام	۴	۴۰
بی‌کلام	۲	۲۰
هردو	۴	۴۰
جمع	۱۰	۱۰۰

این دو جدول، توصیفی از توزیع «تک‌تک» متغیرها به دست می‌دهند، ولی اغلب پژوهشگران مایلند از توزیع «ترکیبیِ» این دو متغیر نیز اطلاع یابند. برای مثال آنان می‌خواهند بدانند چه تعدادی از پاسخگویان هم به موسیقی تند علاقه‌مندند و هم به موسیقی باکلام؛ یعنی آن دسته از پاسخگویان که گزینهٔ «باکلام» را در پاسخ به پرسش نوع موسیقی مورد علاقه انتخاب کرده‌اند و گزینهٔ «تند» را در پاسخ به ضرباهنگ موسیقی مورد علاقه. دو جدول فوق قادر نیستند چنین اطلاعاتی را در اختیار ما بگذارند، زیرا توزیع هر متغیر را جداگانه ارائه می‌کنند. اکنون به جدول ۶ نگاه کنید. عدد ۳ در تقاطع سطر اول و ستون اول این جدول بیانگر آن است که ۳ نفر از ۱۰ پاسخگو همزمان به موسیقی تند و موسیقی باکلام علاقه‌مندند.

جدول ۶: توزیع فراوانی توأمِ دو متغیر نوع موسیقی و ضرباهنگ موسیقی

نوع موسیقی / ضرباهنگ موسیقی	باکلام (۱)	بی‌کلام (۲)	هردو (۳)
تند (۱)	$\boxed{۳}$	۲	۱
ملایم (۲)	۱	۰	۳

در واقع فراوانی تقاطع هر سطر با هر ستون از این جدول با شمارش آن دسته از پاسخگویانی به دست می‌آید که پاسخی مطابق با آن سطر و ستون داده‌اند. بنابراین، فراوانی ۲ در جدول ۶ نشان می‌دهد که ۲ پاسخگو، به موسیقی تند و بی‌کلام علاقه دارند (کامران و پویا) یا تنها ۱ پاسخگوست که به هردو نوعِ موسیقی باکلام و بی‌کلام علاقه دارد و از سوی دیگر به موسیقی تند نیز علاقه‌مند است (مهرزاد). همچنین هیچ فردی در میان این ۱۰ پاسخگو وجود ندارد که از یک سو به موسیقی ملایم و از سوی دیگر به موسیقی بی‌کلام علاقه داشته باشد.

جدول ۶ را جدول توافقیِ دو متغیرِ ضرباهنگ و نوع موسیقی مورد علاقه می‌نامند. این جدول توافقی دو بعدی است، زیرا توزیع توأم «دو» متغیر را نمایش می‌دهد. گاهی به متغیر ضرباهنگ موسیقی، متغیر «سطری» جدول و به متغیر نوع موسیقی، متغیر «ستونی» آن گفته می‌شود، زیرا مقوله‌های متغیر ضرباهنگ موسیقی در سطرهای جدول ۶ و مقوله‌های متغیر نوع موسیقی در ستون‌های آن جای گرفته‌اند. این جدول توافقیِ دو بعدی، ۳×۲ نیز نامیده می‌شود، زیرا دارای ۲ سطر و ۳ ستون است. همان طور که دیده می‌شود، جدولی با ۶ خانه از ترکیب دو مقولهٔ متغیر سطری و ۳ مقولهٔ متغیر ستونی به دست می‌آید.

هر خانه از جدول ۶ شامل «فراوانی» ترکیبی از مقوله‌های دو متغیر ضرباهنگ موسیقی و نوع موسیقی با یکدیگر است. جمع فراوانی‌های تمام خانه‌ها نیز برابر با تعداد کل پاسخگویان (اندازه نمونه) یعنی ۱۰ است. جمع فراوانی‌های هر سطر و هر ستون می‌تواند به توزیع تک‌تک متغیرها یعنی همان نتایج جدول ۵ منجر شود (خانه‌های خاکستری جدول زیر). برای مثال، جمع فراوانی سطر اول برابر با ۶ است و از طرفی خانه‌های این سطر در ویژگیِ علاقه‌مندی به موسیقی تند مشترک هستند، پس جمع فراوانی‌های خانه‌های سطر اول برابر با تعداد پاسخگویانی است که به موسیقی تند علاقه‌مندند. عدد ۶ حاصل از جدول توافقی درست با اطلاعات جدول ۵ سازگار است (به همین ترتیب عدد ۴).

ضرباهنگ موسیقی / نوع موسیقی	باکلام (۱)	بی‌کلام (۲)	هردو (۳)	جمع
تند (۱)	۳	۲	۱	۶
ملایم (۲)	۱	۰	۳	۴
جمع	۴	۲	۴	۱۰

توزیع فراوانی ضرباهنگ موسیقی

توزیع فراوانی نوع موسیقی

تعداد نمونه

سطر و ستون مربوط به جمع فراوانی‌ها را که به جدول ۶ اضافه شده است، یعنی همان خانه‌های خاکستری جدول قبل، توزیع فراوانی حاشیه‌ای جدول توافقی می‌نامند. استفاده از واژهٔ «حاشیه‌ای» برای این نامگذاری از آن روست که این سطر و ستون در کناره‌های جدول توافقی قرار دارند.

این امکان وجود دارد که به جای فراوانی هر خانه، درصد آن در جدول توافقی ارائه شود. سه نوع درصد برای هر خانه از یک جدول توافقی دو بعدی بر اساس فراوانی‌های آن قابل محاسبه است. درصد «کل» هر خانه از تقسیم فراوانی هر خانه بر مجموع «کل» فراوانی‌های جدول، یعنی همان اندازه نمونه به دست می‌آید. جدول زیر درصدهای کل را برای تک‌تک خانه‌ها نشان می‌دهد. درصد «کل»، سهم هر خانهٔ جدول را از کل نمونه به دست می‌دهد. برای مثال ۲۰ درصد بیان می‌کند که ۲۰ درصد از کل نمونه را پاسخگویانی تشکیل می‌دهند که هم به موسیقی تند علاقه‌مندند و هم به موسیقی بی‌کلام. توجه کنید که جمع درصدهای هر ۶ خانه برابر با ۱۰۰ می‌شود.

نوع موسیقی / ضرباهنگ موسیقی	باکلام (۱)	بی‌کلام (۲)	هردو (۳)	جمع
تند (۱)	$\frac{۳}{۱۰}\times ۱۰۰=۳۰\%$	$\frac{۲}{۱۰}\times ۱۰۰=۲۰\%$	$\frac{۱}{۱۰}\times ۱۰۰=۱۰\%$	٪۶۰
ملایم (۲)	$\frac{۱}{۱۰}\times ۱۰۰=۱۰\%$	۰	$\frac{۳}{۱۰}\times ۱۰۰=۳۰\%$	٪۴۰
جمع	٪۴۰	٪۲۰	٪۴۰	٪۱۰۰

در جدول فوق نیز می‌توان درصدهای کل یک سطر یا ستون را جمع کرد و به درصدهای حاشیه‌ای سطری یا ستونی دست یافت. برای مثال، جمع درصدهای سطر اول (۳۰، ۲۰ و ۱۰) به رقم ۶۰ می‌رسد. این بدان معناست که ۶۰ درصد از کل پاسخگویان به موسیقی تند علاقه‌مندند. درصدهای حاشیه‌ای نیز به همان توزیع تک‌تک متغیرها یعنی جدول ۵ می‌انجامند.

دو نوع دیگر از درصدهایی که برای خانه‌های جدول توافقی محاسبه می‌شوند، درصدهای سطری و ستونی هستند. مبنای درصدهای ستونی، جمع فراوانی هر ستون است، در حالی که مبنای درصدهای سطری، جمع درصدهای هر سطر است. جدول‌های پیش رو درصدهای سطری و ستونی مربوط به جدول ۶ را نمایش می‌دهند. برای مثال، ۵۰ درصد

در جدول درصدهای «سطری»، از تقسیم عدد ۳ بر جمع فراوانی سطر اول جدول ۶ یعنی ۶، به صورت $\frac{3}{6}\times 100 = 50\%$ به دست آمده است. همچنین عدد ۲۵ درصد در جدول درصدهای «ستونی»، از تقسیم عدد ۱ بر جمع فراوانی ستون سوم جدول ۶ یعنی ۴، به صورت $\frac{1}{4}\times 1004 = 25\%$ به دست آمده است. توجه کنید که در جدول درصدهای ستونی، جمع درصدهای هر ستون برابر با ۱۰۰ و در جدول درصدهای سطری، جمع درصدهای هر سطر برابر با ۱۰۰ است (به قسمت خاکستری جدول‌های زیر نگاه کنید).

جدول درصدهای ستونی جدول ۶

ضرباهنگ / نوع	باکلام	بی‌کلام	هردو
تند	٪۷۵	٪۱۰۰	٪۲۵
ملایم	٪۲۵	۰	٪۷۵
جمع	٪۱۰۰	٪۱۰۰	٪۱۰۰

جدول درصدهای سطری جدول ۶

ضرباهنگ / نوع	باکلام	بی‌کلام	هردو	جمع
تند	٪۵۰	٪۳۳	٪۱۷	٪۱۰۰
ملایم	٪۲۵	۰	٪۷۵	٪۱۰۰

تفسیر درصدهای ستونی کاملاً متفاوت با درصدهای سطری است. عدد ٪۷۵ در ستون «باکلام» جدول درصدهای ستونی فوق نشان می‌دهد که ۷۵ درصد از «علاقه‌مندان به موسیقی باکلام» را علاقه‌مندان به موسیقی تند تشکیل می‌دهند، در حالی که عدد ٪۵۰ در سطر «تند» جدول درصدهای سطری نشان می‌دهد که ۵۰ درصد از «علاقه‌مندان به موسیقی تند» را علاقه‌مندان به موسیقی باکلام تشکیل می‌دهند. در واقع آنچه چگونگی تفسیر درصدها را تعیین می‌کند، مبنای محاسبه درصدهاست.

اگر مانند درصدهای سطری، جمع فراوانی‌های سطری مبنای محاسبه باشد، درصدهای خانه‌های هر سطر در همان سطر تفسیر می‌شوند و می‌توان این درصدها را بین سطرها مقایسه کرد. برای مثال می‌توان گفت، علاقه‌مندانِ موسیقی تند بیشتر از علاقه‌مندانِ موسیقی ملایم، به موسیقی باکلام تمایل دارند، زیرا ۵۰ درصد بیشتر از ۲۵ درصد است. به همین ترتیب، اگر مانند درصدهای ستونی، جمع فراوانی ستون‌ها مبنای محاسبه باشد، درصدهای خانه‌های هر ستون در همان ستون تفسیر می‌شوند و می‌توان این درصدها را بین ستون‌ها مقایسه کرد. پس می‌توان ادعا کرد که علاقه‌مندانِ موسیقی بی‌کلام بیش از علاقه‌مندانِ موسیقی باکلام، به موسیقی تند تمایل نشان می‌دهند، زیرا ۱۰۰ درصد بیشتر از ۷۵ درصد است.

۱ـ۸ـ آزمون نیکویی برازش

تاکنون ابزاری با عنوان جدول توافقی برای «توصیف» همزمان داده‌های دو متغیر مقوله‌ای معرفی شد. اکنون می‌خواهیم به استنباط آماری درباره رابطهٔ دو متغیر مقوله‌ای بر اساس داده‌های نمونه در جامعه آماری دست بزنیم. به یاد داشته باشید معمولاً هنگام صحبت از متغیرهای مقوله‌ای به جای اصطلاح رابطهٔ دو متغیر از اصطلاح «پیوند» دو متغیر استفاده می‌شود.

برای آغاز بحث، فرض کنید پژوهشگری از ۵۰۰ فرد نمونه می‌پرسد «تا چه حد از اخبار سایت‌های خبری استفاده می‌کنید؟» و آنان در پاسخ، به یکی از گزینه‌های «کم، متوسط یا زیاد» اشاره می‌کنند. همچنین او از آنان می‌پرسد که «تا چه حد به اخبار سایت‌های خبری اعتماد دارید؟» و آنان در پاسخ این پرسش نیز به یکی از گزینه‌های «کم، متوسط یا زیاد» اشاره می‌کنند. اطلاعات این دو پرسش در جدول ۷ که جدولی ۳×۳ است، نمایش داده شده است (اعداد داخل پرانتز درصدهای ستونی هستند).

این پژوهشگر بر این اعتقاد است که میزان استفادهٔ افراد از اخبار سایت‌های خبری با میزان اعتماد آنان به این سایت‌ها در ارتباط (پیوند) است. به عبارت دیگر، افرادی که اعتماد بیشتری به سایت‌های خبری دارند، بیشتر از اخبار آنها استفاده می‌کنند. مقایسهٔ درصدهای ستونی جدول ۷ نیز این ادعا را تأیید می‌کند، زیرا ۵۶ درصد از افرادی که در حد «زیاد» به سایت‌های خبری اعتماد دارند، در حد «زیاد» نیز از آنها استفاده می‌کنند، در حالی که این رقم برای افراد با اعتماد «متوسط» برابر با ۲۰ درصد و برای افراد با اعتماد «کم» برابر با ۸ درصد است. اکنون این پرسش مطرح است که آیا تفاوت درصدهای ستونی مشاهده شده در این جدول بر اساس نمونه ۵۰۰ تایی، نشانه‌ای از پیوند این دو متغیر در جامعه است؟ پاسخ این پرسش در آزمون فرضیه صفرِ «دو متغیر مستقل هستند» در مقابل فرضیه جانشینِ «دو متغیر وابسته‌اند» نهفته است.

آزمون نیکویی برازش آزمونی ناپارامتری است که استقلال دو متغیر مقوله‌ای را می‌آزماید. این آزمون فراوانی‌های مشاهده‌شده در یک جدول توافقی را با فراوانی‌هایی مقایسه می‌کند که در صورتِ استقلال دو متغیر انتظار دیدن آنها را داریم. اگر فراوانی‌های

مشاهده‌شده با فراوانی‌های مورد انتظار اختلاف چشمگیری داشته باشد، آنگاه نتیجه خواهیم گرفت که دو متغیر به یکدیگر وابسته‌اند[1].

جدول ۷: جدول توافقی میزان اعتماد به اخبار سایت‌های خبری و میزان استفاده از آنها

اعتماد / استفاده	کم	متوسط	زیاد	جمع
کم	۸۵ (٪۷۱)	۵۰ (٪۲۵)	۲۰ (٪۱۱)	۱۵۵
متوسط	۲۵ (٪۲۱)	۱۱۰ (٪۵۵)	۶۰ (٪۳۳)	۱۹۵
زیاد	۱۰ (٪۸)	۴۰ (٪۲۰)	۱۰۰ (٪۵۶)	۱۵۰
جمع	۱۲۰	۲۰۰	۱۸۰	۵۰۰

فراوانی مورد انتظار هر خانهٔ جدول از ضرب فراوانی ستونی آن خانه در فراوانی سطری آن و تقسیم نتیجهٔ این حاصل‌ضرب بر اندازه نمونه (جمع فراوانی‌های کل جدول) به دست می‌آید. جدول زیر فراوانی‌های مورد انتظار خانه‌های جدول ۷ را نشان می‌دهد.

جدول فراوانی‌های مورد انتظار مربوط به جدول ۷

اعتماد / استفاده	کم	متوسط	زیاد	جمع
کم	۳۷/۲	۶۲	۵۵/۸	۱۵۵
متوسط	۴۶/۸	۷۸	۷۰/۲	۱۹۵
زیاد	۳۶	۶۰	۵۴	۱۵۰
جمع	۱۲۰	۲۰۰	۱۸۰	۵۰۰

برای مثال عدد ۶۰ در جدول فوق، فراوانیِ مورد انتظار خانهٔ مربوط به عدد ۴۰ در جدول ۷ است. عدد ۴۰ در تقاطع سطر سوم و ستون دوم جدول ۷ قرار دارد، پس برای محاسبه فراوانی مورد انتظار آن به حاصل‌ضرب جمع سطر سوم (۱۵۰) و جمع ستون دوم (۲۰۰) نیاز است تا نتیجهٔ آنها بر اندازه نمونه (۵۰۰) تقسیم شود یعنی $\frac{۱۵۰\times۲۰۰}{۵۰۰}=۶۰$. به همین ترتیب فراوانی مورد انتظار خانهٔ مربوط به عدد ۱۰ که در تقاطع سطر سوم و ستون اول جدول ۷ واقع است، به صورت $\frac{۱۵۰\times۱۲۰}{۵۰۰}=۳۶$ به دست می‌آید.

۱. اگرچه در این قسمت از آزمون نیکویی برازش برای بررسی استقلال دو متغیر استفاده می‌شود، ولی این آزمون کاربرد عام‌تری دارد. به زبان آماری، از این آزمون برای بررسی برازش توزیع خاصی بر داده‌ها استفاده می‌شود.

از این پس از n_{ij} برای نمایش فراوانی خانهٔ سطر iام و ستون jام جدول توافقی، از $n_{i.}$ برای نمایش جمع فراوانی‌های خانه‌های سطر iام، از $n_{.j}$ برای نمایش جمع فراوانی‌های خانه‌های ستون jام این جدول و از n برای نمایش اندازه نمونه آن (جمع فراوانی‌های کل خانه‌های جدول) استفاده می‌کنیم. بنابراین برای جدول ۷ داریم:

اعتماد (X) / استفاده (Y)	کم	متوسط	زیاد	فراوانی حاشیه‌ای ستونی
کم	$n_{۱۱}=۸۵$	$n_{۱۲}=۵۰$	$n_{۱۳}=۲۰$	$n_{۱.}=۱۵۵$
متوسط	$n_{۲۱}=۲۵$	$n_{۲۲}=۱۱۰$	$n_{۲۳}=۶۰$	$n_{۲.}=۱۹۵$
زیاد	$n_{۳۱}=۱۰$	$n_{۳۲}=۴۰$	$n_{۳۳}=۱۰۰$	$n_{۳.}=۱۵۰$
فراوانی حاشیه‌ای سطری	$n_{.۱}=۱۲۰$	$n_{.۲}=۲۰۰$	$n_{.۳}=۱۸۰$	$n=۵۰۰$

با این نمادگذاری، فراوانی مورد انتظار متناظر با n_{ij} به صورت $e_{ij}=\frac{n_{i.}n_{.j}}{n}$ به دست می‌آید، پس برای مثال $e_{۳۲}=\frac{n_{۳.}n_{.۲}}{n}=\frac{۱۵۰\times ۲۰۰}{۵۰۰}=۶۰$ و $e_{۳۱}=\frac{n_{۳.}n_{.۱}}{n}=\frac{۱۵۰\times ۱۲۰}{۵۰۰}=۳۶$. حال که فراوانی‌های مورد انتظار تک‌تک خانه‌ها محاسبه شدند، باید از طریق آمارهٔ آزمون با فراوانی‌های مشاهده‌شده مقایسه شوند. آمارهٔ آزمونِ نیکویی برازش به صورت زیر بیان می‌شود:

$$\chi^{۲}=\sum_{i,j}\frac{(n_{ij}-e_{ij})^{۲}}{e_{ij}}$$

این آماره می‌گوید، ابتدا باید تفاضل فراوانی مورد انتظار هر خانه را از فراوانی مشاهده‌شدهٔ همان خانه محاسبه و نتیجه را به توان ۲ رساند و سپس بر فراوانی مورد انتظار آن خانه تقسیم کرد، یعنی $\frac{(n_{ij}-e_{ij})^{۲}}{e_{ij}}$. در این صورت، آمارهٔ آزمون برابر با جمع مقادیر محاسبه‌شده برای تک‌تک خانه‌هاست؛ برای مثال، برای خانهٔ سطر سوم و ستون دوم $\frac{(n_{۳۲}-e_{۳۲})^{۲}}{e_{۳۲}}=\frac{(۴۰-۶۰)^{۲}}{۶۰}=۶/۶۷$. این کمّیت‌ها برای فراوانی‌های جدول ۷ که پیشتر فراوانی‌های مورد انتظار خانه‌های آن نیز محاسبه شد، در جدول زیر مشاهده می‌شود:

اعتماد / استفاده	کم	متوسط	زیاد
کم	۶۱/۴۲	۲/۳۲	۲۲/۹۷
متوسط	۱۵/۱۰	۱۳/۱۳	۱/۴۸
زیاد	۷۸/۱۸	$\boxed{۶/۶۷}$	۳۹/۱۹

بر اساس اطلاعات این جدول، آزمون فرضیه صفرِ «میزان استفاده از اخبار سایت‌های خبری، مستقل از میزان اعتماد به آنهاست» در مقابل فرضیه جانشین «میزان استفاده از اخبار سایت‌های خبری با میزان اعتماد به آنها پیوند دارد»، آمارهٔ زیر را برای جدول ۷ به دست می‌دهد:

$$\chi^{۲} = ۶۱/۴۲ + ۲/۳۲ + ۲۲/۹۷ + \cdots + ۶/۶۷ + ۳۹/۱۹ = ۱۷۶/۱۱$$

هرچه مقدار این آماره بزرگ‌تر باشد، بین فراوانی مورد انتظار و فراوانی مشاهده‌شده اختلاف بیشتری وجود دارد؛ به طوری که می‌توان فرضیه صفر را رد کرد. نقطهٔ بحرانی آزمون برای جدولی که دارای تعداد r سطر و تعداد c ستون است، از جدول چندک‌های توزیع $\chi^{۲}$ به ازای α مورد نظر و $df = (r-۱)(c-۱)$ به دست می‌آید. اگر مقدار آماره از نقطهٔ بحرانی بزرگ‌تر باشد، فرضیه صفر به نفع فرضیه جانشین رد می‌شود.

نقطه بحرانی برای جدول ۷ با تعداد $r = ۳$ سطر و $c = ۳$ ستون به ازای $\alpha = ۰/۰۵$ برابر با $\chi^{۲}_{۴,۰/۰۵} = ۹/۴۸۸$ است. بنابراین فرضیه صفر رد می‌شود و نتیجه خواهیم گرفت بین میزان استفاده از سایت‌های خبری و میزان اعتماد به آنها پیوند وجود دارد.

> آزمون نیکویی برازش دارای دو پیش‌فرض است: نخست آنکه نباید فراوانی‌های مورد انتظار هیچ خانه‌ای کمتر از ۱ باشند و دوم آنکه حداکثر ۲۰ درصد از فراوانی‌های مورد انتظار می‌توانند کمتر از ۵ باشند.

آزمون نیکویی برازش تنها فرضیه صفر «استقلال دو متغیر مقوله‌ای» را می‌آزماید و در صورت رد شدن این فرضیه یعنی اثبات وجود پیوندِ بین دو متغیر، اطلاعی درباره شدت این پیوند به دست نمی‌دهد. بنابراین به معیارهایی مانند ضریب همبستگی پیرسون نیاز است که قادر باشند شدت پیوند دو متغیر مقوله‌ای را اندازه بگیرند. در ادامه به بررسی این معیارها خواهیم پرداخت[1].

۲ـ۸ـ معیارهای اندازه‌گیری پیوند

در بحث رگرسیون با متغیر مستقل و متغیر وابسته (پاسخ) آشنا شدید. پیوند متغیر مستقل با متغیر وابسته به شکلی است که تغییرات متغیر وابسته به تغییرات متغیر مستقل باز

---

۱. فرض بر این است که آزمون نیکویی برازش پیوند معناداری را بین دو متغیر نشان داده است و این معیارها تنها به منظور اندازه‌گیری شدت این پیوند به کار می‌روند و در واقع معیارهای «تعقیبی» آزمون نیکویی برازش هستند. با وجود این، هر یک از این معیارها آزمون آماری خاص خود را نیز دارد که می‌تواند جدا از آزمون نیکویی برازش برای بررسی پیوند دو متغیر مقوله‌ای استفاده شود، ولی از معرفی این آزمون‌ها صرفنظر شده است.

می‌گردد. این تمایز بین دو متغیر مقوله‌ای نیز می‌تواند مطرح شود. بـرای مثـال، اگـر معتقـد باشیم تفاوت در میزان استفاده از شبکه‌های خبری به تفاوت در میزان اعتماد به این شبکه‌ها باز می‌گردد، پس میزان استفاده را متغیر وابسته و میزان اعتماد را متغیر مسـتقل بـه حسـاب آورده‌ایم. هنگام بررسی پیوند بین متغیرهای مقوله‌ای، توجه به تمایز بـین متغیـر مسـتقل و متغیر وابسته می‌تواند به سنجش دقیق‌تری از شدت پیوند بینجامد. از ایـن نظـر، معیارهـای اندازه‌گیری پیوند به متقارن و نامتقارن تقسیم می‌شوند. معیارهای متقـارن، تمـایزی بـین دو متغیر مقوله‌ای قایل نمی‌شوند و صرفاً شدت پیوند آنها را انـدازه مـی‌گیرنـد، در حـالی کـه معیارهای نامتقارن در محاسبه شدت پیوند می‌توانند یکی از متغیرها را مستقل و دیگـری را وابسته قلمداد کنند.

معیارهای اندازه‌گیری پیوند دو متغیر مقوله‌ای، به معیارهای اسـمی و معیارهـای ترتیبـی نیز تقسیم می‌شوند. معیارهای اسمی فرض می‌کنند شدت پیوند دو متغیر مقوله‌ای اسـمی را اندازه می‌گیرند، حتی اگر هر دو یا یکی از آنهـا ترتیبـی باشـد. از سـوی دیگـر، معیارهـای ترتیبی فرض می‌کنند شدت پیوند دو متغیر مقوله‌ای ترتیبی را اندازه می‌گیرند، حتی اگر هر دو یا یکی از آنها اسمی باشد. بنـابراین، تنهـا هنگـامی مجـاز هسـتیم از معیارهـای ترتیبـی استفاده کنیم که هر دو متغیر ترتیبی باشند، در غیر ایـن صـورت بایـد از معیارهـای اسـمی استفاده کنیم.

۱ـ۲ـ۸ـ معیارهای وی کرامر، فی و ضریب تتراکوریک

معیار V کرامر[1] معیاری اسمی و متقارن است که بـر اسـاس مقـدار آمـارهٔ آزمـون نیکـویی برازش محاسبه می‌شود. فرض کنید دو متغیر X و Y جدولی توافقی بـا r سـطر و c سـتون تشکیل داده‌اند و فرضیه صفر استقلال این دو متغیر با آزمون نیکویی برازش رد شده اسـت و می‌دانیم پیوندی بین این دو متغیر وجود دارد. اگر مقدار آمارهٔ ایـن آزمـون برابـر بـا $\chi^۲$، تعداد نمونه (جمع فراوانی‌های کل جدول) برابر بـا n و کوچـک‌تـرین مقـدار بـین $r-۱$ و $c-۱$ برابر با m باشد، معیار V کرامر به صورت زیر محاسبه می‌شود:

$$V=\sqrt{\frac{\chi^{۲}}{nm}}$$

1. Cramér

مقدار V کرامر می‌تواند از ۰ تا ۱ باشد. هرچه این مقدار به ۱ نزدیک‌تر باشد، شدت پیوند دو متغیر مقوله‌ای قوی‌تر است.

مقدار آمارهٔ آزمون نیکویی برازش برای جدول ۷ با ۳ سطر و ۳ ستون برابر با $\chi^{۲} = ۱۷۶/۱۱$ بود و این آزمون نشان داد که بین متغیر میزان اعتماد به شبکه‌ها و میزان استفاده از آنها پیوند وجود دارد. با توجه به اینکه $n = ۵۰۰$ و $m = ۲$ است، شدت این پیوند بر اساس معیار V کرامر به صورت زیر اندازه‌گیری می‌شود که بر شدت پیوندی در حد متوسط دلالت دارد.

$$V = \sqrt{\frac{۱۷۶/۱۱}{۵۰۰ \times ۲}} = ۰/۴۱۹$$

معیار V کرامر تفسیر جالبی دارد. به یاد بیاورید که آمارهٔ آزمون نیکویی برازش $\chi^{۲}$ اختلافی را اندازه‌گیری می‌کند که بین فراوانی‌های مشاهده‌شده در جدول توافقی با فراوانی‌های مورد انتظار تحت استقلال دو متغیر وجود دارد. از این رو، هر چه این اختلاف بیشتر باشد، با اطمینان بیشتری می‌توان ادعا کرد دو متغیر از هم مستقل نیستند و پیوندی قوی با یکدیگر دارند. حاصل‌ضرب nm که در مخرج کسر معیار V کرامر قرار دارد، برابر با «حداکثر» مقداری است که آمارهٔ آزمون $\chi^{۲}$ می‌تواند از یک جدول توافقی $r \times c$ با نمونه n تایی داشته باشد. به عبارت دیگر مقدار $\chi^{۲}$ نمی‌تواند در یک جدول توافقی $r \times c$ با نمونه n تایی از nm بیشتر شود. پس هر چه $\chi^{۲}$ به nm نزدیک‌تر باشد، اختلاف فراوانی مشاهده‌شده از فراوانی مورد انتظار نیز بیشتر است و دو متغیر دارای پیوند قوی‌تری هستند. بنابراین $\frac{\chi^{۲}}{nm}$ می‌تواند میزان نزدیکی مقدار $\chi^{۲}$ را به حداکثر مقداری اندازه‌گیری کند که می‌توانست داشته باشد.

برای مثال، مقدار $\frac{۱۷۶/۱۱}{۵۰۰\times۲} = ۱۷/۶\%$ در جدول ۷ بیان می‌کند که اختلاف فراوانی مشاهده‌شده و فراوانی مورد انتظار این جدول که از طریق آمارهٔ $\chi^{۲} = ۱۷۶/۱۱$ اندازه‌گیری شده است، ۱۷/۶ درصد به حداکثر مقدار این اختلاف نزدیک است. به عبارت دیگر، این دو متغیر ۸۲/۴ درصد با پیوندی «کاملاً قوی» فاصله دارند. در واقع، اگر مقدار $\chi^{۲}$ درست برابر با nm باشد، یعنی $\frac{\chi^{۲}}{nm} = ۱۰۰\%$ یا $V = ۱$، آنگاه «بیشترین» اختلاف بین فراوانی مشاهده‌شده و فراوانی مورد انتظار رخ‌داده است و دو متغیر «کاملاً» به هم وابسته‌اند، به

طوری که با دانستن مقدار یکی، به‌درستی می‌توان مقدار دیگری را حدس زد. ضعف $\frac{\chi^2}{nm}$ در کوچک جلوه‌دادن شدت پیوند است که تا حدی با جذر گرفتن جبران می‌شود، بنابراین معیار V کرامر به صورت $V=\sqrt{\frac{\chi^2}{nm}}$ محاسبه می‌شود.

در شرایطی که یکی از متغیرهای مقوله‌ای جدول توافقی تنها دارای دو مقوله باشد، $m=1$ خواهد بود و معیار V کرامر به صورت زیر تبدیل می‌شود که به آن معیار فی[1] می‌گویند.

$$\varphi=\sqrt{\frac{\chi^2}{n}}$$

با وجود این، اغلب از فی برای تحلیل جدول‌های توافقی ۲×۲ استفاده می‌شود. برای این منظور، جدول توافقی زیر را با فراوانی‌های a، b، c و d در نظر بگیرید، به طوری که $n=a+b+c+d$.

جمع	۱	۰	X / Y
a+b	b	a	۰
c+d	d	c	۱
n	b+d	a+c	جمع

شیوه ساده‌تر محاسبه فی برای جدول توافقی ۲×۲ فوق به صورت زیر است:

$$\varphi=\frac{bc-ad}{\sqrt{(a+b)(c+d)(a+c)(b+d)}}$$

در برخی موارد، متغیرهای دوحالتی جدول فوق از دسته‌بندی متغیرهای پیوسته به دو وضعیت ایجاد شده‌اند (ضریب همبستگی دو رشته‌ای را به یاد بیاورید). برای مثال، فرض کنید افراد بر اساس درآمد به دو وضعیتِ زیر خط فقر و بالای خط فقر و بر اساس ساعات روزانه کار به دو وضعیت عادی (حداکثر ۸ ساعت در روز) و پرکار (بیش از ۸ ساعت در روز) دسته‌بندی شده‌اند. پس دو متغیر مقوله‌ای دوحالتی وجود دارد که مقوله‌های هر یک برگرفته از دسته‌بندی یک متغیر پیوسته است. بار دیگر می‌توان جدول توافقی ۲×۲ فوق را برای این دو متغیر تشکیل داد، ولی به جای استفاده از ضریب فی، از ضریب همبستگی

1. Phi

تتراکوریک به صورت زیر برای بررسی پیوند این دو متغیر مقوله‌ای استفاده می‌شود (Cos نشاندهنده تابع مثلثاتی کسینوس است):

$$r_{tet} = \mathrm{Cos}\left(\frac{۱۸۰}{۱+\sqrt{\frac{bc}{ad}}}\right)$$

۲ـ۲ـ۸ـ معیار لاندا

هنگامی که بین دو متغیر پیوند وجود دارد، از وضعیت یکی می‌توان وضعیت دیگری را برای یک فرد پیش‌بینی کرد. برای مثال اگر بین میزان اعتماد به سایت‌های خبری و میزان استفاده از آنها پیوند وجود داشته باشد، به طوری که با افزایش میزان اعتماد، میزان استفاده نیز افزایش یابد، آنگاه با دانستن اینکه فردی اعتماد بالایی به سایت‌های خبری دارد، پیش‌بینی می‌کنیم که او از چنین سایت‌هایی زیاد استفاده می‌کند.

این قابلیت راهکاری را برای اندازه‌گیری شدت پیوند دو متغیر در اختیار می‌گذارد، بویژه اگر به دنبال معیاری نامتقارن برای اندازه‌گیری شدت پیوند باشیم. در این صورت کافی است برای تشخیص اینکه بین متغیر مستقلِ X و متغیر وابستهٔ Y پیوند وجود دارد، خطای پیش‌بینیِ وضعیت متغیر Y را در دو حالت بررسی کنیم: حالتی که پیش‌بینی وضعیت Y به تنهایی بر اساس توزیع متغیر Y صورت می‌گیرد و حالتی که پیش‌بینی وضعیت Y با درنظرگرفتن وضعیت متغیر X صورت می‌گیرد. اگر خطای حالت دوم که از متغیر X در پیش‌بینی استفاده می‌کند، کمتر از خطای حالت اول باشد، منطقی است نتیجه بگیریم بین متغیر X و Y پیوند وجود دارد، زیرا این متغیر توانسته است پیش‌بینی وضعیت متغیر Y را دقیق‌تر کند. همچنین هر چه خطای حالت دوم نسبت به خطای حالت اول کمتر باشد، پیوند دو متغیر نیز قوی‌تر است.

فرض کنید مقدار خطای حالت اول را با $E_۱$ و مقدار خطای حالت دوم را با $E_۲$ نشان دهیم. در این صورت هرچه $E_۱ - E_۲$ بزرگ‌تر باشد، خطای حالت دوم نسبت به حالت اول کمتر و در نتیجه متغیر X توانایی بیشتری در پیش‌بینی دقیق‌تر وضعیت متغیر Y دارد. بنابراین هرچه $E_۱ - E_۲$ بزرگ‌تر باشد، پیوند X و Y نیز قوی‌تر است. در عمل به جای آنکه $E_۱ - E_۲$

به عنوان معیار اندازه‌گیری پیوند به کار رود، از $\frac{E_1-E_2}{E_1}$ برای این منظور استفاده می‌شود، زیرا اولاً بیانگر کاهش نسبی در خطا[1] است و ثانیاً همواره مقداری از ۰ تا ۱ دارد.

برای مثال اگر خطای پیش‌بینی وضعیت متغیر Y به تنهایی $E_1 = \%70$ و خطای پیش‌بینی وضعیت آن با درنظر گرفتن متغیر X برابر با $E_2 = \%25$ باشد، شدت پیوند این دو متغیر برابر با $\frac{E_1-E_2}{E_1} = \frac{70-25}{70} = 0/64$ است. بنابراین، متغیر X، ۶۴ درصد در پیش‌بینی متغیر Y مؤثر است. مقدار ۱ هنگامی به دست می‌آید که X بتواند بدون هیچ خطایی وضعیت Y را پیش‌بینی کند؛ یعنی $E_2 = 0$ و مقدار ۰ هنگامی به دست می‌آید که X اصلاً نتواند در پیش‌بینی دقیق‌تر وضعیت Y کمکی کند، یعنی مقدار خطای پیش‌بینی وضعیت Y با استفاده از X و بدون استفاده از X یکسان باشد ($E_1 = E_2$).

لاندای گودمن و کروسکال[2] از معیارهای اسمی نامتقارنی است که از منطق کاهش نسبی در خطا (PRE) برای اندازه‌گیری پیوند دو متغیر مقوله‌ای استفاده می‌کند. فرض کنید به دنبال تعیین شدت پیوند دو متغیر X و Y هستید، به طوری که متغیر X متغیر مستقل و متغیر Y متغیر وابسته است. اگر مقوله‌های متغیر X ستون‌های جدول توافقی و مقوله‌های متغیر Y سطرهای آن را تشکیل دهند، برای محاسبه معیار لاندا ابتدا باید سه کمّیت را بر اساس فراوانی‌های خانه‌های این جدول توافقی به دست آورید: ۱. بزرگ‌ترین فراوانی حاشیه‌ای ستونی یعنی $\max(n_{i.})$، ۲. مجموع بزرگ‌ترین فراوانی ستون‌های جدول یعنی $S_{max} = \sum_j \max(n_{ij})$ و ۳. اندازه نمونه یا همان جمع کل فراوانی‌های خانه‌های جدول (n). در این صورت مقدار لاندا از رابطهٔ زیر به دست می‌آید. توجه کنید که نماد Y|X بیان می‌کند که متغیر X به عنوان متغیر مستقل و متغیر Y به عنوان متغیر وابسته در محاسبه لاندا در نظر گرفته شده است.

$$\lambda_{Y|X} = \frac{S_{max} - \max(n_{i.})}{n - \max(n_{i.})}$$

اکنون معیار لاندا را برای جدول ۷ محاسبه می‌کنیم که فراوانی‌های آن در جدول توافقی مقابل نیز مشاهده می‌شود. بزرگ‌ترین فراوانی حاشیه‌ای ستونی برابر با $\max(n_{i.}) = 195$،

1. Proportional Reduction in Error (PRE)
2. Goodman and Kruskal’s Lambda

بزرگ‌ترین فراوانی خانه‌های ستون اول برابر با ۸۵، بزرگ‌ترین فراوانی خانه‌های ستون دوم برابر با ۱۱۰ و بزرگ‌ترین فراوانی خانه‌های ستون سوم برابر با ۱۰۰ است، پس $S_{max}=۸۵+۱۱۰+۱۰۰=۲۹۵$. اندازه نمونه یا جمع کل فراوانی‌های خانه‌های جدول نیز $n=۵۰۰$ است، پس مقدار لاندا عبارت است از:

$$\lambda_{Y|X}=\frac{۲۹۵-۱۹۵}{۵۰۰-۱۹۵}=۰/۳۲۷$$

این مقدار بر شدت پیوند تا اندازه‌ای متوسط دلالت دارد و بیان می‌کند متغیر میزان اعتماد حدود ۳۲/۷ درصد در پیش‌بینی وضعیت متغیر میزان استفاده مؤثر است.

استفاده (Y) \ اعتماد (X)	کم	متوسط	زیاد	فراوانی حاشیه‌ای ستونی
کم	۸۵	۵۰	۲۰	۱۵۵
متوسط	۲۵	۱۱۰	۶۰	۱۹۵
زیاد	۱۰	۴۰	۱۰۰	۱۵۰
فراوانی حاشیه‌ای سطری	۱۲۰	۲۰۰	۱۸۰	۵۰۰

اگر بخواهید $\lambda_{X|Y}$ را با توجه به فراوانی‌های جدول فوق محاسبه کنید، یعنی X را به عنوان متغیر وابسته و Y را به عنوان متغیر مستقل درنظر بگیرید، آنگاه باید بزرگ‌ترین فراوانی حاشیه‌ای سطری (عدد ۲۰۰) و مجموع بزرگ‌ترین فراوانی سطرها را (که بر حسب اتفاق همان اعداد ۸۵، ۱۱۰ و ۱۰۰ هستند) برای محاسبه به صورت زیر استفاده کنید:

$$\lambda_{X|Y}=\frac{۲۹۵-۲۰۰}{۵۰۰-۲۰۰}=۰/۳۱۶$$

این عدد بیان می‌کند متغیر میزان استفاده حدود ۳۱/۶ درصد در پیش‌بینی وضعیت متغیر میزان اعتماد مؤثر است. توجه کنید که نزدیکی مقدار $\lambda_{X|Y}$ به $\lambda_{Y|X}$ در این مثال کاملاً اتفاقی است و به دلیل نزدیکی فراوانی‌های به کار رفته در محاسبات این دو معیار است، ولی لزوماً چنین نزدیکی در جداول توافقی رخ نمی‌دهد.

اگرچه لاندا معیاری نامتقارن برای اندازه‌گیری پیوند دو متغیر X و Y قلمداد می‌شود، ولی می‌توان با محاسبه میانگین $\lambda_{X|Y}$ و $\lambda_{Y|X}$ به صورت زیر به معیار متقارنی برای اندازه‌گیری پیوند دو متغیر X و Y نیز دست یافت.

$$\lambda_{XY} = \frac{\lambda_{X|Y} + \lambda_{Y|X}}{۲}$$

در مثال پیشین، معیار متقارن لاندا برابر با $\frac{۰/۳۱۶+۰/۳۲۷}{۲} = ۰/۳۲۱۵$ به دست می‌آید. بنابراین نتیجه خواهیم گرفت «به طور متوسط» دو متغیر میزان اعتماد و میزان استفاده ۳۲ درصد در پیش‌بینی یکدیگر مؤثر هستند.

۳ـ۲ـ۸ـ معیار گاما

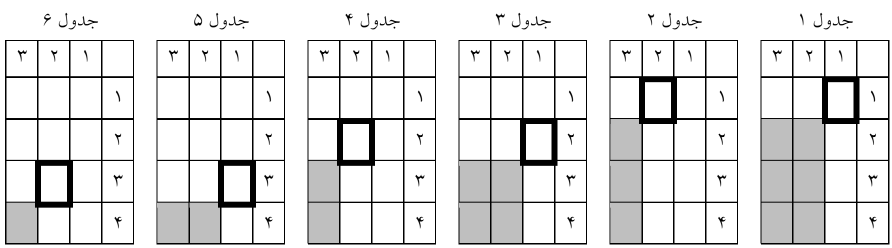
تاکنون، به آن دسته از معیارهای اسمی اندازه‌گیری پیوند اشاره شد که تنها قادرند «شدت» پیوند دو متغیر را به دست دهند. اکنون فرض کنید جدول توافقی از دو متغیر مقوله‌ایِ ترتیبی تشکیل شده است، مانند جدول ۷ که شامل دو متغیر مقوله‌ای میزان اعتماد و میزان استفاده است. معیارهای V کرامر و لاندا شدت پیوند این دو متغیر را به‌خوبی اندازه می‌گیرند و به این ترتیب پژوهشگر درمی‌یابد که این پیوند تا چه حد قوی است. با وجود این، جهت پیوند این دو متغیر را مشخص نمی‌کنند، زیرا V کرامر و لاندا معیارهایی «اسمی» هستند و دو متغیر مقوله‌ای جدول توافقی را متغیرهای اسمی قلمداد می‌کنند، حتی اگر هر دوی آنها ترتیبی باشند. بنابراین، معیارهای اسمی می‌توانند برای جدول‌های توافقی با دو متغیر ترتیبی ناکارآمد باشند.

پژوهشگرانی که جدولی توافقی با دو متغیر ترتیبی در دست دارند، هم می‌خواهند از شدت پیوند این دو متغیر اطلاع یابند، و هم تمایل دارند بدانند این پیوند می‌تواند افزایشی یا کاهشی باشد؟ برای مثال، آیا با افزایش میزان اعتماد به سایت‌های خبری، میزان استفاده از آنها نیز افزایش می‌یابد؟ بنابراین باید معیارهایی برای جدول‌های توافقی با «دو» متغیر ترتیبی یافت که هم شدت پیوند و هم جهت آن را تعیین کنند. گاما[1] معیار ترتیبی متقارنی است که مانند ضریب کندال بر اساس زوج‌های هماهنگ و ناهماهنگ محاسبه می‌شود.

1. Gamma

فرض کنید دو متغیر ترتیبی، جدول توافقی ۳×۴ تشکیل داده‌اند. بنابراین هم متغیر سطری جدول و هم متغیر ستونی آن ترتیبی هستند. اگر مقوله‌های این دو متغیر از کمترین به بیشترین به سطرها و ستون‌های این جدول اختصاص یافته باشند، آنگاه خانه‌هایی با خانهٔ مفروضی از این جدول هماهنگ هستند که رتبهٔ هر دو مقولهٔ آنها بزرگ‌تر از رتبهٔ هر دو مقولهٔ آن خانه باشد. برای مثال، خانهٔ واقع‌شده در سطر اول و ستون اول جدول ۱ از شکل ۲ را در نظر بگیرید. تمام خانه‌های خاکستری این جدول با این خانه هماهنگ هستند، زیرا این خانه به مقولهٔ اول متغیر سطری و مقولهٔ اول متغیر ستونی تعلق دارد، در حالی که مقولهٔ متغیر سطری خانه‌های خاکستری ۲، ۳ یا ۴ و مقولهٔ متغیر ستونی خانه‌های خاکستری ۲ یا ۳ است. پس مقوله‌های هر دو متغیر خانه‌های خاکستری از رتبهٔ بزرگ‌تری نسبت به مقوله‌های این خانه برخوردارند. به همین ترتیب، دو خانهٔ هماهنگ با خانهٔ واقع‌شده در سطر سوم و ستون اول جدول ۵ از شکل ۲ وجود دارد. رتبهٔ مقولهٔ متغیر سطری این دو خانه ۴ و رتبهٔ مقولهٔ متغیر ستونی آنها نیز ۲ و ۳ است که از رتبهٔ این خانه بزرگ‌تر است.

شکل ۲: خانه‌های هماهنگ در یک جدول توافقی ۳×۴

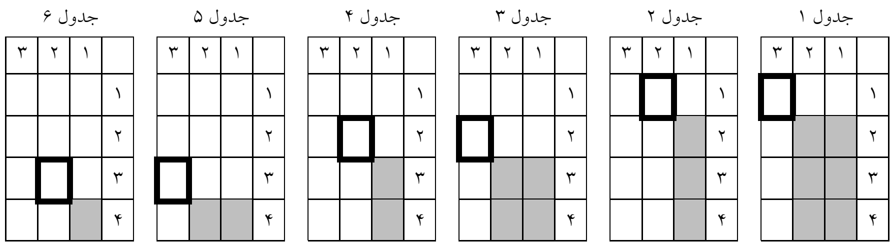


به‌راحتی می‌توان از جدول‌های شکل ۲ به روشی برای تشخیص خانه‌های هماهنگ با یک خانه از جدول توافقی دست یافت. برای یک خانهٔ مفروض از جدول، تمام خانه‌هایی هماهنگ به حساب می‌آیند که هم سطر آنها «پایین‌تر» از سطر این خانه و هم ستون آنها «پس» از ستون این خانه باشد. این همان ویژگی خانه‌هایی است که به صورت خاکستری در شکل ۲ نشان داده شده‌اند. برای مثال، در جدول ۶ از این شکل، تنها یک خانه وجود دارد که هم سطری پایین‌تر و هم ستونی پس از ستون خانهٔ مشخص‌شده دارد. برای آنکه تمام خانه‌های هماهنگ جدول را شناسایی کنید، بهتر است مانند جدول ۱ شکل ۲، از گوشهٔ سمت «راست» جدول شروع کنید. فراموش نکنید که برای استفاده از این روش

ضروری است مقوله‌های دو متغیر ترتیبی از کمترین به بیشترین بـه سـطرها و سـتون‌هـای جدول توافقی اختصاص یافته باشند.

تعیین خانه‌های ناهماهنگ با یک خانهٔ مفروض از جدول توافقی نیز کار ساده‌ای اسـت. خانه‌ای ناهماهنگ با یک خانهٔ مفروض است که اگر رتبهٔ متغیر سطری آن بزرگ‌تر از رتبـهٔ متغیر سطری این خانهٔ مفروض بود، آنگاه رتبهٔ ستونی آن کوچـک‌تـر از رتبـهٔ سـتونی ایـن خانه باشد. در عمل کافی است مانند جدول ۱ از شـکل ۳ از گوشـهٔ سـمت چـپ جـدول شروع کنید. برای یک خانهٔ مفروض از جدول، تمام خانه‌هایی ناهماهنگ به حساب می‌آیند که سطر آنها «پایین‌تر» از سطر این خانه، ولی ستون آنها «پیش» از سـتون ایـن خانـه باشـد. این همان ویژگی خانه‌هایی است که به صورت خاکستری در شکل ۳ نشان داده شده‌اند.

شکل ۳: خانه‌های ناهماهنگ در یک جدول توافقی ۳×۴



اکنون باید تعداد زوج‌های هماهنگ و ناهماهنگ یک جدول توافقی تعیین شود[1]. بـرای یافتن تعداد زوج‌های هماهنگ، ابتدا فراوانی هر خانه را در مجموع فراوانی‌های خانـه‌هـای هماهنگ با آن ضرب کنید و سپس این حاصل‌ضـرب‌هـا را بـا هـم جمـع کنیـد تـا تعـداد زوج‌های هماهنگ (n_c) تعیین شود. همچنین، فراوانی هر خانه را در مجموع فراوانی‌هـای خانه‌های ناهماهنگ با آن ضرب کنید و سپس این حاصل‌ضرب‌ها را با هم جمع ببندیـد تـا تعداد زوج‌های ناهماهنگ (n_d) تعیین شود. اگـر $n_c - n_d$ مثبـت باشـد، تعـداد زوج‌هـای هماهنگ بیشتر از تعداد زوج‌های ناهماهنگ است، پس جهت پیونـد افزایشـی اسـت و بـر عکس، اگر $n_c - n_d$ منفی باشـد، تعـداد زوج‌هـای ناهماهنـگ بیشـتر از تعـداد زوج‌هـای هماهنگ است، پس جهت پیوند کاهشی است. معیار گاما با توجه به این تفاضل به صورت مقابل تعریف می‌شود:

---

۱. هر فرد از یک خانهٔ مفروض، با تک‌تک افراد خانه‌های هماهنگ با این خانه تشکیل زوج هماهنگ مـی‌دهـد. بـه همـین ترتیب هر فرد از یک خانهٔ مفروض با تک‌تک افراد خانه‌های ناهماهنگ با این خانه نیز تشکیل زوج ناهماهنگ می‌دهد.

$$\gamma = \frac{n_c - n_d}{n_c + n_d}$$

اکنون این معیار را برای جدول ۷ محاسبه می‌کنیم که فراوانی‌های آن در جدول زیر مشاهده می‌شود. همان طور که دیده می‌شود، مقوله‌های این دو متغیر از کمترین به بیشترین به سطرها و ستون‌ها اختصاص یافته‌اند تا بتوان به درستی از روش تعیین خانه‌های هماهنگ و ناهماهنگ استفاده کرد. برای مثال، پایین‌ترین مقولهٔ متغیر میزان استفاده (کم) به نخستین سطر، مقولهٔ وسطی به دومین سطر و بالاترین مقوله (زیاد) به سومین سطر جدول اختصاص یافته است. توجه کنید که مثلاً فردی که در خانهٔ سطر اول و ستون دوم قرار دارد، نسبت به فردی که در خانهٔ سطر سوم و ستون سوم قرار دارد، یک زوج هماهنگ تشکیل می‌دهند و مجموع زوج‌های هماهنگ حاصل از این دو خانه ۵۰۰۰=۵۰×۱۰۰ زوج است.

استفاده (Y) \ اعتماد (X)	کم	متوسط	زیاد
کم	۸۵	۵۰	۲۰
متوسط	۲۵	۱۱۰	۶۰
زیاد	۱۰	۴۰	۱۰۰

تعداد زوج‌های هماهنگ و ناهماهنگ کل این جدول با روش پیشگفته به قرار زیر است:

$$n_c = ۸۵\times(۱۱۰+۶۰+۴۰+۱۰۰)+۵۰\times(۶۰+۱۰۰)+۲۵\times(۴۰+۱۰۰)+۱۱۰\times۱۰۰=۴۸۸۵۰$$

$$n_d = ۲۰\times(۱۱۰+۲۵+۴۰+۱۰)+۵۰\times(۲۵+۱۰)+۶۰\times(۴۰+۱۰)+۱۱۰\times۱۰=۹۵۵۰$$

بنابراین:

$$\gamma = \frac{۴۸۸۵۰-۹۵۵۰}{۴۸۸۵۰+۹۵۵۰} = ۰/۶۷۲$$

مقدار گاما از ۱- تا ۱+ است. علامت مثبت بر پیوند افزایشی و علامت منفی بر پیوند کاهشی دلالت دارد. هر قدر گاما به ۱- یا ۱+ نزدیک‌تر باشد، شدت پیوند دو متغیر ترتیبی قوی‌تر است. پس مقدار ۰/۶۷۲ که در مثال فوق به دست آمد، پیوند افزایشی نسبتاً قوی را بین دو متغیر میزان استفاده و میزان اعتماد نشان می‌دهد.

۴ـ۲ـ۸ـ معیار دی سامرز

گاما معیاری است که شدت و جهت پیوند دو متغیر را تنها با توجه به تعداد زوج‌های هماهنگ و ناهماهنگ جدول می‌سنجد و زوج‌های دارای همرتبگی را نادیده می‌گیرد. از این رو، شدت پیوند دو متغیر را قوی‌تر از آنچه هست، جلوه می‌دهد. معیار D سامرز[1] معیار ترتیبی نامتقارنی است که هم مانند گاما مبتنی بر زوج‌های هماهنگ و ناهماهنگ است و هم به تعداد کل زوج‌های موجود در جدول و زوج‌های همرتبه نسبت به متغیر مستقل توجه دارد. پیش از معرفی معیار D سامرز درباره همرتبگی در جدول توافقی بحث می‌کنیم.

فرض کنید مقوله‌های متغیر مستقل X ستون‌های جدول توافقی و مقوله‌های متغیر وابستهٔ Y سطرهای آن را تشکیل می‌دهند. در این صورت، تمام افرادی که در یک ستون قرار دارند، نسبت به متغیر مستقل (X) همرتبه هستند، زیرا مقدار متغیر مستقل برای همهٔ آنها یکسان است. به همین ترتیب، تمام افرادی که در یک سطر قرار دارند، نسبت به متغیر وابسته (Y) همرتبه‌اند، زیرا مقدار متغیر وابسته برای همهٔ آنها یکسان است. این همرتبگی‌ها در جدول توافقی ۵×۷ زیر مشاهده می‌شود. برای مثال، تمام افرادی که در خانه‌های سطر دوم جدول قرار دارند، در متغیر وابسته (Y) همرتبه هستند، زیرا برای همگی آنها $Y=۲$ است. همچنین، تمام افرادی که مثلاً در خانه‌های ستون پنجم جدول قرار دارند، در متغیر مستقل (X) همرتبه هستند، زیرا برای همگی آنها $X=۵$ است.

همرتبگی در جدول توافقی

Y \ X	۱	۲	۳	۴	۵	۶	۷
۱							
۲							
۳							
۴							
۵							

همرتبه در متغیر وابسته
(همرتبه در سطر دوم)

همرتبه در متغیر مستقل
(همرتبه در ستون پنجم)

1. Somers

علاوه بر همرتبگی در سطر یا همرتبگی در ستون، همرتبگی در هر دو متغیر X و Y نیز می‌تواند در جدول توافقی رخ دهد. برای مثال، تقاطع سطر دوم و ستون پنجم جدول قبل شامل افرادی است که هم در سطر و هم در ستون اشتراک دارند، یعنی هم دارای مقدار X یکسان و هم دارای مقدار Y یکسان هستند، زیرا برای همگی آنها $X=۵$ و $Y=۲$ است.

از آنچه گذشت می‌توان دریافت که در یک جدول توافقی با اندازه نمونه n، تعداد $\frac{n(n-۱)}{۲}$ زوج می‌توان یافت که از این میان، تعدادی هماهنگ (n_c)، تعدادی ناهماهنگ (n_d)، تعدادی همرتبه در ستون (T_X)، تعدادی همرتبه در سطر (T_Y) و تعدادی همرتبه در سطر و ستون (T_{XY}) هستند. رابطهٔ زیر بین این کمّیت‌ها برقرار است:

$$\frac{n(n-۱)}{۲}=n_c+n_d+T_X+T_Y-T_{XY}$$

اکنون معیار گاما را دقیق‌تر بررسی می‌کنیم تا ضرورت طرح معیار D سامرز آشکار شود. گاما مانند معیار لاندا از منطق کاهش نسبی در خطا (PRE) پیروی می‌کند. دو فرد را در نظر بگیرید که می‌خواهیم پیش‌بینی کنیم مقدار متغیر وابستهٔ Y برای کدام‌یک از آنها بیشتر است. یک بار این پیش‌بینی را تنها بر اساس توزیع متغیر Y و بار دیگر، با اطلاع از مقدار متغیر X صورت می‌دهیم. کاهش نسبی در خطایِ این دو شیوهٔ پیش‌بینی، مبنای معیار گاماست (اگر مثلاً پیوندی افزایشی بین دو متغیر X و Y برقرار باشد، آنگاه دانستن اینکه مقدار X برای فردی کمتر از فرد دیگر است، به این پیش‌بینی کمک می‌کند که مقدار Y نیز برای آن فرد کمتر از فرد دیگر است).

مقدار PRE در محاسبه گاما، تنها بر اساس افرادی محاسبه می‌شود که نسبت به هم، زوجی هماهنگ یا ناهماهنگ تشکیل می‌دهند، یعنی تمام افراد دارای همرتبگی نادیده گرفته می‌شوند؛ مثلاً کسانی که مقدار متغیر مستقل آنها برابر است یا افرادی که مقدار متغیر وابستهٔ یکسانی دارند. پس گاما نمی‌تواند ارزیابی دقیقی از نقش متغیر مستقل X در پیش‌بینی وضعیت متغیر Y ارائه دهد. به عبارت دیگر، معیار گاما تفاضل تعداد زوج‌های هماهنگ و ناهماهنگ را مشخص می‌سازد؛ یعنی n_c-n_d را بر n_c+n_d تقسیم می‌کند و از میان $\frac{n(n-۱)}{۲}$ زوجی که در جدول توافقی وجود دارد، تنها تعدادِ n_c+n_d زوج را در نظر می‌گیرد و به این ترتیب مرتکب بزرگنمایی در اندازه‌گیری شدت پیوند می‌شود.

نادیده‌گرفتن تعداد همرتبگی‌ها در متغیر وابسته منطقی است، زیرا اصلاً امکان ندارد از متغیر مستقل X برای پیش‌بینی وضعیت متغیر Y افرادی کمک گرفت که مقدار Y برابری دارند. بنابراین، می‌توان از میان $\frac{n(n-۱)}{۲}$ زوجی که در جدول توافقی وجود دارد، T_Y زوجی را نادیده گرفت که در متغیر وابستهٔ Y همرتبه هستند. از این رو، معیار D سامرز به صورت زیر در مخرج کسر خود تمام $\frac{n(n-۱)}{۲}$ زوج جدول توافقی را، جز T_Y زوجِ همرتبه در Y، در نظر می‌گیرد تا زوج‌های بیشتری از جدول را در تعیین شدت و جهت پیوند به حساب آورد. بنابراین، مقدار D سامرز هیچ‌گاه بیشتر از گاما نمی‌شود و سنجش واقع‌بینانه‌تری را از شدت پیوند دو متغیر به دست می‌دهد.

$$D_{Y|X}=\frac{n_c-n_d}{\frac{n(n-۱)}{۲}-T_Y}$$

درعمل می‌توان تعداد همرتبگی در متغیر وابستهٔ Y را به کمک فراوانی‌های حاشیه‌ای ستونی به صورت $T_Y=\frac{۱}{۲}\left(\sum_i n_{i.}^{۲}-n\right)$ به دست آورد.

مقدار D سامرز را برای جدول ۷ به دست می‌آوریم که فراوانی‌های آن در جدول زیر نیز مشاهده می‌شود.

اعتماد (X) / استفاده (Y)	کم	متوسط	زیاد	فراوانی حاشیه‌ای ستونی
کم	۸۵	۵۰	۲۰	$n_{۱.}=۱۵۵$
متوسط	۲۵	۱۱۰	۶۰	$n_{۲.}=۱۹۵$
زیاد	۱۰	۴۰	۱۰۰	$n_{۳.}=۱۵۰$
فراوانی حاشیه‌ای سطری	$n_{.۱}=۱۲۰$	$n_{.۲}=۲۰۰$	$n_{.۳}=۱۸۰$	$n=۵۰۰$

پیشتر، مقادیر $n_c=۴۸۸۵۰$ و $n_c=۹۵۵۰$ برای محاسبه گامای این جدول تعیین شده بود. اکنون مقدار T_Y را به صورت زیر محاسبه می‌کنیم:

$$T_Y=\tfrac{۱}{۲}\times((۱۵۵^{۲}+۱۹۵^{۲}-۱۵۰^{۲})-۵۰۰)=۴۲۰۲۵$$

بنابراین مقدار $D_{Y|X}$ برابر است با:

$$D_{Y|X}=\frac{۴۸۸۵۰-۹۵۵۰}{\frac{۵۰۰\times(۵۰۰-۱)}{۲}-۴۲۰۲۵}=\frac{۳۹۳۰۰}{۸۲۷۲۵}=۰/۴۷۵$$

این مقدار بر وجود رابطهٔ افزایشی متوسطی دلالت دارد، زیرا مقدار D سامرز نیز مانند گاما می‌تواند از ۱− تا ۱+ باشد. همان‌طور که مشاهده می‌شود، مقدار D سامرز جدول کوچک‌تر از مقدار $\gamma = ۰/۶۷۲$ است. مقدار $D_{X|Y}$ نیز برای جدول ۷ قابل محاسبه است، ولی باید $T_X = \frac{۱}{۲}\left(\sum_j n_{.j}^{۲} - n\right)$ را به جای T_Y استفاده کرد که بر اساس فراوانی‌های حاشیه‌ای سطری به شکل زیر تعیین می‌شود.

$$T_X = \frac{۱}{۲} \times ((۱۲۰^{۲} + ۲۰۰^{۲} - ۱۸۰^{۲}) - ۵۰۰) = ۴۳۱۵۰$$

پس:

$$D_{X|Y} = \frac{۴۸۸۵۰ - ۹۵۵۰}{\frac{۵۰۰ \times (۵۰۰-۱)}{۲} - ۴۳۱۵۰} = \frac{۳۹۳۰۰}{۸۱۶۰۰} = ۰/۴۸۱$$

معیار D سامرز مانند لاندا معیاری نامتقارن است، ولی می‌توان از میانگین $D_{X|Y}$ و $D_{Y|X}$ به معیار متقارن $D_{XY} = \frac{D_{X|Y} + D_{Y|X}}{۲}$ برای اندازه‌گیری شدت و جهت پیوند دو متغیر X و Y دست یافت. برای مثال پیشین، $D_{XY} = \frac{۰/۴۸۱ + ۰/۴۷۵}{۲} = ۰/۴۷۸$ به دست می‌آید.

۹ـ آزمون‌هایی درباره نسبت‌ها

در فصل پنجم به تفصیل از آزمون‌های مربوط به میانگین‌ها در شرایط مختلف بحث شد. در این قسمت، بررسی مشابهی درخصوص آزمون‌های مربوط به نسبت‌ها (درصدها) خواهیم داشت. گفتنی است آزمون مربوط به تک نسبت و آزمون مقایسه نسبت‌های دو جامعهٔ مستقل، آزمون‌های پارامتری هستند، ولی در کنار آزمون‌های ناپارامتری مربوط به نسبت‌ها مطرح شده‌اند تا بحث آزمون نسبت‌ها به طور کامل در یک قسمت از کتاب ارائه شود.

۱ـ۹ـ آزمون درباره تک نسبت

نسبت (درصد) را می‌توان حالت خاصی از میانگین قلمداد کرد. برای این منظور، یک متغیر مقوله‌ای دو حالتی را مانند وضعیت اشتغال با دو مقولهٔ شاغل و بیکار در نظر بگیرید. فرض کنید وضعیت اشتغال ۱۲ نفر به صورت جدول زیر است. نسبت شاغلان از تقسیم تعداد شاغلان (۷ نفر) بر تعداد کل افراد (۱۲ نفر)، رقم ۰/۵۸۳ (یا ۵۸/۳ درصد) به دست می‌آید.

سارا	امید	شبنم	سعید	شیما	مهناز	رضا	فرید	راد	حسن	ثریا	حمید
بیکار	بیکار	شاغل	شاغل	بیکار	شاغل	شاغل	شاغل	بیکار	شاغل	بیکار	شاغل
۰	۰	۱	۱	۰	۱	۱	۱	۰	۱	۰	۱

اکنون فرض کنید عددهای ۰ و ۱ به ترتیب برای مقولهٔ بیکار و شاغل در نظر گرفته شده‌اند (حتماً عدد ۱ را به مقوله‌ای اختصاص دهید که نسبت آن محاسبه می‌شود؛ در اینجا، هدف ما محاسبه نسبتِ شاغلان است پس عدد ۱ را به مقوله شاغل اختصاص می‌دهیم و عدد ۰ را به مقوله بیکار). میانگین صفرها و یک‌ها را به شکل زیر برای این ۱۲ نفر محاسبه می‌کنیم و به رقمی برابر با آنچه پیشتر به دست آمده بود، می‌رسیم.

$$\frac{۰+۰+۱+۱+۰+۱+۱+۱+۰+۱+۰+۱}{۱۲}=\frac{۷}{۱۲}=۰/۵۸۳$$

این ویژگی کمک می‌کند تا بتوان از برخی آزمون‌های مربوط به میانگین‌ها برای نسبت‌ها نیز استفاده کرد مانند آزمون مربوط به تک نسبت که مشابه آزمون مربوط به تک میانگین است. در اینجا نیز مانند آزمون تک میانگین، سه نوع فرضیه‌بندی درباره نسبت[۱] یک جامعه یعنی P مطرح می‌شود که یکی آزمونی دو دامنه و بقیه آزمونی یک دامنه هستند. فرض کنید P_o مقداری مشخص باشد، در این صورت، سه فرضیه‌بندی زیر قابل طرح است:

۱. آزمون دو دامنه: فرضیه صفر $P=P_o$ در مقابل فرضیه جانشین $P \neq P_o$
۲. آزمون یک دامنه: فرضیه صفر $P \geq P_o$ در مقابل فرضیه جانشین $P < P_o$
۳. آزمون یک دامنه: فرضیه صفر $P \leq P_o$ در مقابل فرضیه جانشین $P > P_o$

برای مثال، آزمون فرضیه صفرِ «نسبت شاغلان حداکثر ۰/۷ است» در مقابل فرضیه جانشینِ «نسبت شاغلان بیش از ۰/۷ است» مانند سومین فرضیه‌بندی، آزمونی یک دامنه است، زیرا به صورت فرضیه صفر $P \leq ۰/۷$ در مقابل فرضیه جانشین $P > ۰/۷$ قابل بیان است، یعنی عدد ۰/۷ به جای P_o قرار گرفته است.

برای اجرای این آزمون، نمونه‌ای n تایی از جامعه آماری تحت بررسی انتخاب می‌شود و نسبت مورد نظر برای متغیر مقوله‌ای دو حالتی به دست می‌آید. اگر نسبت حاصل از نمونه p باشد، آمارهٔ آزمون به شکل زیر محاسبه می‌شود:

$$z=\frac{p-P_o}{\sqrt{\frac{P_o(۱-P_o)}{n}}}$$

اگر n به اندازه کافی بزرگ باشد (حداقل ۳۰)، از جدول توزیع نرمال استاندارد برای تعیین نقاط بحرانی این آزمون استفاده می‌شود.

---

۱. در بحث کنونی، به جای درصد از نسبت که عدد ۱۰۰ در آن ضرب نمی‌شود، استفاده می‌کنیم. این امر هیچ محدودیتی در استنباط آماری درباره درصدها به همراه ندارد. در مثال بالا، ۵۸/۳ بیانگر درصد و ۰/۵۸۳ بیانگر نسبت است.

اگر فرضیه‌بندی اول یعنی آزمون دو دامنه مورد توجه باشد، دو نقطهٔ بحرانی برای رد این فرضیه صفر در نظر گرفته می‌شود؛ یکی در سمت راست دُم توزیع نرمال استاندارد یعنی $z_{\frac{\alpha}{2}}$ و دیگری در سمت چپ دُم توزیع یعنی $-z_{\frac{\alpha}{2}}$. نقاط بحرانی فرضیه‌بندی دوم و سوم نیز به ترتیب $-z_{\alpha}$ و z_{α} هستند.

برای مثال، فرض کنید برای آزمون فرضیه صفر $P \leq ۰/۷$ در مقابل فرضیه جانشین $P > ۰/۷$ در سطح $\alpha = ۰/۰۵$، نمونه‌ای ۳۰۰ نفری انتخاب‌شده و نسبت شاغلان ۰/۷۵ (یعنی ۷۵ درصد) به دست آمده است. آماره آزمون عبارت است از:

$$z = \frac{۰/۷۵ - ۰/۷}{\sqrt{\frac{۰/۷\times(۱-۰/۷)}{۳۰۰}}} = \frac{۰/۰۵}{۰/۰۲۶۴} \cong ۱/۸۹$$

با استفاده از جدول توزیع نرمال استاندارد، مقدار $z_{۰/۰۵} = ۱/۶۵$ به دست می‌آید. از آنجا که مقدار آماره از نقطه بحرانی بزرگ‌تر است، پس فرضیه صفر رد می‌شود. به عبارت دیگر، درصد شاغلان بیش از ۰/۷ (۷۰ درصد) است.

۲ـ۹ـ آزمون نیکویی برازش برای توزیع چند جمله‌ای

آزمون نیکویی برازش در جدول‌های توافقی به بررسی فرضیه استقلال دو متغیر می‌پردازد. کاربرد دیگری از این آزمون به بررسی فرضیه‌ای درباره توزیع تنها یک متغیر مقوله‌ای اختصاص دارد. برای مثال، اگر سه نامزد «آ»، «ب» و «پ» برای انتخابات ریاست جمهوری وجود داشته باشند (یک متغیر با ۳ مقوله)، آیا می‌توان به عنوان فرضیه صفر، ادعا کرد که رأی‌دهندگان به صورت زیر بین این سه نامزد توزیع شده‌اند؛ یعنی ۲۵ درصد به فرد «آ»، ۶۰ درصد به فرد «ب» و ۱۵ درصد به فرد «پ» گرایش دارند.

نامزد	آ	ب	پ	جمع
درصد رأی‌دهندگان	۲۵	۶۰	۱۵	۱۰۰

برای پاسخ به این پرسش، نمونه‌ای انتخاب و گرایش آنان به نامزدهای انتخاباتی تعیین می‌شود. فرض کنید از یک نمونه ۵۰۰ نفری، ۱۱۵ نفر به نامزد «آ»، ۳۲۰ نفر به نامزد «ب» و ۶۵ نفر به نامزد «پ» گرایش دارند.

نامزد	آ	ب	پ	جمع
فراوانی رأی‌دهندگان	۱۱۵	۳۲۰	۶۵	۵۰۰

اکنون باید بپرسیم که اگر ادعای فرضیه صفر درست باشد، انتظار داشتیم ۵۰۰ پاسخگو چگونه توزیع شده باشند؟ پاسخ در جدول زیر ارائه شده است:

نامزد	آ	ب	پ	جمع
فراوانی مورد انتظار رأی‌دهندگان	$۵۰۰ \times ۲۵\% = ۱۲۵$	$۵۰۰ \times ۶۰\% = ۳۰۰$	$۵۰۰ \times ۱۵\% = ۷۵$	۵۰۰

پس دو دسته فراوانی وجود دارد: فراوانی مشاهده‌شده از نمونه ۵۰۰ تایی و فراوانی مورد انتظار تحت ادعای بیان‌شده در فرضیه صفر. بنابراین، آماره آزمون از همان رابطه آزمون نیکویی برازش به صورت زیر قابل محاسبه است (k نشان‌دهندهٔ تعداد مقوله‌ها و n_i و e_i به ترتیب فراوانی مشاهده‌شده و مورد انتظار در مقوله iام هستند):

$$\chi^۲ = \sum_{i=۱}^{k} \frac{(n_i - e_i)^۲}{e_i}$$

به این ترتیب، مقدار آماره عبارت است از:

$$\chi^۲ = \frac{(۱۱۵-۱۲۵)^۲}{۱۲۵} + \frac{(۳۲۰-۳۰۰)^۲}{۳۰۰} + \frac{(۶۵-۷۵)^۲}{۷۵} = ۳/۴۷$$

مقدار بحرانی آزمون، $\chi^۲_{k-۱,\alpha}$ است که از جدول چندک‌های توزیع $\chi^۲$ به دست می‌آید. در مثال فوق، اگر آزمون در سطح $\alpha = ۰/۰۵$ اجرا شود، با توجه به وجود ۳ مقوله، نقطه بحرانی $\chi^۲_{۲,۰/۰۵} = ۵/۹۹۱$ تعیین می‌شود. بنابراین، فرضیه صفر درباره نحوه توزیع رأی‌دهندگانِ به سه نامزد رد نمی‌شود، زیرا مقدار آماره از نقطه بحرانی بیشتر نیست.

۳ـ۹ـ مقایسه نسبت‌های دو جامعه

دو آزمون پیشگفته به نسبت یا نسبت‌های یک متغیر مقوله‌ای در یک جامعه اختصاص داشتند، چه این آزمون درباره تک نسبت باشد یا درباره چگونگی توزیع نسبت‌های (درصدهای) مقوله‌های یک متغیر. شرایطی وجود دارد که لازم است نسبت‌های دو جامعه مورد مقایسه قرار گیرند. برای مثال، مقایسه نسبتِ شاغلان شهر «آ» با نسبتِ شاغلان شهر «ب». در این قسمت، چنین مقایسه‌هایی را در دو حالت بررسی می‌کنیم: مقایسه نسبت در دو جامعه وابسته و مقایسه نسبت در دو جامعه مستقل.

۱ـ۳ـ۹ـ مقایسه نسبت‌های دو جامعه مستقل

مشابه آزمون فرضیه مقایسه دو میانگین، سه نوع فرضیه‌بندی برای مقایسه نسبت‌های دو جامعه (P_1 و P_2) امکان‌پذیر است:

۱. آزمون دو دامنه: فرضیه صفر $P_1 = P_2$ در مقابل فرضیه جانشین $P_1 \neq P_2$

۲. آزمون یک دامنه: فرضیه صفر $P_1 \geq P_2$ در مقابل فرضیه جانشین $P_1 < P_2$

۳. آزمون یک دامنه: فرضیه صفر $P_1 \leq P_2$ در مقابل فرضیه جانشین $P_1 > P_2$

برای اجرای این آزمون‌ها از فرایند مقایسه میانگین‌های دو جامعه مستقل با فرض «برابری واریانس‌ها» استفاده می‌شود. ابتدا نمونه‌ای به اندازه n_1 از جامعه اول و نمونه‌ای به اندازه n_2 از جامعه دوم انتخاب و نسبت‌های حاصل از این دو نمونه محاسبه می‌شوند. اگر p_1 و p_2 نسبت‌های محاسبه‌شده باشند، واریانس آمیخته عبارت است از:

$$s_p^2 = \frac{(\frac{n_1p_1+n_2p_2}{n_1+n_2})(1-\frac{n_1p_1+n_2p_2}{n_1+n_2})}{n_1+n_2-1}$$

به این ترتیب، آماره آزمون از رابطه زیر به دست می‌آید:

$$t = \frac{p_1 - p_2}{s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$$

نقطه بحرانی برای آزمون دو دامنه، $-t_{n_1+n_2-2,\frac{\alpha}{2}}$ و $t_{n_1+n_2-2,\frac{\alpha}{2}}$ است که از جدول چندک‌های توزیع t مشخص می‌شود. نقاط بحرانی فرضیه‌بندی‌های دوم و سوم نیز به ترتیب $-t_{n_1+n_2-2,\alpha}$ و $t_{n_1+n_2-2,\alpha}$ هستند.

برای مثال، مقایسه نسبت شاغلان دو شهر را در نظر بگیرید. از شهر اول، نمونه‌ای ۳۰۰ نفری انتخاب و نسبت شاغلان آن ۰/۸ به دست آمده است. از شهر دوم، نمونه‌ای ۴۰۰ نفری انتخاب و نسبت شاغلان ۰/۸۳ به دست آمده است. هدف، آزمون فرضیه صفر «برابری نسبتِ شاغلان دو شهر» در مقابل فرضیه جانشین «نابرابری نسبتِ شاغلان دو شهر» است (یعنی فرضیه‌بندی اول که آزمونی دو دامنه است). مقدار واریانس آمیخته بر اساس این اطلاعات به صورت زیر به دست می‌آید:

$$n_1 = ۳۰۰, p_1 = ۰/۸, n_2 = ۴۰۰, p_2 = ۰/۸۳$$

$$s_p^2 = \frac{(\frac{۳۰۰\times ۰/۸+۴۰۰\times ۰/۸۳}{۳۰۰+۴۰۰})(1-\frac{۳۰۰\times ۰/۸+۴۰۰\times ۰/۸۳}{۳۰۰+۴۰۰})}{۳۰۰+۴۰۰-۱} = \frac{(\frac{۵۷۲}{۷۰۰})(1-\frac{۵۷۲}{۷۰۰})}{۶۹۹} = \frac{۰/۸۱۷\times ۰/۱۸۳}{۶۹۹} \cong ۰/۰۰۰۲$$

بر این اساس، آماره آزمون عبارت است از:

$$t = \frac{۰/۸ - ۰/۸۳}{۰/۰۱۴\sqrt{\frac{۱}{۳۰۰} + \frac{۱}{۴۰۰}}} = ۳۰$$

اگر آزمون در سطح $\alpha = ۰/۰۵$ اجرا شود، نقاط بحرانی سمت راست و سمت چپ از جدول چندک‌های توزیع t به ترتیب $t_{۶۹۸, ۰/۰۲۵} = ۱/۹۶$ و $-t_{۶۹۸, ۰/۰۲۵} = -۱/۹۶$ خواهند بود. از آنجا که مقدار آماره از نقطه بحرانی سمت راست بزرگ‌تر است، پس فرضیه صفر رد می‌شود؛ یعنی می‌توان نتیجه گرفت نسبت (درصد) شاغلان دو شهر نابرابر هستند (مقدار احتمال برای این آماره برابر با ۰/۰۱۶ که از $\frac{\alpha}{۲} = ۰/۰۲۵$ کوچک‌تر است).

۲ـ۳ـ۹ـ مقایسه نسبت‌های دو جامعه وابسته (آزمون مک‌نمار)

گاهی با مسئله مقایسه نسبت‌های دو جامعه وابسته روبه‌رو هستیم. برای مثال، نمونه‌ای nتایی از شهری انتخاب و وضعیت اشتغال آنها ثبت و سپس، نسبت شاغلان محاسبه می‌شود. پس از آن، به مدت شش ماه برنامه‌های آموزش کارآفرینی و حمایت‌های مالی برای ایجاد اشتغال به اجرا در می‌آید و بار دیگر به همان افراد نمونه n تایی مراجعه و وضعیت اشتغال آنان ثبت و نسبت شاغلان محاسبه می‌شود. اکنون می‌خواهیم بدانیم نسبت شاغلان پس از اجرای آموزش‌های کارآفرینی و حمایت‌های مالی نسبت به گذشته تغییر کرده است. در اینجا با نمونه ثابتی رو به رو هستیم که دو بار برای گردآوری داده‌ها به آن مراجعه شده است. بنابراین با دو جامعه مستقل سروکار نداریم.

آزمون مَک‌نِمار[1] برای مقایسه نسبت‌های دو جامعه وابسته به کار می‌رود. برای استفاده از این آزمون، مثال فوق را در قالب یک جدول توافقی ۲×۲ به صورت جدول ۸ نمایش می‌دهیم. این جدول نشان می‌دهد a نفر از نمونه n تایی، هم پیش از اجرا و هم پس از اجرای برنامه آموزشی کارآفرینی و حمایت مالی شاغل بوده‌اند. همچنین، در وضعیت بیکاریِ d نفر هیچ تغییری رخ نداده است. از سوی دیگر، b نفر پیش از اجرا شاغل بوده‌اند، ولی پس از اجرا بیکار شده‌اند در حالی که c نفر پیش از اجرا بیکار، ولی پس از اجرا شاغل شده‌اند. بر اساس فراوانی‌های حاشیه‌ای جدول ۸، نسبت شاغلانِ پیش از اجرا $p_۱ = \frac{a+b}{n}$ و نسبت شاغلانِ پس از اجرا $p_۲ = \frac{a+c}{n}$ است. توجه کنید که $a+b+c+d = n$.

1. McNemar

جدول ۸: جدول توافقی ۲×۲ برای آزمون مک‌نمار

پیش از اجرا \ پس از اجرا	شاغل	بیکار
شاغل	a	b
بیکار	c	d

هدف، آزمون فرضیه صفر $P_۱ = P_۲$ در مقابل فرضیه جانشین $P_۱ \neq P_۲$ است. از آنجا که فراوانی خانه‌هایی از جدول ۸ که هیچ تغییر وضعیتی را نشان نمی‌دهند (a و d) نمی‌توانند نقشی در ایجاد تفاوت بین $P_۱$ و $P_۲$ داشته باشند، آماره آزمون مک‌نمار تنها بر اساس b و c به صورت زیر محاسبه می‌شود. نقطه بحرانی آزمون، $\chi^۲_{۱,\alpha}$ است که از جدول چندک‌های توزیع $\chi^۲$ تعیین می‌شود.

$$\chi^۲ = \frac{(b-c)^۲}{b+c}$$

برای روشن شدن کاربرد آزمون مک‌نمار، مثال جدول ۸ را به صورت عددی در قالب جدول زیر برای نمونه‌ای ۴۰۰ نفری بررسی می‌کنیم. این جدول بیان می‌کند که ۰/۶۵ (۶۵ درصد) از افراد نمونه پیش از اجرای برنامه آموزشی کارآفرینی و حمایت مالی شاغل بوده‌اند، در حالی که این رقم پس از اجرای آن به ۰/۷۷۵ (۷۷/۵ درصد) می‌رسد. اکنون با آزمون مک‌نمار، برابری این نسبت‌ها را می‌آزماییم.

پیش از اجرا \ پس از اجرا	شاغل	بیکار
شاغل	۲۵۰	۱۰
بیکار	۶۰	۸۰

آماره آزمون عبارت است از:

$$\chi^۲ = \frac{(۱۰-۶۰)^۲}{۱۰+۶۰} = ۳۵/۷$$

که از نقطه بحرانی آزمون $\chi^۲_{۱,۰/۰۵} = ۳/۸۴$ بزرگ‌تر است. بنابراین فرضیه صفر رد می‌شود و نتیجه می‌گیریم نسبت شاغلانِ پیش از اجرا با نسبت شاغلانِ پس از اجرا تفاوت دارد.

۴ـ۹ـ مقایسه نسبت‌های بیش از دو جامعه مستقل

همان طور که مشاهده شد، علاوه بر استقلال، از آزمون نیکویی برازش می‌توان برای بررسی فرضیه‌های دیگری درباره نسبت‌ها (درصدها) استفاده کرد. این آزمون قادر است درصدهای ستونی را با یکدیگر و درصدهای سطری را نیز با یکدیگر مقایسه کند. برای مثال، جدول ۹ نمایانگر جدول توافقی متغیرهای مقطع تحصیلی و میزان علاقه‌مندی به رشته تحصیلی است. یکی از تحلیل‌های معمول برای چنین جدولی، بررسی این پرسش است که آیا تحصیل در مقطعی خاص از یک رشته با میزان علاقه‌مندی به آن رشته در پیوند است؟ به عبارت دیگر، آیا می‌توان ادعا کرد دانشجویانی که در مقاطع بالاتر تحصیل می‌کنند به رشته خود نیز بیشتر علاقه دارند؟ پاسخ به این پرسش از طریق آزمون فرضیه استقلال دو متغیر یعنی استفاده از آزمون نیکویی برازش امکان‌پذیر است.

جدول ۹: جدول توافقی مقطع تحصیلی و میزان علاقه‌مندی به رشته تحصیلی

علاقه‌مندی / مقطع	کارشناسی	کارشناسی ارشد	دکترا
کم	۲۰	۵	۵
متوسط	۳۰	۲۰	۱۵
زیاد	۵۰	۷۵	۸۰
جمع	۱۰۰	۱۰۰	۱۰۰

جدول ۹ را به جای «پیوند» از طریق «درصدها» نیز می‌توان تحلیل کرد، هر چند در هر دو حالت از آزمون نیکویی برازش استفاده می‌شود. برای این منظور، می‌توان درصدهای ستونی این جدول را مبنا قرار داد و برای مثال از خود پرسید آیا درصد افرادی که در حد زیاد به رشته خود علاقه‌مند هستند در مقاطع مختلف یکسان است؟ در جدول ۹، ۵۰ درصد از دانشجویان کارشناسی در حد زیاد به رشته خود علاقه دارند، در حالی که به ترتیب ۷۵ و ۸۰ درصد از دانشجویان مقطع کارشناسی ارشد و دکترا در حد زیاد به رشته خود علاقه‌مندند. پرسش‌هایی از این دست به اختلاف درصدها در جامعه‌های مستقل می‌پردازد و در واقع، توزیع یک متغیر مقوله‌ای را در جامعه‌های مستقل (مقوله‌های متغیر دیگر) مقایسه می‌کند (تجزیه واریانس را به یاد بیاورید).

اگر آزمون نیکویی برازش در این جدول معنادار شود، یعنی وجود پیوند دو متغیر را نشان دهد، به راحتی می‌توان نتیجه گرفت «حداقل» برخی از درصدهای ستونی با یکدیگر اختلاف دارند (به همین ترتیب می‌توان گفت «حداقل» برخی از درصدهای سطری نیز با یکدیگر اختلاف دارند). این نتیجه‌گیری دور از انتظار نیست، زیرا بین «اختلاف» درصدها و «پیوند» دو متغیر مقوله‌ای رابطه نزدیکی وجود دارد. در واقع، اگر درصدهای مقوله‌های متغیری در بین مقوله‌های متغیر دیگر تغییر کند، به وجود پیوندی بین آن دو متغیر اشاره دارد. برعکس، اگر بین دو متغیر پیوند وجود داشته باشد، به آن معناست که تغییر وضعیت در یک متغیر می‌تواند به تغییر وضعیت در متغیر دیگر بینجامد یعنی برخی درصدهای یک متغیر در مقوله‌های متغیر دیگر دارای اختلاف هستند.

در مثال جدول ۹، تغییر وضعیت از مقطع کارشناسی به کارشناسی ارشد با تغییر وضعیت میزان علاقه‌مندی مقولهٔ زیاد از ۵۰ درصد به ۷۵ درصد همراه شده است، پس می‌توان انتظار داشت که بین مقطع و علاقه‌مندی «پیوند» وجود دارد، به طوری که این پیوند، خود را در «اختلاف» بین درصدها نشان داده است. از همین رو، اگر آزمون نیکویی برازش بر وجود پیوند دلالت داشت، به طور منطقی می‌توان وجود اختلاف بین برخی از درصدها را نیز نتیجه گرفت.

اصطلاحات فصل ششم

Association	پیوند	Median Test	آزمون میانه
Cell	خانه جدول توافقی	Nonparametric Test	آزمون ناپارامتری
Column Percent	درصد ستونی	Observed Frequency	فراوانی مشاهده‌شده
Concordant	(زوج) سازگار	Parametric Test	آزمون پارامتری
Contingency Table	جدول توافقی	Phi Coefficient	ضریب فی
Cramér’s V	معیار V کرامر	Row Percent	درصد سطری
Discordant	(زوج) ناسازگار	Sign Test	آزمون علامت
Expected Frequency	فراوانی مورد انتظار	Sommers’ D	معیار D سامرز
Gamma	معیار گاما	Spearman’s Coefficient	ضریب اسپیرمن
Goodness of Fit Test	آزمون نیکویی برازش	Tetrachoric Correlation	همبستگی تتراکوریک
Kendall's Coefficient	ضریب کندال (تاو)	Tied Rank	همرُتبگی
Kruskal-Wallis Test	آزمون کروسکال-والیس	Total Percent	درصد کل
Lambda Coefficient	ضریب لاندا	Rank Sum Test	آزمون مجموع رتبه‌ای
Mann-Whitney Test	آزمون من-ویتنی	Signed Rank Test	آزمون رتبه علامت‌دار
Marginal Percent	درصد حاشیه‌ای	Parametric Test	آزمون پارامتری
McNemar’s Test	آزمون مک‌نمار	Proportional Reduction in Error (PRE)	کاهش نسبی در خطا

تمرین‌های فصل ششم

۱. بانکی مایل است فرایند پرداخت وام خود را طوری تغییر دهد که حداکثر ظرف مدت ۱۰ روز به متقاضیان داده شود. این بانک انتظار دارد پس از شش ماه از اعمال تغییرات، حداقل نیمی از وام‌ها ظرف مدت ۱۰ روز پرداخت شده باشند. از این رو، ۲۰ مورد از وام‌های پرداخت‌شده را به تصادف انتخاب و زمان سپری‌شده برای اعطای آنها را مطابق جدول زیر ثبت می‌کند. آیا می‌توان ادعا کرد که هنوز فرایند پرداخت وام بیش از ۱۰ روز طول می‌کشد؟

۱۳	۲۰	۹	۱۲	۱۵	۹	۱۲	۱۱	۱۰	۱۴
۲۰	۱۹	۱۲	۱۹	۹	۲۸	۱۶	۱۱	۲۷	۹

۲. هدف پژوهشگری مقایسه ماندگاری تأثیر دو نوع دارو بر تسکین درد ناشی از بیماری خاصی است. برای این منظور، ۱۲ بیمار انتخاب و داروی اول به آنها داده و مدت زمان تسکین درد آنان ثبت می‌شود. روز بعد، بار دیگر به همان افراد داروی دوم داده و مدت زمان تسکین درد آنان ثبت می‌شود. جدول زیر مدت زمان تسکین درد بر حسب ساعت را نشان می‌دهد.

بیمار	۱	۲	۳	۴	۵	۶	۷	۸	۹	۱۰	۱۱	۱۲
داروی ۱	۲	۳/۶	۲/۶	۲/۶	۷/۳	۳/۴	۱۴/۹	۶/۶	۲/۳	۲	۶/۸	۸/۵
داروی ۲	۳/۵	۵/۷	۲/۹	۲/۴	۹/۹	۳/۳	۱۶/۷	۶	۳/۸	۴	۹/۱	۲۰/۹

آ) فرضیه‌های آماری مناسب برای مقایسه این دو دارو را بنویسید.

ب) با استفاده از آزمون رتبه علامت‌دار ویلکاکسون، فرضیه صفر را آزمون کنید (در دو حالت تعداد نمونه کوچک و بزرگ).

۳. روان‌شناسی آزمون درک مطلبی را در دو گروه عقب‌افتاده و دارای آسیب‌های مغزی اجرا کرده است. نمرات این آزمون به شرح جدول زیر است. او می‌خواهد بررسی کند که آیا توانایی درک مطلب دو گروه به یک اندازه است.

گروه	نمرات آزمون	تعداد
عقب‌افتادگان ذهنی	۷۷، ۷۸، ۷۰، ۷۲، ۶۵ و ۷۴	۶
دارای آسیب مغزی	۶۰، ۶۲، ۷۰، ۷۶، ۶۸، ۷۲، ۷۰	۷

آ) فرضیه‌های صفر و جانشین مربوط به این آزمون را بنویسید.

ب) آزمون مجموع رتبه‌ای ویلکاکسون را برای بررسی فرضیه صفر به کار بگیرید.
پ) آزمون فوق را با این فرض که تعداد نمونه‌ها به اندازه کافی بزرگ است، اجرا کنید.
ت) آزمون من‌ـویتنی را در دو حالت تعداد نمونه کوچک و بزرگ برای بررسی این فرضیه صفر به کار بگیرید.

۴. روشی برای کاهش حشرات مضر در کشاورزی، نصب صفحه‌های چسبناک در مزرعه‌هاست تا حشرات را به سمت خود جلب کنند و به دام بیندازند. پژوهشی در پی بررسی این موضوع است که چه رنگی این صفحات را برای حشرات جذاب‌تر می‌کند تا به این ترتیب بتوان حشرات بیشتری را به دام انداخت. چهار رنگ برای این بررسی انتخاب شد و ۶ صفحه از هر رنگ در مزرعه قرار گرفت تا در یک دوره زمانی، تعداد حشرات به دام افتاده در هر صفحهٔ چسبناک شمارش شود. جدول زیر داده‌های مربوط به این پژوهش را نشان می‌دهد.

رنگ	تعداد حشرات به دام افتاده در هر صفحه					
لیمویی	۴۵	۵۹	۴۸	۴۶	۳۸	۴۷
سفید	۲۱	۱۲	۱۴	۱۷	۱۳	۱۷
سبز	۳۷	۳۲	۱۵	۲۵	۳۹	۴۱
آبی	۱۶	۱۱	۲۰	۲۱	۱۴	۷

آ) فرضیه‌های صفر و جانشین را برای رسیدن به هدف این پژوهش بیان کنید.
ب) آزمون کروسکال‌ـوالیس را برای مقایسه این چهار رنگ اجرا و نتیجه‌گیری کنید.
پ) آزمون کروسکال‌ـوالیس را با این فرض که تعداد نمونه‌های هر گروه کم است، اجرا کنید.
ت) چگونه می‌توان از آزمون نیکویی برازش برای مقایسه میانه چهار گروه استفاده کرد؟
(آزمون میانه: میانه کل داده‌ها را بیابید و برای هر گروه فراوانی داده‌های بیشتر از میانهٔ کل و فراوانی داده‌های کمتر یا برابر با آن را تعیین کنید. اکنون جدولی ۴×۲ تشکیل دهید و آزمون نیکویی برازش را اجرا کنید)

۵. جدول زیر درآمد سرانه ۱۱ کشور (برحسب دلار) و تعداد نوزادان مرده (از هر ۱۰۰۰ نوزاد) را نشان می‌دهد.

کشور	A	B	C	D	E	F	G	H	I	J	K
درآمد	۱۳۰	۵۹۵۰	۵۶۰	۲۰۱۰	۱۸۷۰	۱۷۰	۳۹۰	۵۸۰	۸۲۰	۶۶۲۰	۳۸۰۰
نوزادان مرده	۱۵۰	۴۳	۱۲۱	۵۳	۴۱	۱۶۹	۱۴۳	۵۹	۷۵	۲۰	۳۹

آ) نمودار پراکنش متغیر تعداد نوزادان مرده را در مقابل درآمد رسم کنید. چه رابطه‌ای بین دو متغیر مشاهده می‌کنید؟

ب) ضریب همبستگی اسپیرمن دو متغیر را محاسبه و تفسیر کنید.

ت) ضریب تاو کندال دو متغیر را محاسبه و تفسیر کنید.

۶. مقدار آماره آزمون برای یک جدول توافقی با ۲ سطر و ۳ ستون برابر با ۱۳/۵ است. چه تصمیمی در سطح معناداری ۵ درصد درباره فرضیه صفرِ استقلال دو متغیر می‌گیرید؟ اگر اندازه نمونه ۲۵۰ باشد، مقدار V کرامر این جدول چقدر است؟ شدت پیوند را بر اساس این مقدار، چگونه تفسیر می‌کنید؟

۷. فراوانی‌های مورد انتظار خانه‌های جدول توافقی زیر را محاسبه کنید. آیا می‌توان از آزمون نیکویی برازش برای بررسی استقلال دو متغیر X و Y در این جدول استفاده کرد؟

Y \ X	۱	۲	۳	جمع
۱	۱۷۵	۱۸۱	۱۸۷	۵۴۳
۲	۳۶	۳۳	۲۴	۹۳
۳	۲	۷	۱	۱۰
۴	۱	۱	۰	۲
۵	۰	۲	۰	۲
جمع	۲۱۴	۲۲۴	۲۱۲	۶۵۰

۸. یک شرکت خدماتی که در چهار منطقه فعالیت می‌کند، قصد تغییر شیوه خدمات‌رسانی خود را دارد. این شرکت به اجرای یک نظرسنجی با نمونه ۱۳۵۰ نفری دست می‌زند تا دریابد نظر دریافت‌کنندگان خدمات درباره ایجاد این تغییر در مناطق مختلف چگونه است. جدول زیر اطلاعات حاصل از نظرسنجی را نشان می‌دهد.

نظر \ منطقه	اول	دوم	سوم	چهارم	جمع
موافق	۱۷۶	۱۲۱	۳۱۶	۸۶	۶۹۹
مخالف	۱۲۴	۱۲۹	۲۸۴	۱۱۴	۶۵۱
جمع	۳۰۰	۲۵۰	۶۰۰	۲۰۰	۱۳۵۰

آ) فرضیه صفر و فرضیه جانشینی را بیان کنید که می‌تواند به پرسش این شرکت پاسخ دهد.
ب) اگر جدول زیر شامل فراوانی‌های مورد انتظار جدول قبل باشد، ابتدا خانه‌های خالی آن را تکمیل و سپس مقدار آماره آزمون نیکویی برازش را محاسبه کنید؟

نظر / منطقه	اول	دوم	سوم	چهارم	جمع
موافق	۱۵۵/۳	۱۲۹/۴		۱۰۳/۶	
مخالف	۱۴۴/۷	۱۲۰/۶	۲۸۹/۳	۹۶/۴	
جمع					

ت) نتیجه آزمون چیست؟ شدت پیوند این دو متغیر چقدر است؟

۹. پژوهشگری مایل است بداند که آیا بین نوع صدمات ناشی از تصادف و تخطی از سرعت مجاز رابطه‌ای وجود دارد. جدول زیر داده‌های ۲۶۵ تصادف را بر حسب سرعت خودرو و نوع صدمات نشان می‌دهد.

نوع / سرعت	مجاز	۲۰ کیلومتر تخطی	بیش از ۲۰ کیلومتر
صدمات مالی	۵۰	۲۰	۵
صدمات جرحی	۲۰	۷۰	۵۵
منجر به مرگ	۵	۱۰	۳۰

آ) مقدار لاندا را با این فرض که متغیر نوع صدمات، متغیر وابسته باشد، محاسبه و تفسیر کنید.
ب) نوعی ترتیب می‌توان برای نوع صدمات در نظر گرفت، به طوری که صدمات مالی پایین‌ترین سطح و صدمات منجر به مرگ بالاترین آن باشد. بر این اساس، مقدار گاما و D سامرز (نوع صدمه، متغیر وابسته است) را محاسبه و تفسیر کنید.

فصل هفتم

۷ نمونه‌گیری

نمونه‌گیری از کاربردی‌ترین حوزه‌های دانش آمار است که علوم مختلف، داده‌های پژوهشی خود را با استفاده از روش‌های آن گردآوری می‌کنند. این فصل با معرفی نمونه‌گیری و نقش آن در فراهم‌آوردنِ داده‌هایی معتبر از جامعه آماری آغاز می‌شود. سپس روش‌های نمونه‌گیری احتمالی و نااحتمالی معرفی می‌شوند؛ هرچند روش‌های احتمالی محور اصلی مباحث این فصل هستند و از همین رو نیز با تأکید و تفصیل بیشتری بررسی شده‌اند. در پایان فصل نیز مروری گذرا بر روش‌های احتمالی پیچیده‌تر خواهیم داشت.

۱ـ نمونه‌گیری و جایگاه آن

در فصل نخست، جامعه آماری به عنوان قلمرویی معرفی شد که پژوهش در آن صورت می‌گیرد؛ به این معنا که پژوهشگر تلاش می‌کند متغیرهای تحت بررسی خود را در میان اعضای این جامعه اندازه بگیرد و مطالعه کند و به این ترتیب آن جامعه را از نظر آن متغیرها بشناسد. اغلب تعداد اعضای جامعه آماری چنان زیاد است که به پژوهشگر اجازه نمی‌دهد همهٔ آنها را بررسی کند.

برای مثال، دسترسی به اعضای جامعه‌ای مانند یک کلان‌شهر یا حتی شهری با صدهزار شهروند، امری هزینه‌بر و زمان‌بر است. دشواری‌های دسترسی و مراجعه به تمام اعضای جامعه آماری آن‌قدر پیچیده، فراوان و غیرقابل پیش‌بینی است که پژوهشگر در همان مراحل ابتدایی از پژوهش خود صرفنظر می‌کند. برای مثال ممکن است در زمان مراجعه، عضوی در خانه نباشد یا مریض باشد و نتواند به شما پاسخ دهد، یا از آن دسته افراد بدقلق و وسواسی باشد که به سختی به کسی اعتماد می‌کنند و بنابراین حاضر نیست با شما همکاری کند و مشکلات متعدد دیگر. حتی اگر پژوهشگر از عهدهٔ هزینهٔ مراجعه به تک‌تک شهروندان برآید و آن‌قدر پیگیر و سمج باشد که بر فرض محال بتواند به همگان دسترسی پیدا کند و همکاری آنان را نیز جلب نماید، زمان مراجعه و گردآوری داده‌ها از تک‌تک اعضای یک جامعه آماری بزرگ چنان می‌تواند به درازا بکشد که در بسیاری از موارد، اطلاعات حاصل از این داده‌ها با گذشت این زمان طولانی دیگر بهنگام و قابل استفاده نیست. بنابراین، جز در موارد خاص و نادر، مراجعه به تمام اعضای جامعه آماری و گردآوری داده‌ها از تک‌تک آنها (سرشماری جامعه آماری) پیشنهاد نمی‌شود.

محدود کردنِ «گردآوری» داده‌های پژوهش به بخشی از جامعه آماری، راهکاری است که پژوهشگران در عمل به آن متوسل می‌شوند. توجه کنید که تنها «گردآوری» داده‌ها به بخشی از جامعه محدود می‌شود، ولی همچنان هدف پژوهشگر شناخت تمام جامعه آماری است. آن دسته از اعضای جامعه آماری را که برای گردآوری داده‌ها انتخاب می‌شوند، «نمونه» می‌نامند. نمونه‌گیری، فرایند انتخاب نمونه را از جامعه به دست می‌دهد؛ به طوری که از داده‌های آن بتوان به اطلاعاتی «قابل قبول» از جامعه آماری دست یافت. واژهٔ «قابل قبول» بر این واقعیت تأکید دارد که نباید همان انتظاری را که از اطلاعات حاصل از سرشماری داریم، از اطلاعات حاصل از نمونه‌گیری نیز داشته باشیم، زیرا در هر صورت محدود شدن به نمونه‌ای از جامعه

آماری و شناخت جامعه از طریق داده‌های حاصل از این نمونه با «خطا» همراه است، ولی این خطا تا آن اندازه نیست که نتوان به اطلاعات حاصل از نمونه اطمینان کرد.

امروزه می‌توان ادعا کرد، با سازوکاری که نمونه‌گیری در اختیار پژوهشگران می‌گذارد و همچنین با کمک روش‌های آمار استنباطی، بویژه برآوردیابی، خطای اطلاعات حاصل از نمونه‌ها نگران‌کننده نیستند؛ به طوری که بسیاری از اطلاعات ما از جامعه‌های آماری تنها با داده‌های همین نمونه‌ها حاصل می‌شوند. در واقع، سرشماری از جامعه‌های بزرگ تنها به برخی متغیرهای پایه‌ای محدود می‌شود و از سوی یک سازمان دولتی خاص با بودجه و توان کافی مانند مرکز آمار صورت می‌گیرد؛ بنابراین، تقریباً تمام پژوهشگران با استفاده از نمونه‌گیری به گردآوری داده‌های مورد نیاز برای پژوهش خود اقدام می‌کنند.

هنگامی که پژوهشگر برای گردآوری داده‌های پژوهش خود به نمونه‌گیری رو می‌آورد، با سه پرسش اساسی زیر روبه‌روست که فرایند نمونه‌گیری باید به آنها پاسخ دهد:

۱. با توجه به متغیرهای پژوهش و «بافتِ» جامعه آماری، از چه روشی برای نمونه‌گیری استفاده می‌شود؟
۲. چه تعداد از اعضای جامعه آماری و چگونه به عنوان نمونه انتخاب شوند؟
۳. چگونه از داده‌های نمونه درباره جامعه استنباط کنیم؟

پاسخ پرسش نخست به روش نمونه‌گیری، پاسخ پرسش دوم به اندازه نمونه و سازوکار انتخاب افراد نمونه و سرانجام پاسخ پرسش سوم به برآوردیابی باز می‌گردد. این پاسخ‌ها موضوع این فصل را تشکیل می‌دهند.

پیش از ورود به بحث باید بین داده‌های مشاهده‌ای و داده‌های آزمایشی تمایز قایل شویم، زیرا نمونه‌گیری بیشتر در گردآوری داده‌های مشاهده‌ای نقش دارد. از فصل نخست به یاد دارید که داده، محصول اندازه‌گیری متغیرها از افراد، و متغیرها محصول روی‌دادن پدیده‌ها هستند. اگر پژوهشگر، خود شرایط روی‌دادن پدیده را فراهم کرده باشد، داده‌های حاصل، آزمایشی و در غیر این صورت، مشاهده‌ای نامیده می‌شود.

پژوهشگر در گردآوری داده‌های مشاهداتی تنها مشاهده‌گر (اندازه‌گیری‌کننده) است، مانند متغیر هزینهٔ ماهانه خانوار. پدیده یا پدیده‌هایی که باعث هزینه‌های مختلف خانوار می‌شوند، در اختیار پژوهشگر نیستند، بلکه مستقل از او روی می‌دهند و پژوهشگر تنها از میزان هزینهٔ ماهانه هر خانوار می‌پرسد و آن را ثبت می‌کند و به این ترتیب داده‌های هزینهٔ ماهانه را گردآوری می‌کند. بنابراین، داده‌های مشاهداتی محصول فعالیت‌هایی هستند که

بدون دخالت «مشاهده‌گر» رخ می‌دهند و او در آنها تنها نقشی انفعالی (اندازه‌گیرنده و ثبت‌کننده) دارد.

اگر پژوهشگر کم و بیش در وقوع پدیدهٔ منجر به ایجاد متغیر نقش داشته باشد، داده‌های حاصل از این متغیرها را آزمایشی می‌گویند. برای مثال، در مطالعهٔ تاثیر تماشای فیلم‌های خشونت‌آمیز بر رفتار پرخاشگرانهٔ کودکان، تعدادی کودک انتخاب و به دو دسته تقسیم می‌شوند؛ برای یک گروه فیلمی حاوی صحنه‌های خشونت‌آمیز و برای گروه دیگر فیلمی بدون صحنه‌های خشونت‌آمیز نمایش داده می‌شود. سپس میزان پرخاشگری کودکان هر دو گروه از طریق پرسشنامه‌ای استاندارد اندازه‌گیری می‌شود. در این مثال، پژوهشگر خود اقدام به گروه‌بندی افراد و نمایش فیلم برای هر گروه کرده و با این کار، پدیده‌ای مانند تماشای فیلم حاوی صحنه‌های پرخاشگری را «ایجاد کرده است» که می‌تواند به وقوع پرخاشگری در کودکان بینجامد. پس داده‌های مربوط به اندازه‌گیری پرخاشگری از کودکان، محصول فعالیتی است که «آزمایشگر» کم و بیش در ایجاد آن دخالت داشته است.

۲ـ نمونه‌گیری احتمالی و نااحتمالی

تمام تلاش‌های فرایند نمونه‌گیری در جهت انتخاب نمونه‌ای از جامعه آماری است که بتوان یافته‌های آن را به جامعه آماری تعمیم داد. چنین نمونه‌ای معرّف (نمایای) جامعه آماری است، زیرا توانسته اطلاعات قابل قبولی را درباره جامعه‌ای به دست دهد که از آن انتخاب شده است.

شرط لازم برای انتخاب نمونه‌ای معرّف از جامعه آماری، دادنِ امکان انتخاب به تمام اعضای جامعه است. این بدان معناست که باید تک‌تک اعضای جامعه بتوانند به عنوان یکی از افراد نمونه انتخاب شوند و هیچ‌یک در فرایند انتخاب افراد نمونه نادیده گرفته نشوند. بنابراین، ضروری است شرایطی در سازوکار انتخاب افراد نمونه فراهم شود که تمام اعضای جامعه آماری حضور داشته باشند و فردی از قلم نیفتد. چنین نمونه‌ای را نمونه‌گیری احتمالی و در غیر این صورت نااحتمالی می‌نامند.

برای مثال، فرض کنید ۵۰۰ نفر از دانشجویان متقاضی دریافت رایانهٔ همراه هستند، درحالی که دانشگاه تنها ۱۰ دستگاه رایانه در اختیار دارد. پس با مسئله انتخاب نمونه‌ای ۱۰ نفری از جامعه‌ای ۵۰۰ نفری روبه‌رو هستیم. قرعه‌کشی می‌تواند روشی «عادلانه» برای انتخاب این ۱۰ نفر باشد. اسم تک‌تک دانشجویان بر روی کارتی نوشته می‌شود، این

کارت‌ها به خوبی به‌هم زده می‌شوند و بُر می‌خورند و در نهایت بدون آنکه انتخاب‌کننده، اسم روی کارت‌ها را ببیند، ۱۰ کارت را از میان ۵۰۰ کارت انتخاب می‌کند. در این صورت هر یک از این ۵۰۰ دانشجو می‌تواند به عنوان فردی انتخاب شود که یکی از ۱۰ رایانه را دریافت می‌کند؛ هرچند درعمل ۴۹۰ نفر از دریافت رایانه محروم می‌شوند. پس قرعه‌کشی سازوکاری است که امکان انتخاب را به همگان می‌دهد.

ممکن بود دانشگاه این ۱۰ رایانه را به ۱۰ نفر نخستی بدهد که پیش از دیگران تقاضا کرده‌اند. پس برخلاف حالت قرعه‌کشی، ۴۹۰ نفر متقاضی دیگر «هیچ» امکانی ندارند تا از دریافت‌کنندگان یکی از ۱۰ رایانه باشند. همچنین، ممکن بود دانشگاه در یک روز چهارشنبه اعلان کند که دانشجویان متقاضی رایانه به دفتر مربوط مراجعه و رایانهٔ خود را دریافت کنند، پس ۱۰ نفری که از این اعلان زودتر با خبر شوند و سریع‌تر خود را به این دفتر برسانند، صاحب یک دستگاه رایانه خواهند بود و افرادی که در روز چهارشنبه در دانشگاه نیستند یا دیرتر از این موضوع اطلاع یافته‌اند، از دریافت این رایانه‌ها محروم می‌شوند. این دو وضعیت مثال‌هایی نااحتمالی از انتخاب ۱۰ نفر از میان ۵۰۰ نفر است، زیرا در هیچ‌یک از این دو مثال، تمام ۵۰۰ نفر امکان انتخاب شدن به عنوان یکی از ۱۰ دریافت‌کنندهٔ رایانه را ندارند.

انتخاب نمونه‌ای احتمالی، به فهرستی از تمام اعضای جامعه‌ای نیاز دارد که نمونه از میان آنها انتخاب می‌شود. قرعه‌کشی ۱۰ نفر از میان ۵۰۰ نفر در صورتی ممکن است که برای مثال بتوان نام تک‌تک آنها را بر روی کارت‌های جداگانه‌ای نوشت و ۱۰ کارت را از میان این کارت‌ها انتخاب کرد. پس باید فهرستی از ۵۰۰ نفر متقاضی رایانه در اختیار باشد تا نوشتن نام آنها بر روی کارت‌ها امکان‌پذیر شود. این فهرست در بحث نمونه‌گیری، چهارچوب نمونه‌گیری نامیده می‌شود. توجه کنید که یک چهارچوب مناسب باید از یک سو کامل باشد، یعنی هیچ عضوی از جامعه آماری را از قلم نیندازد و از سوی دیگر شامل افراد خارج از جامعه آماری یا تکراری نباشد.

علاوه بر چهارچوب، رعایت انتخاب «تصادفیِ» افراد نمونه، از اصول نمونه‌گیری احتمالی است؛ به طوری که نمونه‌گیری احتمالی را نمونه‌گیری تصادفی نیز می‌نامند. در واقع، انتخابی تصادفی است که به تمام اعضای جامعه امکان انتخاب شدن را می‌دهد. برای مثال، تنها در صورتی که قرعه‌کشی بر اساس چهارچوب مناسبی شکل‌گرفته باشد، شرایط

مناسب برای انتخاب تصادفی افراد نمونه فراهم می‌شود، زیرا نه تنها به دلیل مناسب بودن چهارچوب، تمام اعضای جامعه در انتخاب شرکت دارند، بلکه با بُر و به هم زدن کارت‌ها هیچ‌گونه جهت‌گیری خاصی در انتخاب افراد صورت نمی‌گیرد.

به یاد داشته باشید که انتخاب «تصادفی» با انتخاب «اتفاقی» تفاوت دارد. انتخاب اتفاقی به طور ناگهانی و بدون برنامه‌ریزی قبلی صورت می‌گیرد، درحالی که انتخاب تصادفی با تهیهٔ چهارچوب و بدون جهت‌گیری خاصی در انتخاب افراد نمونه از چهارچوب اجرا می‌شود. برای مثال، اگر دانشگاه ۱۰ رایانه را بدون اعلان قبلی صرفاً به ۱۰ دانشجویی بدهد که برای کار اداری در روز سه‌شنبه به آموزش دانشگاه مراجعه کرده‌اند، در این صورت دریافت‌کنندگان رایانه کاملاً اتفاقی انتخاب شده‌اند نه تصادفی، زیرا در انتخاب تصادفی باید همهٔ متقاضیان دریافت رایانه امکان انتخاب شدن داشته باشند، درحالی که در این انتخاب اتفاقی این امکان تنها به دانشجویانی داده شده است که در روز سه‌شنبه به آموزش دانشگاه مراجعه کرده‌اند. پس نمونه‌گیری اتفاقی نمی‌تواند احتمالی باشد.

۳ـ روش‌های نمونه‌گیری نااحتمالی

اگرچه تأکید این فصل بر روش‌های نمونه‌گیری احتمالی است، زیرا قابل تعمیم به جامعه آماری هستند، ولی در این قسمت مروری بر روش‌های نمونه‌گیری نااحتمالی نیز خواهیم داشت تا خواننده با شناسایی این روش‌ها در عمل دچار استفادهٔ ناخواسته و ناآگاهانه از روش‌های نااحتمالی نشود. با وجود این، پژوهش‌هایی را می‌توان یافت که به دلیل محدودیت‌های مالی، زمانی یا دشواری‌های نمونه‌گیری احتمالی به استفاده از نمونه‌گیری نااحتمالی روی می‌آورند[1].

۱ـ۳ـ نمونه‌گیری در دسترس

نمونه‌گیری در دسترس یا سهل‌الوصول ساده‌ترین، ارزان‌ترین و سریع‌ترین روش نمونه‌گیری نااحتمالی است. پژوهشگر آن دسته از اعضای جامعه آماری را برای انتخاب نمونه درنظر می‌گیرد که به‌راحتی می‌تواند به سراغ آنها برود و متغیرهایش را اندازه بگیرد.

نمونه‌گیری اتفاقی را می‌توان نوعی از نمونه‌گیری در دسترس قلمداد کرد که پژوهشگر افراد نمونه را کاملاً اتفاقی و بدون تهیهٔ چهارچوب و رعایت اصل تصادفی‌بودن انتخاب

---

۱. البته استفاده از نمونه‌گیری نااحتمالی در برخی از پژوهش‌های کیفی بر نمونه‌گیری احتمالی اولویت دارد.

می‌کند. برای مثال، گزارشگری برای شناخت افکار عمومی درباره رخداد سیاسی خاصی به یکی از میادین پرتردد شهر مراجعه می‌کند و با درخواست از رهگذران، نظر آنان را درباره آن رخداد سیاسی جویا می‌شود. آیا می‌توان نظر این عده را معرّف دیدگاه‌های عموم مردم آن شهر به حساب آورد؟ پاسخ منفی است؛ زیرا آنها بر حسب اتفاق و تنها از محدودهٔ خاصی از شهر انتخاب شده‌اند که برای آن گزارشگر به‌راحتی قابل دسترس بوده است. بنابراین، بخش قابل توجهی از مردم شهر نادیده گرفته می‌شوند؛ برای مثال آنهایی که در خانه یا محل کار بوده یا در آن لحظه از معابر دیگر شهر عبور کرده‌اند.

نمونه‌گیری داوطلبانه نیز نوع دیگری از نمونه‌گیری در دسترس است. پژوهشگر افرادی را به عنوان نمونه انتخاب می‌کند که علاقه‌مندی خود را به مشارکت در پژوهش اعلام کرده‌اند، یعنی داوطلبِ حضور در نمونه شده‌اند. برای مثال، فرض کنید دانشگاهی در پی ارزیابی عملکرد آموزشی خود باشد و برای این منظور از دانشجویان دعوت می‌کند در زمان مشخصی برای تکمیل پرسشنامه‌ای در این زمینه در سالن اجتماعات دانشگاه جمع شوند. کسانی که در این زمان مشخص حضور می‌یابند، همان علاقه‌مندانی (داوطلبانی) هستند که به مشارکت در ارزیابی دانشگاه تمایل دارند. آیا این گروه را می‌توان نمونه معرّفی از تمام دانشجویان آن دانشگاه دانست؟ پاسخ باز هم منفی است؛ زیرا ممکن است برنامهٔ بخشی از دانشجویان به شکلی باشد که در آن روز به دانشگاه نمی‌آیند یا آن‌قدر حوصله و انگیزه ندارند که به سالن اجتماعات مراجعه کنند و برای این ارزیابی وقت بگذارند.

نمونه‌گیری در دسترس، بویژه نوع داوطلبانهٔ آن، می‌تواند ناآگاهانه به سوگیری جدی در اطلاعات بینجامد. شاید نمونه از افرادی تشکیل شود که عاشق مطرح شدن و اظهار نظر کردن هستند و از همین رو به شرکت در پژوهش علاقه نشان می‌دهند. در این صورت نمونه پژوهشگر شامل افرادی با طرز فکری خاص است که می‌تواند به اطلاعاتی غیرواقعی و بسیار دور از اطلاعات جامعه آماری منجر شود. همچنین ممکن است در نمونه‌گیری اتفاقی، پژوهشگر در منطقه‌ای به گردآوری افراد نمونه بپردازد که برای مثال کارگران عبور می‌کنند. به این ترتیب اکثر افراد نمونه او را طبقهٔ خاصی از جامعه تشکیل می‌دهند که باز می‌تواند به سوگیری اطلاعات حاصل از نمونه منتهی شود.

۲ـ۳ـ نمونه‌گیری هدفمند

پژوهشگر در نمونه‌گیری هدفمند یا قضاوتی، به‌عمد افراد خاصی را به عنوان نمونه انتخاب می‌کند. اغلب، این انتخاب با توجه به نظر شخصی و دیدگاه او نسبت به این افراد صورت می‌گیرد. برای مثال، دانشگاه می‌تواند برای ارزیابی عملکرد خود تنها به سراغ دانشجویان زرنگ دانشگاه برود و سایرین را نادیده بگیرد (آیا نمونه حاصل معرّف جامعه دانشجویان آن دانشگاه است؟).

نمونه‌گیری هدفمند می‌تواند در برخی از موارد، زمینه‌ای برای گردآوری افرادی با دیدگاه‌های ضد و نقیض فراهم کند تا به این ترتیب، طیف گسترده و متنوعی از دیدگاه‌ها درباره پدیدهٔ تحت مطالعه به دست آید. بنابراین، اگر در مرحله‌ای از پژوهش به دنبال طراحی پرسشنامهٔ جامعی هستید، دریافت دیدگاه‌های یک نمونه هدفمندِ کوچک می‌تواند بسیار راهگشا باشد. از همین رو، پژوهش‌هایی که مبتنی بر روش مصاحبه، روش کیو یا گروه‌های متمرکز[1] هستند، از این روش نمونه‌گیری استفاده می‌کنند.

برخی از پژوهشگران با اطلاعات اولیه‌ای که از افراد مختلف در اختیار دارند، نمونه به نسبت بزرگی از آنها تهیه می‌کنند تا مبنای گردآوری داده‌های پژوهشی آنان باشد. این پژوهشگران چنین نمونه‌ای را جامعه‌ای مینیاتوری (تصویری کوچک‌شده) از جامعه آماری پژوهش خود می‌دانند که می‌تواند اطلاعات همان جامعه را در خود منعکس سازد.

نمونه‌گیری هدفمند کاملاً به قضاوت پژوهشگر بستگی دارد و همین امر می‌تواند تردیدهایی را نسبت به سوگیری و پایایی نمونه ایجاد کند. همواره این پرسش مطرح است که قضاوت پژوهشگر تا چه حد پایدار است؟ اگر دوباره از او خواسته شود نمونه‌ای هدفمند برای همان پژوهش انتخاب کند، باز نمونه‌ای مشابه با نمونه پیشین انتخاب خواهد کرد یا ممکن است نمونه جدید تفاوت‌های چشمگیری با نمونه قبلی داشته باشد؟ شاید تفاوت این دو نمونه چنان بزرگ باشد که به اطلاعات کاملاً متفاوتی از جامعه منجر شود.

۳ـ۳ـ نمونه‌گیری گلوله‌برفی

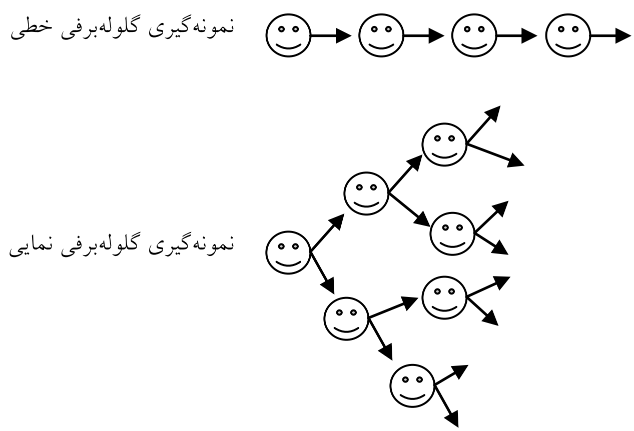
هنگامی که پژوهشگر با اندازه‌گیری متغیرهای حساس روبه‌روست یا پژوهش او در یک جامعه آماری بسیار کوچک، پنهان یا نادر صورت می‌گیرد، از نمونه‌گیری گلوله برفی استفاده

1. Q-method or Focus Groups

می‌کند. برای مثال، اجرای پژوهش درباره معتادان، افراد مبتلا به برخی بیماری‌های روانی نظیر خودآزاری جنسی، هکرها و... با روش نمونه‌گیری گلوله‌برفی امکان‌پذیر است.

انتخاب افراد نمونه در ساده‌ترین شکل می‌تواند با یک فرد آغاز شود که پژوهشگر او را می‌شناسد یا از سوی کسی به او معرفی می‌شود (شکل ۱ را ببینید). به این ترتیب، این فرد به پژوهشگر اعتماد می‌کند و حاضر می‌شود در پژوهش او شرکت نماید. همین فرد، فرد دیگری مانند خود را به پژوهشگر معرفی می‌کند و به همین ترتیب، فرد دوم، فرد سومی را به پژوهشگر معرفی می‌کند تا افراد نمونه به تعداد مورد نیاز پژوهشگر گردآوری شوند. برای مثال، پژوهشگر به معتادی که او را می‌شناسد، مراجعه می‌کند، سپس این معتاد یکی از دوستان معتاد خود را به پژوهشگر معرفی می‌کند و این معتاد دوم نیز پژوهشگر را با فرد دیگری که از دوستان معتاد خود اوست، آشنا می‌کند تا سومین فرد نمونه نیز به دست آید و به همین ترتیب کار گردآوری افراد نمونه زنجیره‌وار پیش می‌رود.

شکل ۱: نمونه‌گیری گلوله‌برفی خطی و نَمایی



اگرچه نمونه‌گیری گلوله‌برفی دسترسی به افراد مورد مطالعه را آسان می‌سازد، ولی پژوهشگر اختیارِ کمی بر جریان انتخاب افراد نمونه دارد. بنابراین بر حسب اینکه نمونه با چه فرد یا افرادی آغاز شود، افراد نمونه گلوله‌برفی می‌توانند به قشر خاصی تعلق داشته باشند، زیرا معمولاً هر فردِ نمونه، فردی را به پژوهشگر معرفی می‌کند که با او نزدیک و صمیمی است، پس انتظار می‌رود زنجیرهٔ افراد نمونه هم سنخ، هم‌فکر و هم‌رفتار باشند و به این ترتیب پژوهشگر نتواند به نمونه‌ای «متنوع» از افراد جامعه تحت مطالعه دست یابد. بنابراین نمونه‌گیری گلوله‌برفی با سوگیری افراد نمونه تهدید می‌شود.

۴ـ۳ـ نمونه‌گیری سهمیه‌ای

نمونه‌گیری سهمیه‌ای صرفاً تلاشی است تا ترکیب نمونه از برخی از ابعاد به ترکیب جامعه آماری آن شبیه باشد. در این نمونه‌گیری نیز انتخاب افراد نمونه با استفاده از یکی از روش‌های پیشگفته صورت می‌گیرد، ولی پژوهشگر خود را ملزم می‌داند که ترکیب نمونه را از برخی ابعاد مانند ترکیب جامعه آماری حفظ کند.

فرض کنید دانشگاهی که ۷۰ درصد از دانشجویان آن در مقطع کارشناسی و ۳۰ درصد آنان در مقطع کارشناسی ارشد به تحصیل مشغولند، از دانشجویان برای ارزیابی عملکرد خود برای حضور در سالن اجتماعات دعوت می‌کند. پس تنها داوطلبان در این ارزیابی مشارکت دارند. اگر دانشگاه ۷۰ درصد از نمونه پژوهش خود را از بین داوطلبان مقطع کارشناسی و ۳۰ درصد را از میان داوطلبان مقطع کارشناسی ارشد انتخاب کند، نمونه‌ای سهمیه‌ای برای پژوهش گردآوری کرده است، زیرا ترکیب نمونه به لحاظ مقطع تحصیلی درست مشابه ترکیب جامعه دانشجویان دانشگاه است.

۴ـ روش‌های کلاسیک نمونه‌گیری احتمالی

تاکنون باید خواننده از ضعف‌های روش‌های نمونه‌گیری نااحتمالی، به اهمیت روش‌های احتمالی در پژوهش پی برده باشد. در این قسمت درباره سه روش نمونه‌گیری تصادفی ساده، طبقه‌ای و خوشه‌ای بحث خواهد شد. این سه روش را می‌توان از قدیمی‌ترین روش‌های نمونه‌گیری به حساب آورد که پایهٔ روش‌های احتمالیِ پیچیده‌تر نیز هستند.

۱ـ۴ـ نمونه‌گیری تصادفی ساده

روش نمونه‌گیری تصادفی ساده را باید مادر روش‌های نمونه‌گیری احتمالی (تصادفی) دانست. این روش فرض می‌کند، چهارچوبی از جامعه آماری پژوهش در دست است که شامل فهرستی از تمام اعضای این جامعه همراه با نشانی آنها یا شیوه‌ای برای تماس با آنهاست. پژوهشگر مایل است نمونه‌ای از این جامعه انتخاب کند تا بر اساس داده‌های آن به استنباط‌های آماری درباره جامعه دست بزند. از این رو از فرایندی مانند قرعه‌کشی برای انتخاب نمونه استفاده می‌شود.

برای قرعه‌کشی، نام هر فرد بر روی کارتی نوشته می‌شود و پس از به هم زدن کارت‌ها، بدون آنکه نام روی کارت‌ها مشاهده شود، تعدادی کارت به عنوان نمونه انتخاب می‌گردد. درعمل به جای نوشتن نام افراد بر روی کارت، به هر فرد، یک شماره نسبت داده می‌شود. برای مثال برای انتخاب نمونه‌ای ۴۰۰ تایی از جامعه‌ای با صد هزار جمعیت، شماره‌ای از ۱ تا ۱۰۰۰۰۰ به هر فرد نسبت داده می‌شود (توجه کنید که نباید دو فرد از جامعه، شماره یکسانی داشته باشند یا فردی دارای دو شماره باشد). سپس به کمک یک ماشین‌حساب یا رایانه، ۴۰۰ عدد تصادفی بین ۱ تا ۱۰۰۰۰۰ تولید می‌شود[1]. این اعداد شماره افرادی را مشخص می‌کنند که باید به عنوان نمونه از میان صد هزار نفر انتخاب شوند. اکنون با مراجعه به چهارچوب، نشانی این افراد تعیین می‌گردد و با مراجعه به آن نشانی، به فرد نمونه دسترسی پیدا خواهیم کرد.

روش نمونه‌گیری تصادفی ساده، شانس برابری را به همه اعضای جامعه آماری برای انتخاب شدن به عنوان یکی از افراد نمونه می‌دهد تا هیچ‌یک نسبت به دیگری امکان بیشتری برای انتخاب شدن نداشته باشد[2]. این قابلیت، به انتساب تنها یک کارت به هر عضو و استفاده از قرعه‌کشی برای انتخاب افراد باز می‌گردد. نمونه‌گیری تصادفی ساده برای جامعه‌هایی مانند یک دانشگاه یا سازمان مناسب است که فهرستی از اعضای آنها وجود دارد و چگونگی دسترسی به آنها ممکن و به نسبت آسان است. توجه کنید که لزومی ندارد چهارچوب، شامل نشانی هر عضو جامعه باشد، بلکه اگر بتوان از طریق ای‌میل یا تلفن با تک‌تک اعضای جامعه تماس گرفت، آنگاه می‌توان از ای‌میل یا شماره تلفن آنان به جای نشانی استفاده کرد.

---

۱. پیش از اختراع ماشین‌حساب و رایانه از «جدول اعداد تصادفی» برای انتخاب شماره‌ها به شکلی تصادفی استفاده می‌شد. چنین جدولی شامل مجموعه‌ای از رقم‌های مختلف بود که به شکلی کاملاً تصادفی در قالب سطرها و ستون‌های این جدول جای گرفته بودند. پژوهشگر می‌توانست از تقاطع سطر و ستون دلخواهی که انتخاب کرده بود، به یک شماره کاملاً تصادفی دست یابد. البته ظهور ماشین‌حساب‌ها و رایانه‌هایی که خود به تعداد دلخواهِ پژوهشگر عدد تصادفی ارائه می‌کنند، پژوهشگران را از مراجعه به این جدول‌ها بی‌نیاز کرده است.

۲. باجایگذاری یا بدون‌جایگذاری بودن انتخاب‌ها در این امر تأثیر دارد، ولی در جوامع پرجمعیت می‌توان از آن تأثیر صرفنظر کرد. مراد ما در سرتاسر این فصل، نمونه‌گیری بدون جایگذاری است.

۱ـ۱ـ۴ـ اندازه نمونه

اکنون به موضوع تعیین اندازه نمونه در روش نمونه‌گیری تصادفی ساده می‌پردازیم. برای این منظور فرض کنید پژوهشگری در پی برآورد میانگین توزیع متغیر X برای جامعه آماری با N عضو از طریق داده‌های یک نمونه است. او از خود می‌پرسد، چه تعداد نمونه باید انتخاب شود تا با اطمینان معینی خطای استفاده از میانگین این نمونه به عنوان برآورد میانگین جامعه در حدی قابل قبول باشد؟ به یاد بیاورید که خطای برآورد برابر با اختلاف مقدار میانگینِ نمونه $\overline{X}$ از میانگینِ نامعلوم جامعه μ است، یعنی $\overline{X}-\mu$. بنابراین، پژوهشگر مایل است $\overline{X}-\mu$ درحدی باشد که بتواند از $\overline{X}$ به عنوان برآورد μ استفاده کند. اگر d حداکثر اختلافی (خطایی) باشد که پژوهشگر حاضر است در برآورد خود تحمل کند، آنگاه اندازه نمونه‌ای که $(۱-\alpha)\times ۱۰۰$درصد تضمین می‌کند که خطای برآورد μ با $\overline{X}$ از d تجاوز نمی‌کند، از رابطهٔ زیر به دست می‌آید:

$$n=\frac{Nz_{\frac{\alpha}{۲}}^{۲}\,\sigma_x^{۲}}{Nd^{۲}+z_{\frac{\alpha}{۲}}^{۲}\,\sigma_x^{۲}}$$

در رابطهٔ فوق، $\sigma_X^{۲}$ واریانس متغیر X، N تعداد اعضای جامعه آماری و $z_{\frac{\alpha}{۲}}$ چندک توزیع نرمال استاندارد است که از جدول پیوست به دست می‌آید.

برای مثال فرض کنید پژوهشگر می‌خواهد میانگین مدت زمانی را برآورد کند که ساکنان شهری با $N=۱۰۰۰۰$ شهروند، در انتظار وسایل حمل‌ونقل عمومی صرف می‌کنند. اگر او بخواهد ۹۵ درصد اطمینان داشته باشد که برآورد او از نمونه بیشتر از $d=۵$ ثانیه با مقدار واقعی اختلاف نداشته باشد، آنگاه به چه تعداد نمونه نیاز دارد؟ از آنجا که $(۱-۰/۰۵)\times ۱۰۰=۹۵$ پس $\alpha=۰/۰۵$ و $z_{۰/۰۲۵}=۱/۹۶$ است. اگر بدانیم واریانس مدت زمان انتظار در این جامعه آماری $\sigma_X^{۲}=۵۴۰$ است، در این صورت اندازه نمونه لازم برای این برآورد عبارت است از:

$$n=\frac{۱۰۰۰۰\times ۱/۹۶^{۲}\times ۵۴۰}{۱۰۰۰۰\times ۵^{۲}+۱/۹۶^{۲}\times ۵۴۰}=۸۱/۴۶$$

در عمل باید حداقل ۸۲ نفر با قرعه‌کشی انتخاب شوند تا با ۹۵ درصد اطمینان تضمین شود خطای برآورد میانگین مدت زمان انتظار با داده‌های این نمونه بیش از ۵ ثانیه نیست.

اندازه نمونه با توجه به رابطهٔ پیشین به چهار مؤلفه بستگی دارد: واریانس متغیر (σ_x^2)، ضریب اطمینان ($z_{\frac{\alpha}{2}}$)، خطای قابل پذیرش (d) و جمعیت جامعه آماری (N).

۱. واریانس، میزان تنوع مقادیر متغیر را در جامعه آماری نشان می‌دهد، پس هرچه واریانس بزرگ‌تر باشد، تنوع بیشتری نسبت به مقدار متغیر در جامعه وجود دارد، بنابراین باید نمونه بیشتری انتخاب کرد تا تنوع جامعه بهتر در آن منعکس شود. بنابراین واریانس رابطهٔ مستقیم با اندازه نمونه دارد.

۲. ضریب اطمینان، میزان اطمینان به کنترل خطای برآورد را در همان سطح قابل تحمل برای پژوهشگر نشان می‌دهد. هرچه پژوهشگر بخواهد با اطمینان بیشتری این خطا را کنترل کند، به تعداد نمونه بیشتری نیاز دارد. پس ضریب اطمینان نیز رابطهٔ مستقیم با اندازه نمونه دارد.

۳. به طور معمول، خطای برآورد از محدود شدن به نمونه‌ای از جامعه (و نه کل آن) ناشی می‌شود. بنابراین، اگر پژوهشگر مایل باشد فاصلهٔ بین نمونه و جامعه کمتر شود و برآوردی نزدیک‌تر به مقدار واقعی ارائه کند، باید تعداد بیشتری از اعضای جامعه را به عنوان نمونه انتخاب کند. به این ترتیب، پذیرش خطای بیشتر در برآورد به تعداد نمونه کمتر می‌انجامد، پس بین خطا و اندازه نمونه رابطهٔ معکوس وجود دارد.

۴. جمعیتِ جامعه آماری نیز با اندازه نمونه رابطهٔ مستقیم دارد، زیرا طبیعی است که انتظار داشته باشیم تعداد نمونه جامعه پرجمعیت‌تر، بیشتر باشد. گفتنی است سهم جمعیت جامعه آماری نسبت به سه مؤلفه دیگر اندک است، به طوری که می‌توان گفت تأثیر این مؤلفه برای جمعیت‌های بزرگ برای مثال بیش از ۵۰۰۰، قابل چشم‌پوشی است، مشروط بر آنکه سایر مؤلفه‌ها تغییری نکنند. در واقع، برای جامعه‌هایی با جمعیت بزرگ، به جای رابطهٔ پیشگفته، از رابطهٔ ساده‌تر زیر برای تعیین اندازه نمونه استفاده می‌شود که در آن جمعیت هیچ نقشی ندارد.

$$n = \frac{z_{\frac{\alpha}{2}}^2}{d^2}\sigma_x^2$$

بر اساس رابطهٔ پیشین، تعداد نمونه برای مثال قبل، بـه ازای $z_{۰/۰۲۵} = ۱/۹۶$، $\sigma_X^۲ = ۵۴۰$ و $d = ۵$ برابر است با:

$$n = \frac{۱/۹۶^۲}{۵^۲} \times ۵۴۰ = ۸۳$$

همان‌طور که دیده می‌شود چشم‌پوشی از جمعیت در محاسبه اندازه نمونه به رقم ۸۳ منجـر می‌شود، در حالی که با احتساب جمعیت، اندازه نمونه برابر با ۸۲ است. پس تأثیر جمعیت، تنها به اختلاف ۱ واحد منتهی شده است. این موضوع دور از انتظار نبود، زیرا جامعـه‌ای بـا ۱۰۰۰۰ جمعیت را باید پرجمعیت قلمداد کرد. پژوهشگر بایـد توجـه کنـد کـه بـه‌حسـاب آوردن جمعیت به کاهش اندازه نمونه منجر می‌شود. بنـابراین، نادیـده‌گـرفتن آن خللـی بـه کیفیت برآوردهای حاصل از نمونه وارد نمی‌کند، بلکه هزینـه‌ای اضـافی صـرف گـردآوری نمونه‌های بیشتر می‌شود.

پرسشی که باید تاکنون ذهن خواننـده را درگیـر کـرده باشـد، بـه مقـدار واریـانس بـاز می‌گردد؛ زیرا اغلب، واریانس متغیر تحت بررسی در جامعه آماری، خود پارامتری نـامعلوم است. به این ترتیب یکی از مؤلفه‌های تعیین اندازه نمونه را در اختیار نداریم تا اندازه نمونه را از رابطه‌های پیشگفته محاسبه کنیم. دو راهکار زیر برای این مسئله پیش روی پژوهشـگر است:

۱. استفاده از اطلاعات پیمایش‌های مشابه: اگر در گذشتهٔ نه چنـدان دور، متغیـر تحـت بررسی در پیمایشی با جامعه آماری به نسبت مشابه مطالعه شده باشد، از بـرآورد واریـانس این متغیر در آن پیمایش در رابطهٔ تعیین اندازه نمونه استفاده می‌شود.

۲. پژوهشگر می‌تواند نمونه کوچکی، برای مثال ۲۰ تایی، انتخـاب کنـد و بـرآوردی از واریانس متغیر بر اساس داده‌های این نمونه کوچک به دست آورد تا در رابطهٔ تعیین انـدازه نمونه به کار بگیرد.

رابطه‌هایی که برای تعیین اندازه نمونه در روش نمونه‌گیری تصادفی ساده معرفـی شـد، برای برآورد «میانگین» توزیع متغیر مناسب هستند، ولی ممکن است پژوهشگر با متغیرهای مقوله‌ای سروکار داشته باشد و بخواهد پارامتری مانند «درصـد» را بـرآورد کنـد. در چنـین حالتی نیز می‌توان از همان رابطه‌ها استفاده کرد، با این تفاوت کـه بایـد $\sigma_X^۲ = p(۱-p)$ را

در این رابطه‌ها به کار بگیرد (p همان نسبتی[1] است که پژوهشگر در پی برآورد آن اســت!). درعمل باید p را با یکی از دو روشی که در بالا اشاره شد، برآورد کنید تا به تخمینی از σ_X^2 دست یابید، ولی معمولاً $p=0/5$ اختیار می‌شود و به این ترتیـب $\sigma_X^2=0/25$ بـه دسـت می‌آید. اندازه نمونه‌ای که از به کارگیری $\sigma_X^2=0/25$ حاصل می‌شود، «بزرگ‌تـرین» تعـداد نمونه‌ای است که واقعاً لازم است. بنابراین پژوهشگر می‌تواند اطمینان داشته باشد که حتـی اگر مقـدار واقعـی p را هـم مـی‌دانسـت، بـه انـدازه نمونـه‌ای بـزرگ‌تـر از آنچـه بـه ازای $\sigma_X^2=0/25$ به دست می‌آید، نمی‌رسید. به این ترتیب، اسـتفاده از $\sigma_X^2=0/25$ خللـی بـه کیفیت برآوردهای حاصل از نمونه وارد نمی‌کند و تنها ممکن است هزینه‌ای اضافی صـرف گردآوری نمونه‌های بیشتر شود.

۲ـ۱ـ۴ـ برآورد میانگین و نسبت و خطای معیار آنها

در نمونه‌گیری تصادفی ساده، میانگین حاصل از نمونه n تایی یعنی $\overline{X}=\frac{1}{n}\sum X_i$ به عنوان برآوردگر نااریب میانگین جامعه آماری به کار می‌رود. انحراف معیار (جذر واریـانس) ایـن برآوردگر که خطای معیار آن نیز نامیده می‌شود، به صورت زیر محاسبه می‌گردد:

$$SE_{\overline{X}}=\sqrt{(1-\tfrac{n}{N})\tfrac{\sigma_X^2}{n}}$$

σ_X^2 واریانس (نامعلوم) متغیر X است که در عمل از واریانس داده‌های نمونه n تایی یعنـی S_X^2 به عنوان تخمینِ آن استفاده می‌شود. $\frac{n}{N}$ را کسـرِ نمونـه‌گیـری مـی‌نامنـد؛ زیـرا نشـان می‌دهد چه نسبتی از جامعه به عنوان نمونه انتخاب شده است. اگر با جامعه‌های پرجمعیت روبه‌رو باشیم، $\frac{n}{N}$ تقریباً برابر با صفر است و می‌توان آن را نادیده گرفت. در ایـن صـورت $SE_{\overline{X}}=\sqrt{\frac{S_X^2}{n}}$ که همانند انحراف معیار $\overline{X}$ از قضـیهٔ حـد مرکـزی اسـت (بـه فصـل سـوم مراجعه کنید). خطای معیار نسبت نیز از رابطهٔ زیر محاسبه می‌شود:

$$SE_p=\sqrt{(1-\tfrac{n}{N})\tfrac{p(1-p)}{n-1}}$$

در جامعه‌های پرجمعیت، این کمیت نیز به شکل سادهٔ $SE_p=\sqrt{\frac{p(1-p)}{n-1}}$ تبدیل می‌شود کـه در فصل چهارم معرفی شد.

---

۱. به یاد بیاورید که درصد از ضرب «نسبت» در عدد ۱۰۰ به دست می‌آید. در واقع نسبت، عـددی از ۰ تـا ۱ اسـت، درحالی که درصد عددی از ۰ تا ۱۰۰ است.

برای مثال، فرض کنید نمونه‌ای $n = ۶۰۰$ تایی از جامعه‌ای $N = ۲۰۰۰۰$ عضوی انتخاب شده است و میانگین و واریانس مدت زمان انتظار برای وسایل حمل‌ونقل عمومی از داده‌های این ۶۰۰ نمونه، به ترتیب برابر با $\overline{X} = ۵/۶$ دقیقه و $S_X^۲ = ۱۰$ است. پژوهشگر می‌خواهد با توجه به این اطلاعات، برآوردی فاصله‌ای با اطمینان ۹۵ درصد برای میانگین مدت زمان انتظار جامعه ارائه کند. از فصل چهارم به یاد بیاورید که حد بالا و پایین برآورد فاصله‌ای میانگین به صورت زیر است:

$$LL = \overline{X} - t_{df,\frac{\alpha}{۲}} SE_{\overline{X}} \qquad UL = \overline{X} + t_{df,\frac{\alpha}{۲}} SE_{\overline{X}}$$

پس باید $SE_{\overline{X}}$ به صورت زیر محاسبه شود:

$$SE_{\overline{X}} = \sqrt{(۱ - \frac{۶۰۰}{۲۰۰۰۰})\frac{۱۰}{۶۰۰}} = ۰/۱۲۷$$

از جدول چندک‌های توزیع t در پیوست به ازای $\frac{\alpha}{۲} = ۰/۰۲۵$ و $df = n - ۱ = ۵۹۹$، مقدار $t_{۵۹۹, ۰/۰۲۵} = ۱/۹۶۲$ به دست می‌آید (در این جدول، عدد ۱۰۰۰ نزدیک‌ترین مقدار d به ۵۹۹ است). بنابراین حدهای بالا و پایین برآورد فاصله‌ای عبارتند از:

$$LL = ۵/۶ - ۱/۹۶۲ \times ۰/۱۲۷ = ۵/۳۵۰ \qquad UL = ۵/۶ + ۱/۹۶۲ \times ۰/۱۲۷ = ۵/۸۴۹$$

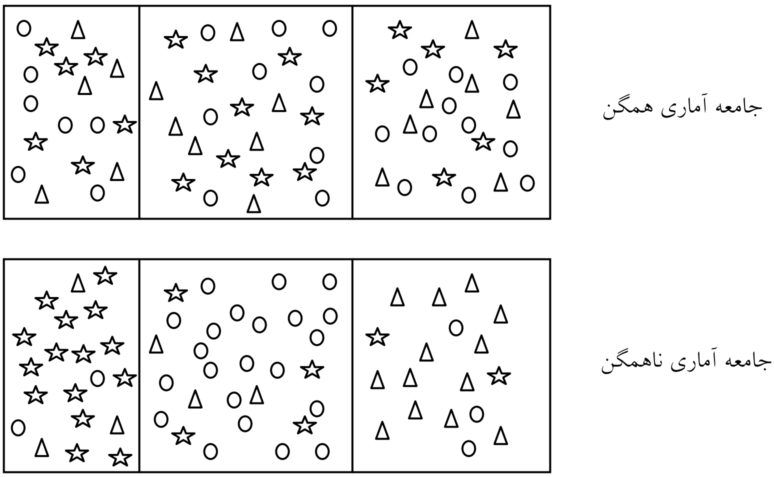
۲ـ۴ـ نمونه‌گیری طبقه‌ای

نمونه‌گیری تصادفی ساده با این پیش‌فرض به‌کار گرفته می‌شود که جامعه آماری نسبت به متغیر تحت بررسی یکدست و همگِن است، بنابراین نمونه از طریق قرعه‌کشی از فهرست کل اعضای جامعه آماری انتخاب می‌شود. وضعیت‌های مختلف متغیر در جامعه آماری همگن به‌طور یکدست در سراسر جامعه پراکنده شده است. بنابراین نمی‌توان طبقه‌ای (قشری) از جامعه را یافت که به وضعیت خاصی از متغیر گرایش بیشتری داشته باشد، بلکه تمام طبقات کم و بیش شامل وضعیت‌های گوناگون متغیر هستند. در مقابل، تفاوت طبقات در جامعه آماری ناهمگن کاملاً آشکار است. در واقع، جامعه آماری از طبقاتی تشکیل می‌شود که هر طبقه به وضعیت خاصی از متغیر گرایش بیشتری دارند.

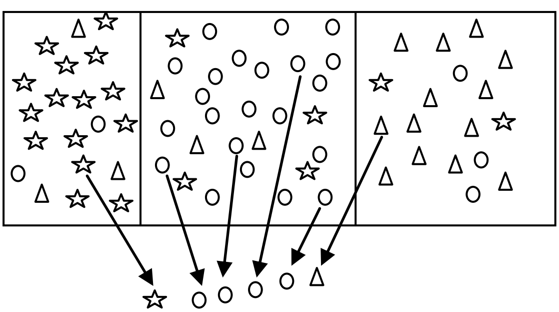
تفاوت جامعه آماری همگن و ناهمگن در شکل ۲ به خوبی مشاهده می‌شود. جامعه ناهمگن شامل سه طبقه است؛ طبقه‌ای به وضعیت مثلث، طبقه‌ای به وضعیت دایره و

طبقه‌ای به وضعیت ستاره گرایش بیشتری دارد. در جامعه همگن چنین تمایزی دیده نمی‌شود. تک‌تک طبقات هم شامل مثلث، هم دایره و هم ستاره هستند. در واقع، هر سه وضعیت در تمام طبقات دیده می‌شود و در عمل طبقه‌بندی بی‌فایده است، زیرا تمایزی بین آنها نسبت به وضعیت‌های دایره، مثلث و ستاره وجود ندارد.

شکل ۲: جامعه آماری همگن و ناهمگن

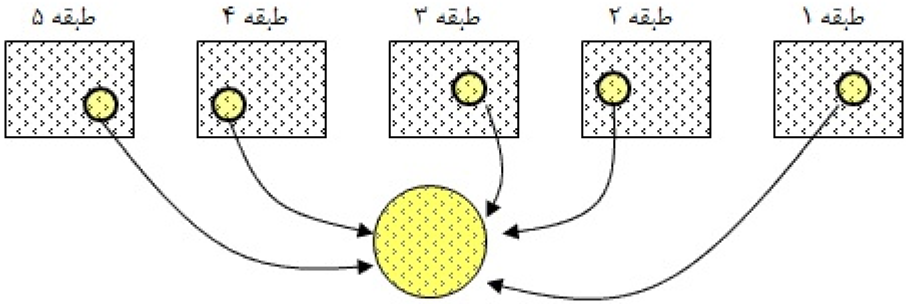


اگر پژوهشگر، نمونه‌گیری تصادفی ساده را برای جامعه دارای طبقه‌بندی به‌کار بگیرد، ممکن است با مشکلی جدی در انتخاب افراد نمونه روبه‌رو شود؛ زیرا قرعه‌کشی از کل جامعه کاملاً «بی‌طرفانه» صورت می‌گیرد و مشخص نیست که افراد نمونه او از کجای جامعه انتخاب می‌شوند. پس این امکان نیز وجود دارد که بیشتر افراد نمونه برحسب تصادف از طبقهٔ خاصی از جامعه انتخاب شوند. برای مثال در انتخاب یک نمونه ۶ تایی با وضعیت مقابل روبه‌رو می‌شود که چهار فرد آن از طبقهٔ دوم انتخاب شده‌اند که بیشتر شامل دایره است. پس انتخاب نمونه با روش تصادفی ساده از جامعه دارای طبقه‌بندی، لزوماً به نمونه معرّفی از آن جامعه منجر نخواهد شد.



روش نمونه‌گیری طبقه‌ای برای برخورد با چنین چالشی ابداع شده است. این روش به جای آنکه انتخاب نمونه‌ها را تنها به قرعه‌کشی از «کل» جامعه واگذار کند، طبقه‌بندی جامعه را نیز در فرایند انتخاب افراد نمونه در نظر می‌گیرد تا به طور حتم افرادی از هر طبقه در نمونه انتخاب شوند. برای این منظور، ابتدا تعداد نمونه‌ای که قرار است از جامعه انتخاب گردد، بین طبقات تقسیم می‌شود و سپس برای انتخاب تعداد اختصاص‌یافته به هر طبقه، در میان اعضایِ آن طبقه قرعه‌کشی صورت می‌گیرد (یعنی نمونه‌گیری تصادفی سادهٔ جداگانه‌ای در هر طبقه اجرا می‌شود).

برای مثال اگر جامعه‌ای با جمعیت $N = ۱۰۰۰۰۰$ نفر شامل پنج طبقه با جمعیتِ $N_۱ = ۱۰۰۰۰$، $N_۲ = ۲۵۰۰۰$، $N_۳ = ۵۰۰۰۰$، $N_۴ = ۱۰۰۰۰$ و $N_۵ = ۵۰۰۰$ باشد، برای انتخاب نمونه‌ای $n = ۷۰۰$ تایی از این جامعه باید این تعداد بین این پنج طبقه قسمت شود تا سهم هر طبقه از نمونه مشخص گردد. فرض کنید نمونه ۷۰۰ تایی به صورتِ $n_۱ = ۷۰$، $n_۲ = ۱۷۵$، $n_۳ = ۳۵۰$، $n_۴ = ۷۰$ و $n_۵ = ۳۵$ بین پنج طبقه تقسیم شود. اکنون با پنج نمونه‌گیری تصادفی سادهٔ جداگانه روبه‌رو هستیم: انتخاب نمونه‌ای ۷۰ تایی از ۱۰۰۰۰ عضو طبقهٔ نخست، انتخاب نمونه‌ای ۱۷۵ تایی از ۲۵۰۰۰ عضو طبقهٔ دوم و. . . .



نمونه‌گیری از هر طبقه به چهارچوبی از اعضای آن طبقه شامل نشانی هر عضو نیاز دارد تا قرعه‌کشی افراد نمونه را از میان اعضای طبقه ممکن سازد. پس از انتخاب نمونه هر طبقه، با روی هم ریختن این نمونه‌ها، به نمونه معرّفی از «جامعه» دست خواهیم یافت.

تاکنون خواننده دریافته است که نمونه‌گیری طبقه‌ای از یک جامعه، مستلزم آن است که ابتدا جامعه آماری طبقه‌بندی شود، سپس اندازه نمونه‌ای برای آن جامعه تعیین گردد، بعد این اندازه نمونه بین طبقات تقسیم شود، آنگاه چهارچوب جداگانه‌ای برای هر طبقه تهیه

گردد و سرانجام از هر طبقه به اندازه نمونه تخصیص‌یافته به آن، نمونه‌گیری تصادفی ساده (قرعه‌کشی) صورت بگیرد. بنابراین، نخستین پرسش درباره نمونه‌گیری طبقه‌ای به چگونگی طبقه‌بندی جامعه آماری باز می‌گردد.

۱ـ۲ـ۴ـ طبقه‌بندی

جامعه باید به طبقات دوبه‌دو مجزا و فراگیر تقسیم شود؛ طبقاتی که از یک سو هیچ اشتراکی با یکدیگر نداشته باشند و از سوی دیگر تمام جامعه آماری را با هم پوشش دهند. بنابراین، هر عضو از جامعه تنها در یکی از طبقات جای می‌گیرد و هیچ عضوی از جامعه خارج از این طبقات باقی نمی‌ماند.

معمولاً از یک متغیر مقوله‌ای یا متغیری که از طریق دسته‌بندی به متغیر مقوله‌ای تبدیل شده است، برای تعیین طبقات استفاده می‌شود. برای مثال، مقطع تحصیلی متغیری با سه مقوله (کارشناسی، کارشناسی ارشد و دکترا) است که می‌تواند در نمونه‌گیری از دانشجویان دانشگاه به کار گرفته شود. در این صورت، جامعه دانشجویان به سه طبقه تقسیم می‌شود: طبقهٔ دانشجویان مقطع کارشناسی، طبقهٔ دانشجویان مقطع کارشناسی ارشد و طبقهٔ دانشجویان مقطع دکترا. گاه از طبقه‌بندی‌هایی که به صورت طبیعی یا قراردادی وجود دارند، برای طبقه‌بندی جامعه استفاده می‌شود؛ طبقه‌بندی کشور به استان‌های آن، طبقه‌بندی اهالی یک شهر به مهاجر و بومی یا طبقه‌بندی جامعه نوزادان به زایمان‌های طبیعی یا سزارین از این دست است[1].

اگرچه شیوه‌های فوق راهی برای طبقه‌بندی جامعه آماری در اختیار می‌گذارند، ولی باید طبقه‌بندی بر مبنای متغیری صورت بگیرد که جامعه آماری ناهمگن را به طبقاتی با همگنی «درونی» طبقه‌بندی می‌کند. با این دیدگاه است که می‌توان با نمونه‌گیری از این طبقات به نمونه معرّفی از جامعه دست یافت. بهترین متغیر برای طبقه‌بندی همان متغیر تحت بررسی در پژوهش است، ولی اغلب اطلاعاتی از این متغیر در دست نیست تا مبنای طبقه‌بندی قرار گیرد. از این رو از متغیر دیگری برای طبقه‌بندی استفاده می‌شود که با متغیر تحت بررسی رابطه دارد. برای مثال در برآورد میانگین متغیر درآمد، متغیر وضعیت اشتغال می‌تواند مبنای مناسبی برای طبقه‌بندی جامعه آماری باشد.

---

۱. شیوه‌های دیگری نظیر «دالینیوس $\sqrt{f}$ انباشتی» نیز وجود دارد که در این کتاب درباره آنها بحث نمی‌شود.

۲ـ۲ـ۴ـ برآورد میانگین و خطای معیار آن

فرض کنید جامعه آماری دارای k طبقه و برآورد میانگینِ توزیع متغیر X در جامعه از طریق نمونه‌گیری طبقه‌ای مورد توجه پژوهشگر باشد. اگر N نشان‌دهنده جمعیت جامعه آماری و N_1، N_2، ...، N_k جمعیت‌های طبقه‌ها باشند، آنگاه $W_1 = \frac{N_1}{N}$، $W_2 = \frac{N_2}{N}$، ...، $W_k = \frac{N_k}{N}$ را وزن‌های طبقه‌ها می‌نامیم[1]. مانند قبل، از n برای نمونه کل جامعه و از n_1، n_2، ...، n_k برای نمونه‌های تخصیص‌یافته به طبقه‌ها استفاده می‌کنیم.

از آنجا که از هر طبقه، نمونه‌گیری تصادفی صورت می‌گیرد، پس داده‌های موجود از هر طبقه دارای میانگینی است که با $\bar{X}_1$، $\bar{X}_2$، ...، $\bar{X}_k$ نشان داده می‌شوند. بر این اساس، برآورد نااریب میانگین نامعلوم متغیر X در جامعه آماری عبارت است از:

$$\bar{X}_{st} = \sum_{i=1}^{k} W_i \bar{X}_i$$

اندیسِ st نشان می‌دهد که این برآورد از داده‌های حاصل از نمونه‌گیری طبقه‌ای محاسبه شده است. توجه کنید که روش برآورد میانگین در نمونه‌گیری طبقه‌ای متفاوت با روش برآورد در نمونه‌گیری تصادفی ساده است. خطای معیار این برآورد نیز از رابطهٔ زیر تعیین می‌شود که در آن σ_i^2 نمایانگر واریانس متغیر X در طبقهٔ iاُم است.

$$SE_{\bar{X}_{st}} = \sqrt{\sum_{i=1}^{k} W_i^2 \left(1-\frac{n_i}{N_i}\right)\frac{\sigma_i^2}{n_i}}$$

در عمل، به دلیل نامعلوم بودن σ_i^2ها، از واریانس حاصل از داده‌های نمونه‌ای هر طبقه در رابطهٔ فوق استفاده می‌شود تا تخمینی از خطای معیار برآورد به دست آید.

برای مثال، کارخانه‌ای با $N = ۵۰۰۰$ شاغل می‌خواهد میانگین رضایت شغلی آنان را به دست آورد. کارکنان این کارخانه شامل $N_1 = ۳۵۰۰$ کارگر، $N_2 = ۱۰۰۰$ کارمند و $N_3 = ۵۰۰$ نیرویِ قراردادی است. از آنجا که نوع کار، حقوق و مزایای این سه دسته متفاوت است، این کارخانه با روش نمونه‌گیری طبقه‌ای، تعداد $n = ۳۵۰$ نفر را از میان کارکنان طوری انتخاب می‌کند تا رضایت شغلی $n_1 = ۲۴۵$ کارگر، $n_2 = ۷۰$ کارمند و $n_3 = ۳۵$ نیروی

---

۱. وزن هر طبقه نشان می‌دهد چه سهمی از جمعیت جامعه آماری به آن طبقه تعلق دارد.

قراردادی اندازه‌گیری شود. میانگین و واریانس رضایت شغلی ۲۴۵ کارگر به ترتیب $\overline{X}_۱ = ۱۷$ و $S_۱^۲ = ۳/۵$، میانگین و واریانس رضایت شغلی ۷۰ کارمند به ترتیب $\overline{X}_۲ = ۱۶/۵$ و $S_۲^۲ = ۱/۵$ و میانگین و واریانس رضایت شغلی ۳۵ نیروی قراردادی نیز به ترتیب $\overline{X}_۳ = ۱۵$ و $S_۳^۲ = ۲/۵$ به دست آمد. بر این اساس، وزن هر طبقه و برآورد میانگین رضایت شغلی ۵۰۰۰ کارمند به قرار زیر است (با توجه به بحث قسمت بعد، اگر از تخصیص متناسب استفاده شود، محاسبه میانگین با روش طبقه‌ای همانند روش تصادفی ساده است):

$$W_۱ = \frac{۳۵۰۰}{۵۰۰۰} = ۰/۷ \qquad W_۲ = \frac{۱۰۰۰}{۵۰۰۰} = ۰/۲ \qquad W_۳ = \frac{۵۰۰}{۵۰۰۰} = ۰/۱$$

$$\overline{X}_{st} = ۰/۷ \times ۱۷ + ۰/۲ \times ۱۶/۵ + ۰/۱ \times ۱۵ = ۱۶/۷$$

خطای معیار این برآورد عبارت است از:

$$SE_{\overline{X}_{st}} = \sqrt{۰/۷^۲ \times (۱ - \frac{۲۴۵}{۳۵۰۰}) \times \frac{۳/۵}{۲۴۵} + ۰/۲^۲ \times (۱ - \frac{۷۰}{۱۰۰۰}) \times \frac{۱/۵}{۷۰} + ۰/۱^۲ \times (۱ - \frac{۳۵}{۵۰۰}) \times \frac{۲/۵}{۳۵}} = ۰/۰۲۶$$

بنابراین، حد بالا و پایین برآورد فاصله‌ای ۹۵ درصدی برای میانگین رضایت شغلی کارکنان این کارخانه به صورت زیر به دست می‌آید:

$$LL = ۱۶/۷ - ۱/۹۶ \times ۰/۰۲۶ = ۱۶/۶۴۹ \qquad UL = ۱۶/۷ + ۱/۹۶ \times ۰/۰۲۶ = ۱۶/۷۵۰$$

اگر پژوهشگر در پی برآورد درصد با روش نمونه‌گیری طبقه‌ای باشد، از $p_{st} = \sum_{i=۱}^{k} W_i p_i$ با خطای معیار $SE_{p_{st}} = \sqrt{\sum_{i=۱}^{k} W_i^۲ (۱ - \frac{n_i}{N_i}) \frac{p_i(۱-p_i)}{n_i - ۱}}$ استفاده می‌کند. p_i نسبت حاصل از داده‌های نمونه از طبقهٔ iاُم است (به یاد بیاورید که با ضربِ «نسبت» در عدد ۱۰۰ به درصد خواهید رسید).

۳ـ۲ـ۴ـ اندازه نمونه و شیوه تخصیص

تخصیص متناسب با جمعیتِ طبقات، طبیعی‌ترین و معمول‌ترین راهی است که برای تقسیم نمونه n تایی بین طبقات به ذهن می‌رسد. در این شیوه، طبقه‌ای که جمعیت بیشتری داشته باشد، سهم بیشتری از نمونه را نیز به خود اختصاص می‌دهد. پس اگر برای مثال، طبقه‌ای شامل ۴۰ درصد از جمعیت جامعه آماری است، ۴۰ درصد از نمونه کل نیز از آن طبقه انتخاب

می‌شود. پس برای تعیین تعداد نمونه هر طبقه، وزن آن طبقه در n ضرب می‌شود یعنی $n_i = W_i n$، بنابراین همواره در این شیوهٔ تخصیص $W_i = \frac{N_i}{N} = \frac{n_i}{n}$ است.

در مثال کارخانه‌ای با $N = ۵۰۰۰$ شاغل، اندازه نمونه $n = ۳۵۰$ تایی با شیوهٔ تخصیص متناسب بین سه طبقه تقسیم شده بود، زیرا:

$$W_۱ = \frac{۳۵۰۰}{۵۰۰۰} = ۰/۷ \Rightarrow n_۱ = ۰/۷ \times ۳۵۰ = ۲۴۵$$
$$W_۲ = \frac{۱۰۰۰}{۵۰۰۰} = ۰/۲ \Rightarrow n_۲ = ۰/۲ \times ۳۵۰ = ۷۰$$
$$W_۳ = \frac{۵۰۰}{۵۰۰۰} = ۰/۱ \Rightarrow n_۳ = ۰/۱ \times ۳۵۰ = ۳۵$$

در این مثال، $n_۱ = ۲۴۵$، $n_۲ = ۷۰$ و $n_۳ = ۳۵$ دارای مقدار اعشاری نیستند، بنابراین جمع آنها یعنی $n_۱ + n_۲ + n_۳$ درست برابر با $n = ۳۵۰$ می‌شود. در عمل ممکن است حاصل‌ضربِ وزن طبقه و اندازه نمونه، دارای اعشار باشد که در این صورت جمع نمونه‌های تخصیص‌یافته به طبقات کمی بیش از اندازه نمونه خواهد شد[۱]. برای مثال، اگر $n = ۳۵۵$ باشد، آنگاه $n_۱ = ۲۴۸/۵$، $n_۲ = ۷۱$ و $n_۳ = ۳۵/۵$ خواهد بود که در عمل به جای ۲۴۸/۵ از ۲۴۹ و به جای ۳۵/۵ از ۳۶ استفاده می‌شود. به این ترتیب، جمع نمونه‌های تخصیص‌یافته به طبقات برابر با ۳۵۶ خواهد شد که یکی بیشتر از $n = ۳۵۵$ است.

تخصیص نِیمن[۲] شیوهٔ دیگری است که برای تقسیم نمونه n تایی در بین طبقات به کار می‌رود. برآوردی که از نمونه‌ای با تخصیص نیمن به دست می‌آید، دارای کمترین خطای معیار نسبت به برآوردهایی است که از داده‌هایی با همان اندازه نمونه، ولی با شیوه‌های دیگر تخصیص به دست می‌آیند. این شیوه به جای آنکه مانند تخصیص متناسب، تنها جمعیت طبقات را مبنای تخصیص نمونه قرار دهد، واریانس طبقات را نیز از طریق رابطهٔ زیر دخالت می‌دهد:

$$n_i = \frac{N_i \sigma_i}{\sum_{j=۱}^{k} N_j \sigma_j} n$$

---

۱. با این مسئله در انواع دیگر شیوه‌های تخصیص نیز برخورد خواهید کرد.

2. Neyman

استفاده از واریانس در این رابطه محدودیتی برای این شیوۀ تخصیص قلمداد می‌شود، زیـرا واریانس‌های طبقات نامعلوم هستند و باید بـرآوردی از آنهـا در دسـت باشـد تـا بتـوان از تخصیص نیمن برای تقسیم اندازه نمونه در بین طبقات استفاده کرد.

برای نشان دادن فرایند محاسبات از مثال مربوط به کارخانـه کمـک مـی‌گیـریم. هـدف، تخصیص نمونه $n = ۳۵۰$ تایی بین ۳ طبقۀ کارگران با جمعیـت $N_۱ = ۳۵۰۰$، کارمنـدان بـا جمعیت $N_۲ = ۱۰۰۰$ و نیروهای قراردادی با جمعیت $N_۳ = ۵۰۰$ است. فرض کنیـد بـرای برآورد واریانس‌ها از یک نمونه‌گیـری مقـدماتی کوچـک ۳۰ تـایی بـا تخصـیص متناسـب استفاده کرده و به برآوردهای $S_۱ = ۱/۹$، $S_۲ = ۲$ و $S_۳ = ۰/۹$ رسـیده‌ایـم. بـر ایـن اسـاس، تخصیص نمونه $n = ۳۵۰$ تایی با روش نیمن به صورت زیر است:

$$N_۱S_۱ = ۳۵۰۰ \times ۱/۹ = ۶۶۵۰, N_۲S_۲ = ۱۰۰۰ \times ۲ = ۲۰۰۰, N_۳S_۳ = ۵۰۰ \times ۰/۹ = ۴۵۰$$

$$N_۱S_۱ + N_۲S_۲ + N_۳S_۳ = ۹۱۰۰$$

$$n_۱ = \frac{۶۶۵۰}{۹۱۰۰} \times ۳۵۰ = ۲۵۵/۷۶, n_۲ = \frac{۲۰۰۰}{۹۱۰۰} \times ۳۵۰ = ۷۶/۹۲, n_۳ = \frac{۴۵۰}{۹۱۰۰} \times ۳۵۰ = ۱۷/۳$$

در عمل، $n_۱ = ۲۵۶$، $n_۲ = ۷۷$ و $n_۳ = ۱۸$ اختیار و ۳۵۱ فرد به عنوان نمونـه انتخـاب مـی‌شـود. توجه کنید تخصیص نیمن به نتیجه‌ای متفاوت با تخصیص متناسب انجامید، بویژه در طبقۀ سوم.

اغلب، محدودیت‌های مالی پژوهش در گردآوری داده‌ها به عنوان عاملی محـدودکننـده در نمونه‌گیری به حساب می‌آید که تأثیر خود را بر اندازه نمونـه و شـیوۀ تخصـیص نشـان می‌دهد. به همین دلیل، پژوهشگر تلاش می‌کند با توجه به توان مالی پـژوهش بـه انتخـاب «بهینه‌ای» از اندازه نمونه و تخصیص آن دست بزند. بر این اساس، اگر C_i حـداکثر مبلغـی باشد که بودجۀ پژوهش می‌تواند بابت انتخاب فردی از طبقۀ iاُم هزینه کند، شیوۀ تخصیص بهینه، نمونه n تایی را طوری بین طبقات تقسیم می‌کند کـه بـرآورد دارای کمتـرین خطـای معیار نسبت به شیوه‌های دیگر تخصیص همین اندازه نمونه باشد:

$$n_i = \frac{\frac{N_i \sigma_i}{\sqrt{C_i}}}{\sum_{j=۱}^{k} \frac{N_j \sigma_j}{\sqrt{C_j}}} n$$

شیوه تخصیص بهینه نیز به اطلاع از واریانس‌های طبقات نیاز دارد و مانند شـیوۀ تخصـیص نیمن در عمل از برآورد حاصل از نمونه‌گیری مقدماتی استفاده می‌کند.

چگونگی تخصیص بهینه را از طریق مثال کارخانه بیان می‌کنیم که پیشتر چگونگی تخصیص نیمن را برای آن ارائه کردیم. فرض کنید هزینهٔ گردآوری داده از یک فرد در طبقهٔ کارگران $C_۱ = ۱۰۰۰$ تومان، در طبقهٔ کارمندان $C_۲ = ۵۰۰$ تومان و در طبقهٔ نیروهای قراردادی $C_۳ = ۹۰۰$ تومان باشد. در این صورت:

$$\frac{N_۱S_۱}{\sqrt{C_۱}} = \frac{۶۶۵۰}{\sqrt{۱۰۰۰}} = ۲۱۰/۳, \frac{N_۲S_۲}{\sqrt{C_۲}} = \frac{۲۰۰۰}{\sqrt{۵۰۰}} = ۸۹/۴, \frac{N_۳S_۳}{\sqrt{C_۳}} = \frac{۴۵۰}{\sqrt{۹۰۰}} = ۱۵$$

$$\frac{N_۱S_۱}{\sqrt{C_۱}} + \frac{N_۲S_۲}{\sqrt{C_۲}} + \frac{N_۳S_۳}{\sqrt{C_۳}} = ۳۱۴/۷$$

$$n_۱ = \frac{۲۱۰/۳}{۳۱۴/۷} \times ۳۵۰ = ۲۳۳/۸۹, n_۲ = \frac{۸۹/۴}{۳۱۴/۷} \times ۳۵۰ = ۹۹/۴, n_۳ = \frac{۱۵}{۳۱۴/۷} \times ۳۵۰ = ۱۶/۶$$

در عمل، $n_۱ = ۲۳۴$، $n_۲ = ۱۰۰$ و $n_۳ = ۱۷$ اختیار و ۳۵۱ فرد به عنوان نمونه انتخاب می‌شوند. نتایج این تخصیص نیز با تخصیص متناسب و تخصیص نیمن تفاوت می‌کند.

شیوه‌های تخصیص، چگونگی تقسیم نمونه n تایی بین طبقات را مشخص می‌کنند، ولی هنوز معلوم نشده است که اندازه n در نمونه‌گیری طبقه‌ای چگونه تعیین می‌شود. یک شیوه، تعیین اندازه نمونه با روش نمونه‌گیری تصادفی ساده و تخصیص آن با یکی از شیوه‌های سه‌گانه است. با وجود این، راهکارهای دیگری نیز در دست است که ویژهٔ نمونه‌گیری طبقه‌ای هستند. این راهکارها مطابق جدول ۱، دو عامل را در تعیین اندازه نمونه دخالت می‌دهند: خطای معیار برآورد و هزینهٔ گردآوری داده‌ها.

فرض کنید پژوهشگر مایل است خطای معیار برآورد او بیشتر از se نباشد، یعنی $SE_{\bar{X}_{st}} \le se$ برقرار باشد. در این صورت اندازه نمونه برای تخصیص نیمن از رابطهٔ زیر تعیین می‌شود:

$$n = \frac{(\sum_{i=۱}^{k} N_i\sigma_i)^۲}{N^۲ se^۲ + \sum_{i=۱}^{k} N_i\sigma_i^۲} = \frac{(\sum_{i=۱}^{k} W_i\sigma_i)^۲}{se^۲ + \sum_{i=۱}^{k} W_i \frac{\sigma_i^۲}{N_i}}$$

و برای تخصیص متناسب از رابطهٔ زیر استفاده می‌شود:

$$n = \frac{\sum_{i=۱}^{k} N_i\sigma_i^۲}{N se^۲ + \frac{۱}{N}\sum_{i=۱}^{k} N_i\sigma_i^۲} = \frac{\sum_{i=۱}^{k} W_i\sigma_i^۲}{se^۲ + \frac{۱}{N}\sum_{i=۱}^{k} W_i\sigma_i^۲}$$

معمولاً پژوهشگر علاوه بر تعیین مرزی برای خطای معیار برآورد، مایل است نمونـه طـوری تعیین شود که کمترین هزینه را نیز بر بودجهٔ پژوهش تحمیل کند. اگر C_i حداکثر مبلغـی باشـد که بودجهٔ پژوهش می‌تواند بابت انتخاب فردی از طبقـهٔ iاُم هزینـه کنـد، آنگـاه انـدازه نمونـه از رابطهٔ زیر به دست می‌آید که باید با شیوهٔ تخصیص بهینه بین طبقات تقسیم شود.

$$n=\frac{(\sum_{i=1}^{k}\frac{N_i\sigma_i}{\sqrt{C_i}})(\sum_{i=1}^{k}N_i\sigma_i\sqrt{C_i})}{N^2 se^2+\sum_{i=1}^{k}N_i\sigma_i^2}=\frac{(\sum_{i=1}^{k}\frac{W_i\sigma_i}{\sqrt{C_i}})(\sum_{i=1}^{k}W_i\sigma_i\sqrt{C_i})}{se^2+\frac{1}{N}\sum_{i=1}^{k}W_i\sigma_i^2}$$

گاهی پژوهشگر تنها به عامل هزینه در تعیین اندازه نمونه اهمیت می‌دهـد؛ یعنـی مایـل است هزینهٔ کل گردآوری افرادِ نمونه از تمام طبقات بیشتر از مبلغِ C نشود، در این صورت اندازه نمونه برای تخصیص بهینه از رابطهٔ زیر به دست می‌آید:

$$n=\frac{C(\sum_{i=1}^{k}\frac{N_i\sigma_i}{\sqrt{C_i}})}{\sum_{i=1}^{k}N_i\sigma_i\sqrt{C_i}}=\frac{C(\sum_{i=1}^{k}\frac{W_i\sigma_i}{\sqrt{C_i}})}{\sum_{i=1}^{k}W_i\sigma_i\sqrt{C_i}}$$

و برای تخصیص متناسب از رابطهٔ زیر تعیین می‌گردد:

$$n=\frac{NC}{\sum_{i=1}^{k}N_iC_i}=\frac{C}{\sum_{i=1}^{k}W_iC_i}$$

اکنون فرایند محاسباتی حـالات پـنج‌گانـهٔ فـوق را در قالـب مثـالی بررسـی مـی‌کنـیم. دانشکده‌ای دارای $N=۲۵۱$ دانشجوست کـه $N_1=۱۳۲$ نفـر آنهـا در مقطـع کارشناسـی، $N_2=۹۲$ نفر در مقطع کارشناسی ارشد و $N_3=۲۷$ نفر در مقطع دکترا در حـال تحصـیل هستند. این دانشکده، میانگین روزانهٔ مدت زمان استفاده از اینترنت دانشجویان خود را هـر دو سال یک بار اندازه می‌گیرد. فرض کنید بر اسـاس نمونـه‌گیـری دو سـال قبـل، بـرآورد انحراف معیار مدت زمان استفاده از اینترنـت بـرای دانشـجویان ایـن سـه مقطـع $S_1=۵۵$، $S_2=۳۰$ و $S_3=۲۰$ دقیقه باشد. امسال دانشکده مایل است خطای معیار بـرآورد میـانگین استفاده از اینترنت بیش از $se=۱۰$ دقیقه نباشد، در این صـورت انـدازه نمونـه لازم بـرای برآورد میانگین روزانهٔ مدت زمان استفاده از اینترنت دانشجویان این گونه محاسبه می‌شود:

$$N_۱S_۱ = ۱۳۲ \times ۵۵ = ۷۲۶۰, N_۲S_۲ = ۹۲ \times ۳۰ = ۲۷۶۰, N_۳S_۳ = ۲۷ \times ۲۰ = ۵۴۰$$

$$N_۱S_۱ + N_۲S_۲ + N_۳S_۳ = ۱۰۵۶۰$$

$$N_۱S_۱^۲ = ۱۳۲ \times ۵۵^۲ = ۳۹۹۳۰۰, N_۲S_۲^۲ = ۹۲ \times ۳۰^۲ = ۸۲۸۰۰, N_۳S_۳^۲ = ۲۷ \times ۲۰^۲ = ۱۰۸۰۰$$

$$N_۱S_۱^۲ + N_۲S_۲^۲ + N_۳S_۳^۲ = ۴۹۲۹۰۰$$

$$n = \frac{۱۰۵۶۰^۲}{۲۵۱^۲ \times ۱۰^۲ + ۴۹۲۹۰۰} = ۱۶/۴۱ \cong ۱۷$$

$$n_۱ = \frac{۷۲۶۰}{۱۰۵۶۰} \times ۱۷ = ۱۱/۶ \cong ۱۲, n_۲ = \frac{۲۷۶۰}{۱۰۵۶۰} \times ۱۷ = ۴/۴ \cong ۵, n_۳ = \frac{۵۴۰}{۱۰۵۶۰} \times ۱۷ = ۰/۸۹ \cong ۱$$

پس بر مبنای تخصیص نیمن، باید ۱۲ دانشجوی مقطع کارشناسی، ۵ دانشجوی مقطع کارشناسی ارشد و ۱ دانشجوی مقطع دکترا یعنی در مجموع ۱۸ دانشجو از ۲۵۱ دانشجوی دانشکده انتخاب شود. اگر دانشگاه بخواهد از تخصیص متناسب استفاده کند، اندازه نمونه و نمونه‌های تخصیص‌یافته عبارتند از:

$$n = \frac{۴۹۲۹۰۰}{۲۵۱ \times ۱۰^۲ + \frac{۴۹۲۹۰۰}{۲۵۱}} = ۱۸/۲۱ \cong ۱۹$$

$$n_۱ = \frac{۱۳۲}{۲۵۱} \times ۱۹ = ۹/۹ \cong ۱۰, n_۲ = \frac{۹۲}{۲۵۱} \times ۱۹ = ۶/۹ \cong ۷, n_۳ = \frac{۲۷}{۲۵۱} \times ۱۹ = ۲/۰۴ \cong ۳$$

بنابراین، باید ۱۰ دانشجوی مقطع کارشناسی، ۷ دانشجوی مقطع کارشناسی ارشد و ۳ دانشجوی مقطع دکترا یعنی در مجموع ۲۰ دانشجو انتخاب شود.

از آنجا که اتاق‌های خاصی به دانشجویان دکترای دانشکده اختصاص یافته است، دسترسی به آنان برای گردآوری داده‌ها، کم‌هزینه‌تر از دسترسی به دانشجویان کارشناسی و کارشناسی ارشد است؛ به طوری که هزینهٔ گردآوری داده از یک دانشجوی دکترا $C_۳ = ۴۰۰$ تومان و هزینهٔ گردآوری داده از یک دانشجوی کارشناسی یا کارشناسی ارشد به طور برابر $C_۱ = C_۲ = ۹۰۰$ تومان است. از این رو، دانشکده تصمیم می‌گیرد علاوه بر اینکه خطای معیار برآورد از $se = ۱۰$ بیشتر نباشد، هزینهٔ مراجعه به دانشجویان نمونه و گردآوری داده از آنان نیز در نظر گرفته شود.

بر این اساس، اندازه نمونه و نمونه‌های تخصیص یافته به طبقات این گونه محاسبه می‌شود:

$$\frac{N_۱S_۱}{\sqrt{C_۱}}=\frac{۱۳۲\times۵۵}{\sqrt{۹۰۰}}=۲۴۲,\ \frac{N_۲S_۲}{\sqrt{C_۲}}=\frac{۹۲\times۳۰}{\sqrt{۹۰۰}}=۹۲,\ \frac{N_۳S_۳}{\sqrt{C_۳}}=\frac{۲۷\times۲۰}{\sqrt{۴۰۰}}=۲۷$$

$$\frac{N_۱S_۱}{\sqrt{C_۱}}+\frac{N_۲S_۲}{\sqrt{C_۲}}+\frac{N_۳S_۳}{\sqrt{C_۳}}=۳۶۱$$

$$N_۱S_۱\sqrt{C_۱}+N_۲S_۲\sqrt{C_۲}+N_۳S_۳\sqrt{C_۳}=۳۱۱۴۰۰$$

$$N_۱S_۱^۲+N_۲S_۲^۲+N_۳S_۳^۲=۴۹۲۹۰۰$$

$$n=\frac{۳۶۱\times۳۱۱۴۰۰}{۲۵۱^۲\times۱۰^۲+۴۹۲۹۰۰}=۱۶/۵\cong۱۷$$

$$n_۱=\frac{۷۲۶۰}{۱۰۵۶۰}\times۱۷=۱۱/۶\cong۱۲,\ n_۲=\frac{۲۷۶۰}{۱۰۵۶۰}\times۱۷=۴/۴\cong۵,\ n_۳=\frac{۵۴۰}{۱۰۵۶۰}\times۱۷=۰/۸۹\cong۱$$

در این مثال، اندازه نمونه با احتساب هزینه نیز مانند حالت تخصیص نیمن است که نتایج آن پیشتر ارائه شد. هزینهٔ کل نمونه‌گیری $۱\times۴۰۰+۵\times۹۰۰+۱۲\times۹۰۰=۱۵۷۰۰$ تومان است.

اکنون فرض کنید دانشگاه تنها به عامل هزینه اهمیت می‌دهد و نمی‌خواهد هزینهٔ کل مراجعه به دانشجویان نمونه و گردآوری داده از آنان بیشتر از $C=۲۰۰۰۰$ تومان شود. در این صورت اندازه نمونه بر اساس تخصیص بهینه عبارت است از:

$$\frac{N_۱S_۱}{\sqrt{C_۱}}+\frac{N_۲S_۲}{\sqrt{C_۲}}+\frac{N_۳S_۳}{\sqrt{C_۳}}=۳۶۱$$

$$N_۱S_۱\sqrt{C_۱}+N_۲S_۲\sqrt{C_۲}+N_۳S_۳\sqrt{C_۳}=۳۱۱۴۰۰$$

$$n=\frac{۲۰۰۰۰\times۳۶۱}{۳۱۱۴۰۰}=۲۳/۱۸\cong۲۴$$

$$\frac{N_۱S_۱}{\sqrt{C_۱}}=\frac{۱۳۲\times۵۵}{\sqrt{۹۰۰}}=۲۴۲,\ \frac{N_۲S_۲}{\sqrt{C_۲}}=\frac{۹۲\times۳۰}{\sqrt{۹۰۰}}=۹۲,\ \frac{N_۳S_۳}{\sqrt{C_۳}}=\frac{۲۷\times۲۰}{\sqrt{۴۰۰}}=۲۷$$

$$n_۱=\frac{۲۴۲}{۳۶۱}\times۲۴=۱۶/۰۸\cong۱۷,\ n_۲=\frac{۹۲}{۳۶۱}\times۲۴=۶/۱\cong۷,\ n_۳=\frac{۲۷}{۳۶۱}\times۲۴=۱/۷\cong۲$$

پس باید ۱۷ دانشجوی مقطع کارشناسی، ۷ دانشجوی مقطع کارشناسی ارشد و ۲ دانشجوی مقطع دکترا یعنی در مجموع ۲۶ دانشجو انتخاب شود. اندازه نمونه با شیوهٔ تخصیص متناسب نیز به صورت زیر به دست می‌آید:

$$n=\frac{۲۵۱\times۲۰۰۰۰}{۱۳۲\times۹۰۰+۹۲\times۹۰۰+۲۷\times۴۰۰}=۲۳/۶۳\cong۲۴$$

$$n_۱=\frac{۱۳۲}{۲۵۱}\times۲۴=۱۲/۶\cong۱۳,\ n_۲=\frac{۹۲}{۲۵۱}\times۲۴=۸/۷\cong۹,\ n_۳=\frac{۲۷}{۲۵۱}\times۲۴=۲/۵\cong۳$$

پس باید در مجموع ۲۵ نمونه انتخاب شود. هزینهٔ کل انتخاب این تعداد نمونه عبارت است از $۱۳\times۹۰۰+۹\times۹۰۰+۳\times۴۰۰=۲۱۰۰۰$ تومان که بالاتر از سقف تعیین‌شده $C=۲۰۰۰۰$ است، زیرا تعداد نمونه‌ها به نزدیک‌ترین عدد صحیح بزرگ‌تر گِرد شده‌اند.

جدول ۱: تعیین اندازه نمونه وشیوهٔ تخصیص در نمونه‌گیری طبقه‌ای تحت شرایط مختلف

<table dir="rtl">
<tr><th>خطای معیار برآورد</th><th>هزینه</th><th>شیوهٔ تخصیص</th><th>اندازه نمونه</th></tr>
<tr><td rowspan="3">مهم است پس مقدار se تعیین می‌شود</td><td rowspan="2">مهم نیست</td><td>$n_i = \frac{N_i\sigma_i}{\sum_{j=1}^{k} N_j\sigma_j} n$</td><td>$n = \frac{(\sum_{i=1}^{k} N_i\sigma_i)^2}{N^2 se^2 + \sum_{i=1}^{k} N_i\sigma_i^2}$</td></tr>
<tr><td>$n_i = W_i n$</td><td>$n = \frac{\sum_{i=1}^{k} N_i\sigma_i^2}{N se^2 + \frac{1}{N}\sum_{i=1}^{k} N_i\sigma_i^2}$</td></tr>
<tr><td>مهم است پس مقدار C_iها تعیین می‌شود</td><td>$n_i = \frac{N_i\sigma_i}{\sum_{j=1}^{k} N_j\sigma_j} n$</td><td>$n = \frac{(\sum_{i=1}^{k} \frac{N_i\sigma_i}{\sqrt{C_i}})(\sum_{i=1}^{k} N_i\sigma_i\sqrt{C_i})}{N^2 se^2 + \sum_{i=1}^{k} N_i\sigma_i^2}$</td></tr>
<tr><td rowspan="3">مهم نیست</td><td rowspan="2">مهم است پس مقدار C_iها و C تعیین می‌شود</td><td>$n_i = \frac{\frac{N_i\sigma_i}{\sqrt{C_i}}}{\sum_{j=1}^{k} \frac{N_j\sigma_j}{\sqrt{C_j}}} n$</td><td>$n = \frac{C(\sum_{i=1}^{k} \frac{N_i\sigma_i}{\sqrt{C_i}})}{\sum_{i=1}^{k} N_i\sigma_i\sqrt{C_i}}$</td></tr>
<tr><td>$n_i = W_i n$</td><td>$n = \frac{NC}{\sum_{i=1}^{k} N_i C_i}$</td></tr>
<tr><td>مهم نیست</td><td colspan="2">دلخواه</td></tr>
</table>

۴ـ۲ـ۴ـ کارایی

هنگامی از کارایی صحبت به میان می‌آید که مقایسهٔ دقت دو روش نمونه‌گیری مورد توجه باشد. دقت یک روش نمونه‌گیری در برآورد پارامتری خاص برابر با عکسِ واریانس برآوردگر آن پارامتر یعنی $\frac{1}{SE^2}$ است. بر این اساس، روشی کاراتر است که دقیق‌تر باشد، یعنی برآوردگر آن، واریانس کوچک‌تری داشته باشد. این مقایسه در شرایطی منطقی است که هر دو روش از اندازه نمونه برابری به منظور برآورد پارامتر واحدی استفاده کرده باشند. در این قسمت کارایی روش نمونه‌گیری طبقه‌ای را برای برآورد میانگین جامعه آماری نسبت به روش نمونه‌گیری تصادفی ساده بررسی می‌کنیم.

فرض کنید نمونه‌ای n تایی با روش نمونه‌گیری تصادفی ساده (SRS) از یک جامعه آماری انتخاب شده است تا میانگین این جامعه برآورد شود. به یاد بیاورید این میانگین از طریق میانگینِ نمونه با واریانس زیر برآورد می‌شود.

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \qquad SE_{\bar{X}}^2 = (1-\frac{n}{N})\frac{\sigma_X^2}{n}$$

همچنین، نمونه‌ای n تایی با روش نمونه‌گیری طبقه‌ای از همان جامعه آماری انتخاب شده است تا میانگین این جامعه با روش طبقه‌ای نیز برآورد گردد. برآوردگر میانگین و واریانس آن به صورت زیر است:

$$\overline{X}_{st} = \sum_{i=1}^{k} W_i \overline{X}_i \qquad SE^2_{\overline{X}_{st}} = \sum_{i=1}^{k} W_i^2 (1-\frac{n_i}{N_i})\frac{\sigma_i^2}{n_i}$$

همان‌طور که اشاره شد، دقت روش نمونه‌گیری تصادفی ساده $\frac{1}{SE^2_{\overline{X}}}$ و دقت روش نمونه‌گیری طبقه‌ای $\frac{1}{SE^2_{\overline{X}_{st}}}$ است. بر این اساس، کارایی روش نمونه‌گیری طبقه‌ای نسبت به روش نمونه‌گیری تصادفی ساده از تقسیم دقت روش طبقه‌ای بر دقت روش تصادفی ساده به صورت زیر به دست می‌آید:

$$Eff(stratified \mid SRS) = \frac{\frac{1}{SE^2_{\overline{X}_{st}}}}{\frac{1}{SE^2_{\overline{X}}}} = \frac{SE^2_{\overline{X}}}{SE^2_{\overline{X}_{st}}}$$

اگر Eff برابر با ۱ باشد، دو روش نمونه‌گیری از یک دقت برخوردارند، در حالی که اگر کارایی بیشتر از ۱ باشد، روش طبقه‌ای دقیق‌تر از روش تصادفی ساده و اگر کمتر از ۱ باشد، روش تصادفی ساده دقیق‌تر از روش طبقه‌ای است.

در بهترین حالت، روش نمونه‌گیری تصادفی ساده می‌تواند دقتی حداکثر برابر با روش نمونه‌گیری طبقه‌ای با تخصیص بهینه یا تخصیص نیمن داشته باشد. پس هر گاه پژوهشگر اطلاعاتی در اختیار دارد تا بر مبنای آن نمونه‌گیری طبقه‌ای با تخصیص نیمن یا بهینه صورت دهد، روش طبقه‌ای بر نمونه‌گیری تصادفی ساده با همان اندازه نمونه برتری دارد. شرایط برای نمونه‌گیری طبقه‌ای با تخصیص متناسب لزوماً به نفع روش طبقه‌ای نیست. تنها هنگامی نمونه‌گیری با تخصیص متناسب حداقل به اندازه نمونه‌گیری تصادفی ساده دقیق است که یا طبقات دارای پراکندگی درون طبقه‌ای کم و پراکندگی بین طبقه‌ای زیادی باشند یا چنان پرجمعیت باشند که بتوان از مقدار $\frac{1}{N_i}$ چشم پوشی کرد.

وضعیت نخست هنگامی فراهم است که طبقه‌بندی بر اساس متغیر مناسبی صورت گرفته باشد؛ متغیری که با متغیر تحت بررسی رابطهٔ نسبتاً قوی دارد و می‌تواند جامعه را به طبقاتی تقسیم کند که اعضای درون هر طبقه تقریباً یکدست و اعضای متعلق به طبقات مختلف ناهمگون هستند. به این ترتیب به همگنی درون طبقاتی و ناهمگنی بینِ طبقاتی دست خواهیم

یافت. به یاد بیاورید که دلیل روی آوردن به نمونه‌گیری طبقه‌ای، از میان بردنِ ناهمگنی جامعه آماری از طریق طبقه‌بندی آن بود؛ طبقاتی که در درون خود دارای ناهمگنیِ جامعه نباشند.

۳ـ۴ـ نمونه‌گیری خوشه‌ای

وجود چهارچوبی مناسب از اعضای جامعه آماری، نقشی اساسی در نمونه‌گیری تصادفی ساده و نمونه‌گیری طبقه‌ای ایفا می‌کند. این چهارچوب باید فهرستی از اعضای جامعه و چگونگی دسترسی به آنان از طریق نشانی، تلفن یا چیزی مانند آن فراهم کند. اغلب، چنین چهارچوبی از جامعه آماریِ پرجمعیت، پراکنده یا با گستردگی جغرافیایی مانند یک کلان‌شهر در دست نیست، زیرا از یک سو تهیهٔ چهارچوبی از جامعه‌هایی با این وسعت خود به سرشماری نیاز دارد که امری زمان‌بر و هزینه‌بر است و از سوی دیگر اعضای چنین جامعه‌هایی یا محل استقرار آنان به شکل غیر قابل کنترلی در حال تغییر است، بنابراین چهارچوب تهیه شده با گذشت زمان قدیمی می‌شود و دیگر قابلیت استفاده ندارد (حداقل اینکه چهارچوب باید پیوسته بهنگام شود که این نیز امری هزینه‌بر و زمان‌بر است).

نمونه‌گیری تصادفی ساده یا طبقه‌ای با مشکل دیگری نیز روبه‌روست. انتخاب افراد نمونه از طریق قرعه‌کشی باعث می‌شود این افراد در سطح جامعه آماری بسیار پراکنده باشند، به طوری که برای مثال باید برای دسترسی به تنها یک فرد از نمونه به یک سوی شهر و برای دسترسی به فردی دیگر از نمونه به سوی دیگر شهر مراجعه کرد. به این ترتیب فرایند گردآوری داده‌ها به زحمت زیادی برای گردش فراگیر در سطح جامعه و مراجعه به نقاط متعدد و دور از هم نیاز دارد.

نمونه‌گیری خوشه‌ای از چهارچوبی برای نمونه‌گیری استفاده می‌کند که هر عضو آن شامل چندین عضو از جامعه آماری پژوهش است، پس پژوهشگر با انتخاب تصادفی یک عضو از این چهارچوب به چندین عضو از جامعه آماری دسترسی پیدا می‌کند و به این ترتیب می‌تواند راحت‌تر و سریع‌تر و ارزان‌تر به گردآوری داده‌ها بپردازد. برای مثال، پژوهشگری که در پی بررسی متغیری در جامعه آماری دانش‌آموزان دبستانی از طریق نمونه‌گیری تصادفی ساده یا طبقه‌ای است، به چهارچوبی از تمام دانش‌آموزان دبستانی نیاز دارد تا با قرعه‌کشی از میان آنها افراد نمونه خود را انتخاب و با مراجعه به نشانی آنان داده‌های خود را گردآوری کند. دسترسی به فهرستی از نام و نشانی دانش‌آموزان دبستانی تقریباً برای او ناممکن است، ولی به‌راحتی می‌تواند به فهرستی از نام و نشانی دبستان‌های

سطح شهر دست یابد. سپس تعدادی از دبستان‌ها را به تصادف انتخاب و با مراجعه به این دبستان‌ها به تعدادی از دانش‌آموزان دبستانی دسترسی پیدا کند. در این مثال، چهارچوب نمونه‌گیری به جای آنکه شامل فهرست اعضای جامعه آماری پژوهش، یعنی دانش‌آموزان دبستانی باشد، شامل فهرستی از دبستان‌هاست که هر یک شامل تعدادی دانش‌آموز است.

نخستین تفاوت روش نمونه‌گیری خوشه‌ای با دو روش نمونه‌گیری تصادفی ساده و طبقه‌ای در چهارچوب نمونه‌گیری آنها نهفته است. در روش تصادفی ساده یا طبقه‌ای، نمونه‌گیری به طور مستقیم از فهرست اعضای جامعه آماری پژوهش صورت می‌گیرد؛ یعنی همان کسانی که داده‌ها از آنان گردآوری خواهد شد، در حالی که در روش خوشه‌ای، نمونه‌گیری از فهرستی صورت می‌گیرد که هر عضو آن می‌تواند شامل چندین عضو از جامعه آماری پژوهش باشد. در اینجا باید بین واحد نمونه‌گیری و اعضای جامعه آماری پژوهش تمایز قایل شویم. واحد نمونه‌گیری به هر عضو از چهارچوب نمونه‌گیری گفته می‌شود که در فرایند انتخابِ تصادفی می‌تواند یکی از انتخاب‌های نمونه‌گیری باشد[1]. برای مثال، اگر فهرست دبستان‌های شهر به عنوان چهارچوب نمونه‌گیری به‌کار گرفته شود، هر دبستان یک واحد نمونه‌گیری است، زیرا این دبستان‌ها هستند که به تصادف از چهارچوب انتخاب می‌شوند. اگر پژوهشگر فهرستی از دانش‌آموزان شهر در اختیار داشت، آنگاه انتخاب‌های تصادفی خود را به طور مستقیم از میان دانش‌آموزان صورت می‌داد، پس واحدهای نمونه‌گیری او همان اعضای جامعه آماری پژوهش او بودند.

۱ـ۳ـ۴ـ خوشه‌بندی

خوشه، به هر عضو از چهارچوب نمونه‌گیری گفته می‌شود که در ساده‌ترین شکل خود شامل گروهی از اعضای جامعه آماری پژوهش است. خوشه‌ها نیز باید مانند طبقه‌ها دوبه‌دو مجزا و فراگیر باشند؛ یعنی هیچ دو خوشه‌ای دارای اعضای مشترک نباشند و خوشه‌ها با هم تمام اعضای جامعه را شامل شوند.

با توجه به دو دلیلی که برای نمونه‌گیری خوشه‌ای بیان شد، جامعه آماری باید طوری خوشه‌بندی شود که اولاً دسترسی به فهرستی از خوشه‌ها برای تهیهٔ چهارچوب نمونه‌گیری

۱. اعضای جامعه آماری پژوهش را واحد آماری (Statistical Unit) نیز می‌نامند، زیرا داده‌ها از آنان گردآوری می‌شود و پژوهشگر توصیف‌ها و استنباط‌های آماری را درباره آنان صورت می‌دهد. در نمونه‌گیری تصادفی ساده یا طبقه‌ای، واحد نمونه‌گیری و واحد آماری یکی هستند، در حالی که در نمونه‌گیری خوشه‌ای چنین نیست.

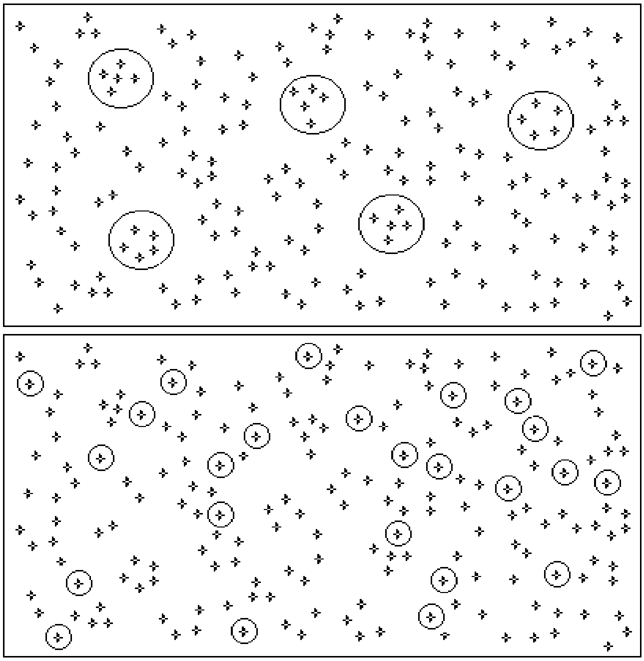
امکان‌پذیر باشد و ثانیاً به سرعت دستخوش تغییر نشوند تا بتوان بر اساس آن به چهارچوب به نسبت بادوامی برای نمونه‌گیری دست یافت. از این رو، خوشه‌بندی یک شهر به بلوک‌های ساختمانی، مناسب به نظر می‌رسد، زیرا جابه‌جایی، تغییر شکل، حذف یا اضافه‌شدن بلوک‌ها در طول زمان با چنان سرعتی رخ نمی‌دهد که تغییرات قابل توجهی در بافت شهر مشاهده شود.

معمولاً می‌توان به نوعی خوشه‌بندی طبیعی در بیشتر جامعه‌های آماری دست یافت. برای مثال، پژوهشگری که در پی بررسی میزان پرداختن روزنامهٔ خاصی به یک موضوع سیاسی در ۱۰ سال گذشته است، با جامعه‌ای شامل ۳۰۰۰ شماره از این روزنامه روبه‌روست که در طول این ۱۰ سال منتشر شده‌اند. اگر کتابخانه‌ای به طور مسلسل هر ۱۰ شماره از این روزنامه را در یک مجلّد نگهداری کند، این ۳۰۰۰ شماره روزنامه به طور طبیعی در ۳۰۰ خوشهٔ ۱۰تایی سازمان یافته‌اند. بلوک‌های شهر نیز مثال دیگری از خوشه‌بندی طبیعی شهروندان یک شهر، و روستاها مثالی از خوشه‌بندی طبیعی روستاییان یک استان است.

گاهی ضروری است پژوهشگر، جامعه آماری پژوهش خود را خوشه‌بندی کند. برای مثال اگر پژوهشگری به دنبال برآورد فروش نوعی دارو باشد و فهرستی از داروخانه‌های شهر را با نشانی آنها در اختیار داشته باشد، آنگاه می‌تواند برای کاهش سفرهای شهری، این داروخانه‌ها را بر حسب نزدیکی به یکدیگر به گروه‌های ۵ تایی خوشه‌بندی کند و سپس به جای انتخاب تصادفی داروخانه‌ها از فهرست، چند گروه از آنها را از فهرست خوشه‌های ساختگی خود انتخاب کند. مراجعه به داروخانه‌های هر خوشه به صرف زمان کمتری نیاز دارد، زیرا هر خوشه از ۵ داروخانهٔ نزدیک به هم تشکیل شده است. بر این اساس، پژوهشگر برای دسترسی به مثلاً ۲۵ داروخانه به انتخاب ۵ خوشه و مراجعه به تقریباً ۵ نقطه از شهر نیاز دارد، در حالی که در نمونه‌گیری تصادفی ساده از همان ابتدا ۲۵ داروخانه انتخاب می‌شود و احتمالاً او باید به ۲۵ نقطهٔ پراکنده در شهر مانند شکل ۳ مراجعه کند.

خوشه‌ها می‌توانند هم‌اندازه یا با اندازه‌های مختلف باشند. در مثال داروخانه‌ها، با خوشه‌های هم‌اندازه ۵ تایی یا در مثال روزنامه‌ها با خوشه‌های هم‌اندازهٔ ۱۰ تایی روبه‌رو هستیم، ولی در خوشه‌بندی شهر به بلوک، با بلوک‌هایی سروکار داریم که شامل تعداد شهروندان متفاوتی هستند یا در خوشه‌بندی روستاییان یک استان بر حسب روستا، به روستاهایی با جمعیت متفاوت برخورد می‌کنیم.

شکل ۳: مقایسهٔ تعداد نقاط مراجعه برای گردآوری داده‌ها در نمونه‌گیری خوشه‌ای و نمونه‌گیری تصادفی ساده



۲ـ۳ـ۴ـ برآورد میانگین و خطای معیار آن

فرض کنید جامعه از N خوشه با اندازه‌های متفاوت تشکیل شده است، به طوری که خوشهٔ iاُم دارای M_i عضو است. هدف نمونه‌گیری، برآورد میانگین توزیع متغیر X در جامعه آماری است. همچنین، n خوشه از میان N خوشه با قرعه‌کشی انتخاب و متغیر X از تمام اعضای هر خوشهٔ انتخابی اندازه‌گیری شده است. مانند قبل میانگینِ متغیر X را از داده‌های این نمونه با $\bar{X}$ نشان می‌دهیم. اگر $\bar{M}$ میانگین اندازه‌های N خوشهٔ جامعه و $\bar{m}$ میانگین اندازه‌های n خوشه‌ای باشد که به عنوان نمونه انتخاب شده‌اند، آنگاه برآورد نااریب میانگین جامعه بر اساس روش نمونه‌گیری خوشه‌ای عبارت است از:

$$\bar{X}_{cl} = \frac{\bar{m}}{\bar{M}} \bar{X}$$

اگر T_i مجموع مقادیر متغیر X از تمام اعضایِ خوشهٔ iاُم و S_T^2 واریانس T_iها باشد، خطای معیار برآورد میانگین از رابطهٔ زیر تخمین زده می‌شود:

$$SE_{\bar{X}_{cl}}=\sqrt{\frac{1}{\bar{M}^2}(1-\frac{n}{N})\frac{S_T^2}{n}}\qquad S_T^2=\frac{1}{n-1}\sum_{i=1}^{n}(T_i-\bar{T})^2\qquad \bar{T}=\bar{m}\bar{X}=\frac{1}{n}\sum_{i=1}^{n}T_i$$

اکنون حالت ساده‌تری را در نظر بگیرید که جامعه از N خوشهٔ هم‌اندازه M عضوی تشکیل شده است. پس در مجموع، این جامعه آماری شامل NM عضو است. مانند قبل n خوشه از N خوشه انتخاب می‌شود، یعنی nM عضو از این جامعه به عنوان نمونـه در اختیـار اسـت. توجه کنید که در جامعه‌ای با خوشه‌های هـم‌انـدازه $\bar{M}=\bar{m}=M$ خواهـد بـود. در ایـن صورت، رابطه‌های فوق، تا اندازه‌ای ساده‌تر می‌شوند:

$$\bar{X}_{cl}=\bar{X}\quad SE_{\bar{X}_{cl}}=\sqrt{\frac{1}{M^2}(1-\frac{n}{N})\frac{S_T^2}{n}}\quad S_T^2=\frac{1}{n-1}\sum_{i=1}^{n}(T_i-\bar{T})^2\quad \bar{T}=M\bar{X}$$

برای مثال، داروخانه‌های شهر به منظور برآورد میانگین فروش نوعی دارو بر حسب نزدیکـی به $N=150$ خوشه با اندازه‌های نابرابر تقسیم و $n=4$ خوشه به تصادف از میان آنهـا انتخـاب شده است. داده‌های جدول ۲ مقدار فروش داروخانه‌های انتخاب شده را نشان می‌دهند.

جدول ۲: مقدار فروش نوعی دارو از ۱۵ داروخانهٔ نمونه

خوشه	اندازه خوشه (M_i)	داده‌ها	مجموع داده‌ها (T_i)
خوشهٔ نخست نمونه	۴	۲۵۰، ۱۰۰، ۱۳۰، ۲۰۰	۶۸۰
خوشهٔ دوم نمونه	۳	۲۰۰، ۱۸۵، ۱۲۰	۵۰۵
خوشهٔ سوم نمونه	۵	۷۰، ۱۰۰، ۲۲۰، ۹۰،۱۱۰	۵۹۰
خوشهٔ چهارم نمونه	۳	۹۰، ۱۵۰، ۱۹۰	۴۳۰
جمع	۱۵	۲۲۰۵	۲۲۰۵

اگر میانگینِ اندازه‌های این ۱۵۰ خوشه برابر با $\bar{M}=4/6$ باشد، بـرآورد میـانگین فـروش این دارو با توجه با اطلاعات ستون T_i به صورت مقابل به دست می‌آید:

$$\bar{m}=\frac{۱}{۴}\times(۴+۳+۵+۳)=۳/۷۵ \quad \bar{M}=۴/۶ \quad \bar{X}=\frac{۲۲۰۵}{۱۵}=۱۴۷$$

$$\bar{X}_{cl}=\frac{۳/۷۵}{۴/۶}\times ۱۴۷=۱۱۹/۸۳$$

خطای معیار این برآورد نیز عبارت است از:

$$\bar{T}=\frac{۲۲۰۵}{۴}=۵۵۱/۲۵$$

$$S_T^۲=\frac{۱}{۴-۱}\times\left[(۶۸۰-۵۵۱/۲۵)^۲+(۵۰۵-۵۵۱/۲۵)^۲+(۵۹۰-۵۵۱/۲۵)^۲+(۴۳۰-۵۵۱/۲۵)^۲\right]=۱۱۶۳۹/۵۸$$

$$SE_{\bar{X}_{cl}}=\sqrt{\frac{۱}{۴/۶^۲}\times(۱-\frac{۴}{۱۵۰})\times\frac{۱۱۶۳۹/۵۸}{۴}}=۱۱/۵۷$$

بنابراین، می‌توان یک برآورد فاصله‌ای ۹۵ درصدی برای میانگین فروش به صورت زیر ارائه داد:

$$LL=۱۱۹/۸۳-۱/۹۶\times ۱۱/۵۷=۹۷/۱۵ \qquad UL=۱۱۹/۸۳+۱/۹۶\times ۱۱/۵۷=۱۴۲/۵۰$$

۳ـ۳ـ۴ـ اندازه نمونه

ابتدا، مسئله تعیین تعداد خوشه‌های نمونه را به لحاظ هزینه بررسی می‌کنیم. فرض کنید C_M هزینهٔ مراجعه به یک خوشهٔ M عضوی و C کلِ هزینه‌ای باشد که بودجهٔ پژوهش برای مراجعه به خوشه‌های نمونه و گردآوری داده از اعضای آنها اختصاص داده است، در این صورت به تعداد زیر می‌توان خوشه به عنوان نمونه انتخاب کرد:

$$n=\frac{C}{C_M}$$

بر این اساس، تعداد افراد نمونه نیز nM نفر خواهد بود، زیرا هر خوشه دارای M عضو است که از همگی آنان داده گردآوری می‌شود. در شرایطی که خوشه‌ها نابرابر هستند، هزینهٔ مراجعه به یک خوشهٔ $\bar{M}$ عضوی را مبنای محاسبات قرار می‌دهیم. در این حالت ممکن است هزینهٔ کل مراجعه به خوشه‌ها و گردآوری داده از اعضای آنها کمی بیشتر یا کمتر از مقدار C شود.

اگر پژوهشگر مایل باشد خطای معیار برآورد میانگین از se بیشتر نباشد، آنگاه باید تعداد خوشه‌های نمونه را از رابطه‌ای که در ادامه می‌آید، تعیین کند. معمولاً $se=\frac{d}{z_{\frac{\alpha}{۲}}}$

اختیار می‌شود تا پژوهشگر قادر باشد مقدار خطای d و ضریب اطمینان $z_{\frac{\alpha}{۲}}$ را نیز در تعیین تعداد خوشه‌ها دخالت دهد.

$$n = \frac{NS_T^۲}{N\bar{M}^۲ se^۲ + S_T^۲}$$

توجه کنید که $S_T^۲$ در رابطهٔ فوق از داده‌های یک نمونه مقدماتی یا اطلاعات نمونه‌گیری‌های پیشین به دست می‌آید. همچنین، اگر جامعه از خوشه‌های هم‌اندازه تشکیل شده باشد، از M به جای $\bar{M}$ استفاده می‌شود. این امکان نیز وجود دارد که در عمل پژوهشگر تنها از تعداد خوشه‌های جامعه یعنی N مطلع باشد، ولی از $\bar{M}$ اطلاعی نداشته باشد که در این صورت $\bar{M}$ نیز باید مانند $S_T^۲$ از همان داده‌های نمونه مقدماتی با $\bar{m}$ تخمین زده شود.

فرض کنید داده‌های جدول ۲ مربوط به یک نمونه مقدماتی باشد و پژوهشگر تنها بداند که جامعه شامل $N = ۱۵۰$ خوشهٔ نابرابر است، ولی از $\bar{M}$ بی‌اطلاع باشد. اگر او مایل باشد خطای برآورد میانگین فروش آن داروی خاص با اطمینان ۹۵ درصد بیش از $d = ۱۰$ نباشد، به چه تعداد خوشه به عنوان نمونه نیاز دارد؟ اطلاعات زیر از محاسبات جدول ۲ در اختیار است:

$$\bar{m} = ۳/۷۵ \qquad S_T^۲ = ۱۱۶۳۹/۵۸ \qquad N = ۱۵۰ \qquad se = \frac{d}{z_{۰/۰۲۵}} = \frac{۱۰}{۱/۹۶} = ۵/۱$$

پس تعداد خوشه‌های نمونه عبارت است از:

$$n = \frac{۱۵۰ \times ۱۱۶۳۹/۵۸}{۱۵۰ \times ۳/۷۵^۲ \times ۵/۱^۲ + ۱۱۶۳۹/۵۸} = ۲۶/۲۵ \cong ۲۷$$

بنابراین، باید ۲۷ خوشه از میان ۱۵۰ خوشه به تصادف انتخاب شود. اگر پژوهشگر تنها به محدودیت مالی پژوهش توجه داشته باشد، به طوری که از یک سو از بودجهٔ پژوهش خود مبلغ $C = ۱۰۰۰۰۰$ تومان را به هزینهٔ مراجعه به خوشه‌ها و گردآوری داده از اعضای آنها اختصاص داده باشد و از سوی دیگر، هزینهٔ مراجعه به یک خوشهٔ ۴ عضوی و گردآوری داده از این ۴ عضو، $C_۴ = ۶۰۰۰$ تومان باشد، آنگاه تعداد خوشه‌های نمونه او عبارتند از:

$$n = \frac{۱۰۰۰۰۰}{۶۰۰۰} = ۱۶/۶ \cong ۱۷$$

کل هزینه‌ای هم که در عمل باید برای این منظور صرف کند، برابر با $۱۰۲۰۰۰ = ۱۷ \times ۶۰۰۰$ تومان است که ۲۰۰۰ تومان بیشتر از هزینهٔ اختصاص یافته است.

۴ـ۳ـ۴ـ کارایی

معمولاً دقت نمونه‌گیری خوشه‌ای به ازای اندازه نمونه مشخص، از دقت روش‌های نمونه‌گیری تصادفی ساده یا طبقه‌ای با همان اندازه نمونه کمتر است و در صورتی که پژوهشگر بخواهد از طریق روش خوشه‌ای به دقتی در حد این روش‌ها دست یابد، به تعداد نمونه بیشتری نیاز دارد. با وجود این، روش نمونه‌گیری خوشه‌ای از دو جنبه پرطرفدار است: اولاً چهارچوب‌های نمونه‌گیری مناسبِ این روش بیشتر از روش‌های دیگر در دسترس هستند و ثانیاً از صرفهٔ اقتصادی ناشی از کاهش هزینهٔ مراجعه به افراد نمونه برخوردار است. به عبارت دیگر، دارای نوعی کارایی اقتصادی نسبت به روش‌های دیگر است.

البته شرایطی وجود دارد که نمونه‌گیری خوشه‌ای می‌تواند به اندازه روش‌های تصادفی ساده یا طبقه‌ای کارآمد یا حتی کاراتر از آنها باشد. اگر جامعه از خوشه‌های همگنی تشکیل شده باشد که در درون خود ناهمگن هستند، نمونه انتخاب شده از آنها می‌تواند معرف‌تر از نمونه هم‌اندازه‌ای باشد که از روش‌های نمونه‌گیری تصادفی ساده یا طبقه‌ای گردآوری شده است[1]. این ویژگی درست عکس ویژگی نمونه‌گیری طبقه‌ای است؛ زیرا در این نوع نمونه‌گیری، طبقات باید دارای همگنی درونی، ولی با یکدیگر ناهمگن باشند (شکل ۴ را ببینید).

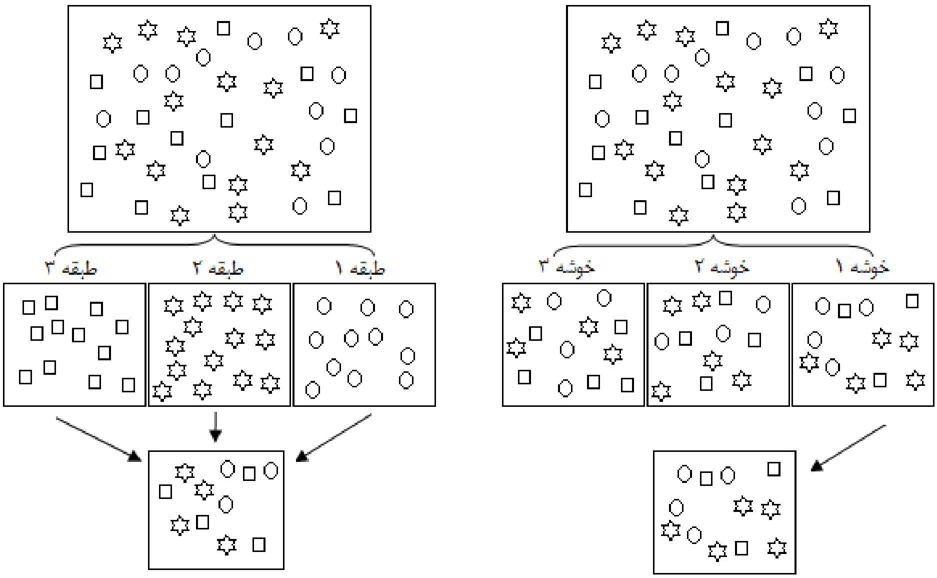
برای روشن شدن این موضوع، جامعه‌ای را با مثلاً خوشه‌های ۱۰ عضوی در نظر بگیرید. انتخاب ۱ خوشه از این جامعه به معنای انتخاب ۱۰ فردِ نمونه است. اگر این خوشه نسبت به متغیر تحت بررسی دارای همگنی درونی باشد، این ۱۰ فرد کم و بیش به یکدیگر شبیه هستند، پس اطلاعاتی که از آنها درباره توزیع متغیر در جامعه به دست می‌آید، فرقی با حالتی ندارد که تنها ۱ فرد از آنها انتخاب شده بود؛ در حالی که اگر ۱۰ نفر به صورت تصادفی ساده از بخش‌های مختلف جامعه و نه ۱ خوشه، انتخاب شده بود، این احتمال بیشتر بود که به افرادی متفاوت برخورد کنیم و اطلاعات متنوع‌تری از جامعه به دست آوریم. پس اگر خوشه‌ها دارای همگنی درونی باشند، انتظار می‌رود دسترسی به نمونه‌ای معرف از جامعه، در نمونه‌گیری تصادفی ساده بیشتر از نمونه‌گیری خوشه‌ای باشد.

در مقابل، اگر خوشهٔ انتخابی دارای ناهمگنی درونی باشد، با انتخاب آن به‌طور همزمان و با هزینهٔ کم، به ۱۰ عضو به نسبت متفاوت از جامعه دست یافته‌ایم که اطلاعات متنوعی

۱. این موضوع از طریق همبستگی درون‌خوشه‌ای به‌خوبی قابل بررسی است، ولی در این کتاب مطرح نمی‌شود. در ادامه از این ویژگی نمونه‌گیری خوشه‌ای به شکل شهودی بحث شده است.

را از توزیع متغیر در اختیار ما می‌گذارند، در صورتی که تضمینی وجود ندارد از طریق نمونه‌گیری تصادفی ساده به چنین تنوعی از افراد دست یابیم. بنابراین، ناهمگنی درونی خوشه‌ها می‌تواند به نمونه‌گیری خوشه‌ای کمک کند تا در مقایسه با نمونه‌گیری تصادفی ساده به نمونه‌ای معرف‌تر برسد.

شکل ۴: مقایسهٔ خوشه‌بندی و طبقه‌بندی



۴ـ۴ـ نمونه‌گیری سیستماتیک

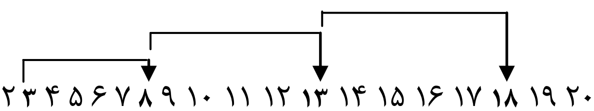
تاکنون، افراد نمونه یا واحدهای نمونه‌گیری از طریق قرعه‌کشی انتخاب می‌شدند. به این ترتیب، انتخاب هیچ واحدی به واحد دیگر ارتباط نداشت. در نمونه‌گیری سیستماتیک (نظام‌مند)، انتخاب نخستین فرد نمونه می‌تواند سرنوشت افراد دیگر نمونه را مشخص کند، زیرا نظام خاصی بین انتخاب افراد نمونه وجود دارد، به طوری که تمام انتخاب‌ها تابعِ نخستین انتخاب هستند.

برای مثال اگر جامعه‌ای شامل ۵۰ عضو باشد و بخواهید با روش سیستماتیک، نمونه‌ای ۵ تایی از آن انتخاب کنید، کافی است اعضای جامعه از ۱ تا ۵۰ شماره‌گذاری شوند و سپس عددی از ۱ تا ۱۰ به تصادف از طریق قرعه‌کشی انتخاب گردد. اگر این عدد ۷ باشد، نخستین فردِ نمونه، عضو شماره ۷ جامعه است. دومین فردِ نمونه عضو شماره

۱۰+۷=۱۷ جامعه، سومین فردِ نمونه عضو شماره ۱۰×۲+۷=۲۷ جامعه، چهارمین فـردِ نمونه عضو شماره ۱۰×۳+۷=۳۷ جامعه و سرانجام پنجمـین فـردِ نمونـه عضـو شـماره ۱۰×۴+۷=۴۷ جامعه است. همان‌طور که مشاهده مـی‌شـود، انتخـاب تصـادفی تنهـا بـه نخستین فرد نمونه باز می‌گردد و سایر افراد نمونه به‌طور خودکار بر اساس قاعدهٔ مشخصی (نظم معینی) تعیین می‌شوند. در واقع در این مثال، از هر ۱۰ عضو جامعه ۱ فـرد بـه عنـوان نمونه انتخاب شده است.

در حالت کلی، اگر N مضربی از n باشد، برای انتخاب نمونه‌ای n تـایی از جامعـه‌ای N عضوی به طوری که از هر $k=\frac{N}{n}$ عضو جامعه، ۱ فـرد بـه عنـوان نمونـه انتخـاب شـود، عددی از ۱ تا k به تصادف (با قرعه‌کشی) تعیین می‌شود. فرض کنید این عـدد b باشـد، در این صورت نخستین فرد نمونه عضو شـماره b جامعـه، دومـین فـرد نمونـه عضـو شـماره $b+k$ جامعه، سومین فرد نمونه، عضو شماره $b+2k$ جامعه، ... و آخـرین فـرد نمونـه عضو شماره $b+(n-1)k$ جامعه است.

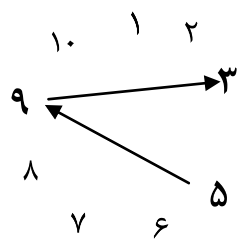
برای مثال در انتخاب یک نمونه $n=4$ تایی از جامعه‌ای $N=20$ تایی، k برابـر بـا ۵ اسـت. پس باید عددی تصادفی از ۱ تا ۵ انتخاب شود. اگر این عدد ۳ باشد، اعضـای شـماره ۳، ۸، ۱۳ و ۱۸ را به عنوان نمونـه از جامعـه انتخـاب مـی‌کنـیم؛ گـویی کـه اعضـای جامعـه بـه ترتیـب شماره‌هایشان در «خط» ردیف شده‌اند و نمونه‌ها با نظام خاصی ۵ تا درمیان انتخاب می‌شوند.



اگر N مضربی از n نباشد، ابتدا عددی تصادفی از ۱ تا N انتخـاب مـی‌شـود تـا شـماره نخستین فرد نمونه مشخص شود. تقسیم N بـر n دارای اعشـار خواهـد بـود، بنـابراین بـه جای $\frac{N}{n}$، نزدیک‌ترین عدد طبیعی بزرگ‌تر از $\frac{N}{n}$ به عنوان k انتخاب مـی‌شـود و ماننـد قبـل افراد نمونه، k در میان انتخاب می‌شوند؛ با این تفاوت که اعضای جامعه بر روی یک دایـره در نظر گرفته می‌شوند.

برای مثال در انتخاب یک نمونه $n=3$ تایی از جامعه‌ای $N=10$ تایی، ابتدا عـددی تصـادفی از ۱ تا ۱۰ انتخاب می‌کنیم تا شماره نخستین فرد نمونه در جامعه مشخص شود. اگـر ایـن عـدد ۵ باشد، عضو شماره ۵ جامعه همان نخستین فرد نمونـه سیسـتماتیک اسـت. از تقسـیم ۱۰ بـر ۳ بـه

۳/۳۳ خواهیم رسید و ۴ نزدیک‌ترین عدد طبیعی بزرگ‌تـر از ۳/۳۳ اسـت، پـس $k = ۴$. بنـابراین، دومـین فـرد نمونـه، عضـو شـماره ۴ + ۵ = ۹ جامعـه و سـومین فـرد نمونـه عضـو شـماره ۴×۲+۵ = ۱۳ جامعه است. ولی جامعه آماری دارای ۱۰ عضو است و عضو شـماره ۱۳ نـدارد. اگر اعضای جامعه به ترتیب شماره‌هایشان دورِ «دایره‌ای» قـرار گرفتـه باشـند فـرد نخسـت نمونـه همان عضو ۵ است، ۴ عضو را رد مـی‌کنـیم تـا بـه عضـو ۹ یعنـی دومین فرد نمونه برسیم و سپس ۴ عضو دیگر را رد می‌کنیم تـا بـه عضو شماره ۳ یعنی همان سومین فرد نمونه برسیم. در واقـع، اگـر اعضای جامعه را دور دایـره‌ای در نظـر بگیریـد، عضـو شـماره ۱۳ همان عضو شماره ۳ است که با گردش دور دایره به آن مـی‌رسـیم. گویی که نمونه‌ها با شروع از نقطهٔ ۵ به طور منظم ۴ تا درمیان از شماره‌های دور این دایره انتخاب می‌شوند. پس شماره نمونه‌ها ۵، ۹ و ۳ خواهد بود. بـر این اساس، اگر نخستین عدد تصادفی برای مثال ۷ بود، شـماره نمونـه‌هـا ۷، ۱ و ۵ یـا اگـر نخستین عدد تصادفی ۶ بود، شماره نمونه‌ها ۶، ۱۰ و ۴ بود.



۱ـ۴ـ۴ـ برآورد

برآورد میانگینِ جامعه از طریق نمونه‌گیری سیستماتیک با میانگینِ نمونه صورت مـی‌گیـرد، درست مانند اینکه یک نمونه‌گیری تصادفی ساده از جامعه صورت گرفته اسـت. بـا وجـود این، خطای معیار برآورد در روش سیستماتیک با روش تصادفی ساده یکسان نیست.

برای یافتن خطای معیار بـرآورد میـانگین توجـه کنیـد کـه روش سیسـتماتیک خطـی می‌تواند k نمونه n تاییِ مختلف از یک جامعـه N عضـوی بـه دسـت دهـد. فـرض کنیـد واریانس متغیر X در این k نمونه n تایی S_1^2، S_2^2، ...، S_k^2 باشد. میانگینِ این واریـانس‌هـا را با $\overline{S^2} = \frac{1}{k}\sum S_i^2$ و واریانس متغیر X را در جامعه مانند قبل با σ_X^2 نشان می‌دهیم. در ایـن صورت، خطای معیار بـرآورد میـانگینِ جامعـه از طریـق روش نمونـه‌گیـری سیسـتماتیکِ «خطی» عبارت است از:

$$SE_{\bar{X}_{sys}} = \sqrt{\frac{N-1}{N}\sigma_x^2 - \frac{k(n-1)}{N}\overline{S^2}}$$

اگرچه تخمینِ σ_x^2 با استفاده از واریانس داده‌های نمونـه n تـایی امکـان‌پـذیر اسـت، ولـی نمی‌توان تخمینی از $\overline{S^2}$ به دست آورد، زیرا تخمینِ $\overline{S^2}$ تنها در شـرایطی ممکـن اسـت کـه حداقل دو نمونه n تایی جداگانه از جامعه در دست باشد[1].

۲ـ۴ـ۴ـ کارایی

از بُعد نظری، هنگامی نمونه‌گیری سیستماتیک کاراتر از نمونه‌گیری تصادفی ساده است کـه $\overline{S^2}$ کوچک‌تر از σ_x^2 باشد. رخ‌دادن این وضعیت کاملاً بـه چگـونگی قرارگـرفتن اعضـای جامعه در چهارچوب نمونه‌گیری باز می‌گردد. اگر اعضا بر حسب مقدار متغیـر بـه ترتیـب صعودی یا نزولی در چهارچوب قرار گرفته باشـند، دقـت روش تصـادفی سـاده بیشـتر از روش سیستماتیک خواهد بود؛ زیرا قرعه‌کشی (روش تصادفی ساده) باعث می‌شـود نمونـه شامل اعضای مختلف جامعه، هم با مقادیر بزرگ هم بـا مقـادیر کوچـک متغیـر باشـد، در حالی که انتخاب سیستماتیک بر حسب اینکه با کدام عضو از جامعه آغاز شود یا به مقـادیر بزرگ یا به مقادیر کوچک گرایش بیشتری خواهد داشت.

برای مثال اگر نمونه با عضوی آغاز شود که مقدار متغیـر بـرای آن بـزرگ اسـت، افـراد بعدی نمونه نیز شامل اعضایی خواهند بود که مقدار متغیر برای آنهـا بـزرگ اسـت. حالـت «الف» در شکل ۵، چهارچوبی ۳۶ عضوی را نشان می‌دهـد کـه بـه ترتیـب صـعودی قـرار گرفته‌اند. انتخاب سیستماتیکِ نمونه‌ای ۴ تایی با آغازی از عضو شماره ۸ به اعضایِ ۸، ۱۷، ۲۶ و ۳۵ منجر می‌شود، در حالی که اگر عضوِ آغازین نمونه شماره ۲ بـود، اعضـایِ ۲، ۱۱، ۲۰ و ۲۹ انتخاب می‌شدند. به‌راحتی می‌توان دید اعضای نمونه نخست شامل مقادیرِ خیلـی بزرگ‌تری نسبت به اعضای نمونه دوم است، زیرا اعضـای جامعـه بـه ترتیـب صـعودی در چهارچوب جای گرفته‌اند.

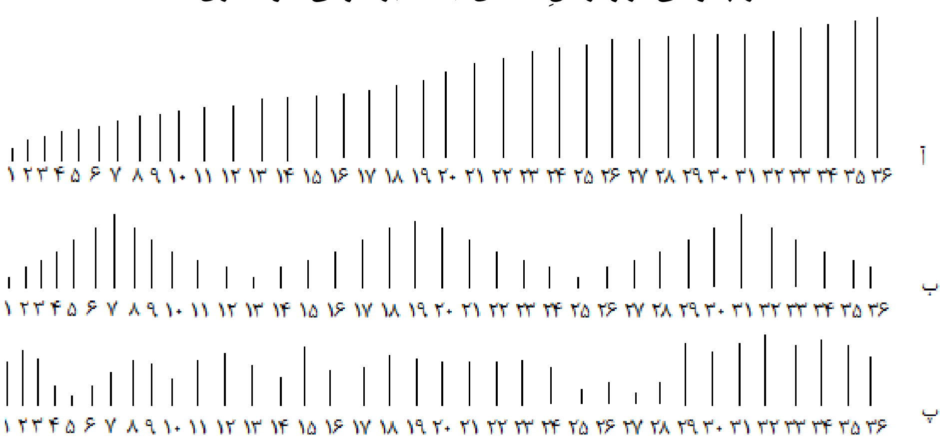
اگر اعضای نمونه به شکل دوره‌ای (نوسانی) در چهارچوب فهرست شده باشـند، ماننـد حالت «ب» در شکل ۵، دقت نمونه‌گیری سیستماتیک بیشتر از تصادفی سـاده اسـت، زیـرا افراد نمونه دارای تنوعی به‌نسبت مشابه با تنوع جامعه خواهنـد بـود. بـرای مثـال، انتخـاب

۱. گاه خطای معیار برآورد میانگین از همان روش تصادفی ساده تخمین زده می‌شود، ولی روش دقیق‌تر اسـتفاده از روش‌هـای بازنمونه‌گیری (Resampling) است که در این کتاب مورد بحث قرار نمی‌گیرند. به زبان ساده، یـک نمونـه n تـایی را مـی‌تـوان برای مثال دو نمونه با اندازه‌های a و b درنظر گرفت، به طوری که $n=a+b$ باشد. در این صورت دو مقدار بـرای واریـانس از این دو نمونه به دست خواهد آمد که بر اساس آنها به تخمینی از میانگینِ واریانس‌ها دست خواهیم یافت.

نمونه‌ای ۴ تایی با عضو آغازینِ ۳، به اعضای ۳، ۱۲، ۲۱ و ۳۰ می‌انجامد. این نمونه، هم شامل اعضایی با مقدار کم یا متوسط مانند ۳ و ۱۲ است و هم شامل اعضایی با مقدار به‌نسبت بزرگ مانند ۲۱ و ۳۰. حتی اگر عضو آغازین نمونه، شماره ۷ باشد، نمونه حاصل (اعضایِ ۷، ۱۶، ۲۵ و ۳۴)، همچنان دارای تنوعی مانند تنوع جامعه است.

اگر برخلاف دو وضعیت قبل، هیچ نظمی در ترتیب قرارگرفتن اعضای جامعه در چهارچوب نمونه‌گیری وجود نداشته باشد (حالت «پ» در شکل ۵)، دقت روش نمونه‌گیری سیستماتیک مانند روش تصادفی ساده است. معمولاً چنین وضعیتی در عمل برای چهارچوب نمونه‌گیری وجود دارد. از این رو، انتخاب سیستماتیک تنها موجب ساده‌تر شدن فرایند انتخاب، در مقایسه با روش قرعه‌کشی می‌شود. با وجود این، برخلاف قرعه‌کشی که هیچ کنترلی روی آن وجود ندارد، انتخاب سیستماتیکِ افراد نمونه باعث می‌شود این افراد به طور یکنواخت‌تری از سرتاسر جامعه آماری انتخاب شوند، بنابراین اطمینان خواهیم داشت نمونه با خطر انتخاب شدن از بخش خاصی از جامعه تهدید نمی‌شود، مگر آنکه قرارگرفتن اعضا در چهارچوب به گونه خاصی باشد که ناخودآگاه به انتخابی جهت‌دار و سوگیرانه بینجامد.

شکل ۵: تأثیر چگونگی قرارگرفتنِ اعضای جامعه بر کارایی نمونه‌گیری سیستماتیک



۵ـ نمونه‌گیری‌های پیچیده

خواننده در این قسمت با نمونه‌گیری با احتمال نابرابر، نمونه‌گیری چند مرحله‌ای و نمونه‌گیری تلفیقی به اجمال و گذرا آشنا خواهد شد. اگرچه این روش‌ها برخاسته از همان روش‌های کلاسیک هستند، ولی با تغییراتی همراهند که به پیچیدگی آنها منجر شده است.

این پیچیدگی در نوع چهارچوب، تعیین اندازه نمونـه، فراینـد انتخـاب نمونـه‌هـا یـا روش برآورد و محاسبه خطای معیار آن منعکس می‌شود.

۱ـ۵ـ نمونه‌گیری با احتمال نابرابر

قرعه‌کشی و حتی روش سیستماتیک به همـۀ اعضـای جامعـه آمـاری، امکـان برابـری را بـرای انتخاب شدن به عنوان یکی از افراد نمونه می‌دهـد. بـرای مثـال، در انتخـاب تصـادفیِ ۲ فـرد از جامعه‌ای ۱۰ عضوی، احتمال اینکه یکی از این ۱۰ عضو به عنوان نخستین فـردِ نمونـه انتخـاب شود، برای تک‌تک اعضا برابر با $\frac{1}{10}$ است[1] و هیچ‌یک از اعضا نسبت به دیگری شانس بیشـتری برای انتخاب شدن ندارد. به همین دلیل این شیوۀ انتخاب را با احتمالِ «برابر» می‌نامند.

اکنون فرض کنید به جای آنکه برای قرعه‌کشی نام هر یک از ۱۰ عضو روی یک کارت جداگانه نوشته شود و از میان این ۱۰ کارت، کارتی به تصادف انتخاب گردد، نـام عضـوی بر روی ۲ کارت نوشته شود و از میان ۱۱ کارت، کارتی به تصادف انتخاب گـردد. در ایـن صورت احتمال انتخاب این عضو $\frac{2}{11}$، ولی احتمال انتخاب ۹ عضو دیگر جامعه $\frac{1}{11}$ اسـت. در این حالت، نمونه‌گیری با احتمال نابرابر صورت گرفته است، زیرا تمام اعضـا از شـانس برابری برای انتخاب شدن به عنوان نمونه برخوردار نیستند. توجه کنید که نمونـه‌گیـری بـا احتمال نابرابر همچنان در دستۀ نمونه‌گیری‌های احتمالی قلمداد می‌شود، زیـرا تمـام اعضـا «احتمالِ» انتخاب شدن دارند، هرچند این احتمال برای همۀ آنها یکسان نیست.

شرایطی وجود دارد که نمونه‌گیری با احتمال نابرابر منطقی‌تر از نمونه‌گیری بـا احتمـال برابر است. برای مثال، منطقی است بانکی که می‌خواهد از طریق قرعه‌کشی به سپرده‌گذاران خود جایزه بدهد، شانسی متناسب با سپردۀ افراد در انتخاب در نظر بگیرد. فرض کنید ایـن بانک بخواهد برای مثال از میان ۵ فرد با سپرده‌های جدول صفحه بعـد، یکـی را بـه عنـوان برندۀ جایزه انتخاب کند و هر سپردۀ ۵۰ تایی یک امتیاز به حساب آید. در این صورت بایـد بیشترین شانس انتخاب به کامران و کمترین شانس انتخاب به کامبیز داده شـود. از ایـن رو، ایـن بانک ۱۷ کارت، به تعداد کل امتیازها، تهیه می‌کند؛ به طوری که ۴ کارت به شاهین، ۲ کـارت بـه سامان، ۱ کارت به کامبیز، ۴ کارت به شاهرخ و ۶ کارت به کامران اختصاص می‌دهد. حال بایـد از طریق قرعه‌کشی، ۱ کارت از میان ۱۷ کارت انتخاب شود تا برندۀ جایزه تعیین گـردد. اگرچـه

---

۱. وقتی نخستین فرد نمونه از جامعه انتخاب شود، ۹ عضو دیگر باقی می‌ماند که هر یک می‌تواند به عنوان دومین فـردِ نمونه انتخاب شود. در اینجا نیز تمامِ ۹ عضو باقیمانده، احتمال برابری برای انتخاب شدن دارند یعنی $\frac{1}{9}$.

انتخاب کارت برنده از طریق قرعه‌کشی صورت می‌گیرد و تک‌تک سپرده‌گذاران کارتی در میان کارت‌ها دارند، ولی دارندهٔ کارت بیشتر (یعنی کامران) شانس بیشتری نیز برای برنده شدن دارد. با وجود این، برنده می‌تواند کامبیز، دارندهٔ تنها ۱ کارت از ۱۷ کارت باشد، هرچند احتمال وقوع آن $\frac{۱}{۱۷}$ و کمتر از سایر سپرده‌گذاران است.

سپرده‌گذار	شاهین	سامان	کامبیز	شاهرخ	کامران
سپرده	۲۰۰	۱۰۰	۵۰	۲۰۰	۳۰۰
امتیاز	۴	۲	۱	۴	۶

مثال مناسب دیگری از نمونه‌گیری با احتمال نابرابر به انتخاب بلوک‌های شهر در نمونه‌گیری خوشه‌ای باز می‌گردد. اگر بلوک‌ها دارای جمعیت‌های بسیار متفاوتی باشند، منطقی نیست به بلوکی کم‌جمعیت همان امکان انتخابی داده شود که به بلوکی پرجمعیت داده می‌شود. در این صورت می‌توان احتمال انتخاب هر بلوک را متناسب با اندازه جمعیت آن درنظر گرفت تا بلوک‌های پرجمعیت شانس بیشتری برای انتخاب شدن داشته باشند. البته این بدان معنا نیست که هیچ بلوک کم‌جمعیتی به عنوان نمونه انتخاب نخواهد شد، زیرا بلوک‌های کم‌جمعیت نیز دارای احتمال انتخابی هرچند ضعیف‌تر هستند.

۲ـ۵ـ نمونه‌گیری چند مرحله‌ای

پژوهشگری که مایل است میزان اعتماد اجتماعی شهروندان شهری را اندازه بگیرد، از طریق خانوارها به شهروندان دسترسی پیدا می‌کند. او می‌تواند از نمونه‌گیری خوشه‌ای از بلوک‌های شهر برای گردآوری داده‌ها استفاده کند. برای این منظور، تعدادی بلوک به تصادف انتخاب و به «تمام» خانوارهای ساکن در آنها مراجعه می‌کند و اعتماد اجتماعی شهروندان هر خانوار را اندازه می‌گیرد.

در نگاه نخست، این نمونه‌گیری خوشه‌ای مناسب به نظر می‌رسد، در حالی که ممکن است با مشکلی جدی روبه‌رو باشد. اگر شهر دارای بلوک‌هایی با جمعیت‌های بسیار متفاوت باشد، به طوری که ۲ یا ۳ بلوکِ نمونه از بلوک‌های پرجمعیت و سایر بلوک‌های نمونه از بلوک‌های کم‌جمعیت انتخاب شده باشند، آنگاه ممکن است برای مثال ۳۰ درصد اندازه نمونه تنها از این ۲ یا ۳ بلوک گردآوری شده باشند! اگر بلوک‌ها از ناهمگنی درونی نیز برخوردار نباشند که اغلب چنین است، نمونه حاصل چندان معرّف جامعه آماری پژوهش نخواهد بود.

اکنون فرض کنید که پژوهشگر به جای گردآوری داده‌ها از تمام خانوارهای بلوک‌های نمونه، بخشی از خانوارهای هر بلوک را به تصادف انتخاب کند و داده‌ها را تنها از شهروندان خانوارهای انتخابی گرد آورد. به این ترتیب، او باید بلوک‌های بیشتری را نسبت به وضعیت قبل انتخاب کند تا در نهایت بتواند اعتماد اجتماعی همان تعداد شهروند را اندازه بگیرد. در واقع، دو مرحله‌ای‌کردنِ نمونه‌گیری به پژوهشگر این امکان را می‌دهد تا با انتخاب بلوک‌های بیشتر تنوع نمونه خود را افزایش دهد و به نمونه معرف‌تری دست یابد.

مثال قبل را نمونه‌گیری دو مرحله‌ای می‌نامند، زیرا شامل مرحلهٔ انتخاب تصادفی بلوک‌ها و مرحلهٔ انتخاب تصادفی خانوارها از هر بلوک است. از این رو، به دو چهارچوب نمونه‌گیری نیاز دارد و به تبع آن، دو نوع واحد نمونه‌گیری برای آن تعریف می‌شود. چهارچوب نخست باید شامل فهرستی از بلوک‌های شهر با نشانی جغرافیایی آنها باشد؛ هر عضوِ این چهارچوب را واحد مرحلهٔ نخست[1] (PSU) می‌نامند، زیرا بلوک‌ها واحدهایی هستند که در نخستین مرحلهٔ نمونه‌گیری انتخاب می‌شوند. چهارچوب دوم شامل فهرست خانوارهای هر بلوک انتخاب‌شده در مرحلهٔ نخست و نشانی این خانوارهاست؛ هر عضوِ این چهارچوب، واحد مرحلهٔ دوم[2] (SSU) نامیده می‌شود، زیرا خانوارها واحدهایی هستند که در دومین مرحلهٔ نمونه‌گیری انتخاب می‌شوند.

نمونه‌گیری از اعضای خانوارها یکی از متداول‌ترین مثال‌های نمونه‌گیری چند مرحله‌ای است. در مثال قبل، معمولاً مرحلهٔ سومی نیز به نمونه‌گیری افزوده می‌شود که شامل انتخاب ۱ فرد از هر خانوار است. در این صورت واحد نمونه‌گیری مرحلهٔ سوم، اعضای خانوارهایی هستند که از هر بلوک به تصادف انتخاب شده‌اند.

1. Primary Stage Unit (PSU)
2. Secondary Stage Unit (SSU)

مرحله‌بندی نمونه‌گیری مستلزم تهیهٔ چهارچوب مناسب برای هر مرحله است. همچنین، باید اندازه نمونه هر مرحله تعیین شود؛ چه تعداد بلوک به تصادف از میان بلوک‌ها انتخاب شود؟ چه تعداد خانوار از هر بلوک انتخاب گردد؟ و داده‌ها از چند عضو هر خانوار گردآوری شود؟ مسئله اساسی‌تر، روش برآورد و محاسبه خطای معیار آن است. هر مرحلهٔ نمونه‌گیری خود منبع بروز خطاست که باید در محاسبه خطای معیار برآورد منعکس شود.

هنگامی نمونه‌گیری چند مرحله‌ای پیچیده‌تر می‌شود که در یک یا بیش از یک مرحلهٔ آن به انتخاب با احتمال نابرابر دست بزنیم؛ برای مثال، بلوک‌ها با احتمالی متناسب با جمعیت، و به همین ترتیب، خانوارهای هر بلوک نیز با احتمالی متناسب با تعداد اعضای خانوار انتخاب شوند. معمولاً انتخاب خانوارها با روش سیستماتیک ـ برای مثال هر ۳ تا در میان ـ صورت می‌گیرد. استفاده از روش انتخاب با احتمال نابرابر می‌تواند روش برآورد پارامترهای نامعلوم و محاسبه خطای معیار آنها را پیچیده‌تر سازد.

۳ـ۵ـ نمونه‌گیری تلفیقی

درعمل، پژوهشگران همزمان از چند روش نمونه‌گیری استفاده می‌کنند تا فرایند گردآوری داده‌ها از مزایای روش‌های مختلف برخوردار باشد. برای مثال، پژوهشگری که در جامعه آماری دانش‌آموزان دبستانی دست به مطالعه می‌زند، نه تنها از روش خوشه‌ای، بلکه همزمان از روش طبقه‌ای نیز برای گردآوری داده‌ها کمک می‌گیرد. او ابتدا دبستان‌ها را به دو طبقهٔ دولتی و خصوصی تقسیم می‌کند و سپس در هر طبقه به نمونه‌گیری خوشه‌ای سه مرحله‌ای می‌پردازد؛ در مرحلهٔ نخست به تصادف، تعدادی دبستان از هر طبقه انتخاب می‌شود. در مرحلهٔ دوم، تعدادی کلاس از هر مدرسه انتخاب می‌کند و در مرحلهٔ سوم، به انتخاب تعدادی دانش‌آموز از هر کلاس دست می‌زند.

نمونه‌گیری تلفیقی می‌تواند علاوه بر تلفیق چند روش، از روش انتخاب با احتمال نابرابر نیز در یک مرحله یا تمام مراحل انتخاب استفاده کند. پژوهشگر باید در نمونه‌گیری تلفیقی نسبت به تعیین اندازه نمونه مراحل مختلف، شیوهٔ تخصیص نمونه و مسائلی از این دست تصمیم بگیرد. نمونه‌گیری‌های تلفیقی از روش‌های پیچیده‌ای برای برآورد پارامتر و محاسبه خطای معیار آن استفاده می‌کنند که تأثیر روش‌های مختلف نمونه‌گیری را در خود منعکس می‌سازند.

اصطلاحات فصل هفتم

Accessible Sampling	نمونه‌گیری در دسترس	Probability Sampling	نمونه‌گیری احتمالی
Accidental/haphazard Sampling	نمونه‌گیری اتفاقی	Proportional Allocation	تخصیص متناسب
Cluster Sampling	نمونه‌گیری خوشه‌ای	Purposive Sampling	نمونه‌گیری هدفمند
Convenience Sampling	نمونه‌گیری سهل‌الوصول	Quota Sampling	نمونه‌گیری سهمیه‌ای
Efficiency	کارایی	Randomized Selection	انتخاب تصادفی
Experimental Data	داده‌های آزمایشی	Reliability	پایایی
Frame	چهارچوب	Representative	معرف، نمایا
Generalization	تعمیم	Sampling Method	روش نمونه‌گیری
Judgmental Sampling	نمونه‌گیری	Sampling Unit	واحد نمونه‌گیری
Mechanism	سازوکار	Secondary Stage Unit	واحد مرحله دوم
Multi-stage Sampling	نمونه‌گیری چندمرحله‌ای	Simple Random Sampling	نمونه‌گیری تصادفی ساده
Neyman Allocation	تخصیص نیمن	Snowball Sampling	نمونه‌گیری گلوله برفی
Nonprobability Sampling	نمونه‌گیری نااحتمالی	Stratified Sampling	نمونه‌گیری طبقه‌ای
Observational Data	داده‌های مشاهداتی	Systematic Sampling	نمونه‌گیری سیستماتیک (نظام‌مند)
Optimum Allocation	تخصیص بهینه	Two-stage Sampling	نمونه‌گیری دو مرحله‌ای
Primary Stage Unit	واحد مرحله نخست	Volunteer Sampling	نمونه‌گیری داوطلبانه
Probability Proportional to Size	احتمال متناسب با اندازه		

تمرین‌های فصل هفتم

۱. روزنامه‌ای مایل است علاقه‌مندی مردم به موضوعات مختلف خبری را شناسایی کند. این روزنامه در یکی از شماره‌هایش از خوانندگان خود می‌خواهد طی هفته آینده با دفتر روزنامه تماس بگیرند و موضوع خبری مورد علاقه خود را بیان کنند. در نهایت ۶۸۰ نفر با این روزنامه تماس می‌گیرند. نمونه حاصل احتمالی است یا نااحتمالی؟ روش آن چیست؟ درباره جامعه آماری این بررسی بحث کنید.

۲. جامعه آماری، واحد نمونه‌گیری و روش نمونه‌گیری را در هر یک از مثال‌های زیر مشخص کنید.

آ) درصد دبیرستان‌های کلان شهری که کلاس‌های فوق برنامه هنری برگزار می‌کنند با انتخاب تصادفی نمونه‌ای ۵۰ تایی از فهرست این دبیرستان‌ها برآورد شده است.

ب) همان مثال «آ» با این تفاوت که دبیرستان‌ها به دو گروه دولتی و غیردولتی تقسیم شدند و ۳۵ دبیرستان دولتی و ۱۵ دبیرستان غیردولتی به تصادف از فهرست انتخاب شد.

پ) همان مثال «ب» با این تفاوت که ابتدا ۳ منطقه به تصادف از میان مناطق آموزش و پرورش انتخاب و سپس از هر منطقه تعدادی دبیرستان دولتی و تعدادی دبیرستان غیردولتی به تصادف انتخاب شد.

ت) برآورد درصد دانش‌آموزان دبیرستانی که در کلاس‌های فوق برنامه هنری شرکت می‌کنند، مورد نیاز است. برای این منظور، دبیرستان‌ها به دو گروه دولتی و غیردولتی تقسیم شدند. هر گروه نیز به دو گروه دارای حداکثر ۲۰۰ دانش‌آموز و دارای بیش از ۲۰۰ دانش‌آموز تقسیم شد. سپس به تصادف از هر یک تعدادی دبیرستان انتخاب و با مراجعه به تمام دانش آموزان، درصد شرکت‌کنندگان در برنامه‌های فوق برنامه هنری برآورد شد.

ث) همان مثال «ت» با این تفاوت که در هر دبیرستان، ۵۰ دانش‌آموز به تصادف انتخاب و وضعیت شرکت آنها پرسیده شد.

۳. شرکتی تصمیم دارد با روش تصادفی ساده، پیمایشی به منظور برآورد میانگین ماهانه هزینه‌های درمانی خانوارهای ۵۰۰۰ کارمند خود اجرا کند. اگر این شرکت مایل باشد میانگین هزینه‌ها را با اطمینان ۹۵ درصد در فاصله ۶۰۰± تومان برآورد کند و از پیمایش مشابه دو سال گذشته خود بداند که واریانس هزینه‌های درمانی ۵۵۰۰ است، حداقل نمونه لازم برای برآورد میانگین چقدر است؟ تعداد نمونه در صورتی که سطح اطمینان ۹۹ درصد باشد، چقدر است؟

۴. در تمرین قبل، اگر قرار باشد کارمندان بر اساس جنس و وضعیت بازنشسته یا شاغل بودن به چهار طبقه تقسیم شوند، به طوری که جمعیت و واریانس هر طبقه به شرح جدول زیر باشد، اندازه نمونه کلی و اندازه تخصیص‌یافته به هر طبقه را به سه روش متناسب، نیمن و بهینه محاسبه کنید (فرض کنید هزینه گردآوری داده از شاغلان ۵۰۰ تومان و از بازنشسته‌ها ۲۰۰۰ تومان است و هزینه کلی نباید از ۱۲۰۰۰۰۰ تومان فراتر رود. در صورت نیاز $se = ۷۰$ در نظر بگیرید).

جنس	وضعیت	بازنشسته	شاغل
زن	جمعیت	۳۰۰	۱۳۰۰
	واریانس	۵۶۰۰	۴۹۰۰
مرد	جمعیت	۴۰۰	۳۰۰۰
	واریانس	۵۸۰۰	۵۱۰۰

فصل هشتم

۸ پیمایش

حوزه‌های سه‌گانهٔ آمار در پژوهش‌هایی با روش‌های مختلف کاربرد دارند، ولی پیمایش روشی است که هر سه حوزه را به طور گسترده در اجرای پژوهش به خدمت می‌گیرد. این فصل، روش پیمایشی را از نگاه آماری معرفی می‌کند تا پیوند حوزه‌های مختلف آمار را با پژوهش آشکار سازد. خواننده از میان بحث‌های آن با پیچیدگی‌های پژوهش پیمایشی تاحدی آشنا خواهد شد و درمی‌یابد که چگونه روش‌های نمونه‌گیری در عمل به چالش کشیده می‌شوند.

۱ـ روش پیمایشی

پیمایش یکی از رایج‌ترین و پرکاربردترین روش‌های کمّیِ پژوهش، بویژه در مطالعهٔ جامعه‌های آماری بزرگ است. در واقع، پیمایش فرایندی نظام‌مند است که طی آن، یک یا چند متغیر از چندین عضو جامعه آماری تحت مطالعه اندازه‌گیری می‌شود تا اطلاعاتی درباره آن جامعه به دست آید. بنابراین، روش‌های نمونه‌گیری می‌توانند نقشی اساسی در چگونگی انتخاب اعضای جامعه آماری به منظور اندازه‌گیری متغیرها ایفا کنند تا اطلاعات حاصل از این اعضا به شکلی معتبر، قابل تعمیم به جامعه باشد. همچنین، روش‌های آمار توصیفی در سازماندهی و گزارش داده‌های حاصل از اندازه‌گیری متغیرها، و روش‌های آمار استنباطی در نتیجه‌گیری از این داده‌ها درباره جامعه آماری سودمندند.

توانایی پیمایش در مطالعهٔ همزمان چندین متغیر، آن هم در جامعه‌های آماری بزرگ، آن را روشی پرطرفدار ساخته است. روش پیمایشی می‌تواند تنها روش به کار رفته در پژوهش یا به عنوانِ مکمل همراه با روش‌های دیگر، در خدمت پژوهش باشد. معمولاً روش پیمایشی در پژوهش‌های چندروشی[1] یا آمیخته‌روش[2] در کنار سایر روش‌ها استفاده می‌شود. به این ترتیب، پیمایش می‌تواند گستره و اعتبار یافته‌های پژوهش را وسعت دهد و ابعادی از پدیدهٔ تحت بررسی را آشکار سازد که سایر روش‌ها قادر به شناخت آن نیستند.

دامنهٔ استفاده از روش پیمایشی بسیار گسترده است. نظرسنجی‌ها و افکارسنجی‌ها با موضوعات بسیار متنوع از یک سو و آمارگیری‌های مورد نیاز دولت‌ها از سوی دیگر، همه با روش پیمایشی صورت می‌گیرند. در واقع، پیمایش‌ها اطلاعات آماری گوناگونی را به دست می‌دهند که مبنای شناخت جامعه‌ها، قضاوت درباره آنها و تصمیم‌گیری و برنامه‌ریزی برای آنهاست. با وجود این، از دیدگاه پژوهشی، پیمایش‌ها تنها به تولید آمارها نمی‌پردازند، بلکه می‌توانند با هدف تبیین پدیده‌ها یا پیش‌بینی وضعیت آنها در آینده نیز به کار گرفته شوند.

۲ـ مراحل پیمایش

از نظر آماری، پیمایش مطابق شکل ۱ با طرح اهداف آن آغاز می‌شود. اهداف، شامل اطلاعاتی درباره جامعه آماری تحت مطالعه است که از پیمایش انتظار می‌رود. بنابراین،

1. Multi-method Research
2. Mixed-method Research

پارامترهایی که باید برآورد شوند، فرضیه‌هایی که باید مورد آزمون قرار گیرند و جدول‌هایی که باید برای نمایش توزیع متغیرها ارائه گردند، از اهداف «آماری» پیمایش قلمداد می‌شوند. در گام دوم، امکان دستیابی به اهداف پیمایش با توجه به جامعه آماری آن، بودجهٔ پژوهش، دقت مورد انتظار، امکانات و توانمندی‌ها و... بررسی می‌شود.

گام سوم، ساختنِ ابزار سنجش متغیرهایی است که اهداف پیمایش معطوف به آنهاست. پرسشنامه، رایج‌ترین وسیله برای اندازه‌گیری متغیرهاست. از آنجا که بیشتر پیمایش‌ها به گردآوری داده‌های مشاهده‌ای می‌پردازند، ابزار سنجش متغیرها وسیله‌ای برای «ثبت» مقادیر متغیرها از افراد نمونه است که در قالب مجموعه‌ای از پرسش‌ها سازمان یافته‌اند. معمولاً هر پرسش برای ثبت مقدار یک متغیر است. پرسش‌ها می‌توانند به صورت باز یا بسته طراحی شوند. پاسخ‌های محتمل پرسش بسته از قبل پیش‌بینی و در کنار آن پرسش، در قالب چند گزینه آورده می‌شوند. پرسش باز دارای هیچ پاسخ از پیش تعیین‌شده‌ای نیست، بلکه درست همان پاسخ فردِ نمونه در مقابل آن ثبت می‌شود[1].

برای مثال، اگر از فردِ نمونه سؤال شود که «مهم‌ترین ویژگی یک آپارتمان خوب کدام است؟» و پاسخ‌های «بزرگ‌بودن»، «شیک‌بودن»، «واقع شدن در طبقات بالای ساختمان»، «کم‌جمعیت بودن» و «واقع شدن در یک محله خوب» به عنوان پاسخ‌های محتمل در نظر گرفته شود تا فردِ نمونه یکی از آنها را انتخاب کند، با پرسش بسته روبه‌رو هستیم. در مقابل، اگر این پرسش بدون هیچ پاسخی برای فردِ نمونه مطرح شود تا او در قالب یک یا چند جمله کوتاه ویژگی آپارتمان مناسب را بیان کند، با پرسشی باز روبه‌رو هستیم.

گام چهارم به تعیین روش نمونه‌گیری باز می‌گردد. اهداف پیمایش، دسترسی به چهارچوب مناسب، محدودیت‌های زمانی، بودجه و امکانات، از عامل‌های مهم در تعیین روش نمونه‌گیری هستند. اکنون پژوهشگر آمادگی دارد ابزار سنجش خود را به‌طور آزمایشی اجرا کند تا ضعف‌های آن مشخص شوند. اجرای آزمایشی در سطحی محدود، نه تنها به بازبینی و اصلاح ابزار سنجش کمک می‌کند، بلکه مشکلات پیش روی اجرا و نمونه‌گیری را نیز تا حدی نمایان می‌سازد. به این ترتیب، پژوهشگر با آگاهی و آمادگی بیشتری وارد مرحلهٔ اجرای اصلی پیمایش خواهد شد.

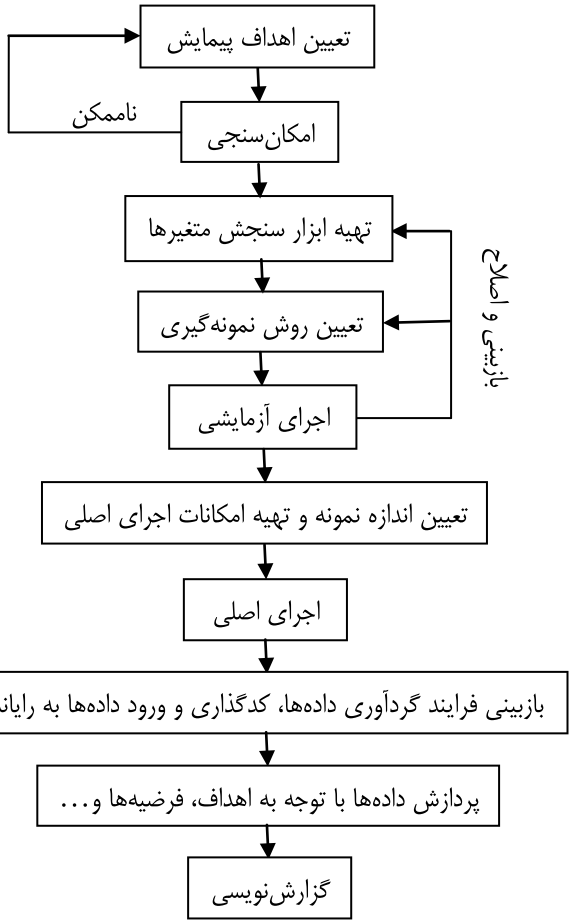
گام ششم، به مهیا کردن شرایط اجرای اصلی پیمایش اختصاص دارد. تعیین اندازه

---

۱. در اینجا از پرسش‌های نیمه‌بسته (نیمه‌باز) صحبتی نشده است.

نمونه متناسب با روش نمونه‌گیری با توجه به خطا، دقت و هزینه، از مهم‌ترین فعالیت‌های آن است. انتخاب نمونه‌ها از چهارچوب، سازماندهی و آموزش پرسشگرانی که باید به افراد نمونه مراجعه و داده‌ها را از آنان گردآوری کنند، تهیهٔ شیوه‌نامه‌های گردآوری و نظارت بر اجرای دقیق فرایند نمونه‌گیری و گردآوری داده‌ها و برنامه‌ریزی و زمان‌بندی فرایند اجرا از جمله فعالیت‌های این گام است.

شکل ۱: گام‌های پژوهش پیمایشی



گام هفتم، شامل مراجعه به افراد نمونه و گردآوری داده‌ها بر اساس زمان‌بندی صورت گرفته است. این گام از پیمایش، به مدیریت و نظارت دقیق و نزدیک نیاز دارد. ممکن است برخی از افراد نمونه در زمان مراجعه حضور نداشته باشند، پس باید پیگیری و مراجعه مجدد در برنامه پیش‌بینی شده باشد. مدیران اجرایی و سرپرست‌ها باید آمادگی

کامل را برای رویارویی با ابهامات و سؤالات پرسشگران و گروه‌های اجرایی، مشکلات احتمالی در یافتن افراد نمونه، غیبت ناگهانی برخی از اعضای گروه و... داشته باشند. نظارت بر فرایند اجرا مانند مراجعه درست به افراد نمونه، تکمیل صحیح پرسشنامه‌ها و رعایت شیوه‌نامه‌ها از وظایف اصلی مدیران و سرپرستان گروه‌های اجرایی است که در نوبت‌های مختلف صورت می‌گیرد.

پس از پایان فرایند مراجعه به تمام افراد نمونه و گردآوری داده‌ها از آنان، به گام هشتم پیمایش وارد می‌شویم. پرسشنامه‌های تکمیل‌شده بار دیگر وارسی می‌شوند. ممکن است لازم باشد به برخی از افراد نمونه دوباره مراجعه شود تا صحت فرایند اجرا نیز مورد بازبینی قرار گیرد. کدگذاری پاسخ‌های سؤالات باز یعنی شماره‌گذاری این پاسخ‌ها برای ورود آنها به رایانه از فعالیت‌های این گام است. سپس تمام داده‌ها وارد رایانه می‌شوند تا مجموعه‌ای از آنها برای پردازش آماری فراهم شود.

گام نهم، شامل پردازش‌های آماری است که از قبل با توجه به اهداف پیمایش، پیش‌بینی شده‌اند. رسم جدول‌های فراوانی، جدول‌های توافقی و نمودارهای مورد نیاز، محاسبه برآوردهای نقطه‌ای و فاصله‌ای و اجرای آزمون‌های آماری همگی در این گام صورت می‌گیرند. اگر پژوهشگر از ابتدا اهداف آماری خود را با تفصیل تعیین کرده باشد، در زمان پردازش به‌راحتی و بدون هیچ سردرگمی می‌داند که به چه اطلاعاتی برای هر متغیر نیاز دارد و از چه روش‌های توصیفی یا استنباطی باید برای دستیابی به این اطلاعات استفاده کند.

گزارش پیمایش در آخرین گام تهیه می‌شود. علاوه بر ارائهٔ یافته‌های آماری و تحلیل‌ها و تفسیرهای آنها، ضروری است پژوهشگر فرایند اجرای پژوهش بویژه نمونه‌گیری را نیز در گزارش شرح دهد. تعریف جامعه آماری، روش نمونه‌گیری و دلیل انتخاب آن، منبع چهارچوب نمونه‌گیری، اندازه نمونه و عوامل تأثیرگذار بر آن، از جمله موضوعاتی هستند که در گزارش بیان می‌شوند. ارائه توزیع افراد نمونه بر اساس متغیرهای جمعیتی نظیر سن و جنس یا متغیرهای مؤثر در نمونه‌گیری ضروری است، زیرا نشان می‌دهد یافته‌های آماری برآمده از نمونه‌ای با چه ترکیبی است.

۳ـ شکل‌های استفاده از پیمایش‌ها

پیمایش‌ها به دو شکل مقطعی یا طولی در پژوهش استفاده می‌شوند. پیمایش مقطعی به استنباط آماری درباره جامعه آماری در مقطعی خاص از زمان محدود می‌شود. برای مثال،

فرض کنید دولت برنامه اصلاحات اقتصادی را در دست اجرا دارد و مایل است پیش از اجرا، از افکار عمومی درباره این برنامه آگاه شود تا تمهیدات لازم را برای تنویر اذهان مردم به کار گیرد. از این رو نظرسنجی گسترده‌ای را اجرا می‌کند تا بر اساس اطلاعات به دست آمده، برای آن مقطع زمانی برنامه‌ریزی کند. معمولاً، پیمایش مقطعی یک بار اجرا و از یافته‌های آن در فاصله کوتاهی پس از اجرا استفاده می‌شود.

پیمایش طولی، متغیرهای ثابتی را در دوره‌هایی با فاصلهٔ زمانی برابر، برای مثال به طور فصلی، اندازه‌گیری می‌کند. دو دلیل اصلی برای تکرار پیمایش وجود دارد. نخست آنکه مطالعهٔ یک یا چند متغیر در دوره‌های زمانی مختلف به شناسایی تغییرات و نوسانات آن در طول زمان کمک می‌کند. برای مثال می‌توان نرخ مهاجرت از روستا به شهر را به منظور شناسایی روند تغییرات سالانه آن، هر ساله برآورد کرد تا پس از چند سال به روند تغییرات سالانه آن دست یافت. نکته اساسی در چنین مطالعاتی، اجرای پیمایش در چندین دوره متوالی با رعایت فاصلهٔ زمانی به‌نسبت یکسان است تا امکان آشکار شدن روند تغییرات به شکلی واقعی فراهم شود.

گاه هدف از پیمایش، برآورد پارامتری در یک دوره زمانی خاص است، ولی به‌دلیل تغییراتی که برای این پارامتر در یک فاصلهٔ زمانی رخ می‌دهد، به اجبار پیمایش در داخل این فاصلهٔ زمانی در چند نوبت تکرار می‌شود تا اثر دوره‌های زمانی مختلف حذف شود. برای مثال، اگر هدف پیمایش، برآورد نرخ بیکاری در یک فاصلهٔ زمانی یک‌ساله باشد، نمی‌توان این پیمایش را تنها در پایان سال اجرا کرد، زیرا نرخ حاصل از آن، نمایانگر نرخ بیکاری کل سال نیست. به عبارت دیگر، از آنجا که نرخ بیکاری در دوره‌های مختلف سال (برای مثال هر فصل) متفاوت است، نرخی که با این روش برآورد می‌شود، تنها نمایانگر نرخ بیکاری زمان پیمایش است. بنابراین، پیمایش در چند نوبت از سال (برای مثال هر فصل) اجرا و از ادغام برآوردهای دوره‌های مختلف، نرخ بیکاری سالانه، برآورد می‌شود.

پیمایش‌های طولی به سه شکل مختلف اجرا می‌شوند. ساده‌ترین شکل آن شامل اجرای پیمایش‌های مستقل از یکدیگر است؛ یعنی انتخاب نمونه‌ای متفاوت در هر دوره زمانی (برای مثال هر فصل یا ماه) از جامعه تحت بررسی. به این ترتیب، پس از چند دوره زمانی، چند نمونه «مستقل» از هم وجود خواهد داشت که بر اساس هر یک می‌توان برای مثال،

پارامتر مورد نظر پژوهش را برای آن دوره زمانی خاص برآورد کرد و در نهایت، روند تغییرات آن پارامتر جامعه را در طول زمان مورد شناسایی قرار داد.

پیمایش پانلی شکل دیگری از پیمایش طولی است که داده‌ها را از نمونه «ثابتی» در طول زمان گردآوری می‌کند. به عبارت دیگر، به افراد ثابتی که در نخستین دوره زمانی انتخاب شده‌اند، در دوره‌های زمانی مختلف، بار دیگر مراجعه می‌شود. این بدان معناست که از یک نمونه پس از چند دوره زمانی، چند مجموعه از داده‌ها در اختیار خواهد بود. بنابراین، در مقایسه با پیمایش مستقل، پس از چند دوره زمانی، چند نمونهٔ «وابسته» وجود خواهد داشت. در پیمایش پانلی، ثابت‌بودن نمونه در دوره‌های زمانی مختلف به حذف «اثرِ نمونه» در شناسایی دقیق‌تر تغییرات کمک می‌کند. به عبارت دیگر، برخلاف پیمایش مستقل، تفاوت مشاهده‌شده در برآوردهای پارامتر در دو دوره زمانی را نمی‌توان ناشی از تفاوت افراد نمونه قلمداد کرد. این شکل پیمایش مکرر، دقت برآورد تغییرات را افزایش می‌دهد.

پیمایش چرخشی، شکل سوم پیمایش طولی و تلفیقی از دو روش پیمایش مستقل و پیمایش پانلی است. در این شکل، افراد نمونه در طول زمان نه کاملاً ثابت هستند و نه کاملاً متفاوت، بلکه بخشی از آنها بر اساس الگویی موسوم به الگوی چرخش تکرار می‌شوند. برای این منظور، ابتدا نمونه‌ای بزرگ‌تر از آنچه برای هر دوره نیاز است از جامعه انتخاب می‌شود تا پایهٔ انتخاب افراد نمونه در هر دوره باشد. سپس، بر اساس الگوی چرخش از میان این نمونه، افرادی به تعداد لازم در هر دوره زمانی انتخاب می‌شوند. نمونه هر دوره با نمونه دوره قبل و دوره بعد خود در برخی از افراد، مشترک (متداخل) و در برخی از افراد متفاوت است. بنابراین، انتظار می‌رود این نمونه، دارای اشتراکاتی (تداخل) با نمونه‌های دوره‌های دورتر نیز باشد.

جدول ۱ مثالی از پیمایش چرخشی ارائه می‌دهد. در این مثال، نمونه پایه، شامل ۱۱ واحد، ولی نمونه لازم در هر دوره، شامل ۵ واحد درنظر گرفته شده است. چرخش افراد نمونه (ورود به نمونه و خروج از آن) برای شش دوره (تکرار) ارائه شده است. در دوره نخست، ۵ واحد نخست از نمونه پایه انتخاب می‌شود. در دوره دوم، دو واحد نخست جای خود را به واحدهای ۶ و ۷ می‌دهند، ولی واحدهای ۳ تا ۵ همچنان در نمونه دوره دوم حضور دارند. در دوره سوم، واحدهای ۵، ۶، و ۷ از نمونه دوره دوم به نمونه دوره سوم منتقل می‌شوند، ولی واحدهای ۸ و ۹ جایگزینِ واحدهای ۳ و ۴ می‌شوند و... . در

پایان دوره چهارم هر یک از ۱۱ واحد نمونه پایه به نحوی در یکی از دوره‌ها انتخاب شده‌اند. در این مثال ساده، نمونه هر دوره، با نمونه دوره قبل و دوره بعد از خود در سه واحد مشترک است (میزان تداخل $۱۰۰\times\frac{۳}{۵}=۶۰$ درصد است). همچنین، نمونه‌های سه دوره متوالی تنها در یک واحد مشترک هستند.

جدول ۱: مثالی از پیمایش چرخشی

واحدهای نمونه پایه	دوره زمانی ۱	۲	۳	۴	۵	۶
۱	✓				✓	✓
۲	✓				✓	✓
۳	✓	✓				✓
۴	✓	✓				✓
۵	✓	✓	✓			
۶		✓	✓			
۷		✓	✓	✓		
۸			✓	✓		
۹			✓	✓	✓	
۱۰				✓	✓	
۱۱				✓	✓	✓

تصمیم درباره اینکه برای اجرای پیمایش مکرر از کدام‌یک از شکل‌های سه‌گانه استفاده شود، به اهداف پیمایش و ویژگی‌های این سه شکل از پیمایش‌های مکرر باز می‌گردد. تعیین فاصلهٔ زمانی مناسب برای دوره‌های پیمایش در هر سه شکل اهمیت دارد. پیمایش پانلی با مشکل دسترسی به افراد یکسان در تمام تکرارها و کاهش همکاری آنها به دلیل مراجعه‌های متعدد یا تغییر نشانی روبه‌روست، در عوض، از مزیت «حذف اثر نمونه» برخوردار است. مشکل دسترسی در پیمایش مستقل وجود ندارد و در پیمایش چرخشی نیز برحسب میزان تداخل، در مقایسه با پیمایش پانلی کمرنگ‌تر است.

۴ـ خطاهای پیمایش

خطاهایی که یافته‌های پیمایش‌ها را تهدید می‌کنند، به دو دسته تقسیم می‌شوند: خطای نمونه‌گیری و خطای غیرنمونه‌گیری. خطای نمونه‌گیری ناشی از تعمیم یافته‌های نمونه

(بخشی از جامعه) به کل جامعه است. در واقع، هرگز نمی‌توان ادعا کرد برآوردی که از نمونه برای پارامتر نامعلوم جامعه به دست می‌آید، درست برابر با آن پارامتر است، بلکه ممکن است بیشتر یا کمتر از آن باشد. حتی این برآوردها پایا نیستند، به طوری که برآورد حاصل از یک نمونه می‌تواند با برآورد حاصل از نمونه دیگر متفاوت باشد. این همان تغییرپذیری نمونه‌ای است که در فصل سوم بحث شد و با خطای معیارِ برآورد اندازه‌گیری می‌شود.

تغییرپذیری نمونه‌ای که موجب بروز خطا در برآورد می‌شود، با روش نمونه‌گیری مناسب و افزایش اندازه نمونه کاهش می‌یابد. نکتهٔ قابل توجه در بحث خطای نمونه‌گیری در مزیت روش‌های احتمالی نسبت به روش‌های نااحتمالی است. خطای نمونه‌گیری در روش‌های احتمالی قابل اندازه‌گیری و کنترل است. حداقل آنکه اگر پژوهشگر نمی‌تواند از خطای نمونه‌گیری خلاصی یابد، از مقدار آن در روش‌های احتمالی با خبر است و خود به نوعی آن را از طریق اندازه نمونه تعیین می‌کند، ولی در روش‌های نااحتمالی هیچ‌گونه اطلاعی از وضعیت خطای نمونه‌گیری در دست نیست.

برخلاف خطای نمونه‌گیری، منبع خطای غیرنمونه‌گیری بسیار متنوع است. ضعف ابزار سنجش (برای مثال پرسشنامه)، خطاهای انسانی یا گِرد کردن ارقام و خطاهای محاسباتی از دلایل عمدهٔ خطای غیرنمونه‌گیری هستند. ممکن است پرسشگر گزینه‌ای را در پرسشنامه به اشتباه علامت بزند، هنگام ورود داده‌ها به رایانه عددی به اشتباه وارد شود یا سؤالی در پرسشنامه مبهم و دوپهلو طراحی شود، به طوری که برداشت افراد نمونه از آن سؤال، با یکدیگر متفاوت باشد. گاه نیز پرسشگر قادر نیست همکاری جدی فرد نمونه را جلب کند و او بی‌توجه و بدون تمرکز به سؤالات پرسشگر پاسخ می‌دهد. این‌ها همه مثال‌هایی از منابع رخ‌دادن خطای غیرنمونه‌گیری هستند.

خطای غیرنمونه‌گیری، تهدیدی جدی برای اعتبار یافته‌های پیمایشی است، بویژه آنکه می‌تواند کاملاً ناخودآگاه و پنهان رخ دهد به طوری که پژوهشگر هرگز از وسعت تأثیرات آن بر داده‌ها و یافته‌های پیمایش آگاه نشود. در شرایطی که خطای نمونه‌گیری تنها ناشی از تغییرات نمونه‌ای است، خطای غیرنمونه‌گیری از منابع متعدد ناشی می‌شود و این چندگانگی می‌تواند منجر به انباشت خطاها بر داده‌ها و یافته‌های پیمایش شود. عملی‌ترین راهِ برخورد با منابع خطای غیرنمونه‌گیری، افزایش نظارت بر تک‌تک مراحل پیمایش است.

همچنین، اجرای آزمایشی پیمایش راهکاری مناسب برای شناسایی منابع بالقوهٔ خطای غیرنمونه‌گیری است.

آموزش دقیق سرپرستان، پرسشگران و نیروهای اجرایی مختلف می‌تواند از یک سو باعث کاهش خطای آنان شود و از سوی دیگر آنان را به افرادی بامهارت تبدیل کند که خود می‌توانند منابع بروز خطای غیرنمونه‌گیری را شناسایی کنند. اجرای دقیق گام هشتم پیمایش، یعنی بازبینی فرایند گردآوری داده‌ها نیز در اطمینان یافتن از صحت فرایند اجرا سودمند است.

۵ـ چالش‌های پیمایش

پیمایش با وجود تمام قابلیت‌ها و طرفدارانی که دارد، در عمل با چالش‌هایی روبه‌روست که هر یک می‌تواند منبع بالقوه‌ای برای رخ‌دادن خطای غیرنمونه‌گیری باشد. در این قسمت به برخی از مهم‌ترین چالش‌ها اشاره می‌شود.

۱ـ۵ـ فرد نمونه (پاسخگو)

برخورد مردم هنگامی که به عنوان فرد نمونه برای شرکت در یک پیمایش دعوت می‌شوند، از چالش‌های اساسی پیمایش است. برخی درک درستی از کارکردهای پیمایش دارند و آن را در برنامه‌ریزی‌ها و تصمیم‌گیری‌های کلان یا پیشبرد اهداف علمی مؤثر می‌دانند، بنابراین مشتاقانه، باحوصله و صادقانه به پرسش‌ها پاسخ می‌دهند. در مقابل، گروهی هم هستند که پیمایش را دخالت در امور شخصی خود می‌دانند، از همین رو یا از همکاری با پرسشگر امتناع می‌ورزند یا در صورت رضایت به همکاری، پاسخ‌های صادقانه و دقیقی به سؤالات نمی‌دهند.

برخی از مردم تمایل به ارائه پاسخ‌هایی دارند که مورد تأیید جامعه یا دلخواه پرسشگر و سازمان متبوع اوست. به این ترتیب پاسخ‌هایی کاملاً متفاوت با دیدگاه‌های واقعی خود ابراز می‌کنند. عده‌ای مبتلا به نوعی خودسانسوری هستند، زیرا گمان می‌کنند پرسشگر در حال تفتیش عقاید آنان است و می‌ترسند از پاسخ‌هایی که می‌دهند، در جایی علیه آنان استفاده شود. حتی درمواردی فرد نمونه به دلیل نداشتن اعتماد به نفس، شرم یا خجالت از پرسشگر، از بیان دیدگاه واقعی خود پرهیز می‌کند. برای مثال ممکن است در پاسخ به سؤال «آیا در یک روز آلوده از خودروی شخصی به صورت تک‌سرنشین استفاده

می‌کنید؟»، فرد نمونه به دروغ بگوید خیر، زیرا احساس می‌کند در ذهن پرسشگر به عنوان شهروندی بی‌تعهد زیر سؤال می‌رود.

پاسخگویانی وجود دارند که پیمایش را به بازی می‌گیرند و سؤالات را بدون توجه و دقت پاسخ می‌دهند. برخی به دلیل بعضی مشکلات و نارضایتی‌ها، به‌عمد پاسخ‌های اغراق‌آمیز به پرسش‌ها می‌دهند. برای مثال، فردی که در حال حاضر بیکار است، ولی مشکل مسکن ندارد، در یک نظرسنجی درباره مسکن و تسهیلات آن تا می‌تواند بد می‌گوید و به این ترتیب خود را خالی می‌کند.

عامل دیگری که می‌تواند بر رفتار پاسخگو تأثیرگذار باشد، زحمتی است که پاسخگو به واسطهٔ مشارکت در پیمایش متحمل می‌شود. او باید برای پاسخگویی به پرسش‌ها فکر کند و زمانی را صرف نماید. پاسخگویی در شرایطی مانند جلوی در منزل، پشت تلفن یا در پیاده‌رو خیابان نیز می‌تواند این فعالیت را برای او خسته‌کننده‌تر سازد. پرسشنامه‌ای با سؤالات فراوان و عمیق، می‌تواند با اشتیاق و مشارکت پاسخگو آغاز شود ولی به مرور او را خسته و مجبور کند با بی‌میلی به پرسش‌ها پاسخ دهد یا از ادامهٔ همکاری منصرف شود. انتخاب زمان و مکان مناسب برای مراجعه و استفاده از مشوق می‌تواند برای افراد نمونه انگیزه‌بخش باشد.

این نکته کلیدی است که پاسخگو باور داشته باشد نقشی مؤثر در پیمایش ایفا می‌کند. او باید احساس امنیت کند و اطمینان داشته باشد پاسخ‌هایش محرمانه باقی می‌ماند و مشخصات فردی او نظیر نام و نشانی در هیچ جا ثبت نمی‌شود. مقدمه‌چینی مناسب در زمان مراجعه به فرد نمونه به نحوی که چگونگی انتخاب او، اهداف پیمایش، متولی و کاربردهای آن را به‌روشنی برای فرد نمونه شرح دهد، الزامی و راهگشاست.

۲ـ۵ـ ابزار سنجش (پرسشنامه)

افراد نمونه از یک سو دارای شخصیت، فرهنگ، سواد و درک متفاوت هستند و از سوی دیگر، همگی باید به پرسشنامه یکسانی پاسخ دهند. این واقعیت بزرگ‌ترین چالشی است که پژوهشگر در ساختنِ ابزار سنجشِ متغیرها با آن روبه‌روست. او باید وسیله‌ای (پرسشنامه‌ای) برای اندازه‌گیری متغیرهایش بسازد که برای افراد گوناگون نمونه قابل استفاده باشد به طوری که همهٔ آنها برداشت و تلقی یکسانی از پرسش‌ها و گزینه‌های آن داشته باشند.

انتخاب واژه‌ها، جمله‌بندی، نوع گزینه‌ها، باز یا بسته بودن سؤالات و حتی ترتیب قرارگرفتن آنها در پرسشنامه می‌تواند بر درک پاسخگو از مفهوم سؤال و چگونگی پاسخگویی او تأثیر بگذارد. پژوهشگر باید به هنگام ساختن ابزار سنجش، به ملاحظات فرهنگی، قومی و دینی جامعه آماری پژوهش توجه کند. پرسشنامه نباید آن‌قدر طولانی یا پیچیده باشد که فرد نمونه از شرکت در پیمایش منصرف شود.

ساختن پرسشنامه درباره موضوعات حساس مانند اعتیاد یا مسائل جنسی، دشوارتر از انواع دیگر پرسشنامه‌هاست. پاسخگو باید نسبت به طرح پرسش‌های حساس کاملاً توجیه و قانع شود تا با آرامش و امنیت به آنها پاسخ دهد. معمولاً، پژوهشگر ترجیح می‌دهد در صورت امکان، سؤالات حساس را با جمله‌بندی غیر مستقیم در پرسشنامه مطرح کند تا پاسخگو راحت‌تر به این پرسش‌ها پاسخ دهد.

اعتبار و روایی پرسشنامه‌هایی (مقیاس‌هایی) که برای اندازه‌گیری ویژگی‌های پنهان (سازه‌ها) ساخته می‌شوند[1]، از معیارهای مهمی هستند که باید پیش از اجرای اصلی پیمایش بررسی شوند. اعتبار معمولاً از طریق داده‌های حاصل از نمونه اجرای آزمایشی بررسی می‌شود، ولی روایی باید در همان گام سوم پیمایش یعنی تهیهٔ ابزار سنجش متغیرها بررسی گردد، هرچند تا پیش از اجرای اصلی هنوز امکان بازبینی و اصلاح پرسشنامه وجود دارد.

پژوهشگر همواره باید متوجه باشد که جبران ضعف‌های پرسشنامه پس از اجرای اصلی یعنی گردآوری داده‌ها امکان‌پذیر نیست. بنابراین در صورت برخورد با ضعفی جدی پس از اجرای اصلی، لازم است یا از کل داده‌های پیمایش یا از داده‌های مربوط به پرسش‌های مشکل‌دار پرسشنامه صرفنظر کند.

۳ـ۵ـ بی‌پاسخی

بی‌پاسخی از اصلی‌ترین چالش‌های پیش روی پیمایش است که با وجود نظارت‌های مختلف بر فرایند اجرا، کم و بیش وجود دارد. بی‌پاسخی به دو نوعِ بی‌پاسخی واحد و بی‌پاسخی جزئی یا قلم اطلاعاتی تقسیم می‌شود. بی‌پاسخی واحد آماری به معنای موفق نشدن در اندازه‌گیری هیچ‌یک از متغیرهای پیمایش از فرد نمونه (پاسخگو) است، در حالی که

۱. سازه (Construct) ویژگی پنهانی است که تنها از طریق نشانه‌های آشکار آن قابل اندازه‌گیری است. برای مثال، اعتماد به نفس ویژگی آشکاری مانند درآمد، قد یا رنگ مورد علاقه نیست تا به‌راحتی اندازه‌گیری شود. سازه‌ای مانند اعتماد به نفس با استفاده از نشانه‌های شناخته‌شده‌ای اندازه‌گیری می‌شود که معمولاً شخص با اعتماد به نفس داراست.

بی‌پاسخی جزئی هنگامی رخ می‌دهد که تنها بخشی از متغیرها بدون اندازه‌گیری باقی می‌مانند. بنابراین، اگر پرسشنامهٔ فرد نمونه کاملاً خالی باز گردد، با بی‌پاسخی واحد و اگر فرد نمونه تنها به تعدادی از سؤالات پرسشنامه پاسخ ندهد، با بی‌پاسخی جزئی روبه‌رو هستیم.

دلایل بروز بی‌پاسخی واحد، به در دسترس نبودن فرد نمونه یا همکاری نکردن او باز می‌گردد. این حالت زمانی رخ می‌دهد که پرسشگر قادر نباشد با فرد نمونه تماس بگیرد، زیرا برای مثال او در زمان مراجعه پرسشگر حضور نداشته یا نشانی‌اش تغییر کرده است. گاه نیز فرد نمونه با پرسشگر همکاری نمی‌کند که این امر می‌تواند به دلایل بسیار متنوع باشد. برخی از معمول‌ترین دلایل عبارتند از: حساسیت موضوع، احساس ناامنی و بی‌اعتمادی فرد نمونه، بی‌اعتقادی به اثربخشی و سودمندی یافته‌های پیمایش، طولانی‌بودن و پیچیدگی پرسشنامه، تجربه‌های ناخوشایند پیشین از مشارکت در پیمایش‌های دیگر و مراجعهٔ زیاد پرسشگران سازمان‌های مختلف به آنان برای شرکت در پیمایش‌های گوناگون.

دلایل همکاری نکردن می‌تواند به بی‌پاسخی جزئی نیز منجر شود. برای مثال، فرد نمونه به سؤالات سادهٔ پرسشنامه پاسخ می‌دهد، ولی سؤالات پیچیده یا حساس را بدون پاسخ رها می‌کند. همچنین ممکن است در میان گزینه‌های یک سؤال بسته، گزینه‌ای را مشاهده نکند که نظر او را به‌درستی منعکس سازد، پس هیچ گزینه‌ای را انتخاب نمی‌کند و سؤال بدون پاسخ باقی می‌ماند.

اگر موارد بی‌پاسخ زیاد باشد، خطای نمونه‌گیری، به شدت افزایش و دقت برآوردها به شدت کاهش می‌یابد؛ زیرا تعداد نمونه کمتر از آن چیزی خواهد شد که برای پیمایش تعیین شده است. از فصل پیش به یاد دارید که پژوهشگر اندازه نمونه را به گونه‌ای تعیین می‌کند که دقت برآوردها در سطح دلخواه او حفظ شوند. حال اگر بی‌پاسخی واحد یا بی‌پاسخی جزئی رخ دهد، افرادِ بی‌پاسخ از نمونه حذف می‌شوند و با اُفت اندازه نمونه روبه‌رو می‌شویم. پس دیگر نمی‌توان اطمینان داشت که دقت برآوردها در همان سطحی باشد که پژوهشگر انتظار دارد، زیرا آن سطح از دقت تنها با داده‌هایی به تعداد اندازه نمونه قابل حصول است.

افزایش خطای نمونه‌گیری در مقایسه با خطای سوگیریِ ناشی از بی‌پاسخی، مسئله به‌نسبت ساده‌ای است. هنگامی که موارد بی‌پاسخ به قشر خاصی از جامعه تعلق داشته باشند که دیدگاه‌های متفاوت با پاسخ‌دهندگان دارند، بی‌پاسخی نه تنها به افزایش خطای نمونه‌گیری، بلکه به برآوردهای اُریب (سوگیرانه) نیز منجر می‌شود. برای مثال در حالتی که

بی‌پاسخی به دلیل پیچیدگی برخی پرسش‌ها باشد، انتظار می‌رود موارد بی‌پاسخ از سوی افراد کم‌سواد نمونه رخ دهد، پس یافته‌های نمونه بیشتر منعکس‌کنندهٔ دیدگاه‌های اعضای باسواد جامعه آماری است، نه افراد کم‌سواد. به عبارت دیگر، برآوردها به بخش باسواد جامعه گرایش بیشتری دارند تا بخش کم‌سواد آن، در حالی که باید اطلاعاتی درباره تمام جامعه آماری به دست دهند.

مقابله با بی‌پاسخی بیش از همه باید پیشگیرانه باشد. پژوهشگر باید سؤالات پرسشنامه را به گونه‌ای طراحی کند که طول آن کوتاه، حساسیت آن کم و درک آن آسان باشد، یا اگر پرسشگر در نخستین مراجعه به فرد نمونه دست نیافت، با پیگیری و مراجعات بعدی به او دسترسی پیدا کند. اگر پژوهشگر اطمینان داشته باشد که موارد بی‌پاسخی به قشر خاصی از جامعه تعلق ندارند، افزایش اندازه نمونه راهکاری برای جلوگیری از افت نمونه است.

برخلاف روش‌های پیشگیرانه، روش تصحیح و روش جانِهی دو راهکار آماری هستند که برای مقابله با بی‌پاسخی پس از اجرا و گردآوری داده‌ها به‌کار می‌روند. روش جانهی، مقداری را به بی‌پاسخی جزئی هر فرد نمونه نسبت می‌دهد، به طوری که این مقدار با توجه به معیاری به مقدار واقعیِ نامعلوم شبیه باشد. مقدار نسبت داده شده به هر فرد نمونه از طریق اطلاعات متغیرهای دیگری به دست می‌آید که از آن فرد نمونه موجود است.

برای مثال، ساده‌ترین شیوه جانهی برای فردی که درآمد خود را در پاسخ به پرسشی از پیمایش بیان نکرده است، انتساب درآمد فرد دیگری از نمونه است که تا حد ممکن شبیه این فرد بی‌پاسخ است. معیارهایی برای شباهت باید برگزیده شوند که با درآمد ارتباطی نزدیک دارند. ممکن است منطقه محل سکونت، تحصیلات، سن یا وضعیت تأهل مفید باشند. همچنین، ممکن است چند فرد مشابه در نمونه وجود داشته باشند که در این صورت می‌توان مقداری را که دارای بیشترین فراوانی است، برگزید یا میانگین مقادیر این افرادِ شبیه را برای جانهی استفاده کرد. روش‌هایی نیز وجود دارند که برای مثال از مدل‌های رگرسیونی برای پیش‌بینی مورد بی‌پاسخ کمک می‌گیرند.

روش تصحیح، به جای جانهی تک‌تک مقادیر بی‌پاسخ، تلاش می‌کند با اصلاح برآورد، به مقداری دقیق‌تر دست یابد. برای مثال، به جای جانهیِ موارد بی‌پاسخِ مربوط به متغیر درآمد، برآورد میانگینِ درآمد که هدف اصلی پیمایش است به شکلی مناسب اصلاح می‌شود به طوری که تاحدی از تأثیر موارد بی‌پاسخ یعنی کاهش دقت یا اُریبی برآورد، جلوگیری شود.

یکی از شیوه‌های تصحیح، وزن‌دهی مجدد است. هر نمونه نمایندهٔ تعداد خاصی از جامعه آماری خود است که در وزن آن نمونه منعکس می‌شود. برای مثال، اگر ۴ نفر از جامعه‌ای ۲۰ عضوی انتخاب شوند، هر فرد نمونه نمایندهٔ ۵ عضوِ این جامعه است. کاهش اندازه نمونه به دلیل بی‌پاسخی باعث می‌شود وزن‌های اولیه که برای یک نمونه کامل معتبر بودند، برای نمونه ناقص معتبر نباشند و برآوردی که به کمک آنها محاسبه می‌شود، با خطا روبه‌رو شود. وزن‌دهی مجدد کمک می‌کند تا افراد نمونه به شکل دقیق‌تری نمایندهٔ اعضای جامعه آماری خود باشند و به این ترتیب از میزان خطای برآوردها کاسته شود.

۴ـ۵ـ تأثیرات پرسشگر

معمولاً، پرسشگران مسئولیت مراجعه به افراد نمونه و گردآوری داده‌ها (اندازه‌گیری متغیرها) را برعهده دارند. در واقع، پیمایش از سوی پرسشگران به افراد نمونه معرفی می‌شود و آنان سؤالات پرسشنامه را از زبان پرسشگران می‌شنوند. بنابراین، چگونگی برخورد و شیوهٔ تعامل آنان با افراد نمونه می‌تواند در جلب همکاری و اعتماد پاسخگویان و دریافت پاسخ‌های دقیق و صادقانه مؤثر باشد.

جنس، قیافه و ظاهر، حرکات بدن، زبان و لهجه، نحوهٔ سخن گفتن و میزان صمیمیت پرسشگران تأثیر مستقیمی بر رفتار و واکنش افراد نمونه دارد. ممکن است افراد نمونه در برخی پیمایش‌ها با پرسشگر هم‌جنس خود راحت‌تر ارتباط برقرار کنند یا اصلاً فرهنگ جامعه اجازهٔ تعامل نزدیک فرد با جنس مخالف خود را ندهد. پرسشگری که خیلی خشک و رسمی با پاسخگو برخورد می‌کند، احتمالاً مانع از ابراز دیدگاه‌های واقعی و درونی او می‌شود، زیرا دریافت چنین دیدگاه‌هایی به برقراری ارتباط دوستانه و صمیمی نیاز دارد.

ضروری است پرسشگر فردی خونگرم، خوش‌برخورد و با بیانی متقاعدکننده باشد، به طوری که افراد نمونه به همکاری با او علاقه‌مند شوند. او باید خود را در نخستین برخورد معرفی و اهداف پیمایش را برای فرد نمونه بیان کند و آمادگی داشته باشد تا به ابهامات یا پرسش‌های او درباره پیمایش پاسخ دهد. پرسشگر باید باحوصله تک‌تک سؤالات پرسشنامه را طرح کند و به پاسخگو فرصت دهد تا در صورت نیاز درباره پاسخ خود فکر کند[۱].

---

۱. مگر در سؤالاتی که پژوهشگر مایل است نخستین و سریع‌ترین پاسخی را شناسایی کند که به ذهن افراد نمونه می‌رسد. مواردی از این دست باید به پرسشگران آموزش داده و در شیوه‌نامه تکمیل پرسشنامه گوشزد شود.

رعایت شیوه‌نامه تکمیل پرسشنامه از سوی پرسشگر ضروری است؛ اینکه نقـاط پـرش[1] پرسشنامه چه هستند؟ آیا پاسخگو مجاز است تمام پرسشنامه را در همان ابتدا مشاهده کنـد یا حتماً باید از زبان پرسشـگر بشـنود و بـه ترتیـب از محتـوای پرسـش‌هـا آگـاهی یابـد؟ درمواردی از پرسشگر خواسته مـی‌شـود رفتـار و حـالات پاسـخگو را دور از چشـم او در مشاهده‌نامه‌ای ثبت کند تا در پاسخگو حساسیت یا کنجکاوی ایجاد نکند.

نگاه پرسشگر یا لحن او در زمان خواندن پرسش‌ها می‌تواند بـه پاسـخ‌هـا جهـت دهـد. برای مثال، خواندن پرسش با لحنی حاکی از تعجب باعث می‌شود فرد نمونه محافظه‌کارانـه پاسخ دهد. همچنین، اگر پرسشگر در حین گفتگو با پاسخگو به بیان دیدگاه شخصی خـود درباره پرسش‌ها بپردازد، حتی با ایـن تصـور کـه مـی‌خواهـد رابطـه خـود را بـا پاسـخگو صمیمی‌تر کند، ناخودآگاه ممکن است دیدگاه پاسخگو را تحت تـأثیر قـرار دهـد. آمـوزش پرسشگران بویژه از طریق برگزاری چند گفتگـوی سـاختگی مـی‌توانـد از چنـین تـأثیرات نامطلوبی تاحدی جلوگیری کند.

۵ـ۵ـ چهارچوب نمونه‌گیری

اغلب، نمونه‌گیری احتمالی بر چهارچوب نمونه‌گیری استوار اسـت. پیشـتر اشـاره شـد کـه چهارچوب مناسب، از یک سو، شامل فهرستی از تمام واحـدهای نمونـه‌گیـری اسـت و از سوی دیگر، هیچ واحد خارج از جامعه آماری در آن جای ندارد. معمولاً چهـارچوب‌هـا از سرشماری به دست می‌آیند و سرشماری‌هـا بـه دلیـل محـدودیت‌هـای هزینـه و زمـان در فاصله‌های زمانی ۵ یا ۱۰ ساله اجرا می‌شـوند. ایـن امـر باعـث مـی‌شـود بـه مـرور زمـان، چهارچوب‌ها قدیمی شوند و مشکلاتی را در نمونه‌گیری به همراه آورند.

استفاده از چهارچوب قدیمی می‌تواند به بی‌پاسخی‌های واحد منجر شود، زیـرا نشـانی افـراد نمونه با گذشت زمان تغییر می‌کند و چهارچوب قدیمی شامل نشانی جدیـد ایـن افـراد نیسـت. نمونه‌گیری خوشه‌ای در شهرها از فهرست بلوک‌ها به عنوان چهارچوب اسـتفاده مـی‌کنـد. ایـن

---

۱. اگر پاسخ به سؤالی با پاسخ به سؤالی پیش از آن مرتبط باشد، سؤال قبلی پرشی در پرسشنامه به شمار می‌آید. بـرای مثال، پرسشگر از فرد نمونه می‌پرسد که «آیا از اینترنت استفاده می‌کنید؟» و اگر او در پاسـخ بگویـد «بلـه»، آنگـاه از او می‌پرسد که «تا چه حد از سرعت اینترنت راضی هستید؟»؛ پاسخ «خیر» باعث می‌شود پرسشگر از پرسیدن ایـن سـؤال صرفنظر کند. بنابراین، سؤال استفاده از اینترنت یک پرش در پرسشنامه است زیرا تعیـین مـی‌کنـد کـه آیـا بایـد سـؤال رضایت از سرعت پرسیده شود یا از آن گذر شود.

بلوک‌ها می‌توانند به مرور تغییر یابند؛ برای مثال بلوکی به دو بلوک تبدیل شود یا کاملاً ویران و به جای آن پارکی ساخته شود. ممکن است در حاشیه‌های شهر بلوک‌های جدیدی احداث شده باشند که هیچ‌یک ازاین موارد در چهارچوب قدیمی بلوک‌ها منعکس نشده است.

نقص چهارچوب چه از ابتدا وجود داشته باشد چه با گذشت زمان و کُهنگی ایجاد شود، به انتخاب واحدهایی در نمونه می‌انجامد که هنگام مراجعه پرسشگران در محل نشانی حضور ندارند. همچنین، اگر دگرگونی و تغییرات جامعه تا آن حد باشد که بخش‌هایی از جامعه ناپدید یا بخش‌های جدیدی به جامعه افزوده شده باشند، چهارچوب نمی‌تواند تصویر کامل و دقیقی از جامعه به دست دهد. این بدان معناست که بخشی از جامعه در نمونه‌گیری نادیده گرفته می‌شود و تمام اعضای آن امکانی برای انتخاب شدن به عنوان نمونه نخواهند داشت. بنابراین، نمونه‌گیری از چهارچوبی با نقص جدی نمی‌تواند به نمونه معرّفی از جامعه بینجامد. برای مثال، اگر چهارچوب، قشر خاصی از جامعه را پوشش ندهد، این نقص می‌تواند باعث سوگیریِ یافته‌های حاصل از نمونه باشد، زیرا این قشر از جامعه امکانی برای انتخاب شدن به عنوان نمونه نخواهند داشت و نمونه نمی‌تواند دیدگاه‌های آنان را منعکس سازد.

۶ـ شیوه‌های تماس

شیوه تماس (ارتباط) با پاسخگو موضوع دیگری است که از بُعد اجرایی در پیمایش‌ها اهمیت می‌یابد، به طوری که می‌تواند بر طراحی پرسشنامه، هزینهٔ اجرا، گزینش روش نمونه‌گیری و اندازه نمونه و حتی دقت برآوردها و استنباط‌های آماری مؤثر باشد. شیوه چهره‌به‌چهره[1] (رو در رو) از قدیمی‌ترین و کارآمدترین شیوه‌هایی است که برای ارتباط با پاسخگو در نظرسنجی‌ها استفاده می‌شده است. در این شیوه فردی به عنوان پرسشگر به پاسخگو مراجعه می‌کند و با پرسیدن تک‌تک سؤالات پرسشنامه از او، پاسخ‌ها را در پرسشنامه ثبت می‌کند؛ یعنی مسئولیت گردآوری داده‌ها را برعهده دارد[2].

حضور پرسشگر از جهات مختلف در اجرای پیمایش‌ها مفید است؛ پاسخگو نظرسنجی را جدی‌تر می‌گیرد، ابهامات پیش‌بینی‌نشده در پرسشنامه با توضیحات پرسشگر برای

1. Face-to-face

۲. این گزینه نیز وجود دارد که پرسشگر به فرد نمونه مراجعه کند و پس از معرفی پیمایش، پرسشنامه را در اختیار فرد نمونه قرار دهد تا او خود آن را تکمیل کند.

پاسخگو روشن می‌شود، امکان ثبت برخی از حالات یا واکنش‌های پاسخگو فراهم است و اگر پرسشنامه طولانی باشد، پرسشگر با ایجاد فضایی صمیمی می‌تواند از خستگی یا بی‌انگیزگی پاسخگو بکاهد.

شیوه چهره‌به‌چهره به دلیل هزینهٔ پرسشگری و رفت‌وآمد به محل استقرار پاسخگو شیوه‌ای هزینه‌بر است. از این رو، شیوه‌های دیگری نیز در اجرای نظرسنجی‌ها به خدمت‌گرفته شده است. با ساخت‌مندی جامعه‌های شهری و روستایی، نظرسنجی پستی نیز پا به عرصه وجود گذاشت. در این شیوه، به جای تماس مستقیم با فرد نمونه، بسته‌ای به نشانی او ارسال می‌شود. این بسته شامل پرسشنامه، راهنمای تکمیل آن و مقدمه‌ای است که اهداف پیمایش، مشخصات متولی و مجری آن و شماره تلفن یا نشانی را برای بررسی صحت پیمایش یا پرسش‌های احتمالی در خود دارد. از فرد نمونه خواسته می‌شود با استفاده از راهنما به تکمیل پرسشنامه اقدام کند و پرسشنامه تکمیل‌شده را تا تاریخ مشخصی به صندوق پست بیندازد. هزینه پست نیز برعهدهٔ مجری پیمایش است.

این شیوه علاوه بر کم‌هزینگی از مزیت مصون ماندن از تأثیرات پرسشگر نیز برخوردار است. همچنین، پاسخگو می‌تواند با فراغت، به تکمیل پرسشنامه در منزل بپردازد. ولی ممکن است حضور اعضای خانواده در کنار پاسخگو در حین تکمیل پرسشنامه بر پاسخ‌های او تأثیر بگذارد. ضمن آنکه معمولاً شیوه پیمایش پستی به پیگیری‌های متعدد نیاز دارد و با وجود این پیگیری‌ها، با بی‌پاسخی‌های واحد زیادی روبه‌روست.

با گسترش شبکه تلفن ثابت و همراه، زمینه استفاده از ارتباط تلفنی نیز برای اجرای پیمایش‌ها فراهم شده است. نمونه‌گیری در پیمایش تلفنی از میان شماره‌های تلفن صورت می‌گیرد. یک دفترچه راهنمای تلفن کامل می‌تواند چهارچوب مناسبی برای انتخاب شماره‌ها باشد. با وجود این، تولید شماره تلفن‌های تصادفی از طریق نرم‌افزار، راهکار مناسب دیگری برای تعیین شماره تلفن افراد نمونه است، ولی می‌تواند به تولید شماره‌هایی نیز منجر شود که در شبکه تلفن وجود ندارند.

پیمایش تلفنی سنتی با کمک پرسشگری (اُپراتوری) اجرا می‌شود که خود شماره‌گیری می‌کند و در صورت برقراری تماس و دسترسی به فرد نمونه، تک‌تک سؤالات پرسشنامه را برای او می‌خواند و پاسخ‌های فرد نمونه را در پرسشنامه ثبت می‌کند. بهره‌گیری از رایانه نیز در اجرای پیمایش تلفنی معمول است. رایانه می‌تواند به‌طور خودکار شماره‌گیری کند و

در صورت برقراری تماس، صدای از پیش ضبط‌شده‌ای تک‌تک سؤالات را برای فرد نمونه بخواند. فرد نمونه نیز با انتخاب شماره گزینه‌ها از طریق شماره‌گیر تلفن، اقدام به پاسخگویی می‌کند. مزیت استفاده از رایانه آن است که پاسخ‌ها (داده‌ها) از همان ابتدا در رایانه ذخیره می‌شوند و دیگر به مرحلهٔ ورود داده‌ها به رایانه نیازی نیست. تلفیقی از همراهی رایانه و پرسشگر نیز در اجرای پیمایش تلفنی امکان‌پذیر است؛ به این صورت که شماره‌گیری بر عهدهٔ رایانه، ولی گفتگوی تلفنی و درج پاسخ‌ها در رایانه، همزمان با گفتگو، بر عهدهٔ پرسشگر باشد.

پیمایش تلفنی نیز مانند پیمایش چهره‌به‌چهره، به معرفی اهداف پیمایش از سوی پرسشگر پیش از طرح پرسش‌ها و در ابتدای تماس نیاز دارد (اگر اجرا کاملاً رایانه‌ای باشد، معرفی پیمایش نیز از طریق صدای ضبط‌شده صورت می‌گیرد). استفاده از پرسشگر می‌تواند به جلب همکاری کمک کند. تماس‌های تلفنی بی‌جواب یا خط‌های اشغال از معمول‌ترین اتفاق‌هایی هستند که در زمان شماره‌گیری رخ می‌دهند. از این رو، تماس مجدد حتی تا ۳ بار اجتناب‌ناپذیر است. در پیمایش تلفنی، لازم است تعداد سؤالات پرسشنامه کم باشد تا پاسخگو حاضر شود در گفتگوی تلفنی شرکت کند. پرهیز از طرح پرسش‌های باز نیز بر سرعت مکالمه و سهولت اجرا می‌افزاید.

همواره در اجرای تلفنی باید از خود پرسید که آیا پوشش تلفن در جامعه آماری تحت مطالعه، امکان دسترسی به تقریباً تمام افراد جامعه را فراهم می‌کند؟ وجود شبکه تلفنی گسترده می‌تواند دسترسی به جامعه آماری وسیعی را ممکن سازد. از این رو، پژوهشگری که به اجبار جامعه آماری پژوهش خود را در اجرای چهره‌به‌چهره تنها به یک شهر محدود می‌سازد تا از پسِ هزینه و مشکلات اجرایی آن برآید، با اجرای تلفنی می‌تواند جامعه آماری را برای مثال به تمام مناطق شهری کشور گسترش دهد، زیرا قادر است از دفتر کار خود با استفاده از چند خط تلفن به تمام این مناطق دسترسی پیدا کند.

استفاده از شبکه جهانی اینترنت نیز برای اجرای پیمایش‌ها مورد توجه قرار گرفته است. در مقایسه با دو شیوه پستی و تلفنی، پیمایش اینترنتی ظرفیت بالقوه و بدون مرزی را در اجرا فراهم آورده است؛ مانند استفاده از امکانات چند رسانه‌ای[1] در پرسشنامه یا دسترسی بدون مرز به افراد تا آنجا که اینترنت اجازه می‌دهد. با وجود این، طراحی پرسشنامه‌های

1. Multi-media

اینترنتی باید به شکلی باشد که پاسخگو به تنهایی بتواند آنها را تکمیل کند. ضروری است حجم و شکل پرسشنامه متناسب با سرعت اتصال به اینترنت باشد تا باز شدن صفحات پرسشنامه به آهستگی صورت نگیرد و پاسخگو را از مشارکت در نظرسنجی منصرف نسازد. هزینه اتصال به اینترنت نیز باید به نحوی برای پاسخگو جبران شود.

موضوع چالش‌برانگیزتر از طراحی پرسشنامه یا هزینهٔ اتصال، یافتن نمونه‌ای معرف برای پیمایش اینترنتی است. دسترسی به اینترنت هنوز بسیار محدودتر از تلفن است. کاربران اینترنت عموماً به افراد تحصیلکرده، نوجوانان یا جوانان محدود می‌شوند. این بدان معناست که نمی‌توان از طریق اینترنت به جامعه‌های عامی مانند افراد با حداقل ۱۲ سال در شهری خاص دسترسی داشت. به عبارت دیگر، یافته‌های پیمایش‌های اینترنتی قابل‌تعمیم به چنین جامعه‌هایی نیستند. حتی این یافته‌ها را به جامعه کاربران اینترنت نیز نمی‌توان تعمیم داد، زیرا هنوز راهی وجود ندارد که بتوان به تمام کاربران اینترنت دسترسی یافت و از این طریق، نمونه‌ای معرف از آنان را انتخاب کرد. شاید در حال حاضر، بهترین استفاده از اینترنت در اجرای پیمایش‌ها در جامعه‌های محدودی مانند یک دانشگاه یا سازمان باشد که اعضای آنها شناخته‌شده و دارای ای‌میل سازمانی هستند.

دو شیوه رایج برای اجرای پیمایش‌های اینترنتی وجود دارد: اجرای وِبی و اجرای ای‌میلی. در اجرای وبی، پاسخگویان با مراجعه به وب‌سایت به پرسشنامه پیمایش دسترسی پیدا می‌کنند. آنان می‌توانند با دعوت یا فراخوان قبلی به سایت مراجعه کنند یا برحسب اتفاق ضمن بازدید از سایت متوجه پیمایش شوند. در پیمایش ای‌میلی، ابتدا فهرستی از ای‌میل‌ها تهیه می‌شود و سپس پرسشنامه از طریق آن برای افراد ارسال می‌گردد، هرچند می‌توان به جای ارسال پرسشنامه، دعوت‌نامه و نشانی سایت پیمایش را از طریق ای‌میل ارسال کرد.

اغلب نمی‌توان هیچ اطمینانی به یافته‌های پیمایش‌هایی داشت که پاسخگویان آنها به صورت خودانتخابی[1] در پیمایش اینترنتی شرکت می‌کنند. در واقع، هیچ کنترل و نظارتی در پیمایش‌های خودانتخابی بر پاسخگویان نیست، به طوری که حتی نمی‌توان اطمینان داشت

---

۱. منظور از خودانتخابی (Self-selected) آن است که پاسخگویان از سوی پژوهشگر و با روش‌های نمونه‌گیری انتخاب نشده‌اند، بلکه خود تصمیم گرفته‌اند در پیمایش شرکت کنند، بدون آنکه هیچ نظارت یا کنترلی از سوی پژوهشگر صورت گرفته باشد.

فردی که برای مثال، خود را زن معرفی کرده است، واقعاً یک پاسخگوی زن باشد. بسیار رایج شده است که تعدادی از وب‌سایت‌های خبری با طرح پرسشی از بازدیدکنندگان سایت خود می‌خواهند با انتخاب یکی از گزینه‌های ایجاد شده، به یک یا چند پرسش پاسخ دهند و به اصطلاح در نظرسنجی شرکت کنند. بدیهی است یافته‌های چنین نظرسنجی‌هایی حتی با تعداد زیادی پاسخگو، برای مثال، چندصد هزار نفر معتبر و قابل‌تعمیم نیستند، زیرا پاسخگویان چنین نظرسنجی‌هایی به بازدیدکنندگان آن سایت محدود می‌شوند که نه تنها نمی‌توانند نمایندهٔ جامعه‌ای عام باشند، بلکه نمایندهٔ کاربران اینترنت یا حتی «تمام» بازدیدکنندگان آن سایت نیز نیستند. از سوی دیگر، معمولاً بازدیدکنندگان این سایت‌ها مشتریان خاص خود را دارند که به سایت و محتوای آن علاقه‌مندند. به عبارت دیگر، انتظار می‌رود پاسخگویان چنین نظرسنجی‌هایی به قشری خاص تعلق داشته باشند که این امر به تنهایی عاملی در جهت سوگیری یافته‌هاست.

شیوه‌های تماس می‌توانند در کاهش هزینه‌های پیمایش‌های پانلی یا چرخشی مؤثر باشند. برای مثال، نخستین تماس با افراد نمونه با شیوهٔ چهره‌به‌چهره صورت می‌گیرد، ولی از آنان خواسته می‌شود تا در صورت تمایل، شیوهٔ تماس را برای مراجعه‌های بعدی خود تعیین کنند. برخی ممکن است شیوهٔ تلفنی، پستی یا اینترنتی را برای تماس‌های بعدی انتخاب کنند. به این ترتیب، در مورد این افراد، هزینهٔ رفت‌وآمد پرسشگر از اجراهای بعدی پیمایش حذف می‌شود. پژوهشگر باید در استفاده از چنین راهکاری برای کاهش هزینه، به تأثیرات شیوه‌های مختلف بر کیفیت داده‌ها توجه کند.

۷ـ زمان‌بندی و بودجه‌بندی

فرایند پیمایش از زمانی که پژوهشگر به اهداف آن فکر می‌کند تا لحظه‌ای که گزارش نهایی آن تدوین و چند نسخه از آن تهیه می‌شود، به طول می‌انجامد. هر گام از فرایند پیمایش که پیشتر در شکل ۱ به آنها اشاره شد، زمان خاص خود را می‌طلبد و باید پیش از آغاز کار مشخص شود تا به این ترتیب زمان کل پیمایش تعیین گردد. قانون کلیدی در زمان‌بندی پیمایش، اختصاص زمانی بیش از آن چیزی است که تصور می‌شود برای هر گام لازم است، زیرا باید در هر حال انتظار مسائل پیش‌بینی نشده را داشت.

معمولاً، جدول زمان‌بندی شامل سه ستون است. ستونی که فعالیت‌ها یعنی جزئیات گام‌های پیمایش در آن فهرست شده‌اند، ستونی که مدت زمان لازم برای هر فعالیت را بر

حسب ساعت نشان می‌دهد و ستونی که مدت زمان هر فعالیت را بر حسب تعداد روزهای کاری ارائه می‌دهد. دلیل اینکه زمان‌بندی در دو قالب ساعت و روز کاری تعیین می‌شود، به فرصتی باز می‌گردد که اعضای گروه پژوهشی به پیمایش اختصاص داده‌اند. بر اساس جدول ۲، برای نخستین فعالیت، به ۱۰ ساعت کار نیاز است. اگر این فعالیت بر عهدهٔ ۱ نفر باشد و او بتواند ۵ ساعت از روز خود را به پیمایش اختصاص دهد، انجام این فعالیت به ۲ روز کاری نیاز دارد. در مقابل، فعالیت ردیف ۲ این جدول، به ۸۰ ساعت کار نیاز دارد ولی از آنجا که ۲ نفر همزمان به این فعالیت می‌پردازند، ظرف ۶ روز کاری پایان می‌یابد.

جدول ۲: مثالی از جدول زمان‌بندی

ردیف	فعالیت	ساعت	روز کاری
۱	تعیین اهداف پیمایش (متغیرها، فرضیه‌ها، جدول‌ها و...)		
۲	مطالعات نظری، جستجو و مشاوره برای طراحی پرسشنامه		
۳	طراحی پرسشنامه		
۴	ارزیابی پرسشنامه از سوی ۳ کارشناس و اصلاح اولیهٔ آن		
۵	تعیین روش نمونه‌گیری و چهارچوب مناسب آن		
۶	اجرای آزمایشی با نمونه‌ای ۵۰ تایی توسط ۵ پرسشگر		
۷	ورود داده‌های اجرای آزمایشی و تحلیل یافته‌های آن		
۸	اصلاح پرسشنامه و تهیهٔ نسخهٔ نهایی آن		
۹	تکثیر ۵۰۰ پرسشنامه برای اجرای اصلی		
۱۰	استخدام ۵۰ پرسشگر، ۲ ناظر و ۱ سرپرست برای اجرای اصلی		
۱۱	انتخاب نمونه‌ها از چهارچوب و تهیه نقشه‌ها و نشانی‌ها		
۱۲	تهیه شیوه‌نامه‌های اجرایی و آموزش نیروهای اجرایی		
۱۳	مراجعه به افراد نمونه و گردآوری داده‌ها		
۱۴	بازبینی پرسشنامه‌های تکمیل‌شده و فرایند اجرا		
۱۵	کدگذاری و ورود داده‌ها		
۱۶	کنترل صحت کدگذاری و ورود داده‌ها		
۱۷	پردازش آماری داده‌ها		
۱۸	گزارش‌نویسی (فصل‌بندی، نگارش، صفحه‌بندی، رسم نمودارها و جدول‌ها و...)		
۱۹	بازبینی و ارزیابی گزارش		
۲۰	تهیه گزارش نهایی (تایپ، ویراستاری، صحافی) و تکثیر چند نسخه از آن		
	جمع		

ممکن است ساعات انجام برخی از فعالیت‌ها مانند ردیف ۴، ۹، ۱۰ و ۱۳ قابل بیان نباشد، بنابراین به ذکر تعداد روزهای کاری لازم اکتفا می‌شود. برای مثال، پژوهشگر نمی‌داند چه زمانی را کارشناسان صرف ارزیابی پرسشنامه می‌کنند (فعالیت ردیف ۴ در جدول ۲)، ولی حداکثر ۷ روز کاری را به آنان فرصت می‌دهد تا نتایج ارزیابی خود را به او اعلام کنند. پس زمان‌بندی این فعالیت تنها به روزکاری محدود شده است. سطر آخر جدول زمان‌بندی شامل جمع زمان اختصاص‌یافته به کل فعالیت‌هاست که معمولاً بر حسب ماه بیان می‌شود. توجه کنید که برخی از فعالیت‌ها می‌توانند به موازات هم صورت بگیرند. برای مثال، پژوهشگر می‌تواند در حین پردازش، به نگارش بخش‌هایی از گزارش بر اساس پردازش‌های صورت گرفته نیز اقدام کند، پس ممکن است در روز دوم پردازش، نخستین روز گزارش‌نویسی نیز آغاز شده باشد.

بودجه‌بندی پیمایش به دو بخش تقسیم می‌شود؛ نخست بودجهٔ اختصاص‌یافته به نیروی انسانی یعنی همان دستمزد که تابعِ نوع فعالیت فرد، مهارت و دانش او و زمانی است که در پیمایش صرف می‌کند. دوم، بودجه مربوط به هزینه‌های غیر نیروی انسانی مانند امکانات، تجهیزات یا هزینهٔ لوازم‌التحریر. برای مثال، فعالیت ردیف ۲ از جدول ۲ می‌تواند هم شامل هزینهٔ نیروی انسانی و هم شامل هزینه غیر نیروی انسانی باشد، زیرا از یک سو لازم است یک یا دو فرد تحصیلکرده زمانی را صرف مطالعه و بررسی منابع علمی مختلف کنند و از سوی دیگر، ممکن است نیاز باشد آنان به دانشگاهی مراجعه یا کتابی را خریداری کنند. بنابراین، هم باید دستمزد زمان صرف‌شده برای مطالعه و رفت‌وآمد آنان در بودجه‌بندی لحاظ شود و هم هزینه‌های سفرهای درون‌شهری برای مراجعه به کتابخانه‌ها و دانشگاه‌ها یا خرید کتاب و مانند آنها. همچنین پردازش آماری داده‌ها در ردیف ۱۷، هم به یک کارشناس آمار و هم به یک دستگاه رایانه، نرم‌افزار آماری، چاپگر، کاغذ و... نیاز دارد.

جدول بودجه‌بندی پیمایش مانند جدول ۳ در دو بخش تنظیم می‌شود که به هزینه‌های نیروی انسانی و هزینه‌های غیر نیروی انسانی اختصاص دارند. در بخش نیروی انسانی، دستمزد هر فرد در یک ساعت و ساعات کل فعالیت او در پیمایش تعیین می‌شود تا از حاصل‌ضرب آنها، کل دستمزد آن فرد تعیین گردد. برای مثال اگر برای فردی به ازای هر ساعت کار، ۱۰۰۰۰ تومان درنظر گرفته شود و او در کل پیمایش ۳۰ ساعت کار کند، کل دستمزد او برابر با ۳۰۰۰۰۰ تومان است. در جدول ۳ نیز برخی از قسمت‌ها مانند ردیف‌های ۴، ۵ و ۸ کامل

نیستند زیرا یا اطلاعی از جزئیات آنها در دست نیست یا نیازی به ارائه جزییات وجود نـدارد. همچنین، هزینه‌های غیر نیروی انسانی تنها به شکلی کلی بیان می‌شوند.

جدول ۳: مثالی از جدول بودجه‌بندی پیمایش

نوع هزینه	ردیف	فعالیت	دستمزد هر ساعت	ساعات کار	نفر	کل
نیروی انسانی	۱	تعیین اهداف پیمایش				
	۲	امور مطالعاتی				
	۳	طراحی پرسشنامه و اصلاح آن				
	۴	کارشناسان و مشاوران				
	۵	اجرای آزمایشی با ۵۰ نمونه				
	۶	کارشناسی‌های آماری (نمونه‌گیری و پردازش)				
	۷	مقدمات اجرای اصلی (استخدام پرسشگران، آموزش، تهیه شیوه‌نامه‌ها، نقشه‌ها و...)				
	۸	اجرای اصلی با نمونهٔ ۵۰۰ تایی				
	۹	بازبینی، نظارت، کدگذاری، ورود داده‌ها				
	۱۰	گزارش‌نویسی، بازبینی و نهایی کردن گزارش				
جمع						
غیر نیروی انسانی	۱۱	هزینه‌های رفت‌وآمد				
	۱۲	لوازم‌التحریر				
	۱۳	تایپ، ویراستاری، تکثیر و صحافی				
	۱۴	رایانه و چاپگر				
	۱۵	اجاره دفتر				
جمع						
کل هزینه‌ها						

اصطلاحات فصل هشتم

Adjustment Method	روش تصحیح	Pilot	پایلوت، اجرای آزمایشی
Closed Question	پرسش بسته	Post Survey	پیمایش پستی
Cross-sectional Survey	پیمایش مقطعی	Qestionnair Reliability	اعتبار پرسشنامه
Imputation	جانهی	Questionnaire Validity	روایی پرسشنامه
Longitudinal Survey	پیمایش طولی	Rotation Survey	پیمایش چرخشی
Nonresponse	بی‌پاسخی	Sampling Error	خطای نمونه‌گیری
Nonsampling Error	خطای غیرنمونه‌گیری	Survey	پیمایش
Open-ended Question	پرسش باز	Telephone Survey	پیمایش تلفنی
Panel Survey	پیمایش پانلی	Unit Nonresponse	بی‌پاسخی واحد (پاسخگو)
Partial/Item Nonresponse	بی‌پاسخی جزئی یا قلم		

تمرین‌های فصل هشتم

۱. کدام عبارت درباره خطای نمونه‌گیری و غیرنمونه‌گیری اشتباه است؟

آ) خطای نمونه‌گیری، اختلاف برآورد حاصل از نمونه با پارامتر نامعلوم جامعه است.

ب) درک نشدن سؤالات پرسشنامه از سوی پاسخگو به خطای نمونه‌گیری می‌انجامد.

پ) وقتی اغلب پاسخگویان مایل هستند خود را موجه و مطابق عرف جامعه نشان دهند، خطای غیرنمونه‌گیری رخ می‌دهد.

ت) خطای نمونه‌گیری را می‌توان به صفر رساند.

ث) خطای نمونه‌گیری ناشی از انتخاب بخشی از جامعه به جای کل آن است.

ج) اگر در ورود داده‌ها به رایانه، عدد ۱۱۰ به صورت ۱۱ وارد شود، خطای غیرنمونه‌گیری رخ داده است.

چ) پاسخ ندادن پاسخگویی به یک پرسش حساس به خطای نمونه‌گیری منجر می‌شود.

ح) اگر در نمونه‌گیری از جوانان ۲۰ تا ۳۰ ساله، اغلب پاسخگویان را جوانان ۲۰ تا ۲۴ ساله دانشجو تشکیل دهند، خطای نمونه‌گیری وجود دارد.

خ) با افزایش اندازه نمونه، خطای نمونه‌گیری کاهش می‌یابد.

د) سرشماری‌ها کمتر از نمونه‌گیری‌ها در معرض خطای غیرنمونه‌گیری هستند.

فصل نهم

۹ ملاحظات آماری در گزارش یافته‌ها

پژوهشگران می‌کوشند داده‌ها را با دقت گردآوری و مناسب‌ترین روش‌های آماری را برای پردازش آنها انتخاب کنند اما اگر یافته‌های حاصل، به شکلی مناسب ارائه نشوند، تمام این تلاش‌ها ناکام خواهد ماند. بنابراین، آنان علاوه بر کسب مهارت در آمار، به مهارت در گزارش کارآمد یافته‌های آماری نیز نیاز دارند. این فصل رهنمودهایی را برای سازماندهی و ارائه یک گزارش آماری منسجم، روان و مؤثر از یافته‌ها به دست می‌دهد.

۱ـ مقدمه

گزارش یافته‌های آماری می‌تواند مرحله‌ای دشوار در پژوهش باشد زیرا شامل اعداد و ارقام، جدول‌ها، نمودارها و در برخی موارد، مدل‌های آماری می‌شود که نیاز به توصیف و تفسیر دارند. نویسنده باید بداند اعداد و ارقام را چگونه در داخل متن به کار برد، چه هنگام از جدول یا نمودار برای بیان مقصود خود بهره گیرد یا نتایج یک آزمون آماری را به چه شکل تفسیر کند. در این مقدمه، درباره چند اصل کلیدی صحبت خواهد شد که توجه به آنها تهیه گزارش آماری را آسان‌تر خواهد ساخت.

هر چند نگارش گزارش یافته‌های آماری می‌تواند پس از پردازشِ همه یا بخشی از داده‌ها آغاز شود، نخستین اصل کلیدی بیان می‌کند که طرح کلی گزارش، در زمان تعیین پرسش‌های «جزئی» پژوهش شکل می‌گیرد زیرا گزارش آماری شامل اطلاعاتی است که به این پرسش‌ها پاسخ می‌دهند. پاسخگویی به پرسش‌هایی که رویکرد توصیفی و اکتشافی دارند، با ارائه جدول‌های توزیع فراوانی، نمودارها و مشخصه‌های توزیع صورت می‌گیرد. در مقابل، پاسخگویی به پرسش‌های با رویکرد تبیینی که گاه نیز در قالب فرضیه بیان می‌شوند، با برآوردها و آزمون‌های آماری امکان‌پذیر است. بنابراین، بیان روشن و کامل پرسش‌های پژوهش می‌تواند هم به ترسیم فضای کلی گزارش آماری و هم به پیش‌بینی روش‌های آماری مناسب برای پردازش داده‌ها کمک کند.

> پژوهشگری که از همان ابتدا به طور هدفمند داده‌ها را برای پاسخ به پرسش‌های مشخصی گردآوری کرده است به راحتی می‌داند که پاسخ پرسش‌هایش را با چه روش‌های آماری در داده‌ها جستجو کند و در مرحله گزارش‌نویسی نیز با دشواری کمتری روبه‌رو است.

پژوهشگری که از پیش، پرسش‌های پژوهش را با جزئیات تعیین کرده است، در زمان پردازش داده‌ها می‌داند که دقیقاً به دنبال چیست و کدام روش آماری را باید برای رسیدن به هدف به کار گیرد. دغدغه او در گزارش‌نویسی، تنها به چگونگی ارائه بهتر یافته‌ها محدود می‌شود. در مقابل، پژوهشگری که در زمان پردازش داده‌ها یا گزارش‌نویسی سردرگم است و دایم از خود می‌پرسد:

آیا لازم است رابطه این متغیر را با آن متغیر بررسی کنم؟
آیا مقایسه میانگین این متغیر در بین گروه‌های متغیر دیگر ضرورت دارد؟
آیا باید جدول توزیع فراوانی این متغیر را نیز ارائه دهم؟

یا پرسش‌های پژوهش را روشن و کامل مشخص نکرده است یا با روش‌های آماری مناسب برای پاسخ به این پرسش‌ها آشنا نیست.

اصل کلیدی دوم، فضایی است که به ارائه یافته‌های آماری اختصاص یافته است. ممکن است یافته‌ها، چند برگ از یک مقاله، فصلی از یک گزارش یا یک گزارش کامل، شامل چند فصل را به خود اختصاص دهند. در شرایطی که فضای بسیار محدودی مانند چند برگ از مقاله در اختیار پژوهشگر است، او اجازه قلم‌فرسایی ندارد و تنها باید به بیان مهم‌ترین یافته‌ها اکتفا و از پرداختن به جزئیات خودداری کند در حالی که در یک گزارش آماری، او قادر است درباره روش‌های خود و حتی چگونگی پردازش داده‌ها نیز بنویسد.

> بهتر است پیش از آنکه خود را نویسنده گزارش بدانید، خواننده آن در نظر بگیرید. به این ترتیب از جایگاه فردی به گزارش نگاه خواهید کرد که ممکن است هیچ‌یک از پیش‌زمینه‌های ذهنی و اطلاعات قبلی شما را نداشته باشد.

اصل کلیدی سوم، به مخاطبان گزارش باز می‌گردد. همواره در نظر داشته باشید که چه کسانی مخاطب گزارش شما هستند و به چه منظوری از یافته‌های آن استفاده می‌کنند. هنگامی که با مخاطبان آشنا با آمار مانند خوانندگان برخی از مجله‌های علمی و پژوهشی روبه‌رو هستید، از بیان برخی جزئیات فنی بی‌نیاز خواهید بود و اطمینان خواهید داشت که آنان قادر به درک یافته‌های آماری هستند. اما گاه، گزارش‌های آماری برای مدیران و به منظور تصمیم‌گیری‌های مدیریتی تهیه می‌شوند، بنابراین، باید تا حد ممکن ساده و به دور از اصطلاحات تخصصی آماری باشند به طوری که یافته‌ها به سادگی در اختیار کاربران قرار گیرند. همواره در ذهن داشته باشید که اگرچه منطقی است پژوهشگر بخشی از قضاوت و تفسیر درباره یافته‌ها را بر عهده خواننده آشنا با آمار بگذارد، هرگز نباید خواننده ناآشنا را در شرایطی قرار دهد که خود به تفسیر و قضاوت درباره داده‌ها بپردازد زیرا در چنین حالتی نمی‌توان اطمینان داشت که او دست به برداشت نادرستی از یافته‌ها نمی‌زند.

۲ـ ساختار گزارش آماری

گزارش آماری ـ کوتاه باشد یا بلند ـ شامل چهار بخش اصلی است[۱]: پرسش‌ها و فرضیه‌های آماری، منابع داده‌ها، معرفی روش‌های تحلیل و سرانجام، یافته‌ها و نتیجه‌گیری از آنها. در ادامه، هر یک از این بخش‌ها به تفصیل بررسی خواهد شد.

---

۱. در گزارش‌های «پژوهشی» برخلاف گزارش‌های مستقل آماری، تنها یک یا دو فصل از گزارش به مباحث آماری اختصاص می‌یابد، در این صورت، این ساختار برای همین دو فصل در نظر گرفته می‌شود.

۱ـ۲ـ پرسش‌ها و فرضیه‌ها

پرسش‌ها و فرضیه‌های آماری شکلی متفاوت با پرسش‌ها و فرضیه‌های پژوهش دارند. در واقع، پرسش‌ها و فرضیه‌های «آماری» در قالب شاخص‌ها و تعبیرهای آماری بیان می‌شوند که به طور مستقیم با داده‌ها و روش‌های توصیفی و استنباطی تحلیل آنها مرتبط هستند. برای مثال، اگر پرسش پژوهش بیان می‌کند که «میزان درآمد شهروندان تهرانی چقدر است؟»، پرسش آماری معادل آن بیان می‌کند که «متوسط درآمد شهروندان تهرانی چقدر است؟». به کارگیری اصطلاح «متوسط» تعبیری آماری از «میزان» است که از یک سو آن را دقیق‌تر می‌سازد و از سوی دیگر، روش آماری پاسخگویی به آن (محاسبه برآورد نقطه‌ای و فاصله‌ای مشخصه میانگین توزیع) را تعیین می‌کند. همچنین، این پرسش نشان از آن دارد که باید متغیری شامل درآمد شهروندان در مجموعه داده‌ها وجود داشته باشد که محاسبه میانگین آن مبنای پاسخگویی به این پرسش است (در واقع، بیان درست و شفاف پرسش‌ها و فرضیه‌های تحقیق به پژوهشگر کمک می‌کند که پرسشنامه‌ای متناسب با آنها طراحی کند و چیزی را از قلم نیندازد).

در بحث آزمون فرضیه نیز بین فرضیه‌های پژوهش و فرضیه‌های آماری تفاوت وجود دارد. نخست آنکه فرضیه‌های آماری مورد آزمون، فرضیه‌های صفر هستند در حالی که فرضیه‌های پژوهش، معادل فرضیه‌های جانشین هستند. دوم آنکه، فرضیه‌های آماری نیز مانند پرسش‌های آماری از شاخص‌ها و تعبیرهای آماری تشکیل می‌شوند. برای مثال، اگر فرضیه پژوهش بیان می‌کند که «رضایت شغلی کارمندان رسمی بیش از کارمندان غیررسمی است»، معادل آماری آن بیان می‌کند که «میانگین نمره رضایت شغلی کارمندان رسمی بیش از کارمندان غیررسمی است». اکنون باید این فرضیه را به شکل فرضیه صفر بیان کرد یعنی «میانگین نمره رضایت شغلی کارمندان غیررسمی حداقل به اندازه کارمندان رسمی است». این فرضیه نیز شامل تعبیرهای آماری است و به راحتی می‌توان تشخیص داد که مقایسه میانگین دو گروه مستقل مطرح است، پس باید برای آزمون این فرضیه صفر از آزمون t یا معادل ناپارامتری آن استفاده شود.

از این دو مثال می‌توان دریافت که بیان کامل و جزئی تک‌تک پرسش‌ها و فرضیه‌ها به عنوان نقطه آغاز گزارش آماری می‌تواند چشم‌اندازی از بخش یافته‌های آماری در اختیار پژوهشگر قرار دهد و او را به روش‌های آماری مناسب برای تحلیل آنها رهنمون سازد.

۲ـ۲ـ منابع داده‌ها

آمارها می‌توانند به شکلی ارائه شوند که به نتیجه مطلوب برای نویسنده بینجامند. به این جمله توجه کنید: «۹۰ درصد از کالای x رضایت دارند». رقم ۹۰ می‌تواند خواننده را نسبت به این کالا کاملاً خوشبین کند. با وجود این، پیش از هر نوع قضاوتی درباره رقم ۹۰، باید درباره اعتبار داده‌ها و رقم حاصل از آنها اطمینان یافت. برای این منظور باید چند پرسش را مطرح کرد: ۹۰ درصد در میان چه کسانی محاسبه شده است، این افراد چگونه انتخاب شده‌اند، رضایت چگونه اندازه‌گیری شده و چه کسی آن را اندازه‌گیری کرده است؟ اگر به همه پرسش‌ها پاسخ اطمینان‌بخشی داده نشود، نمی‌توان به این یافته اطمینان کرد. این بخش از گزارش برای خواننده مشخص می‌کند که یافته‌های آماری از چه داده‌هایی و با چه کیفیتی به دست آمده‌اند[1].

> یک نمونه معرف دارای سه ویژگی مطلوب است: اندازه نمونه کافی، روش نمونه‌گیری مناسب و ترکیبی شبیه به جامعه آماری تحت مطالعه. باید به خواننده اطمینان داده شود که یافته‌های آماری برآمده از یک نمونه معرف هستند.

به طور معمول، داده‌ها با یکی از روش‌های سرشماری، ثبتی و پیمایش صورت گرفته در گذشته یا پیمایش صورت گرفته از سوی پژوهشگر گردآوری می‌شوند[2]. در اغلب موارد، پژوهشگر در گردآوری داده‌های سرشماری یا ثبتی نقشی ندارد؛ از این رو تنها کافی است به متولی گردآوری آنها و اعتبار این متولی اشاره کند. بنابراین، اگر از داده‌های سرشماری در پژوهش استفاده می‌کند، کافی است بیان کند که داده‌ها برای مثال، به سرشماری عمومی نفوس و مسکن ۱۳۸۵ مرکز آمار ایران تعلق دارند و مشخصات گزارش تفصیلی مربوط به آن را شامل سال انتشار، شماره جدول، صفحه و ... ارائه دهد.

> به هیچ یافته‌ای اطمینان نکنید و به همه چیز شک داشته باشید.

هنگامی که داده‌های پیمایش اجرا شده در گذشته یا پیمایشی که خود پژوهشگر در زمان جاری اجرا کرده است، مورد استفاده قرار می‌گیرد، ضروری است جامعه آماری، روش نمونه‌گیری، واحد یا واحدهای آماری، چگونگی تعیین اندازه نمونه، چهارچوب نمونه‌گیری،

---

۱. نگهداری از داده‌ها بویژه هنگامی که تحقیق از حمایت مالی سازمان یا مؤسسه‌ای برخوردار است، موضوعی است که در اینجا به آن پرداخته نمی‌شود هرچند در یک پروژه تحقیقاتی باید به آن اشاره شود. برخی با ذخیره مجموعه داده‌ها بر روی CD یا هر وسیله قابل حمل دیگری، آن را ضمیمه گزارش می‌کنند. برخی نیز از طریق اینترنت آن را قابل دسترس می‌سازند. به این ترتیب، امکان تحلیل‌های آماری بیشتر بر داده‌ها برای سایرین نیز فراهم می‌شود.

۲. البته روش‌های دیگری نظیر آزمایش یا تحلیل محتوا نیز می‌توانند به ایجاد یک مجموعه داده برای تحلیل آماری بینجامند.

شیوهٔ تماس با افراد نمونه و وسیله گردآوری داده‌ها به طور کامل شرح داده شود. پژوهشگر دلیل انتخاب روش نمونه‌گیری، منبعِ چهارچوب یا چهارچوب‌های مورد استفادهٔ خود و زمان اجرای پیمایش را بیان می‌کند. همچنین، متغیر یا متغیرهایی را که مبنای محاسبه اندازه نمونه بوده‌اند، همراه با خطای قابل تحمل، دقت مورد انتظار و ملاحظات هزینه‌ای در تعیین اندازه نمونه پیمایش، ارائه می‌دهد و سپس اندازه نمونه و شیوهٔ محاسبه آن را می‌آورد. اگر از داده‌های نمونه‌گیری مقدماتی برای برآورد واریانس در تعیین اندازه نمونه استفاده شده است، به این داده‌ها و فرایند گردآوری آنها نیز اشاره می‌کند.

ضروری است شیوهٔ انتخاب افراد نمونه (قرعه‌کشی یا سیستماتیک)، برابر یا نابرابر بودن احتمال انتخاب اعضای جامعه و در صورت انتخاب با احتمال نابرابر، چگونگی نابرابری احتمال‌ها شرح داده شود. اگر از نمونه‌گیری چندمرحله‌ای استفاده می‌شود، واحد نمونه‌گیری و اندازه نمونه هر مرحله و چهارچوب مورد استفاده برای آن نیز معرفی می‌شود. همچنین، لازم است هنگام استفاده از نمونه‌گیری طبقه‌ای، تعریف طبقات، شیوهٔ تخصیص و اندازه نمونه هر طبقه و هنگام استفاده از نمونه‌گیری خوشه‌ای، تعریف خوشه‌ها و اندازه آنها ارائه شود.

مثالی از شرح روش و فرایند نمونه‌گیری

روش: پیمایشی

جامعه آماری: روستاییان استان ... با حداقل ۱۸ سال

ابزار سنجش: پرسشنامه

شیوهٔ تماس: گفتگوی رو در رو

روش نمونه‌گیری: خوشه‌ای سه مرحله‌ای به طوری که مرحلهٔ نخست، شامل انتخاب تعدادی روستا از کل روستاهای استان، مرحلهٔ دوم، شامل انتخاب تعدادی خانوار از کل خانوارهای هر روستا و مرحلهٔ سوم، شامل انتخاب ۱ فرد از میان اعضای هر خانوار است.

چهار چوب مرحلهٔ نخست و اندازه نمونه: فهرست ۷۲۰ روستای استان همراه با جمعیت آنها از آخرین سرشماری نفوس‌ومسکن کشور که از سوی مرکز آمار ایران صورت گرفته است. هر روستا، یک واحد مرحلهٔ نخست (PSU) نمونه‌گیری به حساب می‌آید. به دلیل محدودیت‌های مالی، تعداد ۳۰ روستا از میان ۷۲۰ روستای استان با احتمال متناسب با جمعیت، انتخاب می‌شود.

چهارچوب مرحلهٔ دوم و اندازه نمونه: به دلیل آنکه از زمان سرشماری ۸ سال می‌گذرد و روستاها در این فاصله دستخوش تغییراتی شده‌اند، ۳۰ روستای انتخاب‌شده از مرحلهٔ قبل فهرست‌برداری می‌شوند تا تعداد دقیق خانوارهای هر یک در زمان حاضر مشخص

شود. کم‌خانوارترین روستای فهرست‌برداری‌شده شامل ۷ خانوار و پرخانوارترین روستا، شامل ۱۱۰ خانوار بوده است و درمجموع ۱۸۴۰ خانوار شناسایی شده‌اند. از آنجا که روستاها شامل تعداد خانوارهای متفاوتی هستند یعنی خوشه‌های هم‌اندازه‌ای را تشکیل نمی‌دهند و این امر محاسبه برآوردها را دشوار می‌کند، این ۱۸۴۰ خانوار در قالب ۱۸۴ خوشهٔ ۱۰ خانواری سازماندهی شدند تا نمونه‌گیری مرحلهٔ دوم از میان خوشه‌های هم‌اندازه صورت بگیرد. فهرست این ۱۸۴ خوشهٔ ساختگی، چهارچوب نمونه‌گیری مرحلهٔ دوم را تشکیل می‌دهد و هر خانوار نیز یک واحد مرحلهٔ دوم (SSU) نمونه‌گیری قلمداد می‌شود. از هر خوشهٔ ۱۰ تایی، ۴ خانوار به تصادف به عنوان نمونه انتخاب شدند، پس در مجموع به ۷۳۶ خانوار مراجعه شد.

چهارچوب مرحلهٔ سوم و اندازهٔ نمونه: پرسشگر هنگام مراجعه به هر یک از ۷۳۶ خانوار مرحلهٔ قبل، فهرستی از اعضا را با حداقل ۱۸ سال تهیه و یکی را به تصادف انتخاب می‌کند تا پرسشنامه را برای او تکمیل کند. بنابراین، درمجموع ۷۳۶ پرسشنامه تکمیل شد. اگر فرد انتخاب‌شده در آن زمان حضور نداشت، پرسشگر با پرسش از اعضای حاضر در خانوار از ساعت حضور او اطلاع می‌یافت تا با مراجعه مجدد در آن ساعت، به تکمیل پرسشنامه از آن فرد اقدام کند.

به منظور اطمینان از کیفیت داده‌ها، بیان تدبیرهایی که برای حفظ کیفیت آنها اندیشیده شده است نیز از اهمیت برخوردار است. برای مثال، چگونگی آموزش پرسشگران، سازوکار نظارت بر پرسشگری، درصد پرسشنامه‌های بازبینی‌شده، نظارت بر کدگذاری و ورود داده‌ها به رایانه، از جمله مواردی هستند که بهتر است شرح داده شوند. نسبت بی‌پاسخی‌های واحد و سازوکار به کار رفته درباره آنها و چگونگی برخورد با بی‌پاسخی‌های جزئی نیز با کیفیت داده‌ها ارتباط مستقیم دارند.

ضروری است توزیع متغیرهای پایه، بویژه متغیرهای جمعیتیِ افراد نمونه با استفاده از روش‌های آمار توصیفی ارائه شوند تا خواننده دریابد که ترکیب نمونه تا چه حد به ترکیب جامعه آماری خود شبیه است. برای مثال، اگر ۵۰ درصد جامعه آماری را مردان و ۵۰ درصد را زنان تشکیل می‌دهند، انتظار می‌رود توزیع جنسی افراد نمونه «تصادفی» نیز تقریباً چنین باشد (برای مثال ۴۸ درصد مرد و ۵۲ درصد زن). معمول آن است که توزیع افراد نمونه بر حسب متغیرهایی مانند جنس، سن، تحصیلات، منطقهٔ مسکونی، وضعیت فعالیت و وضعیت تأهل ارائه شود. توزیع هر متغیر از داده‌های حاصل از پرسشِ متناظر با آن متغیر در پرسشنامه به دست می‌آید، از این رو، پژوهشگر باید چنین پرسش‌هایی را در پرسشنامه خود گنجانده

باشد. گاه مفید است برخی از مشخصه‌های توزیع نیز مانند نما، میانگین یا چارک‌ها برای بعضی از متغیرها مثل سن یا تعداد سال‌های تحصیل، در کنار توزیع آنها ارائه شود.

۳ـ۲ـ روش‌های تحلیل داده‌ها

روش‌های تحلیل هم شامل روش‌های توصیفی و استنباطی تحلیل داده‌ها و هم شامل معرفی متغیرها و محاسباتی هستند که برای مثال در ایجاد متغیری بر اساس مجموعه‌ای از متغیرهای دیگر به کار می‌روند[۱]. هنگام صحبت از روش‌های تحلیل، ذکر دلیل انتخاب آنها با توجه به نوع و توزیع متغیرها از اهمیت برخوردار است. با وجود این، پژوهشگر باید بداند که هدف از این بخش آموزش آماری خواننده نیست؛ بنابراین هیچ ضرورتی ندارد که برای مثال، بیان کند فرایند اجرای آزمون t چیست یا این آزمون به چه منظوری به کار می‌رود بلکه تنها کافی است اشاره کند که میانگین یک متغیر در دو گروه مستقل با آزمون پارامتری t مقایسه می‌شود.

سطح معناداری آزمون‌ها یا ضریب اطمینان فاصله اطمینان‌ها نیز در این بخش ذکر می‌شود. همچنین، در صورت نیاز، چگونگی گِرد کردن ارقام، نرم‌افزار مورد استفاده، رابطه‌های ریاضی و... نیز معرفی می‌شوند. در برخی از تحلیل‌ها که برای مثال به تبدیل لگاریتمی مقادیر یک متغیر یا محاسبه بَخت یک رویداد نیاز است، داده‌های صفر مشکل‌آفرین هستند، در این شرایط، راهکار مورد استفاده را باید بیان کرد[۲]. علاوه بر این، روش‌های آماری، مقابله با بی‌پاسخی قلم اطلاعاتی نیز در این قسمت شرح داده می‌شود.

۴ـ۲ـ یافته‌ها و نتیجه‌گیری

این بخش می‌تواند شامل سه قسمت باشد. بخش توصیفی، بخش استنباطی و بخش نتیجه‌گیری. بخش توصیفی، از نمودارها، جدول‌ها و مشخصه‌های توزیع برای توصیف وضعیت «نمونه» کمک می‌گیرد. قسمت استنباطی به پرسش‌های استنباطی و آزمون فرضیه‌ها می‌پردازد و در نهایت، در قسمت نتیجه‌گیری، آنچه از تمام تحلیل‌ها به دست

---

۱. سه گونه متغیر در تحلیل‌های آماری به کار گرفته می‌شوند. متغیرهایی که به طور مستقیم اندازه‌گیری شده‌اند و ستونی در مجموعه داده‌ها به آنها اختصاص دارد. متغیرهایی که حاصل پردازش یک یا چند متغیر دیگر هستند. برای مثال، نمره‌ای که از جمع امتیازهای چند متغیر به دست می‌آید یا میانگین‌های حاصل از نمونه‌های مربوط به زیرجامعه‌ها (subpopulations). متغیرهایی که از تبدیل (transformation) بر روی متغیرهای گونه اول یا دوم به دست می‌آیند.

۲. ممکن است این موارد از داده‌ها حذف شوند یا مقدار بسیار کوچکی جایگزین آنها شود یا در صورتی که تعداد صفرها زیاد باشد، دو تحلیل آماری جداگانه یکی برای موارد صفر و یکی برای موارد غیر صفر ارائه شود.

آمده است، بیان می‌شود. نتیجه‌گیری از یک سو با یافته‌های توصیفی و استنباطی مرتبط است و از سوی دیگر با اهداف «پژوهش» و گستره‌ای که یافته‌ها به آن قابل تعمیم هستند. پژوهشگر در نتیجه‌گیری به دستاوردهای پژوهش خود اشاره می‌کند؛ این قسمت باید به گونه‌ای تدوین شود که خواننده بدون مراجعه به دو قسمت قبلی آن را درک کند.

۱ـ۴ـ۲ـ قسمت توصیفی

هدف از قسمت توصیفی، نمایش چشم‌اندازی از داده‌هاست زیرا امکان نمایش تمام داده‌ها به شکلی که خواننده به سرعت بتواند آنها را درک کند وجود ندارد. قالب خاصی برای گزارش یافته‌های توصیفی (و حتی استنباطی) در دست نیست. ممکن است تمام آمارها در جدول یا نمودار ارائه یا در متن شرح داده شوند. گاه آمارها در داخل پرانتز جای می‌گیرند و گاه به طور مستقیم در متن به آنها اشاره می‌شود. با وجود این، هر شیوه‌ای که برای توصیف انتخاب می‌شود، سادگی و سرراست بودن آن از اهمیت بیشتری برخوردار است.

> تحلیل اکتشافی داده‌ها ـ چه در قسمت توصیفی گنجانده شود چه در قسمت استنباطی ـ می‌تواند به خواننده اطمینان دهد که زوایای مختلف داده‌ها مورد واکاوی قرار گرفته است. بررسی توزیع متغیرها به عنوان مقدمه آزمون‌های پارامتری بخشی از این تحلیل است. در واقع، تحلیل اکتشافی فعالیتی افزون بر تحلیل‌هایی است که به منظور پاسخگویی به پرسش‌های پژوهش صورت می‌گیرد.

برای مثال، در مقایسه میانگین‌های چهار گروه، می‌توان هم از درج آمارها در پرانتز و هم از استفاده مستقیم از آنها در متن به صورت زیر کمک گرفت:

> میانگین گروه اول برابر با ۱۳ (با انحراف معیار ۳)، میانگین گروه دوم برابر با ۱۵ (با انحراف معیار ۲/۵)، میانگین گروه سوم برابر با ۱۹ (با انحراف معیار ۳/۲) و میانگین گروه چهارم برابر با ۱۵/۶ (با انحراف معیار ۲/۹) است.

با وجود این، شیوه ساده‌تر، استفاده از جدول زیر است که میانگین‌ها و انحراف معیارهای چهار گروه را در خود جای داده است. این جدول، هم کار مقایسه آمارهای چهار گروه را برای خواننده ساده می‌کند و هم در صورت نیاز به مراجعه‌های بعدی، او را از جستجوی متن بی‌نیاز می‌سازد.

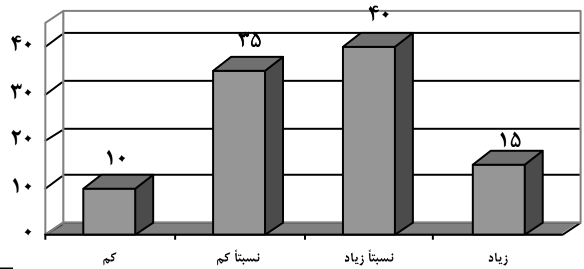
گروه‌ها	اول	دوم	سوم	چهارم
میانگین	۱۳	۱۵	۱۹	۱۵/۶
انحراف معیار	۳	۲/۵	۳/۲	۲/۹

سازماندهی قسمت توصیفی نیز به اولویت‌های پژوهشگر باز می‌گردد. برخی پژوهشگران یافته‌ها را پرسش به پرسش گزارش می‌کنند. برای این منظور، ابتدا به شکلی توصیفی، توزیع هر متغیر در قالب جدول و بویژه نمودار ارائه می‌شود[1]. شرح نوشتاری جدول یا نمودار در بالای آن قرار می‌گیرد و به اطلاعات برجسته، خاص یا دور از انتظار می‌پردازد؛ مانند درصدهای خیلی بزرگ یا خیلی کوچک، نوسان‌های توزیع، چولگی یا برجستگی آن در قسمتی خاص از منحنی. پژوهشگر باید از بیان اطلاعاتی که به‌راحتی از نمودار یا جدول به دست می‌آیند، خودداری کند، مگر آنکه مایل باشد توجه خواننده را بیشتر جلب کند یا او را از درک موضوع ناتوان ببیند. به طور معمول، از مشخصه‌های توزیع در توصیف نوشتاری استفاده می‌شود تا شناخت دقیق‌تری از ویژگی‌های توزیع متغیر به دست آید. اگر بیان چندین مشخصه از توزیع مورد نیاز است، بهتر است همگی آنها در یک جدول جداگانه نمایش داده شوند، در غیر این صورت می‌توان تنها به بیان مشخصه‌ها در توصیف نوشتاری اکتفا کرد. گزارشِ تعداد بی‌پاسخی‌های هر متغیر در جدول یا نمودار توزیع آن ضروری است، هرچند پژوهشگر می‌تواند، بویژه در صورت ناچیز بودن تعداد بی‌پاسخی‌ها، در پاورقی یا متن به آن اشاره کند.

لازم است عنوان جدول یا نمودار آماری به گونه‌ای انتخاب شود که محتوای آنها را منعکس کند. بنابراین، ذکر نام متغیر همراه با نوع اطلاعات جدول یا نمودار راهگشاست. گاهی عنوان نمودار یا جدولِ «توزیع» متغیر، همان پرسشی است که برای سنجش آن متغیر در پرسشنامه به کار رفته است.

مثالی از توصیف توزیع یک متغیر

تا چه حد استفاده از اینترنت را ضروری می‌دانید؟



در پرسشی از افراد نمونه سؤال شد: «تا چه حد استفاده از اینترنت را در زندگی خود ضروری می‌دانید؟». نمودار ستونی نشان می‌دهد، ۵۵ درصد از پاسخگویان استفاده

۱. گاهی چند پرسش از پرسشنامه به یک موضوع می‌پردازند، بنابراین پژوهشگر تصمیم می‌گیرد مجموعه این پرسش‌ها را در کنار هم توصیف کند. به همین دلیل نیز توزیع متغیرهای مربوط به این پرسش‌ها در قالب یک نمودار یا جدول ارائه می‌شود و توصیف‌های نوشتاری آنها نیز در کنار هم و اغلب به صورت مقایسه‌ای صورت می‌گیرد.

از اینترنت را در حد «زیاد» یا «نسبتاً زیاد» ضروری می‌دانند، در حالی که ۴۵ درصد معتقدند استفاده از اینترنت در حد «کم» یا «نسبتاً کم» ضروری است. گفتنی است ۲۰ نفر از نمونه ۴۰۰ نفری به این پرسش هیچ پاسخی نداده‌اند.

۲ـ۴ـ۲ـ قسمت استنباطی

استنباط آماری درباره یک متغیر می‌تواند بلافاصله پس از اطلاعات توصیفی آن ارائه شود[1]. ضروری است پژوهشگر یا از برآورد نقطه‌ای، به همراه «خطای معیار» آن استفاده کند یا از برآورد فاصله‌ای. ارائه برآورد نقطه‌ای بدون اشاره به خطای معیار آن توصیه نمی‌شود، زیرا کاربر هیچ اطلاعی از دقت آن برآورد نخواهد داشت. آزمون فرضیه، با بیان فرضیه «جانشین» آغاز می‌شود[2]. پژوهشگر نام آزمونی را که به کار برده است، همراه با اطلاعات مربوط به آمارهٔ آزمون (مقدار آماره، درجهٔ آزادی و مقدار احتمال) گزارش و بر اساس آن، در سطح معناداری مشخصی نسبت به فرضیه جانشین نتیجه‌گیری می‌کند. اگر آزمون دارای پیش‌فرض‌هایی است که لازم است برقراری آنها پیش از به کارگیری آزمون بررسی شود، نتایج بررسی را باید پیش از ارائه آزمون فرضیه یا در پیوستِ گزارش بیان کرد.

به طور معمول، استنباط‌های آماری که با بیش از یک متغیر سروکار دارند؛ مانند مقایسه میانگین‌ها، همبستگی، تحلیل جدول‌های توافقی یا برازش مدل رگرسیونی، پس از یافته‌های مربوط به تک‌تک متغیرها گزارش می‌شوند. از آنجا که در این استنباط‌ها از متغیر مستقل و متغیر وابسته صحبت می‌شود، پژوهشگر می‌تواند استنباط‌هایی از این دست را نیز بلافاصله پس از یافته‌های توصیفی مربوط به متغیر وابسته ارائه کند، نه در بخشی جداگانه. برای مثال، اگر بررسی رابطهٔ سنوات خدمت به عنوان متغیر مستقل و رضایت شغلی به عنوان متغیر وابسته از اهداف پیمایش باشد، این بررسی پس از ارائهٔ یافته‌های توصیفی متغیر رضایت شغلی صورت می‌گیرد.

گزارش یافته‌های همبستگی دو متغیر می‌تواند از طریق نمودار پراکنش آنها، مقدار ضریب همبستگی و اطلاعات آزمون آماری مربوط صورت بگیرد. یافته‌های برازش مدل رگرسیونی نیز

---

۱. برخی، فصل مربوط به یافته‌های پیمایش را به دو به قسمت تقسیم می‌کنند و در بخش نخست، به ارائه یافته‌های توصیفی و در بخش دوم، به ارائهٔ یافته‌های استنباطی می‌پردازند. بنابراین، یافته‌های توصیفی و استنباطی مربوط به یک متغیر در کنار هم نخواهند بود.

۲. در فصل پنجم، فرضیه صفر، محور بحث آزمون فرضیه بود، در حالی که از نگاه پژوهشی، این فرضیه جانشین است که پژوهش به بررسی آن می‌پردازد. بنابراین پژوهشگر فرضیه‌های جانشین خود را به عنوان فرضیه‌های پژوهش فهرست می‌کند. با وجود این، فرایند از نگاه آماری همان است که در فصل پنجم بیان شد؛ فرضیه جانشین (فرضیه پژوهش) هنگامی پذیرفته می‌شود که آزمون معنادار باشد (یعنی فرضیه صفر رد شده است) و هنگامی رد می‌شود که آزمون معنادار نباشد (یعنی فرضیه صفر رد نشده است).

شامل معادلهٔ مدل برازش‌یافته، ضریب تعیین مدل، آزمون نیکویی برازش مدل و آزمون‌های تک‌تک این ضرایب است. اغلب، این یافته‌ها با تفسیر نوشتاری همراه هستند که جهت رابطه، شدت رابطه و نقش متغیر مستقل را در تبیین تغییرات متغیر وابسته توضیح می‌دهند[۱].

مثالی از تفسیر مدل رگرسیونی

رابطهٔ میزان افسردگی نوجوانان (D) با مدت زمانی که صرف بازی‌های رایانه‌ای (G) و ارتباطات اینترنتی (I) مانند چت می‌کنند، از طریق مدل رگرسیونی با دو متغیر مستقل بررسی شد و معادلهٔ زیر به دست آمد:

$$D = -۳ + ۴/۵G + ۵/۳I$$

آزمون تجزیهٔ واریانس $(F = ۳/۹۱, df_۱ = ۳, df_۲ = ۱۷, p - value = ۰/۰۲۷)$ مناسب‌بودن مدل را تأیید می‌کند و با توجه به جدول زیر، تمام ضرایب مدل از لحاظ آماری در سطح ۵ درصد معنادار هستند.

ضریب	مقدار	خطای معیار	آماره	درجهٔ آزادی	مقدار احتمال
مقدار ثابت	−۳	۰/۴۳	−۶/۹۷	۱	۰/۰۲۲
مدت زمان بازی‌های رایانه‌ای (G)	۴/۵	۰/۷۱	۶/۳۳	۱	۰/۰۲۴
مدت زمان ارتباطات اینترنتی (I)	۵/۳	۰/۹۱	۵/۸۲	۱	۰/۰۲۷

ضریب تعیین نشان می‌دهد، متغیرهای مدت زمان پرداختن به بازی‌های رایانه‌ای و مدت زمان ارتباطات اینترنتی همزمان می‌توانند ۶۰ درصد از تغییرات میزان افسردگی نوجوانان را توضیح دهند. از میان دو متغیر مستقل، متغیر مدت زمان ارتباطات اینترنتی نقش بیشتری در توضیح این تغییرات دارد.

در جدول توافقی، مقوله‌های متغیر مستقل، به عنوان ستون‌های جدول و مقوله‌های متغیر وابسته، به عنوان سطرهای جدول در نظر گرفته می‌شوند، زیرا عموماً پژوهشگر مایل است وضعیت متغیر وابسته را در بین مقوله‌های متغیر مستقل مقایسه کند و این مقایسه، از طریق ستون‌ها، ساده‌تر و سرراست‌تر است. از همین رو، درصدهای ستونی در خانه‌های جدول توافقی جای می‌گیرند و در صورت نیاز، درصدهای حاشیه‌ای سطری یا حاشیه‌ای ستونی نیز ارائه می‌شوند. از آنجا که جدول توافقی، حداقل براساس داده‌های دو متغیر

۱. تمرکز پژوهشگر بر استفاده از مدل‌های آماری پیچیده‌تر با چندین متغیر مستقل، هم بر «فرایند برازش» مدل است و هم بر شرح و تفسیر مدل نهایی. بنابراین، او ابتدا اطلاعات کاملی را درباره پیش‌فرض‌ها، انتخاب مدل و معیارهای نیکویی برازش آن ارائه می‌دهد و سپس به شرح و تفسیر مدل نهایی می‌پردازد.

مقوله‌ای تشکیل می‌شود، با دو پرسش از پرسشنامه مرتبط است. بنابراین، اگر فردی از نمونه به یکی از این دو پرسش پاسخ نداده باشد، در جدول توافقی به حساب نمی‌آید و از موارد بی‌پاسخ قلمداد می‌شود. به همین دلیل، کاملاً طبیعی است تعداد بی‌پاسخی‌های جدول توافقی بیشتر از تعداد بی‌پاسخی‌های تک‌تک متغیرهای جدول باشد[۱].

اگر آزمون رابطهٔ دو متغیر مقوله‌ای جدول معنادار باشد، شدت و جهت رابطه با معیار مناسب اندازه‌گیری و همراه با اطلاعات آمارهٔ آزمون گزارش می‌شود. همچنین در صورت وابستگی دو متغیر، پژوهشگر تفاوت‌های درصدهای ستون‌ها را با یکدیگر با تأکید بر شدت رابطه و تفاوت‌های بزرگ یا دور از انتظار، به طور نوشتاری توصیف می‌کند. توجه کنید که اگر آزمون، وابستگی دو متغیر را نشان ندهد، توصیف اختلاف درصدهای ستون‌ها می‌تواند برای کاربر گمراه‌کننده باشد، زیرا او ناخودآگاه تصور می‌کند این تفاوت‌ها را باید به جامعه تعمیم دهد، در حالی که آزمون نشان‌دهنده آن نبوده است.

مثالی از تفسیر جدول توافقی دو بعدی

آزمون آماری ($\chi^{۲}=۱۱۵/۴,\ df=۶,\ p-value=۰$) بر وجود رابطه بین میزان ضرورت استفاده از اینترنت و میزان تمایل به شرکت در دوره‌های آموزش کاربری اینترنت دلالت دارد به طوری که هرچه اعتقاد به ضرورت استفاده از اینترنت بیشتر باشد، بر تمایل به شرکت در دوره‌های آموزشی کاربری اینترنت نیز افزوده می‌شود ($Sommers'D=۰/۵۴$).

تمایل \ ضرورت		کم	نسبتاً کم	نسبتاً زیاد	زیاد	کل
کم		۷۵	۶۵	۱۰	۵	۲۰
متوسط		۲۰	۲۵	۲۰	۱۵	۳۰
زیاد		۵	۱۰	۷۰	۸۰	۵۰
جمع	درصد	۱۰۰	۱۰۰	۱۰۰	۱۰۰	۱۰۰
	فراوانی	۴۶	۷۱	۹۰	۳۱	۲۳۸

گفتنی است ۸۰ درصد از کسانی که معتقدند استفاده از اینترنت در حد «زیاد» ضرورت دارد، به شرکت در دوره‌های کاربری اینترنت نیز در حد «زیاد» تمایل دارند. در مقابل، تنها ۵ درصد از کسانی که در حد «کم» به این ضرورت معتقدند، به شرکت در این دوره‌ها در حد «زیاد» تمایل دارند.

۱. همین امر می‌تواند باعث اختلاف توزیع هر متغیر در جدول توزیع فراوانی با توزیع حاشیه‌ای متناظر با آن متغیر در جدول توافقی شود.

۳ـ۴ـ۲ـ قسمت نتیجه‌گیری

پیشتر اشاره شد که نتیجه‌گیری، به بخشی از یافته‌های توصیفی و استنباطی نیز می‌پردازد. این بازخوانی باید نوآورانه باشد تا خواننده احساس نکند عیناً با همان مطالب پیشین روبه‌رو است. پژوهشگر می‌تواند بازخوانی یافته‌ها را با تمرکز بر کلیات به جای جزئیات، تازه و جالب کند. برای مثال، تک‌تک فرضیه‌ها و پرسش‌های جزئی در قسمت استنباطی بررسی می‌شوند در حالی که می‌توان در نتیجه‌گیری، جمع‌بندی آنها را ارائه کرد و به تک‌تک آنها نپرداخت. همچنین، جدول‌هایی با جزئیات فراوان می‌توانند جای خود را به نمودارهای ساده ولی مقایسه‌ای بدهند.

تحلیل و تفسیر یافته‌های آماری در این قسمت از گزارش صورت می‌گیرد. فراموش نکنید که آزمون‌های غیرمعنادار به اندزه آزمون‌های معنادار می‌توانند مهم باشند. حذف نتایج آزمون‌ها به دلیل آنکه معنادار نشده‌اند، منطقی نیست. همچنین، یک پژوهش مطمئن هرگز یافته‌های متناقض یا ضعیف را حذف نمی‌کند بلکه مانند مثال زیر، به دنبال تحلیل مناسبی برای آنها می‌گردد.

> هنگام برخورد با این یافته که «تنها ۲ درصد از تصادفات به رانندگان مسن اختصاص دارد»، چه باید کرد؟ یک دیدگاه می‌گوید رانندگان مسن، باتجربه یا محتاط هستند و همین امر آمار تصادفات آنان را کاهش داده است. در مقابل، دیدگاه دیگری می‌گوید رانندگان مسن به دلیل ضعف بینایی یا مشکلات جسمی دیگر، چندان دقیق نیستند و توانایی واکنش سریع را در حوادث رانندگی ندارند پس انتظار می‌رود آمار تصادفات آنان بیشتر باشد. با دو دیدگاه که یکی آمار پایین تصادفات افراد مسن را تأیید و دیگری رد می‌کند، روبه‌رو هستیم. آیا باید این یافته را نادیده بگیریم یا هیچ‌یک از دو دیدگاه را نپذیریم و در جستجوی تحلیل دیگری باشیم؟ ابتدا باید از خود بپرسیم که این یافته می‌تواند مبنای مناسبی برای قضاوت درباره رانندگان مسن باشد؟ دوباره جمله «تنها ۲ درصد از تصادفات به رانندگان مسن اختصاص دارد» را مرور کنید؛ این جمله نمی‌گوید که «تنها ۲ درصد از رانندگان مسن، اخیراً تصادف کرداند». به عبارت دیگر، مبنای محاسبه ۲ درصد در میان «تصادفات» است نه «رانندگان مسن». بنابراین، رقم ۲ درصد هیچ اطلاعی درباره وضعیت رانندگی افراد مسن به دست نمی‌دهد بلکه مربوط به توزیع تمام تصادفات ثبت شده است. بنابراین، تحلیل دیگری را نیز می‌توان مطرح کرد. از آنجا که افراد مسن به دلیل مشکلات جسمی و حرکتی به رانندگی گرایش کمتری دارند، تعداد تصادفات رانندگی مربوط به آنها نیز کمتر از سایر گروه‌های سنی است زیرا آنان به مراتب

کمتر از سایرین رانندگی می‌کنند. با این وصف، ممکن است محاسبه درصد تصادفات افراد مسن در میان تعداد کل رانندگان مسن به رقم بالایی برسد.

در واقع، رمز موفقیت در تحلیل و تفسیر منطقی یافته‌های آماری، در نگاه جامع و پرسشگرانه پژوهشگر نهفته است. هنگام روبه‌رو شدن با یافته‌هایی که دور از انتظار هستند و در نگاه نخست، توجیهی برای آنها پیدا نمی‌کنید، پرسش‌های متعددی را درباره جامعه آماری آن، شیوه محاسبه آماره، برداشت شما از آماره و بستری که یافته‌ها باید در آن تفسیر شوند، از خود بپرسید. به مثال کلاسیک زیر توجه کنید:

در بررسی ویرانی‌های حاصل از آتش‌سوزی‌ها مشخص شد، هر چه گروه‌های بیشتر و مجهزتری از آتش‌نشانان برای خاموش کردن آتش در محل آتش‌سوزی حضور داشته باشند، وسعت ویرانی‌های آتش‌سوزی نیز بیشتر است. بنابراین، بین وسعت ویرانی‌ها و تعداد آتش‌نشانان رابطه مستقیمی وجود دارد و از این رو نباید آتش‌نشانان زیادی را به محل‌های آتش‌سوزی اعزام کرد. اگرچه رابطه وسعت ویرانی‌های ناشی از آتش‌سوزی و تعداد آتش‌نشانان کاملاً درست است، این نتیجه‌گیری منطقی نیست؛ گو اینکه علت وسعت ویرانی‌ها، تعداد زیاد آتش‌نشانان بوده است و با کاهش تعداد آنان می‌توان از وسعت ویرانی‌ها کاست. اشتباه این نتیجه‌گیری در برداشت «علّی» از «رابطه» دو متغیر نهفته است. در واقع، هنگامی که با آتش‌سوزی‌های مهیب و جدی رو به رو هستیم، آتش‌نشانان بیشتری نیز برای خاموش کردن آن اعزام می‌شوند؛ پس «علت» ویرانی‌ها مهیب بودن آتش‌سوزی است که پیش از حضور آتش‌نشانان رخ داده است نه تعداد آتش‌نشانان اعزامی.

همچنین، مراقب باشید که در تحلیل و تفسیر یافته‌ها دچار جانبداری نشوید. گاه روش‌های آماری ابزاری می‌شوند تا پژوهشگر نتیجه دلخواه خود را از آنها ارائه کند. مثال زیر، نمونه‌ای از استفاده نادرست از مشخصه‌های تمرکز توزیع است.

معلمی ۲۱ دانش‌آموز دارد و آزمونی را برای ارزیابی آنان برگزار می‌کند. توزیع فراوانی نمرات دانش‌آموزان به شرح جدول زیر است. مدیر مدرسه با دیدن این نتایج احساس می‌کند که آزمون این معلم خیلی آسان بوده است و او را ملامت می‌کند زیرا میانه نمرات ۱۸ است. اما معلم در دفاع از خود و رد این موضوع، می‌گوید که معدل نمرات دانش‌آموزان ۱۶/۸ است. از سوی دیگر، وقتی پدر یکی از دانش‌آموزان از پسرش می‌پرسد که چه نمره‌ای در آزمون گرفته است، دانش‌آموز پاسخ می‌دهد ۱۵ ولی پدر نگران نباش زیرا بچه‌های کلاس، این نمره را بیشتر از نمره‌های دیگر گرفته‌اند (توجه کنید که نمای توزیع نمرات ۱۵ است).

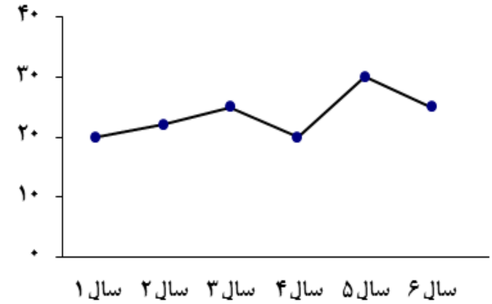
نمره	۲۰	۱۹	۱۸	۱۵	۱۳
فراوانی	۴	۵	۳	۶	۴

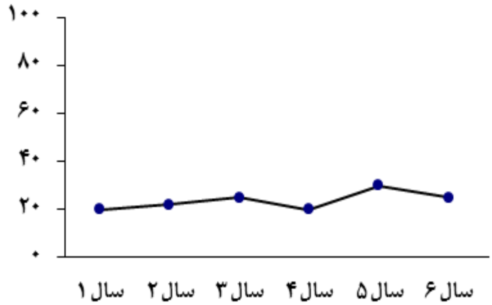
درصدها را نمی‌توان بدون توجه به «مبنای» درصدگیری آنها تفسیر کرد. برای مثال، اگر من به شما بگویم با مطالعه مطالب این فصل، ۳۰ درصد بهتر می‌توانید گزارش آماری پژوهش خود را تدوین کنید، این درصد، تفسیر چندان روشنی نخواهد داشت زیرا شما تصور مشخصی از مبنای درصدگیری ندارید؛ ۳۰ درصد بهتر از چه چیزی؟ این موضوع بویژه در شرایطی که درصدها با یکدیگر مقایسه می‌شوند، از اهمیت بیشتری برخوردار است. به مثال زیر توجه کنید:

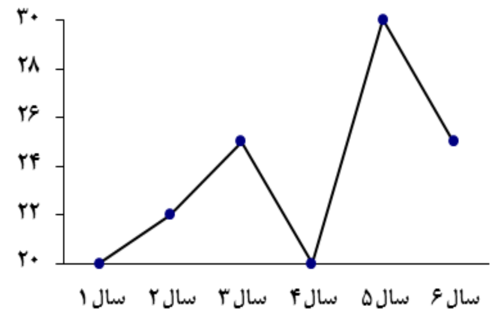
> می‌شنوید که نرخ رشد جرم در شهرهای «آ» و «ب» به ترتیب ۷۵ و ۱۰ درصد است و نتیجه می‌گیرید که وقوع جرم در شهر «آ» جدی‌تر است زیرا ۷۵ بیش از ۱۰ است. اکنون دقیق‌تر به این دو شهر نگاه کنید. شهر «آ» در سال گذشته، شاهد ۴ جرم و در سال جاری، شاهد ۷ جرم بوده است در حالی که شهر «ب» در این دو سال، به ترتیب، شاهد ۳۰ و ۳۳ جرم بوده است. آیا با این اطلاع جدید، نتیجه‌گیری قبلی خود را تغییر می‌دهید؟ ممکن است بگویید تعداد جرم‌های افزایش‌یافته در هر دو شهر برابر با ۳ مورد است پس وضعیت آنها یکسان است. همچنین، ممکن است بگویید تعداد جرم‌های شهر «ب» بیش از شهر «آ» است پس وضعیت این شهر وخیم‌تر از شهر «آ» است. این نتیجه‌گیری‌های متفاوت و حتی متناقض به استفاده از آمارهایی با مبناهای مختلف باز می‌گردد. مقایسه وقوع جرم در دو شهر باید بر اساس معیار یکسانی مانند نرخ «سرانه» صورت بگیرد. برای این منظور، فرض کنید جمعیت شهرهای «آ» و «ب» به ترتیب ۷۰۰ و ۳۳۰۰۰ هزار نفر است. نرخ سرانه وقوع جرم در شهرهای «آ» و «ب» از تقسیم تعداد جرم‌های سال جاری بر جمعیت، به ترتیب ۱ درصد و ۰/۱ درصد به دست می‌آید. این دو رقم بیان می‌کنند در شهر «آ» ۱۰ جرم به ازای هر ۱۰۰۰ نفر و در شهر «ب»، ۱ جرم به ازای هر ۱۰۰۰ نفر رخ می‌دهد. مقایسه این دو رقم کاملاً منطقی است زیرا مبنای هر دو آنها یکسان یعنی ۱۰۰۰ نفر است. بنابراین، می‌توان نتیجه گرفت که وقوع جرم در شهر «آ» جدی‌تر از شهر «ب» است.

نمودارها نیز می‌توانند عامل برداشت‌های نادرست از یافته‌های آماری شوند. همان طور که استفاده از نمودارِ مناسب می‌تواند به درک ساده‌تر یافته‌ها کمک کند، استفاده نادرست از نمودارها نیز می‌تواند مانند مثال زیر به تفسیرهای اشتباه بینجامد.

نمودارهای زیر، درصد فروش نوعی کالا را طی ۶ سال متوالی نشان می‌دهند. داده‌های یکسانی در هر سه نمودار به کار رفته است ولی سه برداشت متفاوت از ظاهر آنها می‌توان کرد زیرا محور عمودی در هر نمودار، متفاوت مدرج شده است. نمودارهای اول و سوم دارای نوسان‌های قابل توجهی هستند در حالی که نمودار دوم هموارتر به نظر می‌رسد. همچنین، نوسان‌ها در نمودار سوم بسیار شدیدتر از نمودار اول است. با وجود اینکه هیچ‌یک از این نمودارها دروغ نمی‌گویند، به لحاظ بصری تأثیر متفاوتی بر روی مخاطب خود می‌گذارند.







ممکن است پژوهشگر مایل نباشد از نمودار دوم استفاده کند زیرا فضای خالی بالای آن جلوه مناسبی ندارد. در مقابل، ظاهر نمودارهای اول و سوم از این نظر مناسب‌تر است. پژوهشگر می‌تواند هر یک از این نمودارها را به کار بگیرد ولی برای جلوگیری از برداشت‌های نادرست خواننده گزارش، ارقام مربوط به نقاط نمودار را در داخل نمودار یا در شرح نوشتاری آن ذکر کند.

۳ـ بایدها و نبایدها

تحت این عنوان، درباره مجموعه‌ای از اشتباه‌های رایج صحبت خواهد شد که در گزارش‌های آماری با آنها برخورد می‌کنید. خواننده می‌تواند از مطالب پیشین و این اشتباه‌ها برای خود فهرست راهنمایی مانند جدول انتهای این فصل تهیه کند تا خطاهای گزارش‌نویسی خود را کاهش دهد.

۱. مستقیم به مسئله مورد بررسی بپردازید و از حاشیه رفتن خودداری کنید. خلاصه‌نویسی را با توجه به دانش مخاطب گزارش مد نظر قرار دهید. اگر او قادر است به درستی از نمودارها یا ارقام جدول‌ها استفاده کند، هیچ ضرورتی ندارد شما نیز در شرح یک جدول یا نمودار، به تمام جزئیات آن اشاره کنید، بلکه کافی است تنها بر نکات برجسته و قابل توجه تأکید نمایید.

۲. اصطلاحات روشن و صریح را به کار بگیرید تا جایی برای ابهام یا برداشت‌های اشتباه باقی نماند. برای مثال، به جای «افراد تحصیلکرده» از «افرادِ دارای مدرک کارشناسی ارشد و دکترا» استفاده کنید. عنوان اخیر به طور دقیق به خواننده می‌گوید که منظور شما از افراد تحصیلکرده چیست.

۳. به طور معمول، با زمان حال به یافته‌ها اشاره می‌شود. برای نمونه، بگویید «این جدول نشان **می‌دهد** که ...» یا «ضریب همبستگی بیانگر آن **است** که ...». تنها اگر رویدادی در گذشته اتفاق افتاده است، از زمان گذشته استفاده کنید مانند «داده‌ها با روش نمونه‌گیری سه مرحله‌ای گردآوری **شدند**».

۴. از ضمیر شخصی «من» یا «ما» کمتر استفاده کنید. بنابراین، به جای «ابتدا آزمون تجزیه واریانس را اجرا کردیم و سپس آزمون تعقیبی شفه را به کار بردیم»، بنویسید «ابتدا آزمون تجزیه واریانس اجرا شد و سپس آزمون تعقیبی شفه به کار رفت».

۵. همان طور که جدول‌ها یا نمودارها را شماره‌گذاری می‌کنید، اگر بیش از یک معادله (برای مثال، رگرسیون) یا رابطه ریاضی در گزارش وجود دارد، آنها را نیز شماره‌گذاری کنید.

۶. لازم نیست خروجی نرم‌افزارها را عیناً در گزارش به کار بگیرید زیرا اغلب این خروجی‌ها شامل جزئیات کم اهمیتی برای مخاطبان غیرحرفه‌ای گزارش هستند (همچنین ظاهر این خروجی‌ها به ویرایش نیاز دارد تا در گزارش بهتر به نظر برسد). در صورت نیاز، می‌توان خروجی‌ها را در پیوست گزارش ارائه داد.

۷. عنوان جدول یا نمودار باید به خوبی گویای محتوای آن باشد. در عین حال بهتر است این عنوان، کوتاه و در قالب یک عبارت و نه جمله بیان شود. پس عنوانِ «توزیع کاربران اینترنت» بهتر از این عنوان است: «توزیع کسانی که از اینترنت استفاده می‌کنند».

۸. دقت در محاسبات، مزیت گزارش است اما همیشه ارائه ارقام با اعشارهای فراوان ضرورت ندارد. به طور معمول، گزارش ارقام حداکثر با دو رقم اعشار کافی است (گاهی

این دو رقم نیز ضرورت ندارد و می‌توان اعشارها را به اعداد صحیح گرد کرد). اگر آماره‌ها به صورت دستی محاسبه می‌شوند، گِرد کردن را در آخرین مرحله انجام دهید تا خطای محاسبات کمتر شود. برای مثال، آماره آزمون t از تقسیم میانگین بر خطای معیار به دست می‌آید. بنابراین، در محاسبه میانگین و خطای معیار از اعشار بیشتری استفاده کنید و تنها در محاسبه نهایی آماره، یعنی تقسیم میانگین بر خطای معیار از گرد کردن تا دو رقم اعشار استفاده نمایید. (استثنا: گرد کردن تا دو رقم اعشار برای اعدادی که ماهیتاً کوچک هستند مانند مقدار احتمالِ آزمون‌ها، به کار نمی‌رود. در چنین حالتی باید با توجه به میزان کوچکی این اعداد، سه یا چهار رقم اعشار و حتی بیشتر از آن را ارائه داد).

۹. اگر از رابطه‌های ریاضی مانند رابطه محاسبه اندازه نمونه یا یک مدل رگرسیون استفاده می‌کنید، تک‌تک نمادها را به طور دقیق تعریف کنید.

۱۰. هرگز از آمار توصیفی، نتیجه‌گیری استنباطی نکنید. همواره توجه داشته باشید با مشاهده تفاوت در جدول‌ها، نمودارها یا مشخصه‌های توزیع که برای توصیف داده‌های نمونه به کار رفته‌اند، ناخودآگاه نتایجی را به جامعه تعمیم ندهید. برای مثال، اگر ۳۵ درصد از مردان و ۳۰ درصد از زنانِ «نمونه» به فوتبال علاقه دارند، فوراً بر اساس این ارقام توصیفی نتیجه نگیرید که مردان بیش از زنان به فوتبال علاقه‌مندند. ضروری است پیش از هر گونه نتیجه‌گیری، با آزمون آماری مناسب، این دو آماره مقایسه شوند و تنها در صورت وجود اختلاف آماری معنادار، نسبت به تفاوت آنها قضاوت شود.

۱۱. برای استنباط آماری از داده‌های حاصل از سرشماری، به روش‌های آمار استنباطی نیازی نیست بلکه تنها با استناد به روش‌های آمار توصیفی می‌توان دست به استنباط زد. بنابراین، اگر رضایت شغلی تمام کارمندان شرکتی سنجیده شده و میانگین نمرات رضایت شغلی آنان ۸۵ از ۱۰۰ است، بدون هیچ آزمون آماری و تنها با استناد به این مشخصه توزیع می‌توان ادعا کرد میانگین رضایت شغلی کارمندان این شرکت دقیقاً برابر با ۸۵ است (خطای غیرنمونه‌گیری را نادیده بگیرید!).

۱۲. اگر لازم است که به توصیف نمونه بپردازید، با استفاده از اصطلاحاتی نظیر «پاسخگویان»، «افراد مشاهده‌شده» یا «آزمودنی‌ها» به خواننده تأکید کنید که اطلاعات در سطح «افراد نمونه» ارائه می‌شود و قابل تعمیم به جامعه نیست. بنابراین وقتی گفته می‌شود «۶۵ درصد از پاسخگویان با وسایل حمل‌ونقل عمومی به سر کار خود می‌روند» یا «میانگین

بهره هوشی آزمودنی‌ها ۱۱۰ است» بر توصیف نمونه تأکید شده است. پس نباید از این جملات چنین برداشت شود که برای مثال، ۶۵ درصد از «مردم» از وسایل حمل‌ونقل عمومی استفاده می‌کنند یا میانگین بهره هوشی دانش‌آموزان ۱۱۰ است.

۱۳. هرگز از روش‌های آماری که با کاربرد و چگونگی تفسیر یافته‌های آنها به خوبی آشنا نیستید، استفاده نکنید. در چنین شرایطی از یک متخصص آماری کمک بگیرید.

۱۴. حتماً قلمرویی را که یافته‌های استنباطی به آن قابل تعمیم هستند، با تعریف دقیق جامعه آماری برای خوانندگان گزارش مشخص کنید تا مانع از برداشت‌ها و تعمیم‌های نادرست شوید.

۱۵. یافته‌های استنباطی را به جای «زبان آماری» با «زبانی قابل فهم‌تر» ارائه دهید. برای نمونه، به جای «آزمون t نشان می‌دهد که دو گروه مردان و زنان دارای اختلاف معنادار هستند»، بگویید «زنان نمرات بیشتری نسبت به مردان کسب کرده‌اند».

۱۶. در آزمون فرضیه، تفسیری اشتباه از «مقدار احتمال» ارائه ندهید. مقدار احتمال، تنها احتمال وقوع آماره آزمون مشاهده‌شده را تحت درستی فرضیه صفر نشان می‌دهد. بنابراین، هرگز آن را به معنای احتمال نادرستی فرضیه صفر به کار نگیرید. برای مثال، اگر آزمون t را برای فرضیه صفرِ «نبود اختلاف بین میانگین‌های دو جامعه» اجرا می‌کنید و اختلاف میانگین‌های نمونه‌های انتخابی از این دو جامعه ۸ و مقدار احتمال آزمون ۰/۰۳ است، این مقدار نشان می‌دهد که اگر اختلافی بین دو میانگین وجود نداشته باشد (فرضیه صفر درست باشد)، ۳ درصد احتمال دارد که میانگین دو نمونه، اختلافی حداقل به بزرگی اختلاف مشاهده‌شده یعنی ۸ داشته باشند.

۱۷. اگرچه بزرگی یا کوچکی مقدار احتمال می‌تواند به رد شدن یا نشدن فرضیه صفر بینجامد، با توجه به مقدار α، اهمیتی ندارد که یک مقدار بسیار کوچک از مقدار احتمال، فرضیه صفر را رد کرده است یا یک مقدار به نسبت کوچک از آن. برای مثال، هر دو مقدار احتمال ۰/۰۰۱ و ۰/۰۲۱ می‌توانند فرضیه صفر را در سطح معناداری ۰/۰۵ رد کنند ولی نباید نتیجه گرفت که مقدار احتمال ۰/۰۰۱ «قوی‌تر» از ۰/۰۲۱ فرضیه صفر را رد می‌کند. در واقع، اگر مقدار احتمال خیلی کوچک باشد، می‌توان گفت «تحت درستی فرضیه صفر»، احتمال وقوع چنین مقداری برای آماره آزمون بسیار کوچک است و از این گفته تنها

می‌توان نتیجه گرفت که یکی از مقادیر نادر (نامحتمل) آماره آزمون تحت درستی فرضیه صفر مشاهده شده است.

۱۸. نرم‌افزارهای آماری به جای ما فکر نمی‌کنند بلکه تنها محاسبات را دقیق و سریع اجرا می‌کنند. پس این رویه را در پیش نگیرید که داده‌ها را از یک سو وارد رایانه کنید و بدون هیچ دخالتی، از سوی دیگر خروجی نرم‌افزار را دریافت کنید. تصمیم درباره روش‌های مناسب برای تحلیل داده‌ها بویژه با توجه به نوع متغیرها، بررسی پیش‌فرض‌های این روش‌ها و تفسیر یافته‌های حاصل همگی بر عهده پژوهشگر است. او باید با فرایند محاسباتی روش‌ها و محدودیت‌های آنها آشنا باشد و از روی آگاهی، روش مناسب را انتخاب کند؛ نه به این دلیل که نرم‌افزار امکان استفاده از روشی را به راحتی برای او فراهم می‌کند.

۱۹. آمار استنباطی چیزی را ثابت نمی‌کند بلکه تنها بخشی از شواهدی است که می‌توانند فرضیه‌ای را حمایت کنند. بنابراین در گزارش‌نویسی، استنادِ صرف به نتیجه یک آزمون آماری نمی‌تواند ادعایی را موجه سازد. از نگاه پژوهشی، فرضیه‌های آماری باید برآمده از نظریه‌ها و دیدگاه‌های علمی باشند. ممکن است آزمون‌های آماری اختلاف‌ها یا رابطه‌های معناداری را نشان دهند ولی تا زمانی که این اختلاف‌ها یا رابطه‌ها به لحاظ علمی تأیید نشوند، پذیرفته نخواهند شد.

پرسش‌ها	توضیحات
آیا پرسش‌ها و فرضیه‌ها جامع و با جزئیات بیان شده‌اند؟	اجرا پیش از طراحی پرسشنامه بیان آماری پرسش‌ها و فرضیه‌ها
آیا از روش‌های آماری مناسب برای پرداختن به پرسش‌ها و فرضیه‌ها استفاده شده است؟	توجه به نوع متغیرها تناسب روش به لحاظ توصیفی یا استنباطی بودن توجه به پیش‌فرض‌های روش بویژه در روش‌های استنباطی انتخاب ساده‌ترین و گویاترین روش با رعایت دقت و اعتبار علمی
آیا اطلاعات کافی در حمایت از معتبر بودن داده‌ها ارائه شده است؟	اعتبار منبع داده‌ها روش نمونه‌گیری و اندازه نمونه کنترل‌های کیفی فرایند گردآوری داده‌ها
آیا تمام یافته‌ها ارائه شده‌اند؟	خودداری از حذف آزمون‌های غیرمعنادار گزارش یافته‌ای ضعیف و حتی متناقض
آیا یافته‌ها به درستی تفسیر شده‌اند؟	پرهیز از تعمیم اشتباه برداشت استنباطی از یافته‌های توصیفی برداشت علّی از یک رابطه آماری توجه به قابل مقایسه بودن یافته‌ها بویژه درباره درصدها
آیا ضرایب اطمینان و حداکثر خطا بیان شده‌اند؟	بیان سطح معناداری آزمون‌ها بیان ضریب اطمینان و خطای فاصله‌های اطمینان خودداری از ارائه برآورد نقطه‌ای بدون اشاره به خطای معیار آن

تمرین‌های فصل نهم

۱. مشکل عبارت «میانگین وقت آزاد پاسخگویان ۳ ساعت و ۱۰ دقیقه بوده است»، چیست؟

۲. پژوهشگری از پاسخگویان یک نظرسنجی دو پرسش زیر را پرسیده است:
پرسش اول: تا چه حد مردم را خوش‌حساب می‌دانید تا به آنان پول قرض دهید؟
پرسش دوم: تا چه حد ممکن است به کسی پول قرض دهید؟

در پاسخ به پرسش اول، ۷۰ درصد اظهار کرده‌اند به افراد اطمینان دارند تا به آنها پول قرض دهند ولی در پاسخ به پرسش دوم، تنها ۴۰ درصد گفته‌اند حاضر هستند به کسی پول قرض دهند. این پژوهشگر احساس می‌کند تناقضی بین این دو یافته وجود دارد، زیرا اگر ۷۰ درصد به مردم اطمینان دارند، پس باید حداقل همین درصد حاضر باشند به مردم پول قرض دهند، در حالی که درصد کمتری از پاسخگویان این گونه عمل می‌کنند. آیا تصور این پژوهشگر درست است؟

۳. اشکالات متن زیر را مشخص کنید:

پژوهشگری قصد دارد تحقیقی درباره میزان استفاده از اینترنت اجرا کند. او یکی از فرضیات تحقیق خود را چنین عنوان می‌کند: «میزان استفاده از اینترنت در میان افراد تحصیلکرده و غیرتحصیلکرده به یک میزان است». خوشبختانه متوجه می‌شود که در سرشماری اخیر، ساعات استفاده از اینترنت هر فرد ثبت شده است. بنابراین، اطلاعات مورد نیاز را از گزارش سرشماری تهیه می‌کند و به کمک آزمون t در می‌یابد به طور متوسط، افراد تحصیلکرده به طور معناداری بیش از افراد غیرتحصیلکرده تلویزیون تماشا می‌کنند. از این رو در گزارش تحقیق خود می‌نویسد: «بر اساس داده‌های موجود فرضیه تحقیق رد می‌شود».

پیوست‌ها

پیوست ۱: فهرست منابع

پیوست ۲: جدول‌های آماری

پیوست ۳: شکل‌های کوچک و بزرگ حروف یونانی

پیوست ۴: تعریف برخی از اصطلاحات

پیوست ۱: فهرست منابع

۱. خوشگویان‌فرد، ع (۱۳۸۷). «قابلیت‌های نمونه‌گیری مکرر»، مجموعه مقالات اولین سمینار تخصصی پیمایش پانلی، تهران، ایسپا.
۲. مدنی، علی (۱۳۷۰). «مفاهیم اساسی آمار»، انتشارات کتابخانه فروردین.
۳. نمازی‌راد، م، نواب‌پور، ح، خوشگویان‌فرد، ع (۱۳۸۶). «مقایسه تجربی چند آمارگیری چرخشی»، مجله علمی و ترویجی گزیده مطالب آماری، سال ۱۸، شماره ۱، صص ۳۵-۴۹.
۴. نواب‌پور، حمیدرضا (۱۳۷۹). «آمار کاربردی و زبان برنامه‌نویسی SAS»، انتشارات ارفع.
5. Agresti, A. (2003). “Categorical Data Analysis”, John Wiley & Sons, Inc.
6. Babbie, E.R. (2010). “The Practice of Social Research”, Wadsworth, Cengage Learning, Inc.
7. Baker, T. L. (1999). “Doing Social Research”, McGraw-Hill Inc.
8. Balanda K.P. and MacGillivray, H.L. (1988). “Kurtosis: A Critical Review”. The American Statistician, 42:2, pp 111–119.
9. Belle, G. (2002). “Statistical Rules of Thumb”, John Wiley & Sons, Inc.
10. Berg, N. (2002). “Non-response Bias”, University of Texas at Dallas. Retrieved from: https://www.utdallas.edu/~nberg/Berg_ARTICLES/BergNonResponseBiasMay2002.pdf
11. Bishop, Y.M.M., Fienberg, S.E. and Holland, P.W. (1975). “Discrete Multivariate Analysis. Theory and Practice”, Cambridge: MIT Press.
12. Bulmer, M.G. (1979). “Principles of Statistics”, Dover Publications, New York: USA.
13. Cargan, L. (2007). “Doing Social Research”, Rowman and Littlefield Publishers, Inc.
14. Conover, W. J. (1999). “Practical Nonparametric Statistics”, John Wiley & Sons, Inc.
15. Corder, G.W. and Foreman, D.I. (2009). “Nonparametric Statistics for Non-Statisticians: A Step-by-Step Approach”, John Wiley & Sons, Inc.
16. Edwards, J.E.,Thomas, M.D., Rosenfeld, P. and Booth-Kewley, B. (1997). “How To Conduct Organizational Surveys: A Step-by-Step Guide”, Sage Publications, Inc.
17. Fowler, F.J. (2009). “Survey Research Methods”, Sage Publications, Inc.
18. Good, P.I. and Hardin, J.W. (2006). “Common Errors in Statistics: (and How to Avoid Them)”, John Wiley & Sons, Inc.
19. Groves, R.M., Fowler, F.J., Couper M.P., Lepkowski, J.M., Singer, E. and Tourangeau, R. (2009). “Survey Methodology”, John Wiley & Sons, Inc.
20. Heckathorn, D.D. (2002). "Respondent-Driven Sampling II: Deriving Valid Estimates from Chain-Referral Samples of Hidden Populations", Social Problems, 49: 11-34.

21. Huff, D., Irving, G. (1993). “How to Lie with Statistics?”, W.W.Norton & Company, Inc.
22. Johnson, R.A. and Bhattacharyya, G.K. (2009). “Statistics: Principles and Methods”, John Wiley & Sons Inc.
23. Kalton, G. (1983). “Introduction to Survey Sampling”, Sage Publications, Inc.
24. Kanji, G.K. (2006). “100 Statistical Tests”, Sage Publications, Inc.
25. Keppel, G., & Wickens, T.D. (2004). “Design and Analysis: A researchers handbook (4rd Edition)”, Upper Saddle River, NJ: Pearson.
26. Khoshgooyanfard, A. (2008). “An Iranian Experience of Internet Survey”, Journal of Website Promotion, Vol. 3, No. 1&2.
27. Liebetrau, A. M. (1983). “Measures of Association”, Newbury Park, CA: Sage Publications. Quantitative Applications in the Social Sciences Series No. 32.
28. Lohr, S.L. (2010). “Sampling: Design and Analysis”, Brooks/Cole, Cengage Learning, Inc.
29. Montgomery, D.C. (2008). “Design and analysis of Experiments”, John Wiley & Sons, Inc.
30. Morgan, G.A. and Gliner, J.A. (2009). “Research Methods in Applied Settings: An Integrated Approach to Design and Analysis”, Lawrence Erlbaum Associates, Inc.
31. Neter, J., Wasserman, W., and Kutner, M.H. (1990). “Applied Linear Statistical Models”. Richard D. Irwin, Inc.
32. NIST/SEMATECH e-Handbook of Statistical Methods, http://www.itl.nist.gov/div898/handbook.
33. Rawlings, J.O., Pantula, S.G., and Dickey, D.A. (1998). “Applied regression analysis: a research tool”, Springer.
34. Särndal, C.E., Swenson, B. and Wreman, J.H. (2003). “Model Assisted Survey Sampling”, Springer-Verlag, New York.
35. Schonlau, M., Fricker, R.D. and Elliott, M.N. (2002). “Conducting Research Surveys via E-mail and the Web”, RAND Monograph/Report Series, MR-1480-RC.
36. Siegel, S. and Castellan Jr., N.J. (1988). “Nonparametric Statistics for the Behavioral Sciences, Second Edition”. McGraw-Hill.
37. Snedecore, G.W. and Cochran, W.G. (1995). “Statistical Methods”. Iowa State University Press, Ames, Iowa.
38. Tabachnick, B. G. and Fidell, L. S. (2007). “Using Multivariate Statistics”. Needham Heights, MA: Allyn and Bacon
39. Tukey, J. W. (1977). “Exploratory Data Analysis”. Addison-Wesley, Reading, MA.
40. Velleman, P.F. and Hoaglin, D.C. (2004). “Applications, Basics and Computing of Explanatory Data Analysis”, Duxbury Press, Boston.

پیوست ۲: جدول‌های آماری

جدول ۱: توزیع نرمال

جدول ۲: توزیع t

جدول ۳: توزیع خی‌دو

جدول ۴: نقاط بحرانی آزمون توکی

جدول ۵: نقاط بحرانی آزمون دانت

جدول ۶: نقاط بحرانی آزمون علامت

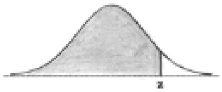
جدول ۷: نقاط بحرانی آزمون رتبه علامت‌دار ویلکاکسون

جدول ۸: نقاط بحرانی آزمون مجموع رتبه علامت‌دار ویلکاکسون

جدول ۹: نقاط بحرانی آزمون کروسکال-والیس

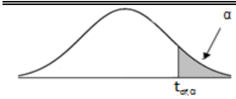
جدول ۱۰: نقاط بحرانی آزمون ضریب اسپیرمن

جدول ۱۱: نقاط بحرانی آزمون ضریب تاو کندال



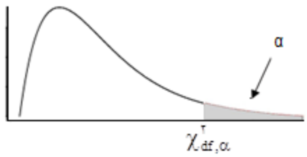
جدول۱: سطح زیر منحنی توزیع نرمال استاندارد تا مقدار z

Z	۰/۰۰	۰/۰۱	۰/۰۲	۰/۰۳	۰/۰۴	۰/۰۵	۰/۰۶	۰/۰۷	۰/۰۸	۰/۰۹
−۳/۴	۰/۰۰۰۳	۰/۰۰۰۳	۰/۰۰۰۳	۰/۰۰۰۳	۰/۰۰۰۳	۰/۰۰۰۳	۰/۰۰۰۳	۰/۰۰۰۳	۰/۰۰۰۳	۰/۰۰۰۲
−۳/۳	۰/۰۰۰۵	۰/۰۰۰۵	۰/۰۰۰۵	۰/۰۰۰۴	۰/۰۰۰۴	۰/۰۰۰۴	۰/۰۰۰۴	۰/۰۰۰۴	۰/۰۰۰۴	۰/۰۰۰۳
−۳/۲	۰/۰۰۰۷	۰/۰۰۰۷	۰/۰۰۰۶	۰/۰۰۰۶	۰/۰۰۰۶	۰/۰۰۰۶	۰/۰۰۰۶	۰/۰۰۰۵	۰/۰۰۰۵	۰/۰۰۰۵
−۳/۱	۰/۰۰۱۰	۰/۰۰۰۹	۰/۰۰۰۹	۰/۰۰۰۹	۰/۰۰۰۸	۰/۰۰۰۸	۰/۰۰۰۸	۰/۰۰۰۸	۰/۰۰۰۷	۰/۰۰۰۷
−۳/۰	۰/۰۰۱۳	۰/۰۰۱۳	۰/۰۰۱۳	۰/۰۰۱۲	۰/۰۰۱۲	۰/۰۰۱۱	۰/۰۰۱۱	۰/۰۰۱۱	۰/۰۰۱۰	۰/۰۰۱۰
−۲/۹	۰/۰۰۱۹	۰/۰۰۱۸	۰/۰۰۱۸	۰/۰۰۱۷	۰/۰۰۱۶	۰/۰۰۱۶	۰/۰۰۱۵	۰/۰۰۱۵	۰/۰۰۱۴	۰/۰۰۱۴
−۲/۸	۰/۰۰۲۶	۰/۰۰۲۵	۰/۰۰۲۴	۰/۰۰۲۳	۰/۰۰۲۳	۰/۰۰۲۲	۰/۰۰۲۱	۰/۰۰۲۱	۰/۰۰۲۰	۰/۰۰۱۹
−۲/۷	۰/۰۰۳۵	۰/۰۰۳۴	۰/۰۰۳۳	۰/۰۰۳۲	۰/۰۰۳۱	۰/۰۰۳۰	۰/۰۰۲۹	۰/۰۰۲۸	۰/۰۰۲۷	۰/۰۰۲۶
−۲/۶	۰/۰۰۴۷	۰/۰۰۴۵	۰/۰۰۴۴	۰/۰۰۴۳	۰/۰۰۴۱	۰/۰۰۴۰	۰/۰۰۳۹	۰/۰۰۳۸	۰/۰۰۳۷	۰/۰۰۳۶
−۲/۵	۰/۰۰۶۲	۰/۰۰۶۰	۰/۰۰۵۹	۰/۰۰۵۷	۰/۰۰۵۵	۰/۰۰۵۴	۰/۰۰۵۲	۰/۰۰۵۱	۰/۰۰۴۹	۰/۰۰۴۸
−۲/۴	۰/۰۰۸۲	۰/۰۰۸۰	۰/۰۰۷۸	۰/۰۰۷۵	۰/۰۰۷۳	۰/۰۰۷۱	۰/۰۰۶۹	۰/۰۰۶۸	۰/۰۰۶۶	۰/۰۰۶۴
−۲/۳	۰/۰۱۰۷	۰/۰۱۰۴	۰/۰۱۰۲	۰/۰۰۹۹	۰/۰۰۹۶	۰/۰۰۹۴	۰/۰۰۹۱	۰/۰۰۸۹	۰/۰۰۸۷	۰/۰۰۸۴
−۲/۲	۰/۰۱۳۹	۰/۰۱۳۶	۰/۰۱۳۲	۰/۰۱۲۹	۰/۰۱۲۵	۰/۰۱۲۲	۰/۰۱۱۹	۰/۰۱۱۶	۰/۰۱۱۳	۰/۰۱۱۰
−۲/۱	۰/۰۱۷۹	۰/۰۱۷۴	۰/۰۱۷۰	۰/۰۱۶۶	۰/۰۱۶۲	۰/۰۱۵۸	۰/۰۱۵۴	۰/۰۱۵۰	۰/۰۱۴۶	۰/۰۱۴۳
−۲/۰	۰/۰۲۲۸	۰/۰۲۲۲	۰/۰۲۱۷	۰/۰۲۱۲	۰/۰۲۰۷	۰/۰۲۰۲	۰/۰۱۹۷	۰/۰۱۹۲	۰/۰۱۸۸	۰/۰۱۸۳
−۱/۹	۰/۰۲۸۷	۰/۰۲۸۱	۰/۰۲۷۴	۰/۰۲۶۸	۰/۰۲۶۲	۰/۰۲۵۶	۰/۰۲۵۰	۰/۰۲۴۴	۰/۰۲۳۹	۰/۰۲۳۳
−۱/۸	۰/۰۳۵۹	۰/۰۳۵۱	۰/۰۳۴۴	۰/۰۳۳۶	۰/۰۳۲۹	۰/۰۳۲۲	۰/۰۳۱۴	۰/۰۳۰۷	۰/۰۳۰۱	۰/۰۲۹۴
−۱/۷	۰/۰۴۴۶	۰/۰۴۳۶	۰/۰۴۲۷	۰/۰۴۱۸	۰/۰۴۰۹	۰/۰۴۰۱	۰/۰۳۹۲	۰/۰۳۸۴	۰/۰۳۷۵	۰/۰۳۶۷
−۱/۶	۰/۰۵۴۸	۰/۰۵۳۷	۰/۰۵۲۶	۰/۰۵۱۶	۰/۰۵۰۵	۰/۰۴۹۵	۰/۰۴۸۵	۰/۰۴۷۵	۰/۰۴۶۵	۰/۰۴۵۵
−۱/۵	۰/۰۶۶۸	۰/۰۶۵۵	۰/۰۶۴۳	۰/۰۶۳۰	۰/۰۶۱۸	۰/۰۶۰۶	۰/۰۵۹۴	۰/۰۵۸۲	۰/۰۵۷۱	۰/۰۵۵۹
−۱/۴	۰/۰۸۰۸	۰/۰۷۹۳	۰/۰۷۷۸	۰/۰۷۶۴	۰/۰۷۴۹	۰/۰۷۳۵	۰/۰۷۲۱	۰/۰۷۰۸	۰/۰۶۹۴	۰/۰۶۸۱
−۱/۳	۰/۰۹۶۸	۰/۰۹۵۱	۰/۰۹۳۴	۰/۰۹۱۸	۰/۰۹۰۱	۰/۰۸۸۵	۰/۰۸۶۹	۰/۰۸۵۳	۰/۰۸۳۸	۰/۰۸۲۳
−۱/۲	۰/۱۱۵۱	۰/۱۱۳۱	۰/۱۱۱۲	۰/۱۰۹۳	۰/۱۰۷۵	۰/۱۰۵۶	۰/۱۰۳۸	۰/۱۰۲۰	۰/۱۰۰۳	۰/۰۹۸۵
−۱/۱	۰/۱۳۵۷	۰/۱۳۳۵	۰/۱۳۱۴	۰/۱۲۹۲	۰/۱۲۷۱	۰/۱۲۵۱	۰/۱۲۳۰	۰/۱۲۱۰	۰/۱۱۹۰	۰/۱۱۷۰
−۱/۰	۰/۱۵۸۷	۰/۱۵۶۲	۰/۱۵۳۹	۰/۱۵۱۵	۰/۱۴۹۲	۰/۱۴۶۹	۰/۱۴۴۶	۰/۱۴۲۳	۰/۱۴۰۱	۰/۱۳۷۹
−۰/۹	۰/۱۸۴۱	۰/۱۸۱۴	۰/۱۷۸۸	۰/۱۷۶۲	۰/۱۷۳۶	۰/۱۷۱۱	۰/۱۶۸۵	۰/۱۶۶۰	۰/۱۶۳۵	۰/۱۶۱۱
−۰/۸	۰/۲۱۱۹	۰/۲۰۹۰	۰/۲۰۶۱	۰/۲۰۳۳	۰/۲۰۰۵	۰/۱۹۷۷	۰/۱۹۴۹	۰/۱۹۲۲	۰/۱۸۹۴	۰/۱۸۶۷
−۰/۷	۰/۲۴۲۰	۰/۲۳۸۹	۰/۲۳۵۸	۰/۲۳۲۷	۰/۲۲۹۶	۰/۲۲۶۶	۰/۲۲۳۶	۰/۲۲۰۶	۰/۲۱۷۷	۰/۲۱۴۸
−۰/۶	۰/۲۷۴۳	۰/۲۷۰۹	۰/۲۶۷۶	۰/۲۶۴۳	۰/۲۶۱۱	۰/۲۵۷۸	۰/۲۵۴۶	۰/۲۵۱۴	۰/۲۴۸۳	۰/۲۴۵۱
−۰/۵	۰/۳۰۸۵	۰/۳۰۵۰	۰/۳۰۱۵	۰/۲۹۸۱	۰/۲۹۴۶	۰/۲۹۱۲	۰/۲۸۷۷	۰/۲۸۴۳	۰/۲۸۱۰	۰/۲۷۷۶
−۰/۴	۰/۳۴۴۶	۰/۳۴۰۹	۰/۳۳۷۲	۰/۳۳۳۶	۰/۳۳۰۰	۰/۳۲۶۴	۰/۳۲۲۸	۰/۳۱۹۲	۰/۳۱۵۶	۰/۳۱۲۱
−۰/۳	۰/۳۸۲۱	۰/۳۷۸۳	۰/۳۷۴۵	۰/۳۷۰۷	۰/۳۶۶۹	۰/۳۶۳۲	۰/۳۵۹۴	۰/۳۵۵۷	۰/۳۵۲۰	۰/۳۴۸۳
−۰/۲	۰/۴۲۰۷	۰/۴۱۶۸	۰/۴۱۲۹	۰/۴۰۹۰	۰/۴۰۵۲	۰/۴۰۱۳	۰/۳۹۷۴	۰/۳۹۳۶	۰/۳۸۹۷	۰/۳۸۵۹
−۰/۱	۰/۴۶۰۲	۰/۴۵۶۲	۰/۴۵۲۲	۰/۴۴۸۳	۰/۴۴۴۳	۰/۴۴۰۴	۰/۴۳۶۴	۰/۴۳۲۵	۰/۴۲۸۶	۰/۴۲۴۷
−۰/۰	۰/۵۰۰۰	۰/۴۹۶۰	۰/۴۹۲۰	۰/۴۸۸۰	۰/۴۸۴۰	۰/۴۸۰۱	۰/۴۷۶۱	۰/۴۷۲۱	۰/۴۶۸۱	۰/۴۶۴۱



جدول ۲: سطح زیر منحنی توزیع t تا نقطهٔ $t_{df,\alpha}$

df \ α	۰/۱۰	۰/۰۵	۰/۰۲۵	۰/۰۱	۰/۰۰۵	۰/۰۰۱	۰/۰۰۰۵
۱	۳/۰۷۸	۶/۳۱۴	۱۲/۷۱	۳۱/۸۲	۶۳/۶۶	۳۱۸/۳۱	۶۳۶/۶۲
۲	۱/۸۸۶	۲/۹۲۰	۴/۳۰۳	۶/۹۶۵	۹/۹۲۵	۲۲/۳۲۷	۳۱/۵۹۹
۳	۱/۶۳۸	۲/۳۵۶	۳/۱۸۲	۴٫۵۴۱	۵/۸۴۱	۱۰/۲۱۵	۱۲/۹۲۴
۴	۱/۵۳۳	۲/۱۳۲	۲/۷۷۶	۳/۷۴۷	۴/۶۰۴	۷/۱۷۳	۸/۶۱۰
۵	۱/۴۷۶	۲/۰۱۵	۲/۵۷۱	۳/۳۶۵	۴/۰۳۲	۵/۸۹۳	۶/۸۶۹
۶	۱/۴۴۰	۱/۹۴۳	۲/۴۴۷	۳/۱۴۳	۳/۷۰۷	۵/۲۰۸	۵/۹۵۹
۷	۱/۴۱۵	۱/۸۹۵	۲/۳۶۵	۲/۹۹۸	۳/۴۹۹	۴/۷۸۵	۵/۴۰۸
۸	۱/۳۹۷	۱/۸۶۰	۲/۳۰۶	۲/۸۹۶	۳/۳۵۵	۴/۵۰۱	۵/۰۴۱
۹	۱/۳۸۳	۱/۸۳۳	۲/۲۶۲	۲/۸۲۱	۳/۲۵۰	۴/۲۹۷	۴/۷۸۱
۱۰	۱/۳۷۲	۱/۸۱۲	۲/۲۲۸	۲/۷۶۴	۳/۱۶۹	۴/۱۴۴	۴/۵۸۷
۱۱	۱/۳۶۳	۱/۷۹۶	۲/۲۰۱	۲/۷۱۸	۳/۱۰۶	۴/۰۲۵	۴/۴۳۷
۱۲	۱/۳۵۶	۱/۷۸۲	۲/۱۷۹	۲/۶۸۱	۳/۰۵۵	۳/۹۳۰	۴/۳۱۸
۱۳	۱/۳۵۰	۱/۷۷۱	۲/۱۶۰	۲/۶۵۰	۳/۰۱۲	۳/۸۵۲	۴/۲۲۱
۱۴	۱/۳۴۵	۱/۷۶۱	۲/۱۴۵	۲/۶۲۴	۲/۹۷۷	۳/۷۸۷	۴/۱۴۰
۱۵	۱/۳۴۱	۱/۷۵۳	۲/۱۳۱	۲/۶۰۲	۲/۹۴۷	۳/۷۳۳	۴/۰۷۳
۱۶	۱/۳۳۷	۱/۷۴۶	۲/۱۲۰	۲/۵۸۳	۲/۹۲۱	۳/۶۸۶	۴/۰۱۵
۱۷	۱/۳۳۳	۱/۷۴۰	۲/۱۱۰	۲/۵۶۷	۲/۸۹۸	۳/۶۴۶	۳/۹۶۵
۱۸	۱/۳۳۰	۱/۷۳۴	۲/۱۰۱	۲/۵۵۲	۲/۸۷۸	۳/۶۱۰	۳/۹۲۲
۱۹	۱/۳۲۸	۱/۷۲۹	۲/۰۹۳	۲/۵۳۹	۲/۸۶۱	۳/۵۷۹	۳/۸۸۳
۲۰	۱/۳۲۵	۱/۷۲۵	۲/۰۸۶	۲/۵۲۸	۲/۸۴۵	۳/۵۵۲	۳/۸۵۰
۲۱	۱/۳۲۳	۱/۷۲۱	۲/۰۸۰	۲/۵۱۸	۲/۸۳۱	۳/۵۲۷	۳/۸۱۹
۲۲	۱/۳۲۱	۱/۷۱۷	۲/۰۷۴	۲/۵۰۸	۲/۸۱۹	۳/۵۰۵	۳/۷۹۲
۲۳	۱/۳۱۹	۱/۷۱۴	۲/۰۶۹	۲/۵۰۰	۲/۸۰۷	۳/۴۸۵	۳/۷۶۸
۲۴	۱/۳۱۸	۱/۷۱۱	۲/۰۶۴	۲/۴۹۲	۲/۷۹۷	۳/۴۶۷	۳/۷۴۵
۲۵	۱/۳۱۶	۱/۷۰۸	۲/۰۶۰	۲/۴۸۵	۲/۷۸۷	۳/۴۵۰	۳/۷۲۵
۲۶	۱/۳۱۵	۱/۷۰۶	۲/۰۵۶	۲/۴۷۹	۲/۷۷۹	۳/۴۳۵	۳/۷۰۷
۲۷	۱/۳۱۴	۱/۷۰۳	۲/۰۵۲	۲/۴۷۳	۲/۷۷۱	۳/۴۲۱	۳/۶۹۰
۲۸	۱/۳۱۳	۱/۷۰۱	۲/۰۴۸	۲/۴۶۷	۲/۷۶۳	۳/۴۰۸	۳/۶۷۴
۲۹	۱/۳۱۱	۱/۶۹۹	۲/۰۴۵	۲/۴۶۲	۲/۷۵۶	۳/۳۹۶	۳/۶۵۹
۳۰	۱/۳۱۰	۱/۶۹۷	۲/۰۴۲	۲/۴۵۷	۲/۷۵۰	۳/۳۸۵	۳/۶۴۶
۴۰	۱/۳۰۳	۱/۶۸۴	۲/۰۲۱	۲/۴۲۳	۲/۷۰۴	۳/۳۰۷	۳/۵۵۱
۶۰	۱/۲۹۶	۱/۶۷۱	۲/۰۰۰	۲/۳۹۰	۲/۶۶۰	۳/۲۳۲	۳/۴۶۰
۸۰	۱/۲۹۲	۱/۶۶۴	۱/۹۹۰	۲/۳۷۴	۲/۶۳۹	۳/۱۹۵	۳/۴۱۶
۱۰۰	۱/۲۹۰	۱/۶۶۰	۱/۹۸۴	۲/۳۶۴	۲/۶۲۶	۳/۱۷۴	۳/۳۹۰
۱۰۰۰	۱/۲۸۲	۱/۶۴۶	۱/۹۶۲	۲/۳۳۰	۲/۵۸۱	۳/۰۹۸	۳/۳۰۰
∞	۱/۲۸۲	۱/۶۴۵	۱/۹۶۰	۲/۳۲۶	۲/۵۷۶	۳/۰۹۰	۳/۲۹۱



جدول ۳: سطح زیر منحنی توزیع χ^2 از نقطه $\chi^2_{df,\alpha}$ به بعد

df \ α	۰/۹۹۵	۰/۹۹۰	۰/۹۷۵	۰/۹۵۰	۰/۹۰۰	۰/۱۰۰	۰/۰۵۰	۰/۰۲۵	۰/۰۱۰	۰/۰۰۵
۱	۰/۰۰۰	۰/۰۰۰	۰/۰۰۱	۰/۰۰۴	۰/۰۱۶	۲/۷۰۶	۳/۸۴۱	۵/۰۲۴	۶/۶۳۵	۷/۸۷۹
۲	۰/۰۱۰	۰/۰۲۰	۰/۰۵۱	۰/۱۰۳	۰/۲۱۱	۴/۶۰۵	۵/۹۹۱	۳/۳۷۸	۹/۲۱۰	۱۰/۵۹۷
۳	۰/۰۷۲	۰/۱۱۵	۰/۲۱۶	۰/۳۵۲	۰/۵۸۴	۶/۲۵۱	۷/۸۱۵	۹/۳۴۸	۱۱/۳۴۵	۱۱/۸۳۸
۴	۰/۲۰۷	۰/۲۹۷	۰/۴۸۴	۰/۷۱۱	۱/۰۶۴	۷/۷۷۹	۹/۴۸۸	۱۱/۱۴۳	۱۳/۲۷۷	۱۴/۸۶۰
۵	۰/۴۱۲	۰/۵۵۴	۰/۸۳۱	۱/۱۴۵	۱/۶۱۰	۹/۲۳۶	۱۱/۰۷۰	۱۲/۸۳۳	۱۵/۰۸۶	۱۶/۷۵۰
۶	۰/۶۷۶	۰/۸۷۲	۱/۲۳۷	۱/۶۳۵	۲/۲۰۴	۱۰/۶۴۵	۱۲/۵۹۲	۱۴/۴۴۹	۱۶/۸۱۲	۱۸/۵۴۸
۷	۰/۹۸۹	۱/۲۳۹	۱/۶۹۰	۲/۱۶۷	۲/۸۳۳	۱۲/۰۱۷	۱۴/۰۶۷	۱۶/۰۱۳	۱۸/۴۷۵	۲۰/۲۷۸
۸	۱/۳۴۴	۱/۶۴۶	۲/۱۸۰	۲/۷۳۳	۳/۴۹۰	۱۳/۳۶۲	۱۵/۵۰۷	۱۷/۵۳۵	۲۰/۰۹۰	۲۱/۹۵۵
۹	۱/۷۳۵	۲/۰۸۸	۲/۷۰۰	۳/۳۲۵	۴/۱۶۸	۱۴/۶۸۶	۱۶/۹۱۹	۱۹/۰۲۳	۲۱/۶۶۶	۲۳/۵۸۹
۱۰	۲/۱۵۶	۲/۵۵۸	۳/۲۴۷	۳/۹۴۰	۴/۸۶۵	۱۵/۹۸۷	۱۸/۳۰۷	۲۰/۴۸۳	۲۳/۲۰۹	۲۵/۱۸۸
۱۱	۲/۶۰۳	۳/۰۵۳	۳/۸۱۶	۴/۵۷۵	۵/۵۷۸	۱۷/۲۷۵	۱۹/۶۷۵	۲۱/۹۲۰	۲۴/۷۲۵	۲۶/۷۵۷
۱۲	۳/۰۷۴	۳/۵۷۱	۴/۴۰۴	۵/۲۶۶	۶/۳۰۴	۱۸/۵۴۹	۲۱/۰۲۶	۲۳/۳۳۷	۲۶/۲۱۷	۲۸/۳۰۰
۱۳	۳/۵۶۵	۴/۱۰۷	۵/۰۰۹	۵/۸۹۲	۷/۰۴۲	۱۹/۸۱۲	۲۲/۳۶۲	۲۴/۷۳۶	۲۷/۶۸۸	۲۹/۸۱۹
۱۴	۴/۰۷۵	۴/۶۶۰	۵/۶۲۹	۶/۵۷۱	۷/۷۹۰	۲۱/۰۶۴	۲۳/۶۸۵	۲۶/۱۱۹	۲۹/۱۴۱	۳۱/۳۱۹
۱۵	۴/۶۰۱	۵/۲۲۹	۶/۲۶۲	۷/۲۶۱	۸/۵۴۷	۲۲/۳۰۷	۲۴/۹۹۶	۲۷/۴۸۸	۳۰/۵۷۸	۳۲/۸۰۱
۱۶	۵/۱۴۲	۵/۸۱۲	۶/۹۰۸	۷/۹۶۲	۹/۳۱۲	۲۳/۵۴۲	۲۶/۲۹۶	۲۸/۸۴۵	۳۲/۰۰۰	۳۴/۲۶۷
۱۷	۵/۶۹۷	۶/۴۰۸	۷/۵۶۴	۸/۶۷۲	۱۰/۰۸۵	۲۴/۷۶۹	۲۷/۵۸۷	۳۰/۱۹۱	۳۳/۴۰۹	۳۵/۷۱۸
۱۸	۶/۲۶۵	۷/۰۱۵	۸/۲۳۱	۹/۳۹۰	۱۰/۸۶۵	۲۵/۹۸۹	۲۸/۸۶۹	۳۱/۵۲۶	۳۴/۸۰۵	۳۷/۱۵۶
۱۹	۶/۸۴۴	۷/۶۳۳	۸/۹۰۷	۱۰/۱۱۷	۱۱/۶۵۱	۲۷/۲۰۴	۳۰/۱۴۴	۳۲/۸۵۲	۳۶/۱۹۱	۳۸/۵۸۲
۲۰	۷/۴۳۴	۸/۲۶۰	۹/۵۹۱	۱۰/۸۵۱	۱۲/۴۴۳	۲۸/۴۱۲	۳۱/۴۱۰	۳۴/۱۷۰	۳۷/۵۶۶	۳۹/۹۹۷
۲۱	۸/۰۳۴	۸/۸۹۷	۱۰/۲۸۳	۱۱/۵۹۱	۱۳/۲۴۰	۲۹/۶۱۵	۳۲/۶۷۱	۳۵/۴۷۹	۳۸/۹۳۲	۴۱/۴۰۱
۲۲	۸/۶۴۳	۹/۵۴۲	۱۰/۹۸۲	۱۲/۳۳۸	۱۴/۰۴۱	۳۰/۸۱۳	۳۳/۹۲۴	۳۶/۷۸۱	۴۰/۲۸۹	۴۲/۷۹۶
۲۳	۹/۲۶۰	۱۰/۱۹۶	۱۱/۶۸۹	۱۳/۰۹۱	۱۴/۸۴۸	۳۲/۰۰۷	۳۵/۱۷۲	۳۸/۰۷۶	۴۱/۶۳۸	۴۴/۱۸۱
۲۴	۹/۸۸۶	۱۰/۸۵۶	۱۲/۴۰۱	۱۳/۸۴۸	۱۵/۶۵۹	۳۳/۱۹۶	۳۶/۴۱۵	۳۹/۳۶۴	۴۲/۹۸۰	۴۵/۵۵۹
۲۵	۱۰/۵۲۰	۱۱/۵۲۴	۱۳/۱۲۰	۱۴/۶۱۱	۱۶/۴۷۳	۳۴/۳۸۲	۳۷/۶۵۲	۴۰/۶۴۶	۴۴/۳۱۴	۴۶/۹۲۸
۲۶	۱۱/۱۶۰	۱۲/۱۹۸	۱۳/۸۴۴	۱۵/۳۷۹	۱۷/۲۹۲	۳۵/۵۶۳	۳۸/۸۸۵	۴۱/۹۲۳	۴۵/۶۴۲	۴۸/۲۹۰
۲۷	۱۱/۸۰۸	۱۲/۸۷۹	۱۴/۵۷۳	۱۶/۱۵۱	۱۸/۱۱۴	۳۶/۷۴۱	۴۰/۱۱۳	۴۳/۱۹۵	۴۶/۹۶۳	۴۹/۶۴۵
۲۸	۱۲/۴۶۱	۱۳/۵۶۵	۱۵/۳۰۸	۱۶/۹۲۸	۱۸/۹۳۹	۳۷/۹۱۶	۴۱/۳۳۷	۴۴/۴۶۱	۴۸/۲۷۸	۵۰/۹۹۳
۲۹	۱۳/۱۲۱	۱۴/۲۵۶	۱۶/۰۴۷	۱۷/۷۰۸	۱۹/۷۶۸	۳۹/۰۸۷	۴۲/۵۵۷	۴۵/۷۲۲	۴۹/۵۸۸	۵۲/۳۳۶
۳۰	۱۳/۷۸۷	۱۴/۹۵۳	۱۶/۷۹۱	۱۸/۴۹۳	۲۰/۵۹۹	۴۰/۲۵۶	۴۳/۷۷۳	۴۶/۹۷۹	۵۰/۸۹۲	۵۳/۶۷۲
۴۰	۲۰/۷۰۷	۲۲/۱۶۴	۲۴/۴۳۳	۲۶/۵۰۹	۲۹/۰۵۱	۵۱/۸۰۵	۵۵/۷۵۸	۵۹/۳۴۲	۶۳/۶۹۱	۶۶/۷۶۶
۵۰	۲۷/۹۹۱	۲۹/۷۰۷	۳۲/۳۵۷	۳۴/۷۶۴	۳۷/۶۸۹	۶۳/۱۶۷	۶۷/۵۰۵	۷۱/۴۲۰	۷۶/۱۵۴	۷۹/۴۹۰
۶۰	۳۵/۵۳۴	۳۷/۴۸۵	۴۰/۴۸۲	۴۳/۱۸۸	۴۶/۴۵۹	۷۴/۳۹۷	۷۹/۰۸۲	۸۳/۲۹۸	۸۸/۳۷۹	۹۱/۹۵۲
۷۰	۴۳/۲۷۵	۴۵/۴۴۲	۴۸/۷۵۸	۵۱/۷۳۹	۵۵/۳۲۹	۸۵/۵۲۷	۹۰/۵۳۱	۹۵/۰۲۳	۱۰۰/۴۲۵	۱۰۴/۲۱۵
۸۰	۵۱/۱۷۲	۵۳/۵۴۰	۵۷/۱۵۳	۶۰/۳۹۱	۶۴/۲۷۸	۹۶/۵۷۸	۱۰۱/۸۷۹	۱۰۶/۶۲۹	۱۱۲/۳۲۹	۱۱۶/۳۲۱
۹۰	۵۹/۱۹۶	۶۱/۷۵۴	۶۵/۶۴۷	۶۹/۱۲۶	۷۳/۲۹۱	۱۰۷/۵۶۵	۱۱۳/۱۴۵	۱۱۸/۱۳۶	۱۲۴/۱۱۶	۱۲۸/۲۹۹
۱۰۰	۶۷/۳۲۸	۷۰/۰۶۵	۷۴/۲۲۲	۷۷/۹۲۹	۸۲/۳۵۸	۱۱۸/۴۹۸	۱۲۴/۳۴۲	۱۲۹/۵۶۱	۱۳۵/۸۰۷	۱۴۰/۱۶۹

جدول ۴: چندک‌های توزیع F

$\alpha =$ ۰/۱

df_2 \ df_1	۱	۲	۳	۴	۵	۶	۷	۸	۹
۱	۳۹/۸۶	۴۹/۵	۵۳/۵۹	۵۵/۸۳	۵۷/۲۴	۵۸/۲	۵۸/۹۱	۵۹/۴۴	۵۹/۸۶
۲	۸/۵۳	۹/۰۰	۹/۱۶	۹/۲۴	۹/۲۹	۹/۳۳	۹/۳۵	۹/۳۷	۹/۳۸
۳	۵/۵۴	۵/۴۶	۵/۳۹	۵/۳۴	۵/۳۱	۵/۲۸	۵/۲۷	۵/۲۵	۵/۲۴
۴	۴/۵۴	۴/۳۲	۴/۱۹	۴/۱۱	۴/۰۵	۴/۰۱	۳/۹۸	۳/۹۵	۳/۹۴
۵	۴/۰۶	۳/۷۸	۳/۶۲	۳/۵۲	۳/۴۵	۳/۴۰	۳/۳۷	۳/۳۴	۳/۳۲
۶	۳/۷۸	۳/۴۶	۳/۲۹	۳/۱۸	۳/۱۱	۳/۰۵	۳/۰۱	۲/۹۸	۲/۹۶
۷	۳/۵۹	۳/۲۶	۳/۰۷	۲/۹۶	۲/۸۸	۲/۸۳	۲/۷۸	۲/۷۵	۲/۷۲
۸	۳/۴۶	۳/۱۱	۲/۹۲	۲/۸۱	۲/۷۳	۲/۶۷	۲/۶۲	۲/۵۹	۲/۵۶
۹	۳/۳۶	۳/۰۱	۲/۸۱	۲/۶۹	۲/۶۱	۲/۵۵	۲/۵۱	۲/۴۷	۲/۴۴
۱۰	۳/۲۹	۲/۹۲	۲/۷۳	۲/۶۱	۲/۵۲	۲/۴۶	۲/۴۱	۲/۳۸	۲/۳۵
۱۱	۳/۲۳	۲/۸۶	۲/۶۶	۲/۵۴	۲/۴۵	۲/۳۹	۲/۳۴	۲/۳	۲/۲۷
۱۲	۳/۱۸	۲/۸۱	۲/۶۱	۲/۴۸	۲/۳۹	۲/۳۳	۲/۲۸	۲/۲۴	۲/۲۱
۱۳	۳/۱۴	۲/۷۶	۲/۵۶	۲/۴۳	۲/۳۵	۲/۲۸	۲/۲۳	۲/۲۰	۲/۱۶
۱۴	۳/۱۰	۲/۷۳	۲/۵۲	۲/۳۹	۲/۳۱	۲/۲۴	۲/۱۹	۲/۱۵	۲/۱۲
۱۵	۳/۰۷	۲/۷۰	۲/۴۹	۲/۳۶	۲/۲۷	۲/۲۱	۲/۱۶	۲/۱۲	۲/۰۹
۱۶	۳/۰۵	۲/۶۷	۲/۴۶	۲/۳۳	۲/۲۴	۲/۱۸	۲/۱۳	۲/۰۹	۲/۰۶
۱۷	۳/۰۳	۲/۶۴	۲/۴۴	۲/۳۱	۲/۲۲	۲/۱۵	۲/۱۰	۲/۰۶	۲/۰۳
۱۸	۳/۰۱	۲/۶۲	۲/۴۲	۲/۲۹	۲/۲۰	۲/۱۳	۲/۰۸	۲/۰۴	۲/۰۰
۱۹	۲/۹۹	۲/۶۱	۲/۴۰	۲/۲۷	۲/۱۸	۲/۱۱	۲/۰۶	۲/۰۲	۱/۹۸
۲۰	۲/۹۷	۲/۵۹	۲/۳۸	۲/۲۵	۲/۱۶	۲/۰۹	۲/۰۴	۲/۰۰	۱/۹۶
۲۱	۲/۹۶	۲/۵۷	۲/۳۶	۲/۲۳	۲/۱۴	۲/۰۸	۲/۰۲	۱/۹۸	۱/۹۵
۲۲	۲/۹۵	۲/۵۶	۲/۳۵	۲/۲۲	۲/۱۳	۲/۰۶	۲/۰۱	۱/۹۷	۱/۹۳
۲۳	۲/۹۴	۲/۵۵	۲/۳۴	۲/۲۱	۲/۱۱	۲/۰۵	۱/۹۹	۱/۹۵	۱/۹۲
۲۴	۲/۹۳	۲/۵۴	۲/۳۳	۲/۱۹	۲/۱۰	۲/۰۴	۱/۹۸	۱/۹۴	۱/۹۱
۲۵	۲/۹۲	۲/۵۳	۲/۳۲	۲/۱۸	۲/۰۹	۲/۰۲	۱/۹۷	۱/۹۳	۱/۸۹
۲۶	۲/۹۱	۲/۵۲	۲/۳۱	۲/۱۷	۲/۰۸	۲/۰۱	۱/۹۶	۱/۹۲	۱/۸۸
۲۷	۲/۹۰	۲/۵۱	۲/۳۰	۲/۱۷	۲/۰۷	۲/۰۰	۱/۹۵	۱/۹۱	۱/۸۷
۲۸	۲/۸۹	۲/۵۰	۲/۲۹	۲/۱۶	۲/۰۶	۲/۰۰	۱/۹۴	۱/۹۰	۱/۸۷
۲۹	۲/۸۹	۲/۵۰	۲/۲۸	۲/۱۵	۲/۰۶	۱/۹۹	۱/۹۳	۱/۸۹	۱/۸۶
۳۰	۲/۸۸	۲/۴۹	۲/۲۸	۲/۱۴	۲/۰۵	۱/۹۸	۱/۹۳	۱/۸۸	۱/۸۵
۴۰	۲/۸۴	۲/۴۴	۲/۲۳	۲/۰۹	۲/۰۰	۱/۹۳	۱/۸۷	۱/۸۳	۱/۷۹
۶۰	۲/۷۹	۲/۳۹	۲/۱۸	۲/۰۴	۱/۹۵	۱/۸۷	۱/۸۲	۱/۷۷	۱/۷۴
۱۲۰	۲/۷۵	۲/۳۵	۲/۱۳	۱/۹۹	۱/۹۰	۱/۸۲	۱/۷۷	۱/۷۲	۱/۶۸
∞	۲/۷۱	۲/۳۰	۲/۰۸	۱/۹۴	۱/۸۵	۱/۷۷	۱/۷۲	۱/۶۷	۱/۶۳

$\alpha = ۰/۱$ (ادامه)

df_1 \ df_2	۱۰	۱۲	۱۵	۲۰	۲۴	۳۰	۴۰	۶۰	۱۲۰	∞
۱	۶۰/۱۹	۶۰/۷۱	۶۱/۲۲	۶۱/۷۴	۶۲	۶۲/۲۶	۶۲/۵۳	۶۲/۷۹	۶۳/۰۶	۶۳/۳۳
۲	۹/۳۹	۹/۴۱	۹/۴۲	۹/۴۴	۹/۴۵	۹/۴۶	۹/۴۷	۹/۴۷	۹/۴۸	۹/۴۹
۳	۵/۲۳	۵/۲۲	۵/۲۰	۵/۱۸	۵/۱۸	۵/۱۷	۵/۱۶	۵/۱۵	۵/۱۴	۵/۱۳
۴	۳/۹۲	۳/۹۰	۳/۸۷	۳/۸۴	۳/۸۳	۳/۸۲	۳/۸۰	۳/۷۹	۳/۷۸	۳/۷۶
۵	۳/۳۰	۳/۲۷	۳/۲۴	۳/۲۱	۳/۱۹	۳/۱۷	۳/۱۶	۳/۱۴	۳/۱۲	۳/۱۰
۶	۲/۹۴	۲/۹۰	۲/۸۷	۲/۸۴	۲/۸۲	۲/۸۰	۲/۷۸	۲/۷۶	۲/۷۴	۲/۷۲
۷	۲/۷۰	۲/۶۷	۲/۶۳	۲/۵۹	۲/۵۸	۲/۵۶	۲/۵۴	۲/۵۱	۲/۴۹	۲/۴۷
۸	۲/۵۴	۲/۵۰	۲/۴۶	۲/۴۲	۲/۴۰	۲/۳۸	۲/۳۶	۲/۳۴	۲/۳۲	۲/۲۹
۹	۲/۴۲	۲/۳۸	۲/۳۴	۲/۳۰	۲/۲۸	۲/۲۵	۲/۲۳	۲/۲۱	۲/۱۸	۲/۱۶
۱۰	۲/۳۲	۲/۲۸	۲/۲۴	۲/۲۰	۲/۱۸	۲/۱۶	۲/۱۳	۲/۱۱	۲/۰۸	۲/۰۶
۱۱	۲/۲۵	۲/۲۱	۲/۱۷	۲/۱۲	۲/۱۰	۲/۰۸	۲/۰۵	۲/۰۳	۲/۰۰	۱/۹۷
۱۲	۲/۱۹	۲/۱۵	۲/۱۰	۲/۰۶	۲/۰۴	۲/۰۱	۱/۹۹	۱/۹۶	۱/۹۳	۱/۹۰
۱۳	۲/۴۰	۲/۱۰	۲/۰۵	۲/۰۱	۱/۹۸	۱/۹۶	۱/۹۳	۱/۹۰	۱/۸۸	۱/۸۵
۱۴	۲/۱۰	۲/۰۵	۲/۰۱	۱/۹۶	۱/۹۴	۱/۹۱	۱/۸۹	۱/۸۶	۱/۸۳	۱/۸۰
۱۵	۲/۰۶	۲/۰۲	۱/۹۷	۱/۹۲	۱/۹۰	۱/۸۷	۱/۸۵	۱/۸۲	۱/۷۹	۱/۷۶
۱۶	۲/۰۳	۱/۹۹	۱/۹۴	۱/۸۹	۱/۸۷	۱/۸۴	۱/۸۱	۱/۷۸	۱/۷۵	۱/۷۲
۱۷	۲/۰۰	۱/۹۶	۱/۹۱	۱/۸۶	۱/۸۴	۱/۸۱	۱/۷۸	۱/۷۵	۱/۷۲	۱/۶۹
۱۸	۱/۹۸	۱/۹۳	۱/۸۹	۱/۸۴	۱/۸۱	۱/۷۸	۱/۷۵	۱/۷۲	۱/۶۹	۱/۶۶
۱۹	۱/۹۶	۱/۹۱	۱/۸۶	۱/۸۱	۱/۷۹	۱/۷۶	۱/۷۳	۱/۷۰	۱/۶۷	۱/۶۳
۲۰	۱/۹۴	۱/۸۹	۱/۸۴	۱/۷۹	۱/۷۷	۱/۷۴	۱/۷۱	۱/۶۸	۱/۶۴	۱/۶۱
۲۱	۱/۹۲	۱/۸۷	۱/۸۳	۱/۷۸	۱/۷۵	۱/۷۲	۱/۶۹	۱/۶۶	۱/۶۲	۱/۵۹
۲۲	۱/۹۰	۱/۸۶	۱/۸۱	۱/۷۶	۱/۷۳	۱/۷۰	۱/۶۷	۱/۶۴	۱/۶۰	۱/۵۷
۲۳	۱/۸۹	۱/۸۴	۱/۸۰	۱/۷۴	۱/۷۲	۱/۶۹	۱/۶۶	۱/۶۲	۱/۵۹	۱/۵۵
۲۴	۱/۸۸	۱/۸۳	۱/۷۸	۱/۷۳	۱/۷۰	۱/۶۷	۱/۶۴	۱/۶۱	۱/۵۷	۱/۵۳
۲۵	۱/۸۷	۱/۸۲	۱/۷۷	۱/۷۲	۱/۶۹	۱/۶۶	۱/۶۳	۱/۵۹	۱/۵۶	۱/۵۲
۲۶	۱/۸۶	۱/۸۱	۱/۷۶	۱/۷۱	۱/۸۰	۱/۶۵	۱/۶۱	۱/۵۸	۱/۵۴	۱/۵۰
۲۷	۱/۸۵	۱/۸۰	۱/۷۵	۱/۷۰	۱/۶۷	۱/۶۴	۱/۶۰	۱/۵۷	۱/۵۳	۱/۴۹
۲۸	۱/۸۴	۱/۷۹	۱/۷۴	۱/۶۹	۱/۶۶	۱/۶۳	۱/۵۹	۱/۵۶	۱/۵۲	۱/۴۸
۲۹	۱/۸۳	۱/۷۸	۱/۷۳	۱/۶۸	۱/۶۵	۱/۶۲	۱/۵۸	۱/۵۵	۱/۵۱	۱/۴۷
۳۰	۱/۸۲	۱/۷۷	۱/۷۲	۱/۶۷	۱/۶۴	۱/۶۱	۱/۵۷	۱/۵۴	۱/۵۰	۱/۴۶
۴۰	۱/۷۶	۱/۷۱	۱/۶۶	۱/۶۱	۱/۵۷	۱/۵۴	۱/۵۱	۱/۴۷	۱/۴۲	۱/۳۸
۶۰	۱/۷۱	۱/۶۶	۱/۶۰	۱/۵۴	۱/۵۱	۱/۴۸	۱/۴۴	۱/۴۰	۱/۳۵	۱/۲۹
۱۲۰	۱/۶۵	۱/۶۰	۱/۵۵	۱/۴۸	۱/۴۵	۱/۴۱	۱/۳۷	۱/۳۲	۱/۲۶	۱/۱۹
∞	۱/۶۰	۱/۵۵	۱/۴۹	۱/۴۲	۱/۳۸	۱/۳۴	۱/۳۰	۱/۲۴	۱/۱۷	۱/۰۰

$\alpha =$ ۰/۰۵

df_2 \ df_1	۱	۲	۳	۴	۵	۶	۷	۸	۹
۱	۱۶۱/۴	۱۹۹/۵	۲۱۵/۷	۲۲۴/۶	۲۳۰/۲	۲۳۴/۰	۲۳۶/۸	۲۳۸/۹	۲۴۰/۵
۲	۱۸/۵۱	۱۹/۰۰	۱۹/۱۶	۱۹/۲۵	۱۹/۳	۱۹/۳۳	۱۹/۳۵	۱۹/۳۷	۱۹/۳۸
۳	۱۰/۱۳	۹/۵۵	۹/۲۸	۹/۱۲	۹/۰۱	۸/۹۴	۸/۸۹	۸/۸۵	۸/۸۱
۴	۷/۷۱	۶/۹۴	۶/۵۹	۶/۳۹	۶/۲۶	۶/۱۶	۶/۰۹	۶/۰۴	۶/۰۰
۵	۶/۶۱	۵/۷۹	۵/۴۱	۵/۱۹	۵/۰۵	۴/۹۵	۴/۸۸	۴/۸۲	۴/۷۷
۶	۵/۹۹	۵/۱۴	۴/۷۶	۴/۵۳	۴/۳۹	۴/۲۸	۴/۲۱	۴/۱۵	۴/۱۰
۷	۵/۵۹	۴/۷۴	۴/۳۵	۴/۱۲	۳/۹۷	۳/۸۷	۳/۷۹	۳/۷۳	۳/۶۸
۸	۵/۳۲	۴/۴۶	۴/۰۷	۳/۸۴	۳/۶۹	۳/۵۸	۳/۵۰	۳/۴۴	۳/۳۹
۹	۵/۱۲	۴/۲۶	۳/۸۶	۳/۶۳	۳/۴۸	۳/۳۷	۳/۲۹	۳/۲۳	۳/۱۸
۱۰	۴/۹۶	۴/۱۰	۳/۷۱	۳/۴۸	۳/۳۳	۳/۲۲	۳/۱۴	۳/۰۷	۳/۰۲
۱۱	۴/۸۴	۳/۹۸	۳/۵۹	۳/۳۶	۳/۲۰	۳/۰۹	۳/۰۱	۲/۹۵	۲/۹۰
۱۲	۴/۷۵	۳/۸۹	۳/۴۹	۳/۲۶	۳/۱۱	۳/۰۰	۲/۹۱	۲/۸۵	۲/۸۰
۱۳	۴/۶۷	۳/۸۱	۳/۴۱	۳/۱۸	۳/۰۳	۲/۹۲	۲/۸۳	۲/۷۷	۲/۷۱
۱۴	۴/۶۰	۳/۷۴	۳/۳۴	۳/۱۱	۲/۹۶	۲/۸۵	۲/۷۶	۲/۷۰	۲/۶۵
۱۵	۴/۵۴	۳/۶۸	۳/۲۹	۳/۰۶	۲/۹۰	۲/۷۹	۲/۷۱	۲/۶۴	۲/۵۹
۱۶	۴/۴۹	۳/۶۳	۳/۲۴	۳/۰۱	۲/۸۵	۲/۷۴	۲/۶۶	۲/۵۹	۲/۵۴
۱۷	۴/۴۵	۳/۵۹	۳/۲۰	۲/۹۶	۲/۸۱	۲/۷۰	۲/۶۱	۲/۵۵	۲/۴۹
۱۸	۱/۴۱	۳/۵۵	۳/۱۶	۲/۹۳	۲/۷۷	۲/۶۶	۲/۵۸	۲/۵۱	۲/۴۶
۱۹	۴/۳۸	۳/۵۲	۳/۱۳	۲/۹۰	۲/۷۴	۲/۶۳	۲/۵۴	۲/۴۸	۲/۴۲
۲۰	۴/۳۵	۳/۴۹	۳/۱۰	۲/۸۷	۲/۷۱	۲/۶۰	۲/۵۱	۲/۴۵	۲/۳۹
۲۱	۴/۳۲	۳/۴۷	۳/۰۷	۲/۸۴	۲/۶۸	۲/۵۷	۲/۴۹	۲/۴۲	۲/۳۷
۲۲	۴/۳۰	۳/۴۴	۳/۰۵	۲/۸۲	۲/۶۶	۲/۵۵	۲/۴۶	۲/۴۰	۲/۳۴
۲۳	۴/۲۸	۳/۴۲	۳/۰۳	۲/۸۰	۲/۶۴	۲/۵۳	۲/۴۴	۲/۳۷	۲/۳۲
۲۴	۴/۲۶	۳/۴۰	۳/۰۱	۲/۷۸	۲/۶۲	۲/۵۱	۲/۴۲	۲/۳۶	۲/۳۰
۲۵	۴/۲۴	۳/۳۹	۲/۹۹	۲/۷۶	۲/۶۰	۲/۴۹	۲/۴۰	۲/۳۴	۲/۲۸
۲۶	۴/۲۳	۳/۳۷	۲/۹۸	۲/۷۴	۲/۵۹	۲/۴۷	۲/۳۹	۲/۳۲	۲/۲۷
۲۷	۴/۲۱	۳/۳۵	۲/۹۶	۲/۷۳	۲/۵۷	۲/۴۶	۲/۳۷	۲/۳۱	۲/۲۵
۲۸	۴/۲۰	۳/۳۴	۲/۹۵	۲/۷۱	۲/۵۶	۲/۴۵	۲/۳۶	۲/۲۹	۲/۲۴
۲۹	۴/۱۸	۳/۳۳	۲/۹۳	۲/۷۰	۲/۵۵	۲/۴۳	۲/۳۵	۲/۲۸	۲/۲۲
۳۰	۴/۱۷	۳/۳۲	۲/۹۲	۲/۶۹	۲/۵۳	۲/۴۲	۲/۳۳	۲/۲۷	۲/۲۱
۴۰	۴/۰۸	۳/۲۳	۲/۸۴	۲/۶۱	۲/۴۵	۲/۳۴	۲/۲۵	۲/۱۸	۲/۱۲
۶۰	۴/۰۰	۳/۱۵	۲/۷۶	۲/۵۳	۲/۳۷	۲/۲۵	۲/۱۷	۲/۱۰	۲/۰۴
۱۲۰	۳/۹۲	۳/۰۷	۲/۶۸	۲/۴۵	۲/۲۹	۲/۱۷	۲/۰۹	۲/۰۲	۱/۹۶
∞	۳/۸۴	۳/۰۰	۲/۶۰	۲/۳۷	۲/۲۱	۲/۱۰	۲/۰۱	۱/۹۴	۱/۸۸

$\alpha = ۰/۰۵$ (ادامه)

df_1 / df_2	۱۰	۱۲	۱۵	۲۰	۲۴	۳۰	۴۰	۶۰	۱۲۰	∞
۱	۲۴۱/۹	۲۴۳/۹	۲۴۵/۹	۲۴۸/۰	۲۴۹/۱	۲۵۰/۱	۲۵۱/۱	۲۵۲/۲	۲۵۳/۳	۲۵۴/۳
۲	۱۹/۴	۱۹/۴۱	۱۹/۴۳	۱۹/۴۵	۱۹/۴۵	۱۹/۴۶	۱۹/۴۷	۱۹/۴۸	۱۹/۴۹	۱۹/۵
۳	۸/۷۹	۸/۷۴	۸/۷۰	۸/۶۶	۸/۶۴	۸/۶۲	۸/۵۹	۸/۵۷	۸/۵۵	۸/۵۳
۴	۵/۹۶	۵/۹۱	۵/۸۶	۵/۸۰	۵/۷۷	۵/۷۵	۵/۷۲	۵/۶۹	۵/۶۶	۵/۶۳
۵	۴/۷۴	۴/۶۸	۴/۶۲	۴/۵۶	۴/۵۳	۴/۵۰	۴/۴۶	۴/۴۳	۴/۴۰	۴/۳۶
۶	۴/۰۶	۴/۰۰	۳/۹۴	۳/۸۷	۳/۸۴	۳/۸۱	۳/۷۷	۳/۷۴	۳/۷۰	۳/۶۷
۷	۳/۶۴	۳/۵۷	۳/۵۱	۳/۴۴	۳/۴۱	۳/۳۸	۳/۳۴	۳/۳۰	۳/۲۷	۳/۲۳
۸	۳/۳۵	۳/۲۸	۳/۲۲	۳/۱۵	۳/۱۲	۳/۰۸	۳/۰۴	۳/۰۱	۲/۹۷	۲/۹۳
۹	۳/۱۴	۳/۰۷	۳/۰۱	۲/۹۴	۲/۹۰	۲/۸۶	۲/۸۳	۲/۷۹	۲/۷۵	۲/۷۱
۱۰	۲/۹۸	۲/۹۱	۲/۸۵	۲/۷۷	۲/۷۴	۲/۷۰	۲/۶۶	۲/۶۲	۲/۵۸	۲/۵۴
۱۱	۲/۸۵	۲/۷۹	۲/۷۲	۲/۶۵	۲/۶۱	۲/۵۷	۲/۵۳	۲/۴۹	۲/۴۵	۲/۴۰
۱۲	۲/۷۵	۲/۶۹	۲/۶۲	۲/۵۴	۲/۵۱	۲/۴۷	۲/۴۳	۲/۳۸	۲/۳۴	۲/۳۰
۱۳	۲/۶۷	۲/۶۰	۲/۵۳	۲/۴۶	۲/۴۲	۲/۳۸	۲/۳۴	۲/۳۰	۲/۲۵	۲/۲۱
۱۴	۲/۶۰	۲/۵۳	۲/۴۶	۲/۳۹	۲/۳۵	۲/۳۱	۲/۲۷	۲/۲۲	۲/۱۸	۲/۱۳
۱۵	۲/۵۴	۲/۴۸	۲/۴۰	۲/۳۳	۲/۲۹	۲/۲۵	۲/۲۰	۲/۱۶	۲/۱۱	۲/۰۷
۱۶	۲/۴۹	۲/۴۲	۲/۳۵	۲/۲۸	۲/۲۴	۲/۱۹	۲/۱۵	۲/۱۱	۲/۰۶	۲/۰۱
۱۷	۲/۴۵	۲/۳۸	۲/۳۱	۲/۲۳	۲/۱۹	۲/۱۵	۲/۱۰	۲/۰۶	۲/۰۱	۱/۹۶
۱۸	۲/۴۱	۲/۳۴	۲/۲۷	۲/۱۹	۲/۱۵	۲/۱۱	۲/۰۶	۲/۰۲	۱/۹۷	۱/۹۲
۱۹	۲/۳۸	۲/۳۱	۲/۲۳	۲/۱۶	۲/۱۱	۲/۰۷	۲/۰۳	۱/۹۸	۱/۹۳	۱/۸۸
۲۰	۲/۳۵	۲/۲۸	۲/۲۰	۲/۱۲	۲/۰۸	۲/۰۴	۱/۹۹	۱/۹۵	۱/۹۰	۱/۸۴
۲۱	۲/۳۲	۲/۲۵	۲/۱۸	۲/۱۰	۲/۰۵	۲/۰۱	۱/۹۶	۱/۹۲	۱/۸۷	۱/۸۱
۲۲	۲/۳۰	۲/۲۳	۲/۱۵	۲/۰۷	۲/۰۳	۱/۹۸	۱/۹۴	۱/۸۹	۱/۸۴	۱/۷۸
۲۳	۲/۲۷	۲/۲۰	۲/۱۳	۲/۰۵	۲/۰۱	۱/۹۶	۱/۹۱	۱/۸۶	۱/۸۱	۱/۷۶
۲۴	۲/۲۵	۲/۱۸	۲/۱۱	۲/۰۳	۱/۹۸	۱/۹۴	۱/۸۹	۱/۸۴	۱/۷۹	۱/۷۳
۲۵	۲/۲۴	۲/۱۶	۲/۰۹	۲/۰۱	۱/۹۶	۱/۹۲	۱/۸۷	۱/۸۲	۱/۷۷	۱/۷۱
۲۶	۲/۲۲	۲/۱۵	۲/۰۷	۱/۹۹	۱/۹۵	۱/۹۰	۱/۸۵	۱/۸۰	۱/۷۵	۱/۶۹
۲۷	۲/۲۰	۲/۱۳	۲/۰۶	۱/۹۷	۱/۹۳	۱/۸۸	۱/۸۴	۱/۷۹	۱/۷۳	۱/۶۷
۲۸	۲/۱۹	۲/۱۲	۲/۰۴	۱/۹۶	۱/۹۱	۱/۸۷	۱/۸۲	۱/۷۷	۱/۷۱	۱/۶۵
۲۹	۲/۱۸	۲/۱۰	۲/۰۳	۱/۹۴	۱/۹۰	۱/۸۵	۱/۸۱	۱/۷۵	۱/۷۰	۱/۶۴
۳۰	۲/۱۶	۲/۰۹	۲/۰۱	۱/۹۳	۱/۸۹	۱/۸۴	۱/۷۹	۱/۷۴	۱/۶۸	۱/۶۲
۴۰	۲/۰۸	۲/۰۰	۱/۹۲	۱/۸۴	۱/۷۹	۱/۷۴	۱/۶۹	۱/۶۴	۱/۵۸	۱/۵۱
۶۰	۱/۹۹	۱/۹۲	۱/۸۴	۱/۷۵	۱/۷۰	۱/۶۵	۱/۵۹	۱/۵۳	۱/۴۷	۱/۳۹
۱۲۰	۱/۹۱	۱/۸۳	۱/۷۵	۱/۶۶	۱/۱۰	۱/۵۵	۱/۵۰	۱/۴۳	۱/۳۵	۱/۲۵
∞	۱/۸۳	۱/۷۵	۱/۶۷	۱/۵۷	۱/۵۲	۱/۴۶	۱/۳۹	۱/۳۲	۱/۲۲	۱/۰۰

$$\alpha = ۰/۰۱$$

df_2 \ df_1	۱	۲	۳	۴	۵	۶	۷	۸	۹
۱	۴۰۵۲	۴۹۹۹/۵	۵۴۰۳	۵۶۲۵	۵۷۶۴	۵۸۵۹	۵۹۲۸	۵۹۸۲	۶۰۲۲
۲	۹۸/۵۰	۹۹/۰۰	۹۹/۱۷	۹۹/۲۵	۹۹/۳۰	۹۹/۳۳	۹۹/۳۶	۹۹/۳۷	۹۹/۳۹
۳	۳۴/۱۲	۳۰/۸۲	۲۹/۴۶	۲۸/۷۱	۲۸/۲۴	۲۷/۹۱	۲۷/۶۷	۲۷/۴۹	۲۷/۳۵
۴	۲۱/۲۰	۱۸/۰۰	۱۶/۶۹	۱۵/۹۸	۱۵/۵۲	۱۵/۲۱	۱۴/۹۸	۱۴/۸۰	۱۴/۶۶
۵	۱۶/۲۶	۱۳/۲۷	۱۲/۰۶	۱۱/۳۹	۱۰/۹۷	۱۰/۶۷	۱۰/۴۶	۱۰/۲۹	۱۰/۱۶
۶	۱۳/۷۵	۱۰/۹۲	۹/۷۸	۹/۱۵	۸/۷۵	۸/۴۷	۸/۲۶	۸/۱۰	۷/۹۸
۷	۱۲/۲۵	۹/۵۵	۸/۴۵	۷/۸۵	۷/۴۶	۷/۱۹	۶/۹۹	۶/۸۴	۶/۷۲
۸	۱۱/۲۶	۸/۶۵	۷/۵۹	۷/۰۱	۶/۶۳	۶/۳۷	۶/۱۸	۶/۰۳	۵/۹۱
۹	۱۰/۵۶	۸/۰۲	۶/۹۹	۶/۴۲	۶/۰۶	۵/۸۰	۵/۶۱	۵/۴۷	۵/۳۵
۱۰	۱۰/۰۴	۷/۵۶	۶/۵۵	۵/۹۹	۵/۶۴	۵/۳۹	۵/۲	۵/۰۶	۴/۹۴
۱۱	۹/۶۵	۷/۲۱	۶/۲۲	۵/۶۷	۵/۳۲	۵/۰۷	۴/۸۹	۴/۷۴	۴/۶۳
۱۲	۹/۳۳	۶/۹۳	۵/۹۵	۵/۴۱	۵/۰۶	۴/۸۲	۴/۶۴	۴/۵۰	۴/۳۹
۱۳	۹/۰۷	۶/۷۰	۵/۷۴	۵/۲۱	۴/۸۶	۴/۶۲	۴/۴۴	۴/۳۰	۴/۱۴
۱۴	۸/۸۶	۶/۵۱	۵/۵۶	۵/۰۴	۴/۶۹	۴/۴۶	۴/۲۸	۴/۱۴	۴/۰۳
۱۵	۸/۶۸	۶/۳۶	۵/۴۲	۴/۸۹	۴/۵۶	۴/۳۲	۴/۱۴	۴/۰۰	۳/۸۹
۱۶	۸/۵۳	۶/۲۳	۵/۲۹	۴/۷۷	۴/۴۴	۴/۲۰	۴/۰۳	۳/۸۹	۳/۷۸
۱۷	۸/۴۰	۶/۱۱	۵/۱۸	۴/۶۷	۴/۳۴	۴/۱۰	۳/۹۳	۳/۷۹	۳/۶۸
۱۸	۸/۲۹	۶/۰۱	۵/۰۹	۴/۵۸	۴/۲۵	۴/۰۱	۳/۸۴	۳/۷۱	۳/۶۰
۱۹	۸/۱۸	۵/۹۳	۵/۰۱	۴/۵۰	۴/۱۷	۳/۹۴	۳/۷۷	۳/۶۳	۳/۵۲
۲۰	۸/۱۰	۵/۸۵	۴/۹۴	۴/۴۳	۴/۱۰	۳/۸۷	۳/۷۰	۳/۵۶	۳/۴۶
۲۱	۸/۰۲	۵/۷۸	۴/۸۷	۴/۳۷	۴/۰۴	۳/۸۱	۳/۶۴	۳/۵۱	۳/۴۰
۲۲	۷/۹۵	۵/۷۲	۴/۸۲	۴/۳۱	۳/۹۹	۳/۷۶	۳/۵۹	۳/۴۵	۳/۳۵
۲۳	۷/۸۸	۵/۶۶	۴/۷۶	۴/۲۶	۳/۹۴	۳/۷۱	۳/۵۴	۳/۴۱	۳/۳۰
۲۴	۷/۸۲	۵/۶۱	۴/۷۲	۴/۲۲	۳/۹۰	۳/۶۷	۳/۵۰	۳/۳۶	۳/۲۶
۲۵	۷/۷۷	۵/۵۷	۴/۶۸	۴/۱۸	۳/۸۵	۳/۶۳	۳/۴۶	۳/۳۲	۳/۲۲
۲۶	۷/۷۲	۵/۵۳	۴/۶۴	۴/۱۴	۳/۸۲	۳/۵۹	۳/۴۲	۳/۲۹	۳/۱۸
۲۷	۷/۶۸	۵/۴۹	۴/۶۰	۴/۱۱	۳/۷۸	۳/۵۶	۳/۳۹	۳/۲۶	۳/۱۵
۲۸	۷/۶۴	۵/۴۵	۴/۵۷	۴/۰۷	۳/۷۵	۳/۵۳	۳/۳۶	۳/۲۳	۳/۱۲
۲۹	۷/۶۰	۵/۴۲	۴/۵۴	۴/۰۴	۳/۷۳	۳/۵۰	۳/۳۳	۳/۲۰	۳/۰۹
۳۰	۷/۵۶	۵/۳۹	۴/۵۱	۴/۰۲	۳/۷۰	۳/۴۷	۳/۳۰	۳/۱۷	۳/۰۷
۴۰	۷/۳۱	۵/۱۸	۴/۳۱	۳/۸۳	۳/۵۱	۳/۲۹	۳/۱۲	۲/۹۹	۲/۸۹
۶۰	۷/۰۸	۴/۹۸	۴/۱۳	۳/۶۵	۳/۳۴	۳/۱۲	۲/۹۵	۲/۸۲	۲/۷۲
۱۲۰	۶/۸۵	۴/۷۹	۳/۹۵	۳/۴۸	۳/۱۷	۲/۹۶	۲/۷۹	۲/۶۶	۲/۵۶
∞	۶/۶۳	۴/۶۱	۳/۷۸	۳/۳۲	۳/۰۲	۲/۸۰	۲/۶۴	۲/۵۱	۲/۴۱

$\alpha = ۰/۰۱$ (ادامه)

df_2 \ df_1	۱۰	۱۲	۱۵	۲۰	۲۴	۳۰	۴۰	۶۰	۱۲۰	∞
۱	۶۰۵۶	۶۱۰۶	۶۱۵۷	۶۲۰۹	۶۲۳۵	۶۲۶۱	۶۲۸۷	۶۳۱۳	۶۳۳۹	۶۳۶۶
۲	۹۹/۴۰	۹۹/۴۲	۹۹/۴۳	۹۹/۴۵	۹۹/۴۶	۹۹/۴۷	۹۹/۴۷	۹۹/۴۸	۹۹/۴۹	۹۹/۵۰
۳	۲۷/۲۳	۲۷/۰۵	۲۶/۸۷	۲۶/۶۹	۲۶/۶۰	۲۶/۵۰	۲۶/۴۱	۲۶/۳۲	۲۶/۲۲	۲۶/۱۳
۴	۱۴/۵۵	۱۴/۳۷	۱۴/۲۰	۱۴/۰۲	۱۳/۹۳	۱۳/۸۴	۱۳/۷۵	۱۳/۶۵	۱۳/۵۶	۱۳/۴۶
۵	۱۰/۰۵	۹/۸۹	۹/۷۲	۹/۵۵	۹/۴۷	۹/۳۸	۹/۲۹	۹/۲۰	۹/۱۱	۹/۰۲
۶	۷/۸۷	۷/۷۲	۷/۵۶	۷/۴۰	۷/۳۱	۷/۲۳	۷/۱۴	۷/۰۶	۶/۹۷	۶/۸۸
۷	۶/۶۲	۶/۴۷	۶/۳۱	۶/۱۶	۶/۰۷	۵/۹۹	۵/۹۱	۵/۸۲	۵/۷۴	۵/۶۵
۸	۵/۸۱	۵/۶۷	۵/۵۲	۵/۳۶	۵/۲۸	۵/۲۰	۵/۱۲	۵/۰۳	۴/۹۵	۴/۸۶
۹	۵/۲۶	۵/۱۱	۴/۹۶	۴/۸۱	۴/۷۳	۴/۶۵	۴/۵۷	۴/۴۸	۴/۴۰	۴/۳۱
۱۰	۴/۸۵	۴/۷۱	۴/۵۶	۴/۴۱	۴/۳۳	۴/۲۵	۴/۱۷	۴/۰۸	۴/۰۰	۳/۹۱
۱۱	۴/۵۴	۴/۴۰	۴/۲۵	۴/۱۰	۴/۰۲	۳/۹۴	۳/۸۶	۳/۷۸	۳/۶۹	۳/۶۰
۱۲	۴/۳۰	۴/۱۶	۴/۰۱	۳/۸۶	۳/۷۸	۳/۷۰	۳/۶۲	۳/۵۴	۳/۴۵	۳/۳۶
۱۳	۴/۱۰	۳/۹۶	۳/۸۲	۳/۶۶	۳/۵۹	۳/۵۱	۳/۴۳	۳/۳۴	۳/۲۵	۳/۱۷
۱۴	۳/۹۴	۳/۸۰	۳/۶۶	۳/۵۱	۳/۴۳	۳/۳۵	۳/۲۷	۳/۱۸	۳/۰۹	۳/۰۰
۱۵	۳/۸۰	۳/۶۷	۳/۵۲	۳/۳۷	۳/۲۹	۳/۲۱	۳/۱۳	۳/۰۵	۲/۹۶	۲/۸۷
۱۶	۳/۶۹	۳/۵۵	۳/۴۱	۳/۲۶	۳/۱۸	۳/۱۰	۳/۰۲	۲/۹۳	۲/۸۴	۲/۷۵
۱۷	۳/۵۹	۳/۴۶	۳/۳۱	۳/۱۶	۳/۰۸	۳/۰۰	۲/۹۲	۲/۸۳	۲/۷۵	۲/۶۵
۱۸	۳/۵۱	۳/۳۷	۳/۲۳	۳/۰۸	۳/۰۰	۲/۹۲	۲/۸۴	۲/۷۵	۲/۶۶	۲/۵۷
۱۹	۳/۴۳	۳/۳۰	۳/۱۵	۳/۰۰	۲/۹۲	۲/۸۴	۲/۷۶	۲/۶۷	۲/۵۸	۲/۴۹
۲۰	۳/۳۷	۳/۲۳	۳/۰۹	۲/۹۴	۲/۸۶	۲/۷۸	۲/۶۹	۲/۶۱	۲/۵۲	۲/۴۲
۲۱	۳/۳۱	۳/۱۷	۳/۰۳	۲/۸۸	۲/۸۰	۲/۷۲	۲/۶۴	۲/۵۵	۲/۴۶	۲/۳۶
۲۲	۳/۲۶	۳/۱۲	۲/۹۸	۲/۸۳	۲/۷۵	۲/۶۷	۲/۵۸	۲/۵۰	۲/۴۰	۲/۳۱
۲۳	۳/۲۱	۳/۰۷	۲/۹۳	۲/۷۸	۲/۷۰	۲/۶۲	۲/۵۴	۲/۴۵	۲/۳۵	۲/۲۶
۲۴	۳/۱۷	۳/۰۳	۲/۸۹	۲/۷۴	۲/۶۶	۲/۵۸	۲/۴۹	۲/۴۰	۲/۳۱	۲/۲۱
۲۵	۳/۱۳	۲/۹۹	۲/۸۵	۲/۷۰	۲/۶۲	۲/۵۴	۲/۴۵	۲/۳۶	۲/۲۷	۲/۱۷
۲۶	۳/۰۹	۲/۹۶	۲/۸۱	۲/۶۶	۲/۵۸	۲/۵۰	۲/۴۲	۲/۳۳	۲/۲۳	۲/۱۳
۲۷	۳/۰۶	۲/۹۳	۲/۷۸	۲/۶۳	۲/۵۵	۲/۴۷	۲/۳۸	۲/۲۹	۲/۲۰	۲/۱۰
۲۸	۳/۰۳	۲/۹۰	۲/۷۵	۲/۶۰	۲/۵۲	۲/۴۴	۲/۳۵	۲/۲۶	۲/۱۷	۲/۰۶
۲۹	۳/۰۰	۲/۸۷	۲/۷۳	۲/۵۷	۲/۴۹	۲/۴۱	۲/۳۳	۲/۲۳	۲/۱۴	۲/۰۳
۳۰	۲/۹۸	۲/۸۴	۲/۷۰	۲/۵۵	۲/۴۷	۲/۳۹	۲/۳۰	۲/۲۱	۲/۱۱	۲/۰۱
۴۰	۲/۸۰	۲/۶۶	۲/۵۲	۲/۳۷	۲/۲۹	۲/۲۰	۲/۱۱	۲/۰۲	۱/۹۲	۱/۸۰
۶۰	۲/۶۳	۲/۵۰	۲/۳۵	۲/۲۰	۲/۱۲	۲/۰۳	۱/۹۴	۱/۸۴	۱/۷۳	۱/۶۰
۱۲۰	۲/۴۷	۲/۳۴	۲/۱۹	۲/۰۳	۱/۹۵	۱/۸۶	۱/۷۶	۱/۶۶	۱/۵۳	۱/۳۸
∞	۲/۳۲	۲/۱۸	۲/۰۴	۱/۸۸	۱/۷۹	۱/۷۰	۱/۵۹	۱/۴۷	۱/۳۲	۱/۰۰

جدول ۵: نقاط بحرانی آزمون مقایسه‌های دوبه‌دوی توکی

df	α	تعداد جامعه‌ها (k)								
		۲	۳	۴	۵	۶	۷	۸	۹	۱۰
۵	۰/۰۵	۳/۶۴	۴/۶۰	۵/۲۲	۵/۶۷	۶/۰۳	۶/۳۳	۶/۵۸	۶/۸۰	۶/۹۹
	۰/۰۱	۵/۷۰	۶/۹۸	۷/۸۰	/۴۲	۸/۹۱	۹/۳۲	/۶۷	/۹۷	۱۰/۲۴
۶	۰/۰۵	۳/۴۶	۴/۳۴	۴/۹۰	۵/۳۰	۵/۶۳	۵/۹۰	۶/۱۲	۶/۳۲	۶/۴۹
	۰/۰۱	۵/۲۴	۶/۳۳	۷/۰۳	/۵۶	۷/۹۷	۸/۳۲	/۶۱	/۸۷	۹/۱۰
۷	۰/۰۵	۳/۳۴	۴/۱۶	۴/۶۸	۵/۰۶	۵/۳۶	۵/۶۱	۵/۸۲	۶/۰۰	۶/۱۶
	۰/۰۱	۴/۹۵	۵/۹۲	۶/۵۴	/۰۱	۷/۳۷	۷/۶۸	۷/۹۴	۸/۱۷	۸/۳۷
۸	۰/۰۵	۳/۲۶	۴/۰۴	۴/۵۳	۴/۸۹	۵/۱۷	۵/۴۰	۵/۶۰	۵/۷۷	۵/۹۲
	۰/۰۱	۴/۷۵	۵/۶۴	۶/۲۰	۶/۶۲	۶/۹۶	۷/۲۴	۷/۴۷	۷/۶۸	۷/۸۶
۹	۰/۰۵	۳/۲۰	۳/۹۵	۴/۴۱	۴/۷۶	۵/۰۲	۵/۲۴	۵/۴۳	۵/۵۹	۵/۷۴
	۰/۰۱	۴/۶۰	۵/۴۳	۵/۹۶	۶/۳۵	۶/۶۶	۶/۹۱	۷/۱۳	۷/۳۳	۷/۴۹
۱۰	۰/۰۵	۳/۱۵	۳/۸۸	۴/۳۳	۴/۶۵	۴/۹۱	۵/۱۲	۵/۳۰	۵/۴۶	۵/۶۰
	۰/۰۱	۴/۴۸	۵/۲۷	۵/۷۷	۶/۱۴	۶/۴۳	۶/۶۷	۶/۸۷	۷/۰۵	۷/۲۱
۱۱	۰/۰۵	۳/۱۱	۳/۸۲	۴/۲۶	۴/۵۷	۴/۸۲	۵/۰۳	۵/۲۰	۵/۳۵	۵/۴۹
	۰/۰۱	۴/۳۹	۵/۱۵	۵/۶۲	۵/۹۷	۶/۲۵	۶/۴۸	۶/۶۷	۶/۸۴	۶/۹۹
۱۲	۰/۰۵	۳/۰۸	۳/۷۷	۴/۲۰	۴/۵۱	۴/۷۵	۴/۹۵	۵/۱۲	۵/۲۷	۵/۳۹
	۰/۰۱	۴/۳۲	۵/۰۵	۵/۵۰	۵/۸۴	۶/۱۰	۶/۳۲	۶/۵۱	۶/۶۷	۶/۸۱
۱۳	۰/۰۵	۳/۰۶	۳/۷۳	۴/۱۵	۴/۴۵	۴/۶۹	۴/۸۸	۵/۰۵	۵/۱۹	۵/۳۲
	۰/۰۱	۴/۲۶	۴/۹۶	۵/۴۰	۵/۷۳	۵/۹۸	۶/۱۹	۶/۳۷	۶/۵۳	۶/۶۷
۱۴	۰/۰۵	۳/۰۳	۳/۷۰	۴/۱۱	۴/۴۱	۴/۶۴	۴/۸۳	۴/۹۹	۵/۱۳	۵/۲۵
	۰/۰۱	۴/۲۱	۴/۸۹	۵/۳۲	۵/۶۳	۵/۸۸	۶/۰۸	۶/۲۶	۶/۴۱	۶/۵۴
۱۵	۰/۰۵	۳/۰۱	۳/۶۷	۴/۰۸	۴/۳۷	۴/۵۹	۴/۷۸	۴/۹۴	۵/۰۸	۵/۲۰
	۰/۰۱	۴/۱۷	۴/۸۴	۵/۲۵	۵/۵۶	۵/۸۰	۵/۹۹	۶/۱۶	۶/۳۱	۶/۴۴
۱۶	۰/۰۵	۳/۰۰	۳/۶۵	۴/۰۵	۴/۳۳	۴/۵۶	۴/۷۴	۴/۹۰	۵/۰۳	۵/۱۵
	۰/۰۱	۴/۱۳	۴/۷۹	۵/۱۹	۵/۴۹	۵/۷۲	۵/۹۲	۶/۰۸	/۲۲	۶/۳۵
۱۷	۰/۰۵	۲/۹۸	۳/۶۳	۴/۰۲	۴/۳۰	۴/۵۲	۴/۷۰	۴/۸۶	۴/۹۹	۵/۱۱
	۰/۰۱	۴/۱۰	۴/۷۴	۵/۱۴	۵/۴۳	۵/۶۶	۵/۸۵	۶/۰۱	۶/۱۵	۶/۲۷
۱۸	۰/۰۵	۲/۹۷	۳/۶۱	۴/۰۰	۴/۲۸	۴/۴۹	۴/۶۷	۴/۸۲	۴/۹۶	۵/۰۷
	۰/۰۱	۴/۰۷	۴/۷۰	۵/۰۹	۵/۳۸	۵/۶۰	۵/۷۹	۵/۹۴	۶/۰۸	۶/۲۰
۱۹	۰/۰۵	۲/۹۶	۳/۵۹	۳/۹۸	۴/۲۵	۴/۴۷	۴/۶۵	۴/۷۹	۴/۹۲	۵/۰۴
	۰/۰۱	۴/۰۵	۴/۶۷	۵/۰۵	۵/۳۳	۵/۵۵	۵/۷۳	۵/۸۹	۶/۰۲	۶/۱۴
۲۰	۰/۰۵	۲/۹۵	۳/۵۸	۳/۹۶	۴/۲۳	۴/۴۵	۴/۶۲	۴/۷۷	۴/۹۰	۵/۰۱
	۰/۰۱	۴/۰۲	۴/۶۴	۵/۰۲	۵/۲۹	۵/۵۱	۵/۶۹	۵/۸۴	۵/۹۷	۶/۰۹
۲۴	۰/۰۵	۲/۹۲	۳/۵۳	۳/۹۰	۴/۱۷	۴/۳۷	۴/۵۴	۴/۶۸	۴/۸۱	۴/۹۲
	۰/۰۱	۳/۹۶	۴/۵۵	۴/۹۱	۵/۱۷	۵/۳۷	۵/۵۴	۵/۶۹	۵/۸۱	۵/۹۲
۳۰	۰/۰۵	۲/۸۹	۳/۴۹	۳/۸۵	۴/۱۰	۴/۳۰	۴/۴۶	۴/۶۰	۴/۷۲	۴/۸۲
	۰/۰۱	۳/۸۹	۴/۴۵	۴/۸۰	۵/۰۵	۵/۲۴	۵/۴۰	۵/۵۴	۵/۶۵	۵/۷۶
۴۰	۰/۰۵	۲/۸۶	۳/۴۴	۳/۷۹	۴/۰۴	۴/۲۳	۴/۳۹	۴/۵۲	۴/۶۳	۴/۷۳
	۰/۰۱	۳/۸۲	۴/۳۷	۴/۷۰	۴/۹۳	۵/۱۱	۵/۲۶	۵/۳۹	۵/۵۰	۵/۶۰
۶۰	۰/۰۵	۲/۸۳	۳/۴۰	۳/۷۴	۳/۹۸	۴/۱۶	۴/۳۱	۴/۴۴	۴/۵۵	۴/۶۵
	۰/۰۱	۳/۷۶	۴/۲۸	۴/۵۹	۴/۸۲	۴/۹۹	۵/۱۳	۵/۲۵	۵/۳۶	۵/۴۵
۱۲۰	۰/۰۵	۲/۸۰	۳/۳۶	۳/۶۸	۳/۹۲	۴/۱۰	۴/۲۴	۴/۳۶	۴/۴۷	۴/۵۶
	۰/۰۱	۳/۷۰	۴/۲۰	۴/۵۰	۴/۷۱	۴/۸۷	۵/۰۱	۵/۱۲	۵/۲۱	۵/۳۰
∞	۰/۰۵	۲/۷۷	۳/۳۱	۳/۶۳	۳/۸۶	۴/۰۳	۴/۱۷	۴/۲۹	۴/۳۹	۴/۴۷
	۰/۰۱	۳/۶۴	۴/۱۲	۴/۴۰	۴/۶۰	۴/۷۶	۴/۸۸	۴/۹۹	۵/۰۸	۵/۱۶

جدول ۶: نقاط بحرانی آزمون دانت

df	α	تعداد جامعه‌ها (k)								
		۲	۳	۴	۵	۶	۷	۸	۹	۱۰
۵	۰/۰۵	۲/۵۷	۳/۰۳	۳/۲۹	۳/۴۸	۳/۶۲	۳/۷۳	۳/۸۲	۳/۹	۳/۹۷
	۰/۰۱	۴/۰۳	۴/۶۳	۴/۹۸	۵/۲۲	۵/۴۱	۵/۵۶	۵/۶۹	۵/۸	۵/۸۹
۶	۰/۰۵	۲/۴۵	۲/۸۶	۳/۱	۳/۲۶	۳/۳۹	۳/۴۹	۳/۵۷	۳/۶۴	۳/۷۱
	۰/۰۱	۳/۷۱	۴/۲۱	۴/۵۱	۴/۷۱	۴/۸۷	۵	۵/۱	۵/۲	۵/۲۸
۷	۰/۰۵	۲/۳۶	۲/۷۵	۲/۹۷	۳/۱۲	۳/۲۴	۳/۳۳	۳/۴۱	۳/۴۷	۳/۵۳
	۰/۰۱	۳/۵	۳/۹۵	۴/۲۱	۴/۳۹	۴/۵۳	۴/۶۴	۴/۷۴	۴/۸۲	۴/۸۹
۸	۰/۰۵	۲/۳۱	۲/۶۷	۲/۸۸	۳/۰۲	۳/۱۳	۳/۲۲	۳/۲۹	۳/۳۵	۳/۴۱
	۰/۰۱	۳/۳۶	۳/۷۷	۴	۴/۱۷	۴/۲۹	۴/۴	۴/۴۸	۴/۵۶	۴/۶۲
۹	۰/۰۵	۲/۲۶	۲/۶۱	۲/۸۱	۲/۹۵	۳/۰۵	۳/۱۴	۳/۲	۳/۲۶	۳/۳۲
	۰/۰۱	۳/۲۵	۳/۶۳	۳/۸۵	۴/۰۱	۴/۱۲	۴/۲۲	۴/۳	۴/۳۷	۴/۴۳
۱۰	۰/۰۵	۲/۲۳	۲/۵۷	۲/۷۶	۲/۸۹	۲/۹۹	۳/۰۷	۳/۱۴	۳/۱۹	۳/۲۴
	۰/۰۱	۳/۱۷	۳/۵۳	۳/۷۴	۳/۸۸	۳/۹۹	۴/۰۸	۴/۱۶	۴/۲۲	۴/۲۸
۱۱	۰/۰۵	۲/۲	۲/۵۳	۲/۷۲	۲/۸۴	۲/۹۴	۳/۰۲	۳/۰۸	۳/۱۴	۳/۱۹
	۰/۰۱	۳/۱۱	۳/۴۵	۳/۶۵	۳/۷۹	۳/۸۹	۳/۹۸	۴/۰۵	۴/۱۱	۴/۱۶
۱۲	۰/۰۵	۲/۱۸	۲/۵	۲/۶۸	۲/۸۱	۲/۹	۲/۹۸	۳/۰۴	۳/۰۹	۳/۱۴
	۰/۰۱	۳/۰۵	۳/۳۹	۳/۵۸	۳/۷۱	۳/۸۱	۳/۸۹	۳/۹۶	۴/۰۲	۴/۰۷
۱۳	۰/۰۵	۲/۱۶	۲/۴۸	۲/۶۵	۲/۷۸	۲/۸۷	۲/۹۴	۳	۳/۰۶	۳/۱
	۰/۰۱	۳/۰۱	۳/۳۳	۳/۵۲	۳/۶۵	۳/۷۴	۳/۸۲	۳/۸۹	۳/۹۴	۳/۹۹
۱۴	۰/۰۵	۲/۱۴	۲/۴۶	۲/۶۳	۲/۷۵	۲/۸۴	۲/۹۱	۲/۹۷	۳/۰۲	۳/۰۷
	۰/۰۱	۲/۹۸	۳/۲۹	۳/۴۷	۳/۵۹	۳/۶۹	۳/۷۶	۳/۸۳	۳/۸۸	۳/۹۳
۱۵	۰/۰۵	۲/۱۳	۲/۴۴	۲/۶۱	۲/۷۳	۲/۸۲	۲/۸۹	۲/۹۵	۳	۳/۰۴
	۰/۰۱	۲/۹۵	۳/۲۵	۳/۴۳	۳/۵۵	۳/۶۴	۳/۷۱	۳/۷۸	۳/۸۳	۳/۸۸
۱۶	۰/۰۵	۲/۱۲	۲/۴۲	۲/۵۹	۲/۷۱	۲/۸	۲/۸۷	۲/۹۲	۲/۹۷	۳/۰۲
	۰/۰۱	۲/۹۲	۳/۲۲	۳/۳۹	۳/۵۱	۳/۶	۳/۶۷	۳/۷۳	۳/۷۸	۳/۸۳
۱۷	۰/۰۵	۲/۱۱	۲/۴۱	۲/۵۸	۲/۶۹	۲/۷۸	۲/۸۵	۲/۹	۲/۹۵	۳
	۰/۰۱	۲/۹	۳/۱۹	۳/۳۶	۳/۴۷	۳/۵۶	۳/۶۳	۳/۶۹	۳/۷۴	۳/۷۹
۱۸	۰/۰۵	۲/۱	۲/۴	۲/۵۶	۲/۶۸	۲/۷۶	۲/۸۳	۲/۸۹	۲/۹۴	۲/۹۸
	۰/۰۱	۲/۸۸	۳/۱۷	۳/۳۳	۳/۴۴	۳/۵۳	۳/۶	۳/۶۶	۳/۷۱	۳/۷۵
۱۹	۰/۰۵	۲/۰۹	۲/۳۹	۲/۵۵	۲/۶۶	۲/۷۵	۲/۸۱	۲/۸۷	۲/۹۲	۲/۹۶
	۰/۰۱	۲/۸۶	۳/۱۵	۳/۳۱	۳/۴۲	۳/۵	۳/۵۷	۳/۶۳	۳/۶۸	۳/۷۲
۲۰	۰/۰۵	۲/۰۹	۲/۳۸	۲/۵۴	۲/۶۵	۲/۷۳	۲/۸	۲/۸۶	۲/۹	۲/۹۵
	۰/۰۱	۲/۸۵	۳/۱۳	۳/۲۹	۳/۴	۳/۴۸	۳/۵۵	۳/۶	۳/۶۵	۳/۶۹
۲۴	۰/۰۵	۲/۰۶	۲/۳۵	۲/۵۱	۲/۶۱	۲/۷	۲/۷۶	۲/۸۱	۲/۸۶	۲/۹
	۰/۰۱	۲/۸	۳/۰۷	۳/۲۲	۳/۳۲	۳/۴	۳/۴۷	۳/۵۲	۳/۵۷	۳/۶۱
۳۰	۰/۰۵	۲/۰۴	۲/۳۲	۲/۴۷	۲/۵۸	۲/۶۶	۲/۷۲	۲/۷۷	۲/۸۲	۲/۸۶
	۰/۰۱	۲/۷۵	۳/۰۱	۳/۱۵	۳/۲۵	۳/۳۳	۳/۳۹	۳/۴۴	۳/۴۹	۳/۵۲
۴۰	۰/۰۵	۲/۰۲	۲/۲۹	۲/۴۴	۲/۵۴	۲/۶۲	۲/۶۸	۲/۷۳	۲/۷۷	۲/۸۱
	۰/۰۱	۲/۷	۲/۹۵	۳/۰۹	۳/۱۹	۳/۲۶	۳/۳۲	۳/۳۷	۳/۴۱	۳/۴۴
۶۰	۰/۰۵	۲	۲/۲۷	۲/۴۱	۲/۵۱	۲/۵۸	۲/۶۴	۲/۶۹	۲/۷۳	۲/۷۷
	۰/۰۱	۲/۶۶	۲/۹	۳/۰۳	۳/۱۲	۳/۱۹	۳/۲۵	۳/۲۹	۳/۳۳	۳/۳۷
۱۲۰	۰/۰۵	۱/۹۳	۲/۰۸	۲/۱۸	۲/۲۶	۲/۳۲	۲/۳۷	۲/۴۱	۲/۴۵	
	۰/۰۱	۲/۶	۲/۷۳	۲/۸۲	۲/۸۹	۲/۹۴	۲/۹۹	۳/۰۳	۳/۰۶	
∞	۰/۰۵	۱/۹۲	۲/۰۶	۲/۱۶	۲/۲۳	۲/۲۹	۲/۳۷	۲/۳۹	۲/۴۶	
	۰/۰۱	۲/۵۶	۲/۶۸	۲/۷۷	۲/۸۴	۲/۸۹	۲/۹۳	۲/۹۷	۳	

جدول ۷: نقاط بحرانی آزمون علامت

α دو دامنه	۰/۱	۰/۰۵	۰/۰۲	۰/۰۱	α دو دامنه	۰/۱	۰/۰۵	۰/۰۲	۰/۰۱
α یک دامنه	۰/۰۵	۰/۰۲۵	۰/۰۱	۰/۰۰۵	α یک دامنه	۰/۰۵	۰/۰۲۵	۰/۰۱	۰/۰۰۵
$n^- + n^+$	نقاط بحرانی				$n^- + n^+$	نقاط بحرانی			
۱	–	–	–	–	۳۱	۱۱	۱۳	۱۵	۱۷
۲	–	–	–	–	۳۲	۱۲	۱۴	۱۶	۱۶
۳	–	–	–	–	۳۳	۱۱	۱۳	۱۵	۱۷
۴	–	–	–	–	۳۴	۱۲	۱۴	۱۶	۱۶
۵	۵	–	–	–	۳۵	۱۱	۱۳	۱۵	۱۷
۶	۶	۶	–	–	۳۶	۱۲	۱۴	۱۶	۱۸
۷	۷	۷	۷	–	۳۷	۱۱	۱۳	۱۷	۱۷
۸	۶	۸	۸	۸	۳۸	۱۲	۱۴	۱۶	۱۸
۹	۷	۷	۹	۹	۳۹	۱۳	۱۵	۱۷	۱۷
۱۰	۸	۸	۱۰	۱۰	۴۰	۱۲	۱۴	۱۶	۱۸
۱۱	۷	۹	۹	۱۱	۴۵	۱۳	۱۵	۱۷	۱۹
۱۲	۸	۸	۱۰	۱۰	۴۶	۱۴	۱۶	۱۸	۲۰
۱۳	۷	۹	۱۱	۱۱	۴۹	۱۳	۱۵	۱۹	۱۹
۱۴	۸	۱۰	۱۰	۱۲	۵۰	۱۴	۱۶	۱۸	۲۰
۱۵	۹	۹	۱۱	۱۱	۵۵	۱۵	۱۷	۱۹	۲۱
۱۶	۸	۱۰	۱۲	۱۲	۵۶	۱۴	۱۶	۱۸	۲۰
۱۷	۹	۹	۱۱	۱۳	۵۹	۱۵	۱۷	۱۹	۲۱
۱۸	۸	۱۰	۱۲	۱۲	۶۰	۱۴	۱۸	۲۰	۲۲
۱۹	۹	۱۱	۱۱	۱۳	۶۵	۱۵	۱۷	۲۱	۲۳
۲۰	۱۰	۱۰	۱۲	۱۴	۶۶	۱۶	۱۸	۲۰	۲۲
۲۱	۹	۱۱	۱۳	۱۳	۶۹	۱۵	۱۹	۲۳	۲۵
۲۲	۰	۱۲	۱۲	۱۴	۷۰	۱۶	۱۸	۲۲	۲۴
۲۳	۹	۱۱	۱۳	۱۵	۷۵	۱۷	۱۹	۲۳	۲۵
۲۴	۱۰	۱۲	۱۴	۱۴	۷۶	۱۶	۲۰	۲۲	۲۴
۲۵	۱۱	۱۱	۱۳	۱۵	۷۹	۱۷	۱۹	۲۳	۲۵
۲۶	۱۰	۱۲	۱۴	۱۴	۸۰	۱۶	۲۰	۲۲	۲۴
۲۷	۱۱	۱۳	۱۳	۱۵	۸۹	۱۷	۲۱	۲۳	۲۷
۲۸	۱۰	۱۲	۱۴	۱۶	۹۰	۱۸	۲۰	۲۴	۲۶
۲۹	۱۱	۱۳	۱۵	۱۵	۹۹	۱۹	۲۱	۲۵	۲۷
۳۰	۱۰	۱۲	۱۴	۱۶	۱۰۰	۱۸	۲۲	۲۶	۲۸

جدول ۸: نقاط بحرانی آزمون رتبهٔ علامت‌دار ویلکاکسون

α یک‌دامنه	۰/۰۲۵	۰/۰۱	۰/۰۰۵
α دو دامنه	۰/۰۵	۰/۰۲	۰/۰۱
$n^{-}+n^{+}$	نقاط بحرانی		
۶	۰		
۷	۲	۰	
۸	۴	۲	۰
۹	۶	۳	۲
۱۰	۸	۵	۳
۱۱	۱۱	۷	۵
۱۲	۱۴	۱۰	۷
۱۳	۱۷	۱۳	۱۰
۱۴	۲۱	۱۶	۱۳
۱۵	۲۵	۲۰	۱۶
۱۶	۳۰	۲۴	۲۰
۱۷	۳۵	۲۸	۲۳
۱۸	۴۰	۳۳	۲۸
۱۹	۴۶	۳۸	۳۲
۲۰	۵۲	۴۳	۳۸
۲۱	۵۹	۴۹	۴۳
۲۲	۶۶	۵۶	۴۹
۲۳	۷۳	۶۲	۵۵
۲۴	۸۱	۶۹	۶۱
۲۵	۸۹	۷۷	۶۸

جدول ۹: نقاط بحرانی آزمون مجموع رتبه‌ای ویلکاکسون

n_{min}	n_{max}	دامنهٔ پایین						دامنهٔ بالا					
		۰/۰۰۵	۰/۰۱	۰/۰۲۵	۰/۰۵	۰/۰۱	۰/۰۲	۰/۰۲	۰/۰۱	۰/۰۵	۰/۰۲۵	۰/۰۰۱	۰/۰۰۵
۴	۴			۱۰	۱۱	۱۳	۱۴	۲۲	۲۳	۲۵	۲۶		
	۵		۱۰	۱۱	۱۲	۱۴	۱۵	۲۵	۲۶	۲۸	۲۹	۳۰	
	۶	۱۰	۱۱	۱۲	۱۳	۱۵	۱۷	۲۷	۲۹	۳۱	۳۲	۳۳	۳۴
	۷	۱۰	۱۱	۱۳	۱۴	۱۶	۱۸	۳۰	۳۲	۳۴	۳۵	۳۷	۳۸
	۸	۱۱	۱۲	۱۴	۱۵	۱۷	۲۰	۳۲	۳۵	۳۷	۳۸	۴۰	۴۱
	۹	۱۱	۱۳	۱۴	۱۶	۱۹	۲۱	۳۵	۳۷	۴۰	۴۲	۴۳	۴۵
	۱۰	۱۲	۱۳	۱۵	۱۷	۲۰	۲۳	۳۷	۴۰	۴۳	۴۵	۴۷	۴۸
	۱۱	۱۲	۱۴	۱۶	۱۸	۲۱	۲۴	۴۰	۴۳	۴۶	۴۸	۵۰	۵۲
	۱۲	۱۳	۱۵	۱۷	۱۹	۲۲	۲۶	۴۲	۴۶	۴۹	۵۱	۵۳	۵۵
۵	۵	۱۵	۱۶	۱۷	۱۹	۲۰	۲۲	۳۳	۳۵	۳۶	۳۸	۳۹	۴۰
	۶	۱۶	۱۷	۱۸	۲۰	۲۲	۲۴	۳۶	۳۸	۴۰	۴۲	۴۳	۴۴
	۷	۱۶	۱۸	۲۰	۲۱	۲۳	۲۶	۳۹	۴۲	۴۴	۴۵	۴۷	۴۹
	۸	۱۷	۱۹	۲۱	۲۳	۲۵	۲۸	۴۲	۴۵	۴۷	۴۹	۵۱	۵۳
	۹	۱۸	۲۰	۲۲	۲۴	۲۷	۳۰	۴۵	۴۸	۵۱	۵۳	۵۵	۵۷
	۱۰	۱۹	۲۱	۲۳	۲۶	۲۸	۳۲	۴۸	۵۲	۵۴	۵۷	۵۹	۶۱
	۱۱	۲۰	۲۲	۲۴	۲۷	۳۰	۳۴	۵۱	۵۵	۵۸	۶۱	۶۳	۶۵
	۱۲	۲۱	۲۳	۲۶	۲۸	۳۲	۳۶	۵۴	۵۸	۶۲	۶۴	۶۷	۶۹
۶	۶	۲۳	۲۴	۲۶	۲۸	۳۰	۳۳	۴۵	۴۸	۵۰	۵۲	۵۴	۵۵
	۷	۲۴	۲۵	۲۷	۲۹	۳۲	۳۵	۴۹	۵۲	۵۵	۵۷	۵۹	۶۰
	۸	۲۵	۲۷	۲۹	۳۱	۳۴	۳۷	۵۳	۵۶	۵۹	۶۱	۶۳	۶۵
	۹	۲۶	۲۸	۳۱	۳۳	۳۶	۴۰	۵۶	۶۰	۶۳	۶۵	۶۸	۷۰
	۱۰	۲۷	۲۹	۳۲	۳۵	۳۸	۴۲	۶۰	۶۴	۶۷	۷۰	۷۳	۷۵
	۱۱	۲۸	۳۰	۳۴	۳۷	۴۰	۴۴	۶۴	۶۸	۷۱	۷۴	۷۸	۸۰
	۱۲	۳۰	۳۲	۳۵	۳۸	۴۲	۴۷	۶۷	۷۲	۷۶	۷۹	۸۲	۸۴
۷	۷	۳۲	۳۴	۳۶	۳۹	۴۱	۴۵	۶۰	۶۴	۶۶	۶۹	۷۱	۷۳
	۸	۳۴	۳۵	۳۸	۴۱	۴۴	۴۸	۶۴	۶۸	۷۱	۷۴	۷۷	۷۸
	۹	۳۵	۳۷	۴۰	۴۳	۴۶	۵۰	۶۹	۷۳	۷۶	۷۹	۸۲	۸۴
	۱۰	۳۷	۳۹	۴۲	۴۵	۴۹	۵۳	۷۳	۷۷	۸۱	۸۴	۸۷	۸۹
	۱۱	۳۸	۴۰	۴۴	۴۷	۵۱	۵۶	۷۷	۸۲	۸۶	۸۹	۹۳	۹۵
	۱۲	۴۰	۴۲	۴۶	۴۹	۵۴	۵۹	۸۱	۸۶	۹۱	۹۴	۹۸	۱۰۰
۸	۸	۴۳	۴۵	۴۹	۵۱	۵۵	۵۹	۷۷	۸۱	۸۵	۸۷	۹۱	۹۳
	۹	۴۵	۴۷	۵۱	۵۴	۵۸	۶۲	۸۲	۸۶	۹۰	۹۳	۹۷	۹۹
	۱۰	۴۷	۴۹	۵۳	۵۶	۶۰	۶۵	۸۷	۹۲	۹۶	۹۹	۱۰۳	۱۰۵
	۱۱	۴۹	۵۱	۵۵	۵۹	۶۳	۶۹	۹۱	۹۷	۱۰۱	۱۰۵	۱۰۹	۱۱۱
	۱۲	۵۱	۵۳	۵۸	۶۲	۶۶	۷۲	۹۶	۱۰۲	۱۰۶	۱۱۰	۱۱۵	۱۱۷
۹	۹	۵۶	۵۹	۶۲	۶۶	۷۰	۷۵	۹۶	۱۰۱	۱۰۵	۱۰۹	۱۱۲	۱۱۵
	۱۰	۵۸	۶۱	۶۵	۶۹	۷۳	۷۸	۱۰۲	۱۰۷	۱۱۱	۱۱۵	۱۱۹	۱۲۲
	۱۱	۶۱	۶۳	۶۸	۷۲	۷۶	۸۲	۱۰۷	۱۱۳	۱۱۷	۱۲۱	۱۲۶	۱۲۸
	۱۲	۶۳	۶۶	۷۱	۷۵	۸۰	۸۶	۱۱۲	۱۱۸	۱۲۳	۱۲۷	۱۳۲	۱۳۵
۱۰	۱۰	۷۱	۷۴	۷۸	۸۲	۸۷	۹۳	۱۱۷	۱۲۳	۱۲۸	۱۳۲	۱۳۶	۱۳۹
	۱۱	۷۳	۷۷	۸۱	۸۶	۹۱	۹۷	۱۲۳	۱۲۹	۱۳۴	۱۳۹	۱۴۳	۱۴۷
	۱۲	۷۶	۷۹	۸۴	۸۹	۹۴	۱۰۱	۱۲۹	۱۳۶	۱۴۱	۱۴۶	۱۵۱	۱۵۴

جدول ۱۰: نقاط بحرانی آزمون کروسکال‌ـوالیس برای $\alpha = ۰/۰۵$ و $\alpha = ۰/۰۱$

k = ۳					k = ۴						k = ۵						
اندازه نمونه			۰/۰۵	۰/۰۱	اندازه نمونه				۰/۰۵	۰/۰۱	اندازه نمونه					۰/۰۵	۰/۰۱
۲	۲	۲	–	–	۲	۲	۱	۱	–	–	۲	۲	۱	۱	۱	–	–
۳	۲	۱	–	–	۲	۲	۲	۱	۵/۶۷۹	–	۲	۲	۲	۱	۱	۶/۷۵۰	–
۳	۲	۲	۴/۷۱۴	–	۲	۲	۲	۲	۶/۱۶۷	۶/۳۳۷	۲	۲	۲	۲	۱	۷/۱۳۳	۷/۵۳۳
۳	۳	۱	۵/۱۴۳	–	۳	۱	۱	۱	–	–	۲	۲	۲	۲	۲	۷/۴۱۸	۸/۲۹۱
۳	۳	۲	۵/۳۶۱	–	۳	۲	۱	۱	–	–	۳	۱	۱	۱	۱	–	–
۳	۳	۳	۵/۶۰۰	۷/۲۰۰	۳	۲	۲	۱	۵/۸۳۳	–	۳	۲	۱	۱	۱	۶/۵۸۳	–
۴	۲	۱	–	–	۳	۲	۲	۲	۶/۳۳۳	۷/۱۳۳	۳	۲	۲	۱	۱	۶/۸۰۰	۷/۶۰۰
۴	۲	۲	۵/۳۳۳	–	۳	۳	۱	۱	۶/۳۳۳	–	۳	۲	۲	۲	۱	۷/۳۰۹	۸/۱۲۷
۴	۳	۱	۵/۲۰۸	–	۳	۳	۲	۱	۶/۲۴۴	۷/۲۰۰	۳	۲	۲	۲	۲	۷/۶۸۲	۸/۶۸۲
۴	۳	۲	۵/۴۴۴	۶/۴۴۴	۳	۳	۲	۲	۶/۵۲۷	۷/۶۳۶	۳	۳	۱	۱	۱	۷/۱۱۱	–
۴	۳	۳	۵/۷۹۱	۶/۷۴۵	۳	۳	۳	۱	۶/۶۰۰	۷/۴۰۰	۳	۳	۲	۱	۱	۷/۲۰۰	۸/۰۷۳
۴	۴	۱	۴/۹۶۷	۶/۶۶۷	۳	۳	۳	۲	۶/۷۲۷	۸/۰۱۵	۳	۳	۲	۲	۱	۷/۵۹۱	۸/۵۷۶
۴	۴	۲	۵/۴۵۵	۷/۰۳۶	۳	۳	۳	۳	۷/۰۰۰	۵/۵۳۸	۳	۳	۲	۲	۲	۷/۹۱۰	۹/۱۱۵
۴	۴	۳	۵/۵۹۸	۷/۱۴۴	۴	۱	۱	۱	–	–	۳	۳	۳	۱	۱	۷/۵۷۶	۸/۴۲۴
۴	۴	۴	۵/۶۹۲	۷/۶۵۴	۴	۲	۱	۱	۵/۸۳۳	–	۳	۳	۳	۲	۱	۷/۷۶۹	۹/۰۵۱
۵	۲	۱	۵/۰۰۰	–	۴	۲	۲	۱	۶/۱۳۳	۷/۰۰۰	۳	۳	۳	۲	۲	۸/۰۴۴	۹/۵۰۵
۵	۲	۲	۵/۱۶۰	۶/۵۳۳	۴	۲	۲	۲	۶/۵۴۵	۷/۳۹۱	۳	۳	۳	۳	۱	۸/۰۰۰	۹/۴۵۱
۵	۳	۱	۴/۹۶۰	–	۴	۳	۱	۱	۶/۱۷۸	۷/۰۶۷	۳	۳	۳	۳	۲	۸/۲۰۰	۹/۸۷۶
۵	۳	۲	۵/۲۵۱	۶/۹۰۹	۴	۳	۲	۱	۶/۳۰۹	۷/۴۵۵	۳	۳	۳	۳	۳	۸/۳۳۳	۱۰/۲۰
۵	۳	۳	۵/۶۴۸	۷/۰۷۹	۴	۳	۲	۲	۶/۶۲۱	۷/۸۷۱							
۵	۴	۱	۴/۹۸۵	۶/۹۵۵	۴	۳	۳	۱	۶/۵۴۵	۷/۷۵۸							
۵	۴	۲	۵/۲۷۳	۷/۲۰۵	۴	۳	۳	۲	۶/۷۹۵	۸/۳۳۳							
۵	۴	۳	۵/۶۵۶	۷/۴۴۵	۴	۳	۳	۳	۶/۹۸۴	۸/۶۵۹							
۵	۴	۴	۵/۶۵۷	۷/۷۶۰	۴	۴	۱	۱	۵/۹۴۵	۷/۹۰۹							
۵	۵	۱	۵/۱۲۷	۷/۳۰۹	۴	۴	۲	۱	۶/۳۸۶	۷/۹۰۹							
۵	۵	۲	۵/۳۳۸	۷/۳۳۸	۴	۴	۲	۲	۶/۷۳۱	۸/۳۴۶							
۵	۵	۳	۵/۷۰۵	۷/۵۷۸	۴	۴	۳	۱	۶/۶۳۵	۸/۲۳۱							
۵	۵	۴	۵/۶۶۶	۷/۸۲۳	۴	۴	۳	۲	۶/۸۷۴	۸/۶۲۱							
۵	۵	۵	۵/۷۸۰	۸/۰۰۰	۴	۴	۳	۳	۷/۰۳۸	۸/۸۷۶							
۶	۱	۱	–	–	۴	۴	۴	۱	۶/۷۲۵	۸/۵۸۸							
۶	۲	۱	۴/۸۲۲	–	۴	۴	۴	۲	۶/۹۵۷	۸/۸۷۱							
۶	۲	۲	۵/۳۴۵	۶/۶۵۵	۴	۴	۴	۳	۷/۱۴۲	۹/۰۷۵							
۶	۳	۱	۴/۸۵۵	۶/۸۷۳	۴	۴	۴	۴	۷/۲۳۵	۹/۲۸۷							
۶	۳	۲	۵/۳۴۸	۶/۹۷۰													
۶	۳	۳	۵/۶۱۵	۷/۴۱۰													
۶	۴	۱	۴/۹۴۷	۷/۱۰۶													
۶	۴	۲	۵/۳۴۰	۷/۳۴۰													
۶	۴	۳	۵/۶۱۰	۷/۵۰													
۶	۴	۴	۵/۶۸۱	۷/۷۹۵													
۶	۵	۱	۴/۹۹۰	۷/۱۸۲													
۶	۵	۲	۵/۳۳۸	۷/۳۷۶													
۶	۵	۳	۵/۶۰۲	۷/۵۹۰													
۶	۵	۴	۵/۶۶۱	۷/۹۳۶													
۶	۵	۵	۵/۷۲۹	۸/۰۲۸													
۶	۶	۱	۴/۹۴۵	۷/۱۲۱													
۶	۶	۲	۵/۴۱۰	۷/۴۶۷													
۶	۶	۳	۵/۶۲۵	۷/۷۲۵													
۶	۶	۴	۵/۷۲۵	۸/۰۰۰													
۶	۶	۵	۵/۷۶۵	۸/۱۲۴													
۶	۶	۶	۵/۸۰۱	۸/۲۲۲													
۷	۷	۷	۵/۸۱۹	۸/۳۷۸													
۸	۸	۸	۵/۸۰۵	۸/۴۶۵													

جدول ۱۱: نقاط بحرانی آزمون ضریب همبستگی اسپیرمن

n	α دو دامنه		α یک دامنه	
	۰/۰۵	۰/۰۱	۰/۰۵	۰/۰۱
۵	۱	–	۰/۹۰۰	۱
۶	۰/۸۸۶	۱	۰/۸۲۹	۰/۹۴۳
۷	۰/۷۸۶	۰/۹۲۹	۰/۷۱۴	۰/۸۹۳
۸	۰/۷۳۸	۰/۸۸۱	۰/۶۴۳	۰/۸۳۳
۹	۰/۷۰۰	۰/۸۳۳	۰/۶۰۰	۰/۷۸۳
۱۰	۰/۶۴۸	۰/۷۹۴	۰/۵۶۴	۰/۷۴۵
۱۱	۰/۶۱۸	۰/۷۵۵	۰/۵۳۶	۰/۷۰۹
۱۲	۰/۵۸۷	۰/۷۲۷	۰/۵۰۳	۰/۶۷۱
۱۳	۰/۵۶۰	۰/۷۰۳	۰/۴۸۴	۰/۶۴۸
۱۴	۰/۵۳۸	۰/۶۷۵	۰/۴۶۴	۰/۶۲۲
۱۵	۰/۵۲۱	۰/۶۵۴	۰/۴۴۳	۰/۶۰۴
۱۶	۰/۵۰۳	۰/۶۳۵	۰/۴۲۹	۰/۵۸۲
۱۷	۰/۴۸۵	۰/۶۱۵	۰/۴۱۴	۰/۵۶۶
۱۸	۰/۴۷۲	۰/۶۰۰	۰/۴۰۱	۰/۵۵۰
۱۹	۰/۴۶۰	۰/۵۸۴	۰/۳۹۱	۰/۵۳۵
۲۰	۰/۴۴۷	۰/۵۷۰	۰/۳۸۰	۰/۵۲۰
۲۱	۰/۴۳۵	۰/۵۵۶	۰/۳۷۰	۰/۵۰۸
۲۲	۰/۴۲۵	۰/۵۴۴	۰/۳۶۱	۰/۴۹۶
۲۳	۰/۴۱۵	۰/۵۳۲	۰/۳۵۳	۰/۴۸۶
۲۴	۰/۴۰۶	۰/۵۲۱	۰/۳۴۴	۰/۴۷۶
۲۵	۰/۳۹۸	۰/۵۱۱	۰/۳۳۷	۰/۴۶۶
۲۶	۰/۳۹۰	۰/۵۰۱	۰/۳۳۱	۰/۴۵۷
۲۷	۰/۳۸۲	۰/۴۹۱	۰/۳۲۴	۰/۴۴۸
۲۸	۰/۳۷۵	۰/۴۸۳	۰/۳۱۷	۰/۴۴۰
۲۹	۰/۳۶۸	۰/۴۷۵	۰/۳۱۲	۰/۴۳۳
۳۰	۰/۳۶۲	۰/۴۶۷	۰/۳۰۶	۰/۴۲۵
۳۱	۰/۳۵۶	۰/۴۵۹	۰/۳۰۱	۰/۴۱۸
۳۲	۰/۳۵۰	۰/۴۵۲	۰/۲۹۶	۰/۴۱۲
۳۳	۰/۳۴۵	۰/۴۴۶	۰/۲۹۱	۰/۴۰۵
۳۴	۰/۳۴۰	۰/۴۳۹	۰/۲۸۷	۰/۳۹۹
۳۵	۰/۳۳۵	۰/۴۳۳	۰/۲۸۳	۰/۳۹۴
۳۶	۰/۳۳۰	۰/۴۲۷	۰/۲۷۹	۰/۳۸۸
۳۷	۰/۳۲۵	۰/۴۲۱	۰/۲۷۵	۰/۳۸۳
۳۸	۰/۳۲۱	۰/۴۱۵	۰/۲۷۱	۰/۳۷۸
۳۹	۰/۳۱۷	۰/۴۱۰	۰/۲۶۷	۰/۳۷۳
۴۰	۰/۳۱۳	۰/۴۰۵	۰/۲۶۴	۰/۳۶۸
۴۱	۰/۳۰۹	۰/۴۰۰	۰/۲۶۱	۰/۳۶۴
۴۲	۰/۳۰۵	۰/۳۹۵	۰/۲۵۷	۰/۳۵۹
۴۳	۰/۳۰۱	۰/۳۹۱	۰/۲۵۴	۰/۳۵۵
۴۴	۰/۲۹۸	۰/۳۸۶	۰/۲۵۱	۰/۳۵۱
۴۵	۰/۲۹۴	۰/۳۸۲	۰/۲۴۸	۰/۳۴۷
۴۶	۰/۲۹۱	۰/۳۷۸	۰/۲۴۶	۰/۳۴۳
۴۷	۰/۲۸۸	۰/۳۷۴	۰/۲۴۳	۰/۳۴۰
۴۸	۰/۲۸۵	۰/۳۷۰	۰/۲۴۰	۰/۳۳۶
۴۹	۰/۲۸۲	۰/۳۶۶	۰/۲۳۸	۰/۳۳۳
۵۰	۰/۲۷۹	۰/۳۶۳	۰/۲۳۵	۰/۳۲۹

جدول ۱۲: نقاط بحرانی برای آزمون ضریب کندال برای $\alpha =$ ۰/۰۵ و $\alpha =$ ۰/۰۱

n	یک دامنه		دو دامنه	
	۰/۰۵	۰/۰۱	۰/۰۵	۰/۰۱
۴	۱			
۵	۰/۸۰۰	۱	۱	
۶	۰/۷۳۳	۰/۸۶۷	۰/۸۶۷	۱
۷	۰/۶۱۹	۰/۸۱۰	۰/۷۱۴	۰/۹۰۵
۸	۰/۵۷۱	۰/۷۱۴	۰/۶۴۳	۰/۷۸۶
۹	۰/۵۰۰	۰/۶۶۷	۰/۵۵۶	۰/۷۲۲
۱۰	۰/۴۶۷	۰/۶۰۰	۰/۵۱۱	۰/۶۴۴
۱۱	۰/۴۱۸	۰/۵۶۴	۰/۴۹۱	۰/۶۰۰
۱۲	۰/۳۹۴	۰/۵۴۵	۰/۴۵۵	۰/۵۷۶
۱۳	۰/۳۵۹	۰/۵۱۳	۰/۴۳۶	۰/۵۶۴
۱۴	۰/۳۶۳	۰/۴۷۳	۰/۴۰۷	۰/۵۱۶
۱۵	۰/۳۳۳	۰/۴۶۷	۰/۳۹۰	۰/۵۰۵
۱۶	۰/۳۱۷	۰/۴۳۳	۰/۳۸۳	۰/۴۸۳
۱۷	۰/۳۰۹	۰/۴۲۶	۰/۳۶۸	۰/۴۷۱
۱۸	۰/۲۹۴	۰/۴۱۲	۰/۳۴۶	۰/۴۵۱
۱۹	۰/۲۸۷	۰/۳۹۲	۰/۳۳۳	۰/۴۳۹
۲۰	۰/۲۷۴	۰/۳۷۹	۰/۳۲۶	۰/۴۲۱
۲۱	۰/۲۶۷	۰/۳۷۱	۰/۳۱۴	۰/۴۱۰
۲۲	۰/۲۶۴	۰/۳۵۹	۰/۳۰۷	۰/۳۹۴
۲۳	۰/۲۵۷	۰/۳۵۲	۰/۲۹۶	۰/۳۹۱
۲۴	۰/۲۴۶	۰/۳۴۱	۰/۲۹۰	۰/۳۷۷
۲۵	۰/۲۴۰	۰/۳۳۳	۰/۲۸۷	۰/۳۶۷
۲۶	۰/۲۳۷	۰/۳۲۹	۰/۲۸۰	۰/۳۶۰
۲۷	۰/۲۳۱	۰/۳۲۲	۰/۲۷۱	۰/۳۵۶
۲۸	۰/۲۲۸	۰/۳۱۲	۰/۲۶۵	۰/۳۴۴
۲۹	۰/۲۲۲	۰/۳۱۰	۰/۲۶۱	۰/۳۴۰
۳۰	۰/۲۱۸	۰/۳۰۱	۰/۲۵۵	۰/۳۳۳

پیوست ۳: شکل‌های کوچک و بزرگ حروف یونانی

کاربرد در آمار	نام فارسی	نام انگلیسی	شکل کوچک	شکل بزرگ
احتمال خطای نوع اول	آلفا	alpha	α	Α
احتمال خطای نوع دوم	بِتا	beta	β	Β
ضریب پیوند در جداول توافقی	گاما	gamma	γ	Γ
تفاضل متغیرها	دِلتا	delta	δ	Δ
جمله خطا در مدل‌ها	اِپسیلُن	epsilon	ε	Ε
	زِتا	zeta	ζ	Ζ
ضریب رابطه (اندازه اثر)	اِتا	eta	η	Η
پارامتر دلخواه	تِتا	theta	θ	Θ
	یوتا	iota	ι	Ι
ضریب توافقی	کاپا	kappa	κ	Κ
ضریب پیوند در جداول توافقی	لاندا	lambda	λ	Λ
پارامتر میانگین	میو/مو	mu	μ	Μ
درجه آزادی	نو	nu	ν	Ν
	ایکسی	xi	ξ	Ξ
	اُمیکرن	omicron	ο	Ο
مقدار احتمال	پی	pi	π	Π
ضریب همبستگی	رُ	rho	ρ	Ρ
انحراف معیار/واریانس	سیگما	sigma	σ یا ς	Σ
ضریب رابطه	تا	tau	τ	Τ
	آپسیلُن	upsilon	υ	Υ
ضریب پیوند در جداول توافقی	فی	phi	φ	Φ
آماره آزمون نیکویی برازش	خی	chi	χ	Χ
	پی‌سای	psi	ψ	Ψ
ضریب رابطه (اندازه اثر)	اُمِگا	omega	ω	Ω

پیوست ۴: تعریف برخی از اصطلاحات

آزمون پارامتری (parametric test)

آزمونی که مبتنی بر پیش‌فرض‌هایی از جمله توزیع متغیر مورد استنباط است. برای مثال، آزمون F در تجزیه واریانس نیازمند آن است که داده‌ها در هر یک از جامعه‌ها دارای توزیع نرمال باشند یا در آزمون ضریب همبستگی پیرسون، ضروری است توزیع یکی از متغیرها نرمال باشد. اغلب، آزمون‌های پارامتری برای متغیرهای فاصله‌ای یا نسبتی به کار می‌روند و از معادل‌های ناپارامتری خود پرتوان‌تر هستند.

آزمون فرضیه (hypothesis test)

فرایندی است که بر اساس داده‌های نمونه، درستی ادعایی (فرضیه‌ای) را درباره جامعه آماری بررسی می‌کند. روش‌های آزمون فرضیه از حوزه‌های آمار استنباطی به شمار می‌آید.

آزمون ناپارامتری (nonparametric test)

آزمونی که هیچ پیش‌فرضی را درباره توزیع متغیر مورد استنباط در نظر نمی‌گیرد، بنابراین به آن آزمون آزاد-توزیع (distribution free) نیز گفته می‌شود. بخشی از آزمون‌های ناپارامتری، معادل‌های آزمون‌های پارامتری هستند، مانند آزمون کروسکال-والیس که معادل ناپارامتری تجزیه واریانس است و بخشی از آزمون‌های ناپارامتری به استنباط درباره متغیرهای اسمی یا ترتیبی اختصاص دارند؛ مانند آزمون نیکویی برازش برای تحلیل جدول‌های توافقی.

آمار (statistics)

دانش گردآوری داده‌ها، سازماندهی و تلخیص آنها و سرانجام «نتیجه‌گیری» مبتنی بر احتمال از داده‌هاست. این دانش دارای سه حوزه آمار توصیفی، آمار استنباطی و نمونه‌گیری است.

آمار استنباطی (inferential statistics)

مجموعه‌ای از روش‌هاست که نتایجی احتمالی، ولی قابل اعتماد درباره جامعه آماری بر اساس داده‌های نمونه به دست می‌دهد. آمار استنباطی دارای دو حوزه برآوردیابی و آزمون فرضیه است.

آمار توصیفی (descriptive statistics)

مجموعه‌ای از روش‌ها که به سازماندهی و خلاصه‌سازی داده‌ها می‌پردازد. این روش‌ها به دو دسته تقسیم می‌شوند. دسته اول، داده‌ها را در قالب جدول یا نمودار و دسته دوم در قالب کمّیت‌هایی موسوم به مشخصه‌های توزیع توصیف می‌کنند.

آماره (statistics)

به زبان ساده، هر رابطه‌ای که تنها بر اساس داده‌های نمونه محاسبه می‌شود. تمام مشخصه‌های توزیع متغیر که برای نمونه به دست می‌آیند، یک آماره هستند.

اُریبی (bias)

اختلاف میانگینِ توزیع برآوردگر از مقدار واقعی پارامتر است. به عبارت دیگر، اُریبی، برابر با میانگینِ خطاهای برآوردهای حاصل از یک برآوردگر است (به تفاوت آن با خطا توجه کنید).

اندازه نمونه (sample size)

تعداد افرادی است که به عنوان نمونه از جامعه آماری انتخاب می‌شود. تعیین اندازه نمونه با توجه به روش نمونه‌گیری از رابطه‌های خاص خود صورت می‌گیرد. عواملی مانند واریانس متغیر تحت استنباط، حداکثر خطای قابل پذیرش، ضریب اطمینان، هزینه و جمعیت بر اندازه نمونه مؤثرند.

برآوردگر (estimator)

آماره‌ای است که به منظور برآورد پارامتری نامعلوم به کار می‌رود. در واقع، برآوردگر «رابطه‌ای» است که با استفاده از داده‌ها به مقداری به عنوان برآورد منجر می‌شود. برای مثال، میانگین نمونه می‌تواند برآورد میانگین جامعه‌ای باشد که نمونه از آن انتخاب شده است. دو نوع برآوردگر وجود دارد: نقطه‌ای و فاصله‌ای.

بی‌پاسخی (nonresponse)

به تلاش ناموفق در اندازه‌گیری یک یا چند متغیر از یک فرد نمونه گفته می‌شود. در پیمایش‌ها، بی‌پاسخی، خود را به صورت بی‌جواب ماندن یک یا چندین پرسش نشان

می دهد. اگر پژوهشگر به طور کامل در گرفتن جواب از فرد نمونه ناموفق باشد، با بی‌پاسخی واحد و در صورتی که تنها یک یا چند پرسش بدون جواب باشد، با بی‌پاسخی جزئی روبه‌روست. بی‌پاسخی می‌تواند به اریبی برآوردها یا افزایش خطای نمونه‌گیری منجر شود.

پارامتر (parameter)

به هر مشخصه توزیع متغیر در جامعه آماری گفته می‌شود. اغلب، مقدار پارامتر نامعلوم است و تنها از طریق سرشماری می‌توان آن را به دست آورد. از این رو، در بحث برآوردیابی به دنبال روش‌های برآورد آنها هستیم.

پیش‌فرض (assumption)

مفروضاتی که تحت برقراری آنها استفاده از یک روش آماری مجاز است. برای مثال، آزمون F در تجزیه واریانس دارای سه پیش‌فرض است: همسانی واریانس‌ها، استقلال نمونه‌ها (جامعه‌ها) و نرمال‌بودن توزیع داده‌ها. تنها در صورت برقراری هر سه پیش‌فرض می‌توان از آزمون F برای مقایسه میانگین چند جامعه استفاده کرد.

پیمایش (survey)

روشی کمّی برای پژوهش است که با اندازه‌گیری یک یا چند متغیر از چندین عضو جامعه آماری تحت مطالعه، اطلاعاتی «قابل تعمیم» درباره آن جامعه به دست می‌دهد. اغلب، این روش به ایجاد مجموعه‌ای از داده‌های حاصل از نمونه منجر می‌شود که با استفاده از روش‌های آمار استنباطی به اطلاعاتی درباره جامعه آماری می‌انجامند.

تغییرپذیری (variability)

به وجود پراکندگی در مقادیر اشاره دارد. منبع تغییرپذیری می‌تواند متفاوت باشد. تغییرپذیری ناشی از نمونه‌گیری، از تفاوت نمونه‌های مختلفی که می‌توان از یک جامعه انتخاب کرد، ایجاد می‌شود. این نوع از تغییرپذیری باعث ایجاد توزیع نمونه‌گیری می‌شود. تفاوت‌های فردی افراد یا اشیا نیز می‌تواند منبع دیگر تغییرپذیری باشد، به طوری که حتی تحت شرایط یکسان نیز رفتار متفاوتی از خود نشان می‌دهند. برای مثال، چند بوته از یک

گونه گیاه حتی تحت آبیاری، نور، خاک و کود یکسان، رشد برابری ندارند. چنین تغییرپذیری در میان انسان‌ها نیز کاملاً مشهود است.

توزیع نمونه‌گیری (sampling distribution)

توزیع یک آماره که ناشی از تفاوت نمونه‌هاست. به عبارت دیگر، آماره بر اساس نمونه‌های متفاوت، مقادیر مختلفی به خود می‌گیرد. بنابراین، آماره نیز مانند متغیر می‌تواند دارای توزیعی باشد که به آن توزیع نمونه‌گیری آماره می‌گویند.

جامعه آماری (statistical population)

مجموعه افراد یا اشیایی است که پژوهشگر مطالعه خود را در میان آنها صورت می‌دهد، می‌تواند به تک‌تک آنها دسترسی پیدا کند و یافته‌های پژوهشی خود را به تمامی آنها تعمیم دهد.

خطا (error)

در بحث برآوردیابی، اختلاف برآورد از مقدار واقعیِ پارامتر است. این اختلاف می‌تواند مثبت یا منفی باشد، ولی اغلب، قدرمطلق آن مورد توجه است.

خطای غیرنمونه‌گیری (nonsampling error)

صرفنظر از نمونه‌گیری، هر نوع خطایی که در طراحی چگونگی گردآوری داده‌ها، اجرای آن یا ارائه یافته‌ها رخ دهد. ابزار گردآوری (پرسشنامه)، خطاهای انسانی، گرد کردن ارقام و... برخی از منابع متنوع بروز چنین خطایی هستند.

خطای معیار (standard error)

به انحراف معیار برآوردگر گفته می‌شود. از آنجا که هدف برآوردگر ارائه مقداری نزدیک به پارامتر نامعلوم جامعه است، انحراف معیار آن، شاخصی برای خطای آن است. از این رو، به آن «خطای» معیار می‌گویند.

خطای نمونه‌گیری (sampling error)

ناشی از تعمیم نتایج مربوط به بخشی (نمونه‌ای) از جامعه به کل آن است. روش‌های مختلف نمونه‌گیری به منظور کنترل این خطا به کار گرفته می‌شوند.

رابطه (relationship)

درباره حداقل دو متغیر مطرح می‌شود و به معنای «هم‌تغییری» است. هرگاه با تغییر در مقادیر متغیری، در مقادیر متغیر دیگری نیز تغییر ایجاد شود، بین آن دو متغیر رابطه آماری وجود دارد. رابطه می‌تواند یکنوا (افزایشی یا کاهشی) یا نایکنوا (درجه ۲، نوسانی و...) باشد. هنگامی که رابطه یکنواست، می‌توان از «شدت» و «جهت» آن صحبت کرد، ولی اگر نایکنوا باشد، تنها از «شدت» آن صحبت می‌شود.

سرشماری (census)

به گردآوری داده‌ها از تمام اعضای جامعه آماری گفته می‌شود. سرشماری در مقابل نمونه‌گیری قرار دارد. اصطلاحات «تمام‌شماری» و «همه‌پرسی» نیز به سرشماری اشاره دارند.

فرضیه (hypothesis)

هر ادعایی درباره توزیع یا مشخصه‌های توزیع متغیر در جامعه است. در بحث آزمون فرضیه از دو نوع فرضیه به نام‌های فرضیه صفر (خنثی) و فرضیه جانشین (مقابل) صحبت می‌شود. فرضیه پژوهش همان فرضیه جانشین است.

قضیه حد مرکزی (central limit theorem)

این قضیه بیان می‌کند که توزیع نمونه‌گیری میانگین با افزایش تعداد نمونه، توزیع نرمال خواهد بود، به طوری که میانگین برابر با میانگین جامعه‌ای است که نمونه از آن انتخاب شده است و واریانس آن برابر با نسبت واریانس جامعه به تعداد نمونه است. این قضیه در بسیاری از استنباط‌های آماری مانند فاصله اطمینان برای میانگین یا آزمون فرضیه درباره میانگین نقش اساسی دارد.

متغیر (variable)

خصیصه‌ای (ویژگی‌ای) در جامعه آماری که همه اعضا به طور یکسان دارا نیستند. متغیر دارای چندین شکل تقسیم‌بندی است؛ دسته‌بندی به پیوسته یا گسسته، دسته‌بندی به اسمی، ترتیبی، فاصله‌ای یا نسبتی و دسته‌بندی به مقوله‌ای یا غیرمقوله‌ای از آن جمله‌اند.

مجموعه داده‌ها (data set)

اغلب، جدولی (ماتریسی) است که هر سطر آن به یک عضو جامعه یا نمونه و هر ستون آن به یک متغیر اندازه‌گیری‌شده از آن عضو اختصاص دارد.

مشخصه‌های توزیع (distribution measures)

کمّیت‌هایی که هر یک ویژگی خاصی از توزیع را نشان می‌دهد. این ویژگی‌ها به سه دستهٔ تمرکز، پراکندگی و شکل تقسیم می‌شوند. ویژگی تمرکز شامل مشخصه‌های نما، چندک‌ها و میانگین است. ویژگی پراکندگی شامل مشخصه‌های دامنه، دامنه میان‌چارکی، دامنه میان‌صدکی، واریانس (انحراف معیار)، قدرمطلق انحراف از میانگین و ضریب تغییرات است. ویژگی شکل، شامل مشخصه‌های ضریب چولگی و برجستگی است.

منحنی توزیع متغیر (variable distribution curve)

منحنی‌ای که با استفاده از سطح زیر منحنی آن می‌توان تعیین کرد که چه نسبتی از جامعه یا نمونه بین مقادیر متغیر توزیع شده است. منحنی توزیع به متغیرهای پیوسته اختصاص دارد و از بافت‌نگار به دست می‌آید. برخی از منحنی‌های توزیع نظیر نرمال استاندارد، t یا F کاملاً شناخته شده‌اند و جدول‌هایی وجود دارد که سطح زیر منحنی آنها را به ازای مقادیر مختلف به دست می‌دهند.

نمونه‌گیری (sampling)

فرایند انتخاب تعدادی از اعضای یک جامعه آماری است. این انتخاب بر اساس روش نمونه‌گیری (sampling method) صورت می‌گیرد و افراد انتخاب شده را نمونه (sample) می‌نامند. نمونه‌ای که با یکی از روش‌های احتمالی انتخاب شود، نمونه تصادفی یا احتمالی گفته می‌شود و در مقابل، نمونه‌ای که با روش‌های نااحتمالی انتخاب شود، نمونه نااحتمالی یا غیرتصادفی نامیده می‌شود.

همبستگی (correlation)

رابطه خطی دو متغیر است که می‌تواند افزایشی یا کاهشی باشد. ضریب همبستگی پیرسون یکی از معیارهای رایج برای اندازه‌گیری «جهت» و «شدت» همبستگی دو متغیر است.