

حقیق در تعلیم و تربیت

این امر با انتخاب یک روش تحقیقاتی معین دنبال می‌شود. برخی از مسائل، مستلزم انجام آزمایش هستند؛ مسائل دیگر نیز ممکن است نیازمند یک راهبرد توصیفی باشند. پس انتخاب یک روش تحقیقاتی، جزئیات طرح تحقیق و روشهای اندازه‌گیری متغیرها را تحت تأثیر قرار می‌دهد. ابزار اندازه‌گیری متغیرها به عنوان وسایلی استاندارد ممکن است از قبل وجود داشته باشد، و یا می‌باید توسط محققان ساخته شود.

گردآوری و تفسیر داده‌ها

نتایج قیاسهای حاصل از فرضیه‌های بررسی باید آزمون شوند. بنابراین در این مرحله گردآوری داده‌ها صورت می‌گیرد که شامل جنبه‌های عادی به اجرا گذاردن ابزار تحقیق، ثبت اطلاعات، تهیه جداول و غیره است. بر خلاف باور عمومی، این مرحله معمولاً نسبت به مراحل برنامه ریزی قبلی طرح، زمان کمتری را به خود اختصاص می‌دهد.

بعد از گردآوری داده‌ها، می‌باید آنها را به طور معمول به شیوه آماری تجزیه و تحلیل کرد و تفسیرهای مناسبی از یافته‌های تحقیق به عمل آورد. در رویکرد علمی، فرد مدعی اثبات فرضیه نیست؛ زیرا این امر مستلزم پی‌جویی حقیقت مطلق است که از ویژگیهای این رویکرد به شمار نمی‌رود. محقق صرفاً نتیجه می‌گیرد که آیا شواهد فرضیه را حمایت می‌کنند یا خیر.

گزارش نتایج

محقق می‌باید روشها، یافته‌ها و نتایج را به شکلی قابل فهم در دسترس دیگرانی که مشتاق آن هستند، قرار دهد. این امر ارائه روشن و مختصر مراحل تحقیق را با جزییات کافی شامل می‌شود؛ به طوری که محقق دیگری نیز می‌تواند آن را تکرار کند.

هر یک از مراحل گفته شده در مطالعه تحقیقاتی به طور مبسوط در فصول بعدی این کتاب بررسی شده است. شاید ضروری باشد که مجدداً این نکته را به خواننده یادآوری کنیم که احتمال بسیار کمی وجود دارد که یک محقق بتواند سلسله مراحل مشخصی را، که در فرازهای قبل توصیف شد، دنبال کند. این مراحل مداوماً با هم تداخل دارند و ممکن است که جای خود را با هم عوض کنند.

// سؤالات مطروحه از سوی محققان تعلیم و تربیت

سؤال ویژه‌ای که محقق انتخاب می‌کند، مسلماً به حیطه مورد علاقه، زمینه قبلی و مسئله رویاروی او بستگی خواهد داشت. به هر حال، ما می‌توانیم مسائل مربوط به تحقیق در تعلیم و تربیت را به مسائل نظری، که با اصول اساسی سر و کار دارد، یا مسائل عملی، که برای حل مشکلات عاجل در شرایط روزمره طراحی شده‌اند، تقسیم کنیم.

/ سؤالات نظریه‌ای

سؤالهایی که ماهیت نظری دارند، از این دست هستند: «این چیست؟» یا «این امر چگونه اتفاق می‌افتد؟» یا «چرا این اتفاق می‌افتد؟» در تحقیق در تعلیم و تربیت، سؤالاتی که با «چه» مرتبط هستند، به طور مشخص این گونه صورت‌بندی می‌شوند: «هوش چیست؟» یا «خلافت چیست؟» سؤالات نوعی درباره «چگونگی» عبارتند از: «کودکان چگونه یاد می‌گیرند؟» یا «شخصیت چگونه شکل می‌گیرد؟» سؤالهای «چرائی» می‌توانند از این قبیل باشند: «انسان چرا فراموش می‌کند؟» یا «چرا بعضی از کودکان بیش از دیگران پیشرفت‌گرا هستند؟»

تحقیقی که سوی‌گیری نظری داشته باشد، می‌تواند هم به سمت ارائه نظریه و هم به سوی آزمون نظریه‌های موجود هدایت شود. حالت اول مطالعه‌ای است که محقق ضمن آن تلاش می‌کند با هدف روشن کردن ماهیت رابطه‌های موجود بین متغیرها، به تعمیم‌هایی درباره رفتار دست یابد. او ممکن است بر این باور باشد که متغیرهای خاصی با یکدیگر رابطه دارند، بنابراین تحقیق را برای توصیف ماهیت رابطه انجام می‌دهد. او می‌تواند بر اساس یافته‌هایش نظریه‌ای را در مورد پدیده‌ای خاص صورت‌بندی کند. نظریه‌های یادگیری بعد از آن که محققان قادر به نشان دادن رابطه بین روشهای خاص، فرد و متغیرهای محیطی، و کارایی فرآیند یادگیری شدند، تکامل یافتند.

بررسیهای متداولتر در تحقیقات آموزشی احتمالاً بررسیهایی با هدف آزمون نظریه‌های موجود هستند. ارائه یک نظریه به عنوان هدف یک بررسی، بویژه برای محققان تازه‌کار، هدفی بلندپروازانه است و معمولاً کوشش برای قیاس فرضیه‌هایی از نظریه‌های موجود در مباحث یادگیری، شخصیت، انگیزش و غیره و نیز آزمون آنها واقع‌بینانه‌تر است. اگر فرضیه‌ها به طور منطقی از نظریه قیاس شده باشند و آزمایشهای تجربی شواهدی در تائید آنها ارائه کنند، این شواهد به نوبه خود از نظریه نیز حمایت خواهند کرد.

/ سوالات کاربردی

بسیاری از سوالات در تحقیق مسائل تعلیم و تربیت، دارای ماهیتی کاربردی و عملی هستند و هدفشان حل مسائلی مشخصی است که کارشناسان تعلیم و تربیت در فعالیتهای روزانه خود با آن دست به گریبان هستند. آنها به این علت سوالاتی مرتبط با تحقیق در زمینه تعلیم و تربیت هستند که با مسائلی واقعی و مبتنی بر تجربه سر و کار دارند. این سوالات از این گونه اند:

آموزش برنامه‌ای در تدریس هجی کردن به دانش‌آموزان کلاس دوم چقدر مؤثر است؟
 بکارگیری روش تدریس بنیادی الفبا در پیشرفت خواندن دانش‌آموزان کلاس اول چه اثری دارد؟
 در تدریس علوم اجتماعی دبیرستان، تاثیر نسبی روش گفتگو پیرامون مسائل، در مقایسه با روش سخنرانی به چه میزان است؟ پاسخگویی به این سوالات ممکن است در یاری رساندن به معلمان برای تصمیم‌گیریهای معقول بسیار ارزشمند باشد.

// تحقیقات بنیادین و کاربردی

گاهی از اوقات یک سیستم طبقه‌بندی دیگر نیز در رابطه با این دو نوع سؤال مورد استفاده قرار می‌گیرد که مبنای آن هدف تحقیق است. اولین نوع تحقیق، که هدفش دستیابی به داده‌های تجربی برای صورتبندی، بسط و ارزشیابی نظریه است، تحقیق بنیادین خوانده می‌شود. سوی‌گیری این نوع مطالعه، چه در طرح تحقیق و چه از نظر هدف، به سمت حل مسائل عملی نیست. مقصود اصلی آن گشودن مرزهای دانش بدون توجه به کاربرد عملی آن است. البته ممکن است یافته‌های تحقیق به طور تصادفی در مسائل عملی که ارزش اجتماعی دارند، به کار گرفته شوند. برای مثال، پیشرفت در تجربیات پزشکی به تحقیق بنیادین در بیوشیمی و میکروبیولوژی وابسته است؛ همچنانکه پیشرفت در امر تعلیم و تربیت به پیشرفت در کشف قوانین عمومی از طریق انجام تحقیق بنیادین در زمینه روان‌شناسی، تعلیم و تربیت و علوم اجتماعی وابسته است.

کار اصلی تحقیق بنیادین کسب علم برای علم است. تأمل در فواید اجتماعی یافته‌های اینگونه تحقیقات، مانع از طرح ریزی آنها نمی‌شود.

دومین نوع تحقیق، که هدف آن حل یک مسئله عملی فوری است، تحقیق کاربردی خوانده می‌شود. تحقیق کاربردی گونه‌ای از تحقیق است که در رابطه با مسائل واقعی و در شرایطی که آن

مسائل در عمل بروز می‌کنند، صورت می‌گیرد. از طریق تحقیق کاربردی کارشناس تعلیم و تربیت اغلب قادر خواهد بود که مسائلش را در سطح مناسبی از پیچیدگی، یعنی در موقعیت آموزش - یادگیری در کلاس، حل کند. محقق ممکن است برای کشف قوانین کلی تر یادگیری به تحقیق بنیادین نیاز داشته باشد، ولی برای تعیین نحوه عملکرد این قوانین در کلاس درس باید از تحقیق کاربردی کمک بگیرد. اگر قرار باشد تغییرات علمی در کار تدریس جامعه عمل بباشد، این رویکرد ضرورتی اساسی خواهد داشت. بجز کارشناسان تعلیم و تربیت فرد دیگری عهده‌دار حل مسائل کاربردی از این دست نخواهد بود. این نکته باید تذکر داده شود که تحقیق کاربردی نیز از روش تحقیق علمی بهره می‌گیرد. ما اذعان داریم که خط مشخصی بین تحقیق بنیادین و کاربردی وجود ندارد. یقیناً کاربردهای حاصله از نظریه به حل مسائل عملی کمک می‌کند. ما سعی می‌کنیم که نظریه‌های یادگیری را در کلاسهای درس به کار ببریم. از سوی دیگر، ممکن است تحقیق بنیادین برای تکمیل صورتبندی نظریه‌ای خود به یافته‌های تحقیق کاربردی وابسته باشد. یک تجربه یادگیری در کلاس درس می‌تواند پرتو جدیدی بر نظریه یادگیری بیفکند. بعلاوه، مشاهده در یک موقعیت عملی، به آزمون نظریه‌ها کمک می‌کند و ممکن است به صورتبندی نظریه‌های جدید منجر شود.

هیچیک از روشهای گفته شده لزوماً بر دیگری رجحان ندارد. روش مورد استفاده در یک مطالعه تحقیقی را ماهیت مسئله و نوع داده‌های مورد نیاز تعیین می‌کند. در این مورد اغلب ترتیبی منطقی وجود دارد که از طریق آن نوعی از تحقیق، نوع دیگر را دنبال می‌کند. گاهی اوقات ممکن است یک محقق جهت روشننگری اتفاقات گذشته، بررسی را با مطالعه‌ای تاریخی آغاز کند و در مرحله بعدی، با استفاده از روش توصیفی اطلاعات مربوط به وضعیت کنونی مسئله را در تعلیم و تربیت تهیه کند. با این زمینه، محقق می‌تواند برای تعیین رابطه بین متغیرهایی که ممکن است از طریق انواع دیگر تحقیق پیشنهاد شده باشد، به آزمایش دست زند.

/// زبان تحقیق

هر رشته علمی برای تشریح و خلاصه کردن مشاهدات زمینه خاص خود به یک زبان علمی نیازمند است. دانشمند در سطح کاربردی به اصطلاحاتی نیاز دارد که با کمک آنها مشاهدات ویژه را توصیف کند؛ او همچنین در سطح نظری برای اشاره به فرآیندهای فرضی که مستقیماً

نمی‌توانند مورد مشاهده قرار گیرند نیز به اصطلاحاتی نیاز دارد. دانشمند می‌تواند از کلماتی برگرفته از زبان روزمره بهره جوید، ولی اغلب به این کلمات معانی تازه و بخصوصی نسبت می‌دهد که به طور معمول در کاربرد رایج آن یافت نمی‌شود و یا ممکن است مجبور شود اصطلاحات فنی جدیدی را که جزو زبان روزمره و رایج نیستند، خلق کند. هدف این بخش معرفی بعضی از اصطلاحات عمومی مورد استفاده در تحقیق آموزشی است.

اصطلاحاتی که یک دانشمند در دو سطح توصیفی و نظری به کار می‌برد، به ترتیب، مفاهیم و سازه‌ها نامیده می‌شوند.

// مفاهیم و سازه‌ها

یک مفهوم، شکل تجریدی رویدادهای مشاهده شده است؛ کلمه‌ای است که تشابهات و جنبه‌های مشترک اشیا و یا رویدادهایی را که بجز آن تشابهات و جنبه‌های مشترک، کاملاً با هم تفاوت دارند، معرفی می‌کند. کلماتی مانند *صندلی*، *سگ*، *درخت*، *مايع* و هزاران کلمه در زبان ما مفاهیم مشترک چیزهای گوناگونی را ارائه می‌کنند. هدف هر مفهوم، ساده کردن تفکر از طریق گردآوری شماری از رویدادها در زیر عنوانی عمومی است. بعضی از مفاهیم کاملاً به رویدادهایی که معرف آنها هستند، نزدیک می‌باشند. بنابراین به عنوان مثال، مفهوم *درخت* به راحتی می‌تواند با اشاره به درختهای بخصوصی روشن شود. مفهوم *درخت*، تجرید کلیه خصوصیات مشترک درختان است؛ خصوصياتی که مستقیماً قابل مشاهده هستند.

به این ترتیب، اصطلاحاتی از قبیل *انگیزه*، *قضاوت* و *توانایی حل مسئله* از طریق اشاره به موضوعات و رویدادهای ویژه براحتی قابل شرح نیستند. این تجریدهای سطح بالا به عنوان سازه‌ها تلقی می‌شوند. انسان از مفاهیم موجود، تجریدهای پیچیده‌تری خلق کرده و در کنار هم قرار داده است. همانگونه که انسان از کنار هم قرار دادن چوب و سایر مواد در طرحی هدفمند، موفق به ساختن خانه شده است، با ادغام مفاهیم و سازه‌هایی ساده‌تر در یک طرح هدفمند، سازه‌های پیچیده‌تری را خلق کرده است. برای مثال، مفاهیمی مانند *قدرت بینایی*، *تفکیک* *نمادها*، *جهت یابی* *چپ و راست*، *گنجینه لغات شنیداری* و غیره، در طرحی هدفمند به منظور خلق سازه آمادگی برای خواندن ادغام شده‌اند. سازه‌ها در تفسیر داده‌های تجربی و در ساختن نظریه مفید هستند. آنها برای توصیف قواعد و روابط مشاهده شده مورد استفاده قرار می‌گیرند.

سازه‌ها برای این خلق می‌شوند که مشاهدات را خلاصه کنند و تبیین‌هایی را فراهم سازند. یک سازه زمانی کنار گذاشته می‌شود که راه بهتری برای تبیین و خلاصه کردن مشاهدات جایگزین آن شود. برای مثال این مشاهدات که (۱) بعضی از مواد می‌سوزند و بعضی نمی‌سوزند (۲) بعضی از مواد سریعتر از دیگر مواد می‌سوزند، برای اولین بار با سازه و یا مفهوم قابلیت اشتعال^(۱) خلاصه شد و این اعتقاد را در برداشت که این امر جزء لاینفک همه مواد قابل اشتعال است که در معرض فرایند سوختن قرار می‌گیرند. این سازه هنگامی که تبیینهای مفیدتر از فرایند اشتعال ارائه شد، کنار گذاشته شد.

// مشخص کردن معنی

هر چه مفاهیم و یا سازه‌های (ذهنی) فرد از واقعیات علمی یا پدیده‌هایی که آن مفاهیم و سازه‌ها معرف آنها هستند، دورتر باشد، امکان سوء تفاهم بیشتر خواهد بود و در نتیجه نیاز بیشتری به تعریف دقیق آنها احساس می‌شود. معنی کلمات موجود در گنجینه لغات دانشمند باید تعریف شود. مفاهیم باید هم با اصطلاحات تجربیدی، که در واقع معنی کلی فرض شده برای این مفاهیم را ارائه می‌دهند و هم بر حسب عملیاتی که در یک مطالعه خاص بر آن مبنا اندازه‌گیری خواهند شد، تعریف شوند. کرلینگر^(۲) نوع اول تعریف را یک تعریف ترکیبی و نوع دوم آنرا تعریف کاربردی می‌خواند.

/ تعریف ترکیبی

به نوع قراردادی تر تعریف، که در آن یک اصطلاح با استفاده از اصطلاحات دیگر تعریف می‌شود، تعریف ترکیبی اطلاق می‌شود. برای مثال، هوش به عنوان توانایی تفکر تجربیدی تعریف شده است. این نوع تعریف به انتقال ماهیت کلی پدیده‌ای که محقق به آن جلب شده است و همچنین نشان دادن رابطه آن با نظریه و سایر مطالعاتی که از مفاهیم مشابه استفاده کرده‌اند، کمک می‌کند. یک تعریف ترکیبی، یک اصطلاح را توضیح می‌دهد و شاید درباره پدیده موصوف، بینشی را در دیگران ایجاد کند. به هر حال، اگر فرد بخواهد تحقیقی انجام دهد، باید مفاهیم مورد نظر خود را به رویدادهای قابل مشاهده تبدیل کند.

/ تعریف کاربردی

یک تعریف کاربردی تعریفی است که با مشخص کردن عملیاتی که انجام آنها برای اندازه‌گیری یک مفهوم یا سازه ضروری است، معنایی را به آن سازه یا مفهوم نسبت می‌دهد، این نوع تعریف در تحقیق اساسی است؛ زیرا داده‌ها باید برحسب رویدادهای قابل مشاهده گردآوری شوند. یک دانشمند ممکن است در سطح نظریه‌ای با سازه‌هایی از قبیل یادگیری، انگیزه، هیجان یا پیشرفت سر و کار داشته باشد، ولی قبل از آنکه بتواند آنها را به طور تجربی مطالعه کند، باید بر روی چند نوع رویداد قابل مشاهده برای معرفی آن سازه‌ها تصمیم‌گیری کند. هنگامی که کسی یک مفهوم یا سازه را از نظر عملیاتی تعریف می‌کند، رویدادهای متمایزی را به عنوان نشانگرهای آن مفهوم مجرد انتخاب و عملیاتی را برای دستیابی به داده‌های مرتبط با آن مفهوم ابداع می‌کند.

دو نوع تعریف کاربردی وجود دارد: اندازه‌گیری شده و آزمایشی. یک تعریف اندازه‌گیری شده کاربردی به عملیاتی اطلاق می‌شود که از طریق آن محققان می‌توانند یک مفهوم را بسنجند. برای مثال، هوشبهر را به طور کاربردی می‌توان به عنوان نمرات به دست آمده از آزمون هوش استانفورد - بینه، و یا خلاقیت را به طور کاربردی می‌توان به عنوان نمرات به دست آمده از آزمون خلاقیت مینه‌سوتا، تعریف کرد. یک تعریف کاربردی آزمایشی به مراحل طی شده از طرف محقق برای خلق شرایط دقیق آزمایش اطلاق می‌شود. برای مثال، تعریف کاربردی از ناکامی، در یک بررسی تحقیقی می‌تواند در قالب باز داشتن آزمودنی‌ها از رسیدن به هدف، بیان شود، و یا ناهماهنگی شناختی را به طور کاربردی می‌توان به درخواست از آزمودنی‌های یک بررسی برای اظهار بیانات عمومی برخلاف باورهای خصوصی و فردی آنها اطلاق کرد.

اگر چه محقق بر اساس تجربه و دانش خود و با استفاده از گزارشهای دیگر محققان هدایت می‌شود، ولی تعریف کاربردی از یک مفهوم در بعضی از پهنه‌ها یک روش اختیاری است. محقق اغلب از بین تعاریف کاربردی ممکن، تعریفی را انتخاب می‌کند که بهترین وجه رویکرد او را نسبت به مسئله بیان کند. یک تعریف کاربردی مطمئناً نمی‌تواند معنای کامل هر مفهوم را در برگیرد. این تعریف در واقع بسیار مشخص است، هدف آن نیز محدود کردن یک اصطلاح، برای اطمینان از این امر است که هر کس با آن سروکار داشته باشد، بتواند شیوه خاص استفاده از آن اصطلاح را در یابد. تعاریف کاربردی در صورتی مکفی تلقی می‌شوند که بر مبنای آنها بتوان

داده‌هایی را گردآوری کرد که نشانگرهای قابل پذیرش مفاهیمی باشند که باید معرفی شوند. اغلب دست یافتن و یا عدم دستیابی به این نتیجه، موضوعی است که به عقیده شخصی فرد مربوط می‌شود.

تعاریف کاربردی به این دلیل در تحقیق اساسی هستند که به محققان اجازه اندازه‌گیری مفاهیم و سازه‌های مجرد را می‌دهند و همچنین برای آنها این امکان را فراهم می‌آورند که بر مبنای علم از سطح ذهنیات و نظریه به سطح مشاهده حرکت کنند. با استفاده از تعاریف کاربردی است که محققان مبادرت به تحقیقاتی می‌کنند که در فقدان این تعاریف انجام آنها غیر ممکن خواهد بود. یادآوری این مسئله مهم است که محقق با اینکه یافته‌های خود را بر اساس سازه‌های تجربی گزارش می‌کند و سپس آنها را به نظریه و تحقیقات دیگر مرتبط می‌کند، اما در واقع آنچه که در عمل یافته است، رابطه بین دو مجموعه داده‌های قابل مشاهده و اندازه‌گیری است که برای معرفی سازه‌ها انتخاب شده بودند. برای مثال، یک تحقیق درباره رابطه بین سازه خلاقیت و سازه هوش، در عمل نمرات یک آزمون هوش و نمرات یک مقیاس خلاقیت را با یکدیگر مرتبط خواهد کرد.

// متغیرها

یک متغیر، چند مفهوم و یا سازه را منعکس و یا تشریح می‌کند. یک متغیر مفهومی است که می‌تواند مقادیر متفاوتی را به خود بگیرد. بلندی قد، مثالی از یک متغیر است، در مورد یک شخص، قد از زمانی به زمان دیگر تغییر می‌کند، در یک زمان بین افراد مختلف فرق می‌کند، بین متوسط گروه‌ها، متفاوت است و الی آخر. طبقه اجتماعی، جنسیت، سطح گنجینه لغات، هوشبهر و نمرات آزمون هجی کردن مثالهای دیگری از متغیرها هستند. محققان تعلیم و تربیت به تعیین رابطه متغیرها با یکدیگر علاقه‌مند هستند. در یک مطالعه در مورد ارتباط سطح گنجینه لغات و پیشرفت علمی در بین دانش‌آموزان کلاس هشتم، متغیرهای مورد توجه، اندازه‌های سطح گنجینه لغات و اندازه‌های پیشرفت علمی خواهد بود. ما متغیرها را با مقادیر ثابت، یعنی مقادیری که ارزش عددی آنها ثابت هستند، مقایسه می‌کنیم. مفهوم سطح کلاس، اگر چه بنا به تعریف یک متغیر است، اما در این مطالعه، یک متغیر محسوب نمی‌شود؛ زیرا همه دانش‌آموزان در کلاس هشتم هستند.

انواع متغیرها /

برای طبقه‌بندی متغیرها راههای متعددی وجود دارد. بعضی از متغیرها می‌توانند مقوله‌ای^(۱) و بعضی دیگر پیوسته^(۲) باشند. زمانی که آزمودنی‌ها از طریق دسته‌بندی به گروه‌هایی تقسیم شوند، آن ویژگی که تقسیم‌بندی بر مبنای آن صورت گرفته است، یک متغیر مقوله‌ای نامیده می‌شود. زبان مادری، محل سکونت، شغل اصلی پدر، و مدرسه‌ای که در آن ثبت نام کرده‌اید، مثالهایی از متغیرهای مقوله‌ای هستند. ساده‌ترین نوع متغیر مقوله‌ای، تنها دو وجه دارد و متغیر دو وجهی^(۳) یا دو مقوله‌ای نامیده می‌شود. زن - مرد، شهری - روستایی، قبول - مردود از جمله متغیرهای دو مقوله‌ای هستند. بعضی از متغیرهای مقوله‌ای دارای بیش از دو وجه هستند، به عنوان مثال سطح آموزش، وابستگی مذهبی، و محل تولد از جمله این متغیرها هستند. هنگامی که یک ویژگی مقادیر عددی نامحدودی را در یک دامنه دارا باشد، آن را یک متغیر پیوسته می‌نامند. هنگامی که قد یک کودک برای مثال از ۴۰ اینچ به ۴۱ اینچ می‌رسد، او در واقع بی‌نهایت واحد اندازه‌گیری قد را طی می‌کند. قد، وزن، سن و نمرات آزمون پیشرفت تحصیلی، نمونه‌هایی از متغیرهای پیوسته هستند.

مهمترین طبقه‌بندی متغیرها بر مبنای نوع استفاده از آنها در جریان تحقیق و با توجه به این مسئله است که در چه زمانی به عنوان متغیر مستقل^(۴) و در چه زمانی به عنوان متغیر وابسته^(۵) طبقه‌بندی می‌شوند.

بعضی از متغیرها نسبت به دیگر متغیرها مقدم هستند. ما ممکن است این مسئله را بر اساس تجربه و یا بر اساس فرضیاتی بر مبنای نظریه دربابیم که یک متغیر نسبت به دیگر متغیرها مقدم است. به عنوان مثال، فرض بر این است که فرد برای خواندن، نیازمند سطحی از هوش است. به این ترتیب، متغیر هوش است که بر متغیر خواندن تقدم دارد. در زمینه‌ای خاص، توانایی خواندن، یکی از نتایج منطقی متغیر هوش است که به سطح هوش فرد بستگی دارد. در تحقیق، متغیرهایی که نتیجه یک متغیر متقدم و یا وابسته به آن هستند، متغیرهای وابسته نامیده می‌شوند. متغیرهایی که نسبت به متغیرهای وابسته مقدم هستند، متغیرهای مستقل نامیده می‌شوند. برای مثال: بلندی قد یک کودک (متغیر وابسته) در حد مشخصی به سن او (متغیر

1- Categorical

2- Continuous

3- Dichotomous

4- Independent

5- Dependent

مستقل) بستگی دارد. اغلب این اصطلاحات را، حتی در فقدان دلایل تجربی و یا نظریه‌ای برای داشتن تقدم و تأخر متغیرها، به کار می‌بریم. این متغیرها برای روشن کردن سیر پیش‌بینی - از وضعیت افراد در متغیر مستقل تا وضعیت آنها در متغیر وابسته - نیز به کار برده می‌شوند.

یک متغیر وابسته پدیده‌ای است که موضوع یک مطالعه و یا تحقیق به حساب می‌آید. متغیر مستقل پارامتری است که به شیوه‌ای قابل محاسبه از متغیر وابسته قابل تفکیک و تشخیص است، ولی ممکن است با متغیرهای وابسته نیز مرتبط باشد. بسیاری از پارامترها می‌توانند به عنوان متغیرهایی مستقل، که از جنبه‌های مختلف محیطی قابل تفکیک هستند، عمل کنند، مانند سطح طبقاتی، محیط خانه و شرایط کلاس درس. بعلاوه، خصیصه‌های فردی، مانند، سن، جنس، هوش و انگیزه نیز خود به خود می‌توانند متغیرهای مستقلی باشند که با متغیر وابسته رابطه دارند.

از این به بعد در هر کجا از روشهای آزمایشی در تحقیق بحث می‌کنیم، متغیر مستقل را به عنوان پارامتری تعریف خواهیم کرد که از طرف آزمایشگر مورد دستکاری^(۱) قرار گرفته است یا تغییر داده می‌شود. هر متغیر بر اساس آنکه کدامیک از اثرات دستکاری شده آن مورد مشاهده قرار گیرد، متغیر وابسته به شمار می‌رود. این متغیر را به آن دلیل وابسته می‌نامند که مقدارش وابسته به متغیر مستقل است و با مقدار آن تغییر می‌کند. برای مثال، به منظور بررسی اثر حل تمرینات با کمک کامپیوتر در پیشرفت ریاضی، محقق روش آموزش، متغیر مستقل را دستکاری می‌کند، و سپس اثر آن را در پیشرفت ریاضی، متغیر وابسته، مورد مشاهده قرار می‌دهد. بعد از آنکه رابطه بین متغیرها در جریان تحقیق معین شد، محقق می‌تواند از یک متغیر مستقل، متغیر وابسته را پیش‌بینی کند. در تحقیق در تعلیم و تربیت، روشهای تدریس و رویه‌های مختلف احتمالاً بیش از همه عوامل به عنوان متغیرهای مستقل مورد استفاده قرار گرفته‌اند. عوامل دیگری از این دست را سن، جنس، طبقه اجتماعی، نگرشها، هوش و انگیزش تشکیل می‌دهند. یکی از متداولترین متغیرهای وابسته پیشرفت تحصیلی و یا یادگیری است.

ممکن است متغیری در یک مطالعه متغیر مستقل و در مطالعه دیگر متغیر وابسته باشد؛ چرا که یک متغیر بر مبنای هدف مطالعه به عنوان متغیر مستقل و یا وابسته مورد توجه قرار می‌گیرد.

اگر اثر انگیزش بر پیشرفت مورد بررسی قرار گیرد، آنگاه انگیزش به عنوان متغیر مستقل به حساب خواهد آمد. به هر حال، اگر کسی به تعیین اثر روشهای آزمون کردن، گروه‌بندی کلاس، یا روشهای نمره‌گذاری بر انگیزه دانش‌آموزان راغب باشد، آن وقت است که انگیزش، یک متغیر وابسته خواهد بود و به طور کلی زمانی که خواهان بررسی اثر هوش بر یادگیری به عنوان متغیری وابسته هستیم، هوش متغیر مستقل قلمداد می‌شود و بر همین منوال در مطالعات تحقیقاتی در مورد اثر رفتن به کودکستان در رشد قوای ذهنی کودکان، هوش یک متغیر وابسته است.

دسته‌بندی دیگری از متغیر مستقل می‌تواند بر اساس قابلیت دستکاری شدن آن متغیرها از طرف محقق صورت گیرد. بر طبق این تقسیم‌بندی، دو نوع متغیر وجود دارد: فعال^(۱) و خصیصه‌ای^(۲). یک متغیر فعال به عنوان متغیری تعریف می‌شود که محقق می‌تواند آنرا مستقیماً دستکاری کند. برای مثال، روش تدریس، روش گروه‌بندی و روشهای تقویت، همگی متغیرهایی هستند که می‌توان آنها را مورد دستکاری قرار داد و بر اساس متغیرهای فعال خوانده می‌شوند.

یک متغیر خصیصه‌ای متغیری است که محقق نمی‌تواند آنرا مورد دستکاری قرار دهد. اینگونه متغیرها، که گاهی اوقات متغیرهای مقرر^(۳) نیز خوانده می‌شوند، جزو خصوصیات افراد محسوب می‌شوند و به خواست محقق قابل دستکاری نیستند. برای مثال، استعداد، جنس، نژاد، سن و طبقه اجتماعی نمونه‌های بارزی از متغیرهای مقرر هستند. یک محقق با تخصیص آزمودنی‌ها به گروههایی مبتنی بر این گونه متغیرهای از پیش تعیین شده، متغیرهای مقرر را در تحقیق خود به کار می‌گیرد.

در برخی شرایط گفتن آنکه کدام متغیر بر دیگری اثر می‌گذارد، ممکن نیست. برای مثال، هنگامی که رابطه بین جزم اندیشی^(۴) و نگرش سیاسی مورد بررسی قرار می‌گیرد، محقق نمی‌تواند تعیین کند کدامیک از متغیرها مستقل و کدامیک وابسته هستند. این گونه موقعیتها برای انجام آزمایش مساعد نیستند. رابطه بین چنین متغیرهایی اغلب در بررسیهای همبستگی مورد تحقیق قرار می‌گیرند.

/// روش شناسی تحقیق در تعلیم و تربیت

روش تحقیق به راهبرد کلی، که گردآوری و تجزیه و تحلیل داده‌های ضروری برای پاسخگویی به سؤال مورد نظر را دنبال می‌کند، اطلاق می‌شود. روش تحقیق، برنامه رویارویی با مسئله مورد بررسی است. پنج طبقه‌بندی متداول از تحقیقات آموزشی عبارتند از:

۱- آزمایشی: یک نوع بررسی علمی است که در جریان آن، محقق یک یا چند متغیر مستقل را دستکاری و کنترل می‌کند و تغییرات حاصله از این دستکاری را در متغیر یا متغیرهای وابسته مورد مشاهده قرار می‌دهد.

مثالها

اثرات تمرین برنامه‌نویسی کامپیوتری لوگو در حل مسئله و توانائی تجسم روابط فضایی
اثر رئوس مطالب و سرفصلها در یادآوری متن از طرف خواننده

۲- علی - مقایسه‌ای: ^(۱) این طبقه‌بندی مشابه روش آزمایشی است، بجز آنکه در این روش محقق نمی‌تواند به طور مستقیم متغیرهای مستقل را دستکاری کند.

مثالها

رابطه اعتماد به نفس و عملکرد تحصیلی با پشتکار
رابطه بلوغ جنسی با توانایی تجسم فضایی و مهارت کارکردی در دو نیمکره مغز در زنان
بزرگسال

۳- توصیفی: هدف اصلی در این روش بیان چگونگی وضعیت است. در تحقیق توصیفی زیرتقسیمات متعددی وجود دارد: (الف) پیمایشها (ب) بررسیهای رشد ^(۲) (ج) بررسیهای پیگیری (د) تجزیه و تحلیل اسنادی ^(۳) (ه) مطالعه روند ^(۴) و (و) بررسیهای همبستگی. ^(۵)

مثالها

یک پیمایش راجع به نیازهای یک معلم علوم دوره دوم

الگوهای تغییر در نمرات آزمونهای استعداد کلاسی در بین نوجوانان دارای استعداد تحصیلی ۴- کیفی: در این نوع بررسی، محقق اشخاص یا رویدادها را در محیط طبیعی آنها مورد مشاهده قرار می‌دهد. هدف اصلی اینگونه بررسی پی‌بردن به اثر شرایط ویژه بر رویدادهاست. این روش شناسی شامل مورد پژوهی‌ها و انواع دیگر روشهای طبیعت‌گرایانه است.

مثالها

یک بررسی در مورد کلاس چهارم موسیقی با استفاده از روش مشاهده. آمادگی حاصل از رفتن به کودکستان، و به خاطر سپاری: یک بررسی کیفی در مورد باورها و شیوه کار آموزگاران.

۵- تاریخی: این مقوله شامل گردآوری داده‌های دوران گذشته برای درک رویدادهای آن زمان و معانی ضمنی آنها برای رویدادهای این زمان است. هدف اصلی آن، بیان آنچه که رخ داده است، می‌باشد.

مثالها

تعلیم و تربیت فرقه مذهبی کویکرها در آمریکای قرن نوزدهم
سیاستگذاریهای آموزشی در دوره زمامداری نیکسون

/// خلاصه

بشر تلاش کرده است تا از طریق تجربه، منبع آگاه و مطلع، استدلال قیاسی، استدلال استقرایی، و روش علمی به سوالات خود پاسخ گوید. هر روشی فرضیات معینی را طلب می‌کند. صحت پاسخها بستگی به فرضیات زیربنایی روشهای مورد استفاده دارد. روش علمی به دو فرض بنیادین متکی است: (۱) حقیقت می‌تواند از مشاهده حاصل شود، و (۲) هر پدیده‌ای از روابط قانونمند پیروی می‌کند.

تحقیقات علمی حقایق مطلق را جستجو نمی‌کنند، بلکه به دنبال نظریه‌هایی هستند که پدیده را طبق روشی معتبر تبیین و پیش‌بینی کنند. این روشها نظریه‌هایی را جستجو می‌کنند که نه تنها موجز، قابل آزمون و پایدار باشند، بلکه به نوبه خود محرکی برای تحقیقات بعدی

به شمار آیند. رویکرد علمی متضمن ویژگی خوداصلاحی^(۱) است؛ زیرا هر نظریه‌ای آزمایشی تلقی می‌شود و احتمالاً می‌تواند با نظریه‌ای جدیدتر و درخورتر با معیارهای موجود، تعویض شود.

رویکرد علمی طی قرون متمادی جهت توضیح، پیش‌بینی، و کنترل پدیده‌های فیزیکی به کار برده شده است، اما در تعلیم و تربیت به تازگی مورد استفاده قرار گرفته است. پیچیدگی متغیرهای آموزشی و مشکلات موجود بر سر راه مشاهدات معتبر، مانع انجام تحقیق علمی در تعلیم و تربیت بوده است؛ هر چند با شروع جنبش در آغاز قرن اخیر، جستجوی علمی در تعلیم و تربیت، هم در تحقیقات نظریه‌ای و هم در بررسی‌هایی عملی، از پذیرش عام برخوردار شد و به موفقیت‌های بیشتری دست یافت.

مراحل ویژه تحقیق در تعلیم و تربیت عبارتند از: (۱) انتخاب یک مسئله، (۲) صورتبندی یک فرضیه، (۳) انتخاب راهبرد تحقیقاتی و تهیه ابزار تحقیق، (۴) گردآوری، تجزیه و تحلیل، و تفسیر داده‌ها و (۵) ارائه یافته‌ها از طریق گزارش نتایج بررسی.

تحقیق در زمینه تعلیم و تربیت را می‌توان به دو دسته اصلی تقسیم کرد: بنیادی و کاربردی. گرایش اصلی تحقیق بنیادی گسترش مرزهای دانش و کشف قوانین عام است. هدف اصلی تحقیق کاربردی حل فوری مسائل عملی است.

در سطح نظریه‌ای، صاحب‌نظران حیطه تعلیم و تربیت از اصطلاحاتی مانند هوش، خلاقیت، توانایی حل مسئله، و انگیزه استفاده می‌کنند که تجربدهای حاصل از مشاهده رفتارهای معین می‌باشند. این اصطلاحات به عنوان سازه‌ها تلقی می‌شوند. در هر بررسی سازه‌ها کمی شده هستند و مقادیر متنوعی به خود می‌گیرند. به این ترتیب آنها در زمره متغیرها تلقی می‌شوند. دو نوع متغیر اصلی وجود دارد: وابسته و مستقل. اگر یک متغیر بر دیگری مقدم باشد، متغیر مستقل خوانده می‌شود، ولی اگر پیامد متغیر دیگر باشد، متغیر وابسته قلمداد می‌شود.

در تحقیق، این موضوع که متغیرها به طور عملی تعریف شوند، امری اساسی به شمار می‌رود. دو راه برای تعریف عملی متغیرها وجود دارد: (۱) به کار بردن بعضی از انواع اندازه‌گیری یا (۲) فهرست کردن مراحل انجام یک آزمایش برای خلق شرایط مطمئنه. تحقیق نوع اول را تعریف کاربردی اندازه‌گیری شده و نوع دوم را تعریف کاربردی آزمایشی می‌نامند.

// مفاهیم اصلی

استدلال استقرایی	استدلال قیاسی
استقرای ناقص	اصل موجز بودن
تحقیق بنیادی	تحقیق کاربردی
تعریف ساختی	تعریف کاربردی
تعریف کاربردی آزمایشی	تعریف کاربردی اندازه گیری شده
ثابت (عدد ثابت)	رویکرد علمی
سازه‌ها	فرضیه
متغیر	متغیر پیوسته
متغیر دو وجهی	متغیر مستقل
متغیر مستقل خصیصه‌ای	متغیر مستقل فعال
مدل	متغیر مقوله‌ای
متغیر وابسته	مفهوم
نظریه	

// تمرینها

- ۱- در تمرینهای ذیل منابع علم - استدلال قیاسی، استدلال استقرایی، یا روش علمی - را که به طور چشمگیری مورد استفاده قرار گرفته‌اند، معین کنید:
 الف - بعد از مشاهده وسیع واکنشها، لاوازیه نتیجه گرفت که سوختن فرایندی است که طی آن یک ماده مشتعل با اکسیژن ترکیب می‌شود. نتیجه کار او ضربت کشنده‌ای بر نظریه قدیمی قابلیت اشتعال^(۱) در این زمینه وارد آورد.
 ب - دالتون بعد از اندیشه بسیار، نتیجه گرفت که اجسام بایستی از ذرات کوچکی به نام اتم تشکیل شده باشند. فرضیات اولیه او پایه‌ای برای نظریه اتمی شد.
 ج - پس از آنکه دانشمندان فرضیات دالتون را پذیرفتند، از آنها قیاس کردند و شروع به گردآوری اطلاعاتی کردند تا این فرضیات را تأیید کنند. مدارکی برای پشتیبانی از نظریه اتمی یافته شد.

د - انیشتین با این شناخت که مواد رادیو اکتیو همواره بدون آنکه ظاهراً از جرمشان کاسته شود ذرات انرژی را بیرون می دهند، فرمول $E = mc^2$ را برای تبدیل اجسام به انرژی ابداع کرد.

ه - فرمی^(۱) با پذیرش نظریه انیشتین، دست به تجربیاتی زد که در شکافتن اتم نتیجه بخش بود.

و - پس از بررسی نظریه تقویت، یک مدرس فرضیه می دهد که استفاده از برنامه کامپیوترهای شخصی باعث مهارت بیشتری در حساب می شود. او یک بررسی را ترتیب داد که طی آن دو گروه از دانش آموزان کلاس شش از کامپیوتر شخصی استفاده می کردند، در حالی که دو دسته دیگر از مواد درسی متداول استفاده می کردند.

۲- نقش نظریه در بررسی علمی چیست؟

۳- چه تفاوتی بین یک نظریه استقرایی و یک نظریه قیاسی وجود دارد؟

۴- بر اساس عنوان بررسی، تحقیقات ذیل را به بنیادی و کاربردی دسته بندی کنید:

الف - اثر تزریق آر آن ای (اسید ریبونوکلیک) در انتقال مهارتها در حیوانات تربیت شده نسبت به حیوانات تربیت نشده

ب - نتایج برنامه اصلاحی درس حساب

ج - شرطی کردن به عنوان یک عملکرد فاصله گذاری بین محرکات شرطی شده و اصلی

د - تدریس هندسه برای پرورش تفکر پویا: یک بررسی آزمایشی

۵ - با استفاده از عناوین فعال، خصیصه ای، یا احتمالاً هر دو، متغیرهای ذیل را طبقه بندی کنید:

الف - مقدار داروی تجویز شده

ب - اضطراب

ج - زمینه اقتصاد اجتماعی

د - روش تدریس

۶ - در بیانیه های ذیل کدامیک از نگرشهای ویژه دانشمندان، نقض شده است:

الف - این بررسی عهده دار اثبات این است که استفاده از ماری جوآنا برای مهارتهای دانشگاهی مضر است.

ب - این موضوع به طور قطع ثابت شده است.

ج - نتایج نشان می‌دهد که ماری جوآنا زیانبار است.

۷- ویژگیهای یک نظریه سودمند کدام است؟

۸- در یک بررسی که برای تعیین اثر انواع مختلف محرومیت از خواب در یادگیری هجاهای

بی معنی طراحی شده است:

الف - متغیر مستقل کدام است؟

ب - متغیر وابسته کدام است؟

۹- متغیرهای ذیل را بر اساس نوع مقوله‌ای یا پیوسته طبقه‌بندی کنید:

الف - پیشرفت تحصیلی

ب - روش هجایی خواندن در مقابل روش بین - بگو در خواندن

ج - مکالمه اسپانیایی، مکالمه انگلیسی، مکالمه فرانسوی

د - مهارت عضلانی

ه - قریحه موسیقی

۱۰- ویژگیهای تعاریف کاربردی کدام است؟

۱۱- کدام نوع از تعریف کاربردی (اندازه‌گیری شده یا آزمایشی) برای هر یک از متغیرهای

ذیل مناسبتر است؟

الف - تقویت

ب - پیشرفت تحصیلی

ج - نگرش

د - روش تدریس

۱۲- شما آموزگاری هستید که به میزان رفتار پرخاشگرانه قابل مشاهده در بین کودکان

علاقه‌مند بوده‌اید. با اولیای دانش‌آموزان درباره روش تربیتی (استفاده از تنبیه، تشویق و امثالهم)

آنها در ارتباط با فرزندشان و مجموع زمانی که کودکان تلویزیون تماشا می‌کنند، مصاحبه

می‌نمایید. اطلاعات زیادی گردآوری کرده، اما احساس می‌کنید که در این موارد یکنواختی

وجود ندارد. طبق روش علمی، به احتمال قوی گام بعدی شما چه خواهد بود؟

۱۳- کدام روش تحقیقاتی (آزمایشی، علی - مقایسه‌ای، یا پیمایشی) به طور مؤثری به هر

یک از سؤالات شما پاسخ می‌گوید؟

- الف - آیا کودکانی که صبحانه می‌خورند، نمرات بهتری در مدرسه به دست می‌آورند؟
- ب - آیا یک تغذیه مناسب، عادات صبحانه خوردن را در کودکان تغییر می‌دهد؟
- ج - چه تعدادی از کودکان در مدرسه گزارش می‌کنند که در خانه صبحانه نمی‌خورند؟
- د - آیا برقرار کردن یک برنامه آزاد برای صرف صبحانه در مدرسه تفاوتی در مهارت‌های دانش‌آموزان ایجاد خواهد کرد؟
- ۱۴- بر اساس عناوین، هر یک از بررسی‌های ذیل را بر اساس مناسبترین روش مورد استفاده در تحقیق طبقه‌بندی کنید:
- الف - تفاوت‌های مبتنی بر جنسیت در عملکرد درس ریاضی خود را نشان می‌دهند.
- ب - تأثیر بازخورد آزمون کامپیوتری در جدی بودن اشتباهات بعدی
- ج - دیدگاه‌های دانش‌آموزان کالج و ارزشیابی یک استاد ایده آل
- د. اثر فقدان زودرس پدر در استعداد‌های تحصیلی
- ه - یک دبیرستان منتخب: یک بررسی عمیق
- ۱۵- مثالهایی از کاربرد منبع مطلع و تجربه به عنوان منابع دانش بیاورید.
- ۱۶- یک مثال از اینکه چگونه تحقیق بنیادی در علوم زیست شناختی باعث رشد حرفه پزشکی شده است، بزنید.
- ۱۷- یک مثال از اینکه چگونه تحقیق بنیادی در یادگیری باعث تکامل شیوه تدریس شده است، بزنید.
- ۱۸- یک مثال از تحقیق کاربردی در زمینه مورد علاقه خود بزنید. زمینه‌های دیگری را نیز فهرست کنید که نیاز به انجام تحقیقات اضافی دارند.
- کدامیک از متغیرها، می‌باید در چنین تحقیقاتی مورد بررسی قرار گیرند.

/// پاسخها

۱- الف - استدلال استقرایی

ب - استدلال قیاسی

ج - روش علمی

د - روشن علمی

۲- نظریه، یافته‌ها را تلفیق و اطلاعات را خلاصه می‌کند، و خط مشی‌هایی برای تحقیقات جدید فراهم می‌آورد، و ما را قادر می‌سازد تا پدیده‌ها را تبیین و پیش‌بینی کنیم.

۳- یک نظریه استقرایی به تبیین مشاهدات پیشین کمک می‌کند، در حالی که یک نظریه قیاسی، پیش از انجام مشاهدات وسیع، ساخته می‌شود.

۴- الف - بنیادی

ب - کاربردی

ج - بنیادی

د - کاربردی

۵- الف - فعال

ب - احتمالاً هر دو

ج - خصیصه‌ای

د - فعال

۶- الف - دانشمند، واقع‌گرا و بی‌طرف است.

ب - دانشمند شکاک است و یافته‌ها را به عنوان نتایجی آزمایشی تلقی می‌کند.

ج - دانشمند با واقعیات و نه ارزشها، سر و کار دارد.

۷- یک نظریه مفید، پدیده را به ساده‌ترین شکل ممکن بیان می‌کند، با مشاهدات مربوط و مجموعه دانش پیشین سازگار است، وسایلی برای اثبات خود فراهم می‌سازد و بررسیهای جدیدی را دامن می‌زند.

۸- الف - میزان محرومیت از خواب

ب - تعداد هجاهای بی‌معنی آموخته شده

۹- الف - پیوسته

ب - مقوله‌ای

ج - مقوله‌ای

د - پیوسته

ه - پیوسته

۱۰- تعاریف قابل قبول، یک طرز عمل واضح برای تعیین جوهر پدیده و گستردگی آن ارائه می‌کنند.

۱۱- الف - آزمایشی

ب - اندازه گیری شده

ج - اندازه گیری شده

د - آزمایشی

۱۲- برای صورتبندی یک نظریه در مورد پرخاشگری در بین کودکان تلاش کنید.

۱۳- الف - علی - مقایسه‌ای

ب - آزمایشی

ج - پیمایشی

د - آزمایشی

۱۴- الف - علی - مقایسه‌ای

ب - آزمایشی

ج - توصیفی

د - علی - مقایسه‌ای

ه - کیفی

۱۵- پاسخها متفاوت خواهند بود.

۱۶- پاسخها متفاوت خواهند بود.

۱۷- پاسخها متفاوت خواهند بود.

۱۸- پاسخها متفاوت خواهند بود.

بخش دوم

زمینه تحقیق

فصل دوم

موضوع تحقیق

هدفهای آموزشی

دانشجو پس از مطالعه این فصل، قادر خواهد بود:

- ۱- یک مسئله تحقیق را بیان کند.
- ۲- منابع بالقوه را در مورد مسائل مربوط به تحقیق در تعلیم و تربیت، تعیین کند.
- ۳- معیارهایی را برای ارزشیابی یک مسئله تحقیق بیان کند.
- ۴- با استفاده از معیارهای قابل قبول، مسئله مورد نظر را ارزشیابی کند.
- ۵- یک مسئله کلی را در زمینه مورد علاقه خود در تعلیم و تربیت انتخاب و آن را در شکل خاصی که برای بررسی تجربی آماده باشد، بیان کند.
- ۶- اصطلاحاتی مانند جامعه آماری و متغیرها را که در تحقیق مورد استفاده قرار می گیرند، تعریف کند.
- ۷- جامعه آماری و متغیرهای مورد استفاده در تحقیق مورد نظر را شناسایی کند.

تحقیق نظام‌مند با یک مسئله آغاز می‌شود. جان دیویی اولین قدم در روش علمی را تشخیص یک مشکل، یک مانع یا مسئله می‌داند که محققان را سردرگم می‌کند. انتخاب و صورتبندی یک مسئله یکی از مهمترین جنبه‌های تحقیق در هر زمینه است. محققان تازه‌کار اغلب از کشف این امر که این مرحله مقدماتی، بخش عمده‌ای از زمان کل پژوهش را در یک پروژه تحقیقاتی به خود اختصاص می‌دهد، شگفت زده می‌شوند. برای انجام تحقیق هیچ راهی بجز تشخیص یک مسئله، تفکر در مورد آن و صورتبندی آن به روشی سودمند وجود نخواهد داشت.

محقق در اولین وهله ضرورتاً می‌باید در مورد حیطه وسیعی که در آن به تحقیق خواهد پرداخت، تصمیم بگیرد. محقق باید در زمینه مورد بررسی، دانش و یا تجربه داشته باشد؛ به این ترتیب که سؤالاتی را طرح کند که با تحقیق بتوان به آنها پاسخ گفت. ما اغلب می‌شنویم که دانشجویان در درسهای مشکل اظهار می‌دارند که: «من به قدر کافی نمی‌دانم تا بتوانم سؤالی مطرح کنم.» همچنین، محقق بدون داشتن دانش یا تجربه مورد نظر در زمینه‌ای خاص نخواهد دانست که دانش اضافی مورد نیاز چیست یا چگونه می‌توان آن را در جریان انجام بررسیهای تجربی کسب کرد. بعلاوه، حیطه موضوع انتخابی برای بررسی باید از جمله موضوعات عمیقاً جالب باشد و یا در مورد آن کنجکاوی واقعی وجود داشته باشد. انتخاب موضوع ضرورتاً باید بسیار شخصی باشد. در غیر این صورت، شاید انگیزه محقق تا پایان کار تحقیق پایدار نماند. یک آموزگار دوره ابتدایی ممکن است به یافتن راه مؤثرتری در تدریس خواندن علاقه‌مند باشد. یک معلم انگلیسی دوره دبیرستان شاید بخواهد بداند که آیا استفاده از دستگاه واژه‌پرداز^(۱) در نگارش دانشجویانش مؤثر خواهد بود یا خیر.

بعد از انتخاب مسئله، موضوع کلی به یک بیان مشخص از موضوع بررسی محدود می‌شود. بیشتر محققان تازه‌کار صورتبندی یک موضوع و یا مسئله قابل بررسی را مشکل می‌یابند. این امر به کمبود مسائل قابل بررسی در تعلیم و تربیت مربوط نمی‌شود. در واقع سؤالات نیازمند پاسخگویی، به قدری زیاد هستند که محقق اغلب در انتخاب یکی از میان آنها دچار زحمت می‌شود. یکی از مشکلات متداول این است که در اولین مرحله می‌باید مسئله تحقیق انتخاب و یک سؤال نیز صورتبندی شود، در حالی که درک محقق تازه‌کار از چگونگی انجام یک بررسی

بسیار محدود است. بعلاوه، عدم اطمینان از ماهیت مسئله‌های تحقیقاتی، محدود کردن مسئله، معیارهای پذیرش آن و چگونگی حل مسئله اغلب توانفرسا به نظر می‌رسد. حتی پژوهشگران با تجربه نیز اغلب این مسئله را ضروری می‌دانند که قبل از رسیدن به یک مسئله تحقیقاتی، برای دستیابی به معیارهای پذیرفته شده تلاشهای متعددی کنند. یک انتخاب یا یک صورتبندی اولیه را می‌توان با بررسی دقیق‌تر، نامعقول یا فاقد ارزش یافت. مهارت در انجام تحقیق تا حد زیادی بستگی به انتخاب عاقلانه زمینه مورد تحقیق دارد. مهارت در انجام تحقیق به زمان و تلاشهای مضاعفی نیاز دارد، ولی محققان مشتاق تازه کار می‌توانند به آن دست یابند.

بیان مسئله می‌باید بروشنی متضمن آن چیزی باشد که مورد بررسی قرار می‌گیرد. گزاره اصلی می‌تواند به صورت اخباری^(۱) یا سؤالی مطرح شود. به سخن دیگر، بیان مسئله باید متضمن متغیرهای مورد نظر و رابطه و ویژه بین متغیرهای مورد مطالعه باشد. برای مثال، مسئله تحقیق معلم دبیرستان یاد شده می‌تواند چنین بیان شود: «مسئله مورد بررسی در این پژوهش اثر بکارگیری واژه‌پرداز در نگارش انگلیسی دانش‌آموزان دوره دبیرستان است» یا «استفاده از واژه‌پرداز چه اثری در نگارش انگلیسی دانش‌آموزان دوره دبیرستان دارد؟» بسیاری از محققان شکل سؤالی را ترجیح می‌دهند، چون به نظر می‌رسد جهت‌گیری روشن‌تری در قبال یافتن پاسخ دارد و قابل قبول نیز هست.

هر چند بعید به نظر می‌رسد، ولی به مجرد آنکه حیطه موضوع برگزیده و یک سؤال یابیانیه بروشنی صورتبندی شود، یکی از مشکل‌ترین مراحل فرایند تحقیق انجام یافته است.

/// ماهیت مسائل

اگر چه در تعلیم و تربیت انواع مسائل قابل تحقیق وجود دارد، ولی همه آنها متضمن جستجو برای یافتن دانش مورد نیاز در این حیطه هستند. مسئله تحقیق روشن می‌کند که محقق چه می‌خواهد بداند و روش مورد استفاده برای کسب این آگاهی را نیز معین خواهد کرد. تحقیق آزمایشی به دنبال سوالات مربوط به روابط موجود بین متغیرهایی است که محقق می‌تواند بر همه آنها کنترل داشته باشد. در تحقیق آزمایشی محقق ضمن دستکاری متغیر مستقل

(علت) همه متغیرهای دیگر را کنترل می‌کند و هر یک از تغییرات بعدی را در متغیرهای وابسته (اثر) مورد مشاهده قرار می‌دهد. کنترل در یک آزمایش، حذف تبیین‌های جایگزین برای نتایج مشاهده شده را برای محقق امکان‌پذیر می‌سازد. اگر محقق بخواهد اثر روش آموزش را در یادگیری ترکیبات جمع، در بین کودکان دبستانی مورد بررسی قرار دهد، می‌تواند ضمن کنترل همه عوامل دیگری که می‌تواند در یادگیری ترکیبات جمع دخالت داشته باشد، روش آموزش را دستکاری کند و سپس اثر این روش را در یادگیری مورد مشاهده قرار دهد.

تحقیق علی - مقایسه‌ای همچنین در پی تعیین روابط موجود بین دو یا چند متغیر است، اما در اینگونه تحقیقات دستکاری متغیر مستقل ممکن نیست؛ چون متغیر مستقل قبلاً رخ داده است. بنابراین محقق قادر نیست در مورد چگونگی بکارگیری متغیر مستقل تصمیم‌گیری کند. اگر یک محقق بخواهد اثر طلاق اولیا را بر پیشرفت تحصیلی کودکان بررسی کند، مجبور خواهد بود که پیشرفت تحصیلی گروه‌های از پیش شکل یافته را با هم مقایسه کند: یک گروه از کودکانی که پدر و مادرشان از هم جدا شده‌اند، با گروه دیگری که پدر و مادرشان با هم زندگی می‌کنند. تحقیق توصیفی سؤالاتی را در مورد ماهیت، وقوع، پراکندگی متغیرهای آموزشی و یا روابط موجود بین این متغیرها مطرح می‌کند.

در این نوع تحقیق هیچ نوع دستکاری در متغیرها صورت نمی‌گیرد - تنها به توصیف متغیرها و روابط بین آنها که به صورت طبیعی اتفاق می‌افتد، پرداخته می‌شود. یک مثال از این نوع مسئله تحقیق این سؤال است که «دانش آموزان دوره اول دبیرستان در مورد فرایندی که طی آن احزاب سیاسی کاندیداهای خود را برای ریاست جمهوری انتخاب می‌کنند، چه می‌دانند؟» یا «چه رابطه‌ای بین تعداد سالهای فراگیری زبان فرانسه در دبیرستان و تعیین دوره درسی دانشجویان درس زبان فرانسه در دانشگاه وجود دارد؟»

/// منابع مسائل و موضوعات

اولین سؤالی که اکثر دانشجویان مطرح می‌کنند، این است که: «چگونه یک مسئله تحقیق پیدا کنیم؟» اگر چه ضابطه مشخص برای تعیین یک مسئله وجود ندارد، ولی پیشنهادهاى مشخصی می‌توان یافت که به این امر کمک می‌کند. مهمترین منابع در تعیین مسائل و موضوعات عبارتند از: تجربه و قیاس از نظریه و متون تحقیقاتی پیشین.

// تجربه

یکی از سودمندترین منابع برای محقق تازه کار، تجربیات او در زمینه کار آموزشی است. او هر روز ملزم به تصمیم‌گیری در مورد اثرات احتمالی تجربیات آموزشی بر رفتار دانش‌آموزان است. اگر قرار است این تصمیمات درست باشد، کارشناسان تعلیم و تربیت می‌باید برای اطمینان از اعتبار مفروضات خود در مورد رابطه بین تجربیات یادگیری و تغییر در دانش‌آموزان، بررسی‌های انتقادی انجام دهند.

درباره روشها و تکنیکهای تدریس، تصمیمات مختلفی باید اتخاذ شود. مطمئناً روشهای تدریس، مستعد و نیازمند بررسیهای علمی است. رویکرد علمی در بررسیهای آموزشی متضمن آن است که تصمیم‌گیری در مورد نحوه انجام امور در تعلیم و تربیت مبتنی بر شواهد تجربی باشد، نه بر حدس و گمان، احساسات و جزم‌اندیشی. برای مثال، یک معلم دوره ابتدایی ممکن است در مورد اثر روش خود در تدریس قرائت سؤال کند. او ممکن است بخواهد روش معمول خود، یا هر روش شناخته شده دیگری را برای آگاهی از آنکه کدام یک مؤثرتر است، ارزشیابی کند، یا معلم دوره دبیرستان سؤال می‌کند که برای تدریس درس اقتصاد، روش مباحثه مؤثر است یا روش مقاله نویسی. استاد زیست‌شناسی ممکن است بخواهد در مورد اثر شبیه‌سازی از طریق کامپیوتر در مهارتهای حل مسئله از طرف دانشجویان تحقیق کند و یا این سؤال را مطرح کند که آیا شبیه‌سازی کامپیوتری از روشهای دیگر مؤثرتر است؟

مشاهده روابط مشخص برای مواردی که هنوز هیچ تبیین رضایت‌بخشی در مورد آنها وجود ندارد، یکی دیگر از منابع مسائل برای تحقیق است. معلم ممکن است دریابد که در زمانهای خاصی علائم آشکاری از اضطراب در دانش‌آموزان بروز می‌کند. این معلم، برای بررسی مسئله، می‌تواند تبیین‌های آزمایشی متعددی را در مورد منشأ این اضطراب طرح کند و سپس آنها را به طور تجربی مورد آزمون قرار دهد. این بررسی نه تنها مسئله آنی را حل می‌کند، بلکه سهم کوچکی نیز در درک عوامل مؤثر در اضطراب مربوط به کلاس درس ایفا می‌کند.

همچنین باید تصمیماتی درباره شیوه‌های معمول در مدارس گرفته شود. بسیاری از این شیوه‌ها اساساً بر سنتها یا نظریات گذشته‌گان بنا شده‌اند؛ در حالی که یا تایید علمی بسیار کمی پشتوانه آنها است و یا اصلاً چنین تاییدی وجود ندارد.

پس چرا نباید برخی از این شیوه‌ها را ارزشیابی کرد. برای مثال، چرا نباید برنامه آزمون

سالیانه را ارزشیابی کرد. آیا آزمونهای جایگزینه^(۱) معتبرتری از آنچه در حال حاضر مورد استفاده قرار می‌گیرد، برای این منظور خاص وجود دارند؟

بنابراین تجربیات روزمره متخصص تعلیم و تربیت می‌تواند مسائل ارزشمندی برای تحقیق فراهم آورد. در واقع، بیشترین ایده‌های تحقیقاتی، که به وسیله پژوهشگر تازه کار در تعلیم و تربیت ارائه می‌شود، ناشی از تجربیات شخصی او است. یک محقق ممکن است درباره روابط تازه و یا شیوه‌های مختلف دستیابی به هدفهای معین تصوّراتی داشته باشد، بنابراین، از طریق نوعی فرآیند شهودی، به ایده‌های خود برای تحقیق دست پیدا می‌کند؛ اگر چه حقیقت این است که این بررسیها عمدتاً نوعی راهبرد برای حل مسائل آتی هستند، با این حال این مسائل اغلب برای محققان تازه کار مفیدتر و پرمعنا تر از مسائلی هستند که از طریق فرایند منطقی قیاس از نظریه به دست می‌آیند. بعلاوه آنها می‌توانند بر اساس سهمشان در پیشرفت تجربیات تعلیم و تربیتی تایید و توجیه شوند.

// قیاس از نظریه

قیاسها که می‌توانند بر مبنای نظریه‌های متعدد آموزشی و رفتاری که محقق با آنها آشناست، به عمل آیند، یک منبع عالی برای مسائل تحقیقاتی فراهم می‌آورند. نظریه‌ها با اصول کلی سروکار دارند که در مسائل خاص تعلیم و تربیت، مادامیکه در عمل مورد تایید قرار نگرفته‌اند، تنها به صورت فرضیه باقی می‌مانند. تنها در جریان تحقیق است که محقق تعیین می‌کند که آیا تعمیم‌های موجود در نظریه‌ها، قابل تبدیل به پیشنهادهای ویژه‌ای برای فعالیتهای آموزشی هستند یا خیر.

محقق، از یک نظریه می‌تواند فرضیاتی را استنتاج کند که یافته‌های مورد انتظار را در یک موقعیت کاری ویژه بیان می‌کنند.

بنابراین، محقق این پرسش را مطرح می‌کند که اگر نظریه بدرستی وضعیت موجود را خلاصه کرده باشد، چه روابطی بین متغیرها مشاهده خواهد شد، و سپس برای یقین از اینکه داده‌های تجربی به دست آمده، فرضیه و به دنبال آن نظریه را تایید می‌کند، یا خیر، دست به تحقیق نظام‌مند خواهد زد.

نظریه‌هایی در زمینه یادگیری، شخصیت، علوم اجتماعی، توسعه اجتماعی و بسیاری زمینه‌های دیگر وجود دارد که اعتبار، گستره و عمل پذیری آنها می‌تواند به طور سودمندی در موقعیتهای آموزشی، آزمون شود. نظریه تقویت بویژه می‌تواند به عنوان نقطه شروع در تحقیقات مربوط به شرایط کلاس درس مورد استفاده قرار گیرد. معانی ضمنی را که از یکی از مفروضه‌های نظریه تقویت می‌توان قیاس کرد، برای آزمون کلاسی در نظر بگیرید، با این معنی که، تقویت پاسخها موجب افزایش میزان و شدت پاسخها می‌شود. ما می‌دانیم که این نظریه محرک بسیاری از تحقیقات اخیر بوده است، ولی همچنان قیاسهای بسیاری وجود دارند که باید ساخته شوند و در شرایط کلاس مورد آزمون قرار گیرند. برای مثال، در مورد اثر فقدان تقویت آشکار و یا عدم تقویت پاسخهای صحیح دانش‌آموزان در شرایط کلاس درس، تحقیق کافی صورت نگرفته است. ما از طریق مطالعات آزمایشگاهی با حیوانات دریافته‌ایم که عدم تقویت، احتمال دادن پاسخ را کاهش می‌دهد و در نتیجه به خاموشی پاسخ منجر می‌شود. آیا این یافته می‌تواند به کلاس درس بسط داده شود؟ یعنی بر این اساس معلم می‌تواند فرض کند که پاسخهای درست دانش‌آموزان اگر به طور آشکار تقویت نشوند، خاموش خواهند شد؟ در حال حاضر، تحقیقات کافی در جهت آزمون این اصول در کلاس درس در دست نداریم.

در بین نظریه‌های دیگری که به نظر می‌رسد منابع سودمندی برای تحقیق باشند، می‌توان از نظریه اضطراب امتحان ساراسون^(۱)، نظریه ناهماهنگی شناختی فستینگر^(۲)، نظریه‌های متعدد خویشتن، نظریه انگیزه پیشرفت، نظریه قضاوت - اجتماعی، نظریه نقشها و نظریه پدیده‌شناختی یاد کرد.

برای مثال کپنیز از نظریه فرایندهای مقایسه - اجتماعی فستینگر به عنوان مبنایی برای تحقیق در مورد رابطه بین ادراک دیگران و فرایند تغییر خود پنداره استفاده کرد.^(۳) نظریه فستینگر این فرض را در بر دارد که خود سنجی‌ها از طریق انجام مقایسه بین خود و دیگران صورتبندی می‌شوند.^(۴) تعدادی فرضیه از این نظریه قیاس شده‌اند که می‌گویند تغییرات مورد انتظار در خود پنداره در رابطه با ادراک و مقایسه با افراد مهم صورت می‌گیرد. آزمودنی‌های بررسی را ۸۷

1- Sarason, I. (1980). Introduction to the Study of Test Anxiety. In: I. G Sarason (ed). Test Anxiety: Theory, Research, And Applications. Hillsdale, NJ: Erlbaum.

2- Festinger, L. A Theory of Cognitive Dissonance (New York, Har Per & Row)

3- Kipnis, D. M. (1961). Change in Self- Concepts in relation to Perceptions of Others. Journal of Personality, 29, 449-465.

4- Festinger, L. (1954). A Theory of Social Comparison Process. Human Relations, 7, 117- 140.

دانشجویی تشکیل می‌دادند که با یکدیگر در خوابگاه دانشگاه زندگی می‌کردند. ارزیابی آنها از ویژگیهای شخصیتی خود در رابطه با ادراک آنها از بهترین دوستانشان مورد بررسی قرار گرفت. همه فرضیات مورد تایید داده‌ها قرار گرفت، و کپینیز نتیجه گرفت که ادراک میان فردی و خودپنداره شدت به یکدیگر نزدیکند.

این رویکرد نسبت به مسئله‌های تحقیقاتی به مطالعاتی منتج می‌شود که به دلیل مبتنی بودن بر یک نظریه مشترک، به آسانی با هم قابل تلفیق هستند. چنین تحقیقات مرتبط با هم بخصوص به عنوان وسیله‌ای برای گسترش دانش در حیطه‌ای مشخص ثمربخش است.

// متون تحقیقاتی پیشین

یکی دیگر از منابع موضوعات پژوهشی، متون تحقیقاتی موجود در زمینه مورد علاقه محقق است. به هنگام مطالعه تحقیقات پیشین، در معرض نمونه‌هایی از مسئله‌های تحقیقاتی و شیوه‌های انجام آن تحقیقات قرار می‌گیریم. راههای متعددی وجود دارند که طی آن مرور تحقیقات پیشین می‌تواند در صورتبندی یک مسئله تحقیقاتی جدید مؤثر افتد.

۱- مطالعه تحقیقات پیشین ما را در صورتبندی سؤالات تحقیقاتی که نسبت به بررسیهای گذشته، گام منطقی بعدی تلقی می‌شوند، کمک می‌کند. بسیاری از اوقات نتایج بخشی از یک تحقیق، ما را به سمت سؤالات جدیدی هدایت می‌کند. در برخی از موارد، محققان در بخش نتیجه‌گیری گزارش تحقیقاتی خود سؤالات جدیدی را که پدید آمده‌اند، مطرح می‌کنند. یک راه ثمربخش برای وسعت بخشیدن به بررسیها، معرفی متغیرهای جدید برای کنترل بیشتر و شناسایی اثرات متقابل بین متغیرها است.

۲- مرور تحقیقات پیشین می‌تواند برای محقق انگیزه‌ای باشد، برای اینکه دریابد شیوه‌هایی که در آن تحقیقات مورد استفاده قرار گرفته‌اند، می‌توانند برای حل مسئله‌های تحقیقاتی دیگر مناسب باشند یا خیر، یا یک بررسی مشابه می‌تواند در زمینه یا حوزه متفاوت و یا با گروههای متفاوت آزمودنی انجام یابد یا خیر. برای مثال، محقق، پژوهشی را که اثر بخشی یک رویکرد چند رسانه‌ای را در آموزش شیمی مورد بررسی قرار داده است، مطالعه می‌کند. شاید یک بررسی مشابه بتواند در حوزه زیست‌شناسی یا حوزه دیگر انجام یابد، یا یک بررسی در مورد دانش‌آموزان دوره متوسطه ممکن است رهنمودی برای یک معلم دوره ابتدایی باشد که

می‌خواهد بداند آیا در سطح ابتدایی نیز همان روابط مابین متغیرها حاکم است یا خیر.

۳- مرور تحقیقات گذشته می‌تواند ضرورت تکرار آنها را برای تایید یافته‌های پیشین خاطر نشان کند. یکی از ویژگیهای اساسی یک تحقیق علمی، قابلیت تکرار آن است؛ به گونه‌ای که یافته‌های آن قابل اثبات باشد. تکرار یک بررسی با ایجاد تغییرات و یا بدون آن احتمالاً برای محقق تازه کار فعالیت ارزشمند و پرسودی است. تکرار یک بررسی گسترده، تعمیم یافته‌های تحقیق را وسعت می‌بخشد و همچنین شواهد اضافی درباره اعتبار آن یافته‌ها فراهم می‌آورد. در بسیاری از آزمایشهای آموزشی، امکان انتخاب آزمودنی‌ها به طور تصادفی وجود ندارد و ما می‌باید از گروههای کلاسی که از قبل سازماندهی شده‌اند، استفاده کنیم. این امر البته ظرفیت تعمیم یافته‌ها را محدود می‌سازد. به هر حال با تکرار تجربیات در زمانها و مکانهای مختلف و با تایید روابط مورد قبول در هر بررسی، اطمینان نسبت به اعتبار علمی یافته‌ها نیز افزایش می‌یابد. تکرار صرف بررسیهای دیگر، مهمترین خاستگاه تلاشهای تحقیقی نیست، ولی در زمینه مسائل تعلیم و تربیت نیاز مداوم به تایید و بسط یافته‌ها وجود دارد.

در بیشتر موارد، تکرار بررسیهای گذشته، دقیقاً شبیه همان تحقیق قبلی نیست. به هنگام تکرار تحقیقات پیشین، برای روشن ساختن برخی از وجوه یافته‌ها، برای آزمون اینکه یافته‌ها را تا چه حد می‌توان تعمیم داد، یا برای بررسی عواملی که در تحقیق اصلی گنجانده نشده بودند، تغییراتی اعمال می‌شود. برای مثال، بررسیهای پیاژه در مورد رشد داوریهی اخلاقی در کودکان، به دفعات در دیگر کشورها تکرار شده است.^(۱) این بررسیها اگر چه از روش بنیادی پیاژه استفاده کرده‌اند، ولی رشد داوریهی اخلاقی را در کودکان طبقات اجتماعی متفاوت، در کودکان با مقطع سنی یکسان ولی با سطح هوش متفاوت، در کودکانی که در گروههای سنی خود میزان مشارکت متفاوتی دارند، در کودکانی که آداب تربیتی متفاوت دارند و در بین دختران و پسران، مورد بررسی قرار داده‌اند. اخیراً محققان دیگری نیز در تلاش برای اثبات یافته‌ها و نتایج پیاژه از روشهای متفاوتی استفاده کرده‌اند. در مجموع، قسمت عمده تحقیقات مشتق از پژوهشهای پیاژه، نتایج اصلی او را تایید کرده‌اند. به این ترتیب، یک بررسی منفرد، اگر با یک موضوع بااهمیت سروکار داشته باشد، و یا اگر یافته‌هایش هیجان‌انگیز باشد، می‌تواند برای بسیاری از بررسیهای دیگر الهام بخش باشد.

۴- تکرار تحقیقاتی که در گذشته انجام شده است، می‌تواند سؤالاتی در مورد قابلیت بکارگیری یافته‌های آن در فرهنگهای دیگر نیز برانگیزد. نتایج به دست آمده از یک تحقیق انجام یافته در فرهنگی خاص نمی‌تواند به طور خودکار در مورد فرهنگهای دیگر به کار گرفته شود و این یک دلیل برای آن است که در دهه‌های اخیر تأکید قابل توجهی بر بررسیهای میان فرهنگی شده است. در حیطه‌هایی مانند روان شناسی کودکان و بزرگسالان، یادگیری اجتماعی، رشد شناختی و زبانی، انگیزه پیشرفت، تکوین شخصیت، تجربیات آموزشی، مثالهای بی‌شماری از تحقیقات میان فرهنگی می‌توان یافت. تعداد رو به تزاید بررسیهای بین‌المللی در دانشگاههای آمریکا، هم علاقه به این نوع بررسی را افزایش داده و هم گردآوری داده‌های میان فرهنگی را تسهیل کرده است. تعدادی از رساله‌ها و پایان نامه‌های این دانشجویان در این زمره‌اند و بنیانی برای بررسیهای میان فرهنگی بیشتر فراهم می‌سازند.

۵- مرور تحقیقات پیشین می‌تواند در شناسایی تضادها و تناقضات یا در ناخرسندی از مفهوم سازی، روش شناسی، وسایل اندازه‌گیری، و تجزیه و تحلیل‌های آماری مورد استفاده، ثمر بخش باشد. محققان اغلب می‌توانند در تحقیقات گذشته چیزی را برای اثبات بیابند. برای مثال، در سال ۱۹۲۸ هارت شورن و می این نظریه موجود را که صداقت یک ویژگی شخصیتی یکپارچه است، مورد چالش قرار دادند.^(۱)

آنها رفتار چند هزار کودک دانش‌آموز را در موقعیتهای متعدد اغواکننده همبسته کردند و از همبستگی‌های پایین بین اندازه‌های اغواگرانه نتیجه گرفتند که صداقت، یک خصوصیت درونی مشترک نیست، بلکه خصوصیتی ویژه و تحت تأثیر موقعیتی است که افراد در آن قرار دارند. برتون در سال ۱۹۶۳ این بررسی را تکرار کرد^(۲)؛ چرا که تأکید هارت شورن و می را درباره تأثیر عوامل محیطی بر رفتار صادقانه، مورد تردید قرار داده بود. برتون از روش تحلیل عوامل استفاده کرد و از تمام تکالیف محوله به کودکان، شواهدی در تایید ثبات درون فردی به دست آورد. او نتیجه گرفت که احتمالاً در فرد خصیصه‌ای بنیادین در مورد صداقت وجود دارد که آن را به‌هنگام مقاومت در برابر موقعیتهای اغواکننده متجلی می‌کند، ولی او نیز با هارت شورن و می در رد یک فرمول همه یا هیچ در مورد شخصیت افراد موافقت داشت. در سال ۱۹۶۵ هانت

1- Hart Shorne, H. and M.A. May (1928). Studies in the Nature of Character. Studies in Deciet (vol. I). New York: Macmillan.

2- Burton, R.V (1963) - Generality of Honesty Reconsidered. Psychological Review, 70, 481-499.

بررسی برتون را تکرار کرد، ولی به عنوان روش آماری، از تحلیل واریانس به جای تحلیل عوامل بهره گرفت.^(۱) هانت اعتقاد داشت که تفاوت‌های شخصیتی، موقعیتها و تعامل بین فرد و موقعیت بایستی به عنوان منابع تغییر در رفتار صادقانه محسوب شوند و در تأیید فرضیه تعامل، شواهدی نیز فراهم کرد.

نلسن، گریندر، و موتر در سال ۱۹۶۹ همه بررسیهای یاد شده را تکرار کردند و آنها را بسط دادند.^(۲) آنها رویکردهای روش شناختی مختلف را با هم مقایسه کردند و نتیجه گرفتند که رفتار اغواگرانه، طی انجام مجموعه‌ای از تکالیف مختلف، تنها تا حدودی باثبات است. بنابراین، یافته‌های آنها با بررسیهای اولیه هارت شورن و می، به رغم بیش از چهل سال فاصله زمانی در جمع‌آوری اطلاعات مربوط به هر دو بررسی و به رغم جامعه‌های آماری متفاوت، تکالیف و شیوه‌های روش شناختی متفاوت، مطابقت داشت.

نمونه دیگری از بررسیهای برگرفته از بررسیهای پیشین، بررسیهایی است که اثرات گروه بندی بر اساس استعداد را در پیشرفت تحصیلی دانش‌آموزان مورد مطالعه قرار می‌دهد. در اولین سالهای قرن اخیر صدها بررسی در مورد این موضوع انجام یافت و در حال حاضر نیز موضوعی برای تحقیق به شمار می‌رود.^{(۳)(۴)(۵)}

اغلب اوقات محقق از فواصل بارز در دانش نظام یافته در حیطه‌ای خاص آگاه است. بنابراین تحقیق باید طوری طراحی شود که به پر کردن این فواصل یاری رساند و به دانش معتبرتری منتج شود. برای دانشجویانی که در این مرحله از تحقیق قرار دارند، تشکیل جلساتی با استادان صاحب‌نظر، می‌تواند مثمر ثمر باشد. بنابراین، دانشجو با تحلیل انتقادی تحقیقات منتشر شده در زمینه مورد نظر و با کمی خلاقیت، قادر خواهد بود که موضوعات متعددی را که بالقوه قابل تحقیق هستند، بیابد. درک جنبه‌های نظریه‌ای و کاربردی حیطه مورد نظر، محقق را قادر می‌سازد تا به شیوه‌ای انتقادی مطالعه کند و موضوعی درست را انتخاب کند.

- 1- Hunt, J. Mev. (1965). Traditional Personality Theory in The Light of Recent Evidence. American Scientist, 53, 80-96.
- 2- Nelsen, E.A., R.E. Grindler, and M.L. Mutterer (1969) Sources of Variance in behavioral measures of honesty in temptation Situations: Methodological analysis. Developmental Psychology, 1, 265-279.
- 3- Slavin, R. E. (1987). Ability Grouping and Student Achievement in Elementary Schools: A Best Evidence Synthesis. Review of Educational Research, 57, 293- 335.
- 4- Hiebert, E. (1987). The Context of Instruction and Student Learning: An Examination of Slavin's Assumption. Review of Educational Research, 57, 337- 340.
- 5- Gamoran, A. (1987). Organization, Instruction, and the Effects of Ability Grouping: Comment on Slavin's "best- evidence Synthesis." Review of Educational Research, 57, 341 - 345.

// منابع غیر آموزشی

تجربیات مشاهدات در کل جهان، به همراه فعالیتهای تخصصی ما، می تواند منابع پرباری در انتخاب موضوع تحقیق باشد. نظریه ها یا شیوه هایی که محقق در سایر زمینه ها با آن برخورد می کند، می تواند برای استفاده در حیطه تعلیم و تربیت منطبق شوند. اغلب جنبشهایی که از خارج زمینه تخصصی ما سرچشمه می گیرند، ما را به سمت راههای جدیدی از پژوهش سوق می دهند. جنبش زنان، ما را به سمت بررسی قالب پردازی های جنسی در مواد آموزشی، اثر مدارس در یادگیری نقشهای جنسی، تفاوت های جنسی در شخصیت و از این قبیل هدایت کرده است. جنبش حقوق مدنی بررسیهای بیشماری را در مورد آموزش کودکان گروههای اقلیت به راه انداخت. شیوع بیماری ایدز در کشور موجب شده است که بررسیهای زیادی برای تعیین بهترین شیوه ها و مواد آموزشی در زمینه آگاه ساختن دانش آموزان از خطرات بیماری و بهترین نحوه حفاظت از آنها در مقابل بیماری، صورت پذیرد.

بیشتر تحقیقات ارزشمند در تعلیم و تربیت از چنین منابع غیر آموزشی الهام می گیرند.

/// ارزشیابی مسئله تحقیق

بعد از آنکه یک مسئله آزمایشی انتخاب شد، می باید مورد ارزشیابی قرار گیرد. محقق باید مطمئن باشد که زمینه مربوط به مسئله از آنچنان اهمیتی برخوردار است که تحقیق درباره آن را توجیه می کند؛ هر چند که تعیین این مسئله همیشه آسان نیست. قضاوت در مورد ارزش یک مسئله غالباً امری است که به ارزشهای فردی و عقاید شخصی مربوط می شود. به هر حال، معیارهای خاصی وجود دارند که در فرآیند ارزشیابی اهمیت مسئله تحقیق باید مورد استفاده قرار گیرند.

۱- در بهترین حالت، مسئله باید به نوعی باشد که حل آن در پیکره دانش سازمان یافته تعلیم و تربیت سهم داشته باشد. محقق باید نشان دهد که بررسی او در پی پر کردن فواصل موجود در دانش کنونی و یا کمک به حل پاره ای تناقضات در تحقیقات گذشته است. شاید یک تحقیق بتواند مطالعات قبلی را به گونه ای غنی تر سازد که دانش معتبرتری حاصل شود. بیشتر دانشمندان بر این باورند که مسائلی که از نظریه سرچشمه گرفته اند، امکان بالقوه بیشتری برای حصول به این معیار دارند. محقق می تواند از خود سؤال کند که آیا بررسی او به شناخت روابط

جدید منجر خواهد شد یا به سادگی، یافته‌های قبلی را تکرار خواهد کرد.

مطالعات خاصی ممکن است هم در نظریه و هم در عمل سهم داشته باشند. به هر حال اگر مسئله‌ای فاقد مفاهیم آشکار نظری باشد، در نهایت باید دارای چند مفهوم عملی باشد. محقق باید با توجه به بررسی‌اش قادر به پاسخگویی به این سؤال باشد که «بسیار خوب، نتیجه چه خواهد بود؟» آیا راه حل مسئله تفاوتی در شیوه‌های تعلیم و تربیت ایجاد خواهد کرد؟ آیا دیگر متخصصان تعلیم و تربیت نیز به این یافته‌ها علاقه‌مند هستند؟ آیا یافته‌ها در موقعیتهای تصمیم‌گیری آموزشی سودمند خواهند بود؟

از آنجا که مسائل زیادی با مفاهیم عملی و نظریه‌ای وجود دارند که نیازمند راه حل هستند، صرف وقت و تلاش برای پاسخگویی به سؤالاتی که فاقد یکی از این دو جنبه مهم هستند، توجه ناپذیر است.

بعضی اوقات، یک دانشجو در تلاش برای تعیین یک مسئله، سؤالی را انتخاب می‌کند که روابط ضعیفی را مطرح می‌کند. برای مثال، اینکه «چه رابطه‌ای بین محبوبیت فرد در میان همسالان و سرعت خواندن وجود دارد؟» یک مسئله جزئی محسوب خواهد شد. این امر برای تجارب آموزشی اهمیت بسیار کمی دارد و یا اصلاً اهمیت ندارد. همچنین سؤالی نظیر «چه رابطه‌ای بین هوش و پیشرفت خواندن وجود دارد؟» نیز سؤالی بی‌اهمیت قلمداد می‌شود؛ چرا که تاکنون اطلاعات معتبری در این زمینه به دست آمده است و تحقیقات بیشتر، غیر ضروری به نظر می‌رسد. «هیچ نیازی برای اختراع مجدد چرخ نیست.»^(۱) بسیاری از مطالعات پیشنهادی باید به سهولت رد شوند؛ چرا که نه در نظریه و نه در کل، سودی برای تعلیم و تربیت ندارند. برای مثال، محقق تازه کاری که می‌خواهد مزیت‌های تطبیقی دو روش متفاوت را در شروع آموزش خواندن ارزشیابی کند، ممکن است بررسی نگرش آموزگاران کلاس اول را در مورد این مسئله پیشنهاد کند. چنین پیمایشی، پاسخ با معنایی در مزایای نسبی این دو سیستم به دست نخواهد داد؛ چرا که تنها از گروه کوچکی از معلمان، که احتمالاً به اندازه محقق در این زمینه جاهل هستند، به دست آمده است. رویکرد بهتر، یک آزمایش کنترل شده برای مقایسه پیشرفت دو گروه از دانش‌آموزان همسان تحت یک سیستم است که هر کدام با یکی از روشهای مذکور با خواندن آشنا شده‌اند. سهولت در پیمایش نظرات، بسیاری از افراد را در بکارگیری این روش،

حتی زمانی که پاسخ به پرسش مطرح شده نابجاست، اغوا می‌کند.

۲- یک مسئله تحقیق باید ما را به مسائل جدید و در نتیجه به تحقیقات بیشتر هدایت کند. این امر زمانی اتفاق می‌افتد که محقق بتواند مسئله تحقیق خود را با دانش سازمان یافته پیوند دهد و توجه خود را به نوع مطالعه‌ای که به طور منطقی می‌تواند پی‌آمد کار او باشد، معطوف کند. هنگامی که یک مطالعه مطلوب به پاسخ یک سؤال دست پیدا می‌کند، معمولاً سؤالات دیگری ایجاد می‌کند که نیاز به بررسی دارند. این مسئله در مورد بررسیهایی که با نظریه تقویت در کلاس درس سروکار دارند، واقعیت داشته است. در مقابل، بسیاری از تحقیقات توصیفی، به این معیار دوم دست نمی‌یابند. در این زمینه، مؤلفان کتاب پیشنهاد می‌کنند که دانشجویان مبتدی در امر تحقیق، در انتخاب موضوع به مسائلی توجه داشته باشند که امکان توسعه و پیگیری آتی آنها در بررسیهای دوره‌های عالی‌تر دانشگاهی وجود داشته باشد.

۳- مسئله باید قابل تحقیق باشد. اگرچه این معیار بدیهی به نظر می‌رسد، ولی در عمل بسیاری از مسائل در زمره سؤالاتی نیستند که بتوان آنها را مورد بررسی علمی قرار داد. یک مسئله، برای آن که قابل تحقیق باشد، باید این قابلیت را داشته باشد که در عمل مورد بررسی قرار گیرد. بسیاری از سؤالات جالب نمی‌توانند موضوع تحقیق تجربی واقع شوند، ولی باید طی بررسیهای فلسفی مورد پژوهش قرارگیرند. یک چنین سؤالاتی همانند «آیا آموزش مسائل جنسی در مدارس ابتدایی واجب است؟» یا «آیا ما باید کمونیزم را در دبیرستان تدریس کنیم؟» همگی جزو مسائل فلسفی هستند که با تحقیق علمی قابل پاسخگویی نیستند. هر چند با سؤالات مذکور نمی‌توان به طور تجربی کار تحقیق را آغاز کرد، اما می‌توان آنها را در قالب سؤالات قابل تحقیق مجدداً صورتبندی کرد. به عنوان مثال، باید سؤال اول را به این شیوه روشن‌تر بیان کنیم که آموزش مسائل جنسی در مدارس ابتدایی چه اثری بر روی نگرش دختران دوره اول دبیرستان در مورد تجربه جنسی قبل از ازدواج دارد. یک بررسی می‌تواند در جهت دستیابی به اطلاعات لازم برای یک چنین سؤالاتی طراحی شود. هر چند که سؤالات فلسفی از این نوع برای تحقیق علمی مناسب نیستند، با وجود این اطلاعاتی که تحقیق فراهم می‌آورد، می‌تواند در یافتن پاسخ سؤالات فلسفی و اخلاقی کاربرد داشته باشد. به این ترتیب اطلاعات به دست آمده در جریان تحقیق بر روی مسئله‌ای خاص می‌تواند در تصمیم‌گیری در مورد اخلاقیات و ارزشها، کارشناسان تعلیم و تربیت را یاری کند.

محقق همچنین باید به تعریف و اندازه‌گیری متغیرهای دخیل در سؤال مطرح شده توجه داشته باشد. مسئله‌ای نظیر «تغییر الویتهای ملی بر آینده تعلیم و تربیت در آمریکا چه تاثیری دارد»، برای تحقیق مناسب نخواهد بود. تعریف عباراتی مانند تغییر الویتهای ملی و آینده تعلیم و تربیت در آمریکا، به گونه‌ای که قابل اندازه‌گیری باشند، مشکل خواهد بود.

۴- مسئله می‌باید با ویژگیهای محقق مناسبت داشته باشد. ممکن است مسئله از نظر معیارهای تبیین شده عالی باشد، ولی شاید برای محقق مهم جلوه نکند. تعدادی از جنبه‌های مشخصی که می‌باید مورد ملاحظه قرار گیرند، عبارتند از:

الف - مسئله باید از نوعی باشد که محقق به آن علاقه‌راستین دارد و مشتاقانه بتواند در مورد آن به تحقیق بپردازد. راه حل مسئله باید برای محقق مهم باشد. از آنجا که راه حل مسئله می‌تواند در افزایش دانش محقق در حیطه‌ای خاص، یا بهبود عملکرد او به عنوان یک متخصص تعلیم و تربیت مؤثر باشد، این راه حل باید برای محقق مهم باشد. اگر مسئله تحقیق پر معنا و جالب نباشد، بعید به نظر می‌رسد که به عنوان وظیفه‌ای تمام و کمال محقق را به صرف وقت و انرژی وادارد.

ب - مسئله می‌باید در زمینه‌ای باشد که محقق در مورد آن صاحب دانش و مقداری تجربه است. او نیازمند آشنایی با نظریه‌ها و مفاهیم موجود و پایه‌ریزی واقعیات به شیوه‌ای است که منجر به تعیین مسائلی پر ارزش شود. بعلاوه محقق باید روشن سازد که آیا از مهارت‌ها و صلاحیتهای لازم برای به اتمام رساندن مطالعه برخوردار است؟ امکان دارد تهیه و اعتباریابی ابزار تحقیق و انجام تحلیل‌های آماری پیچیده ضروری باشد.

ج - مسئله مورد نظر تحقیق باید از محیطی که خود محقق نیز در آن حضور دارد، انتخاب شود. او باید مطمئن باشد که اطلاعات ضروری برای پاسخگویی به سؤال موردنظر قابل دستیابی است یا خیر. او باید بررسی کند تا اطمینان یابد که آزمودنیهای موردنیاز با پیشینه‌های تحصیلی مناسب، قابل دستیابی هستند. مسئولان مدارس اغلب اوقات با انجام تحقیق در مدرسه‌شان مخالف هستند. آنها به سهولت به دانش‌آموزان اجازه شرکت در بررسی را نمی‌دهند، بنابراین جز آنکه محقق در زمان بررسی در مدرسه‌ای مشغول به کار باشد، وسیله‌ای برای حل مشکل خود نخواهد داشت. یکی از نویسندگان کتاب حاضر قبل از آنکه بتواند اجازه انجام یک تجربه را در زمینه تعلیم و تربیت به دست آورد، مجبور شد که به چهار مدرسه مراجعه کند.

د- مسئله باید قابلیت بررسی و تکمیل شدن در زمانی معین را داشته باشد. مسئله‌ای را انتخاب نکنید که بسیار بزرگ و یا درگیر کننده باشد. از برنامه‌ریزی زمان مورد نیاز برای ساخت وسایل اندازه‌گیری و بکارگیری این وسایل، تجزیه و تحلیل اطلاعات و نوشتن گزارش مطمئن باشید.

/// بیان مسئله

بعد از انتخاب و تعیین ویژگیهای مسئله، هنوز وظیفه صورتبندی یا بیان مسئله، به شکلی که پاسخگوی تحقیق باشد، وجود دارد. اگر بخواهیم یک مسئله خوب بیان شود، باید (۱) آن چیزهایی را که قرار است تعیین و یا حل شوند، دقیقاً روشن کنیم و (۲) دامنه بررسی را به سؤالی خاص محدود نماییم. ما نمی‌توانیم بر اهمیت بیان روشن و موجز سؤال تحقیق تاکید بیش از حد داشته باشیم. نیاز به گفتن نیست که فرد قبل از آنکه تواند کار زیادی برای حل مسئله انجام دهد، باید آن را تعیین کند. محقق تازه‌کار دارای یک فکر کلی در مورد مسئله است، اما در صورتبندی آن به عنوان یک سؤال عملی تحقیق دچار مشکل می‌شود. او در می‌یابد که فکر کلی اولیه‌اش، اگر چه برای انتقال آن به دیگران و درک آن از طرف دیگران کافی به نظر می‌رسد، ولی به قدر کافی مشخص نیست تا به او اجازه دهد که به طور تجربی با مسئله مواجه شود و او تا زمانی که به یک بیان عینی از سؤال تحقیق دست نیابد، نمی‌تواند تحقیق خود را به پیش ببرد.

برای روشن شدن مطلب، فرض می‌کنیم یک محقق تازه‌کار اظهار می‌کند که به مطالعه در مورد تأثیر برنامه درسی جدید علوم در مدارس دوره متوسطه، علاقه‌مند است. همانطور که گفته شد، می‌توان منظور محقق را به طور کل درک و با او به صورتی کلی ارتباط برقرار کرد، ولی اگر او خواهان یافتن روشی برای حل مسئله است، باید مسئله را به روشنی تمام مشخص کند.

گام اول با تعریف اصطلاحات تحقیق آغاز می‌شود. منظور از تأثیر، برنامه درسی علوم و مدارس دوره متوسطه چیست؟ تعاریف مورد نیاز تحقیق اکثراً از طریق فرهنگ لغات کامل نمی‌شود. برای مثال، تأثیر، خلق «نتیجه مورد نظر و یا مورد انتظار» تعریف می‌شود. این تعریف ساخت کلی تأثیر را توصیف می‌کند، ولی برای هدف تحقیق بسنده نیست. ما باید بتوانیم دقیقاً مشخص کنیم که محقق از چه چیزی به عنوان شاخص تأثیر استفاده خواهد کرد یا محقق برای سنجش وجود یا فقدان پدیده‌ای که از آن به عنوان مفهوم تأثیر نام می‌برد، چه کاری انجام خواهد داد. این امر در مورد سایر اطلاعات نیز صدق می‌کند. به عبارت دیگر، محقق باید متغیرهای

مسئله را به طور کاربردی تعریف کند. جهت تعریف کاربردی، محقق باید برای ارائه مفاهیم مورد نظر، برخی از رفتارها و یا وقایع آشکار را که مستقیماً به وسیله خود محقق و یا سایر افراد قابل مشاهده و اندازه‌گیری هستند، مشخص کند. همانطور که در فصل اول تشریح شد، تعریف کاربردی، تعریفی است که یک مفهوم را بر حسب عملیات و یا فرآیندهایی که برای اندازه‌گیری مفاهیم به کار می‌روند، تعریف کند.

در این بررسی محقق باید تاثیر را یا به عنوان پیشرفت در نمرات یک آزمون تفکر انتقادی و یا آزمون استاندارد شده علوم تعریف کند. اصطلاح برنامه درسی به مثابه دوره زیست‌شناسی برنامه‌ریزی شده که به دانش‌آموزان سال دوم دبیرستان ارائه می‌شود، تلقی می‌شود. مدارس دوره متوسطه می‌تواند به دبیرستانهایی اطلاق شود که از لحاظ تعداد دانش‌آموز، نحوه فعالیت و غیره، ویژگیهای معینی را دارا باشند. مسئله اولیه اکنون به این سؤال تبدیل می‌شود: «تاثیر دوره زیست‌شناسی برنامه‌ریزی شده در درک مفاهیم زیست‌شناختی دانش‌آموزان سال دوم دبیرستان، که درس زیست‌شناسی را تازه شروع کرده‌اند، چیست؟» تعاریف کاربردی، دامنه سؤال کلی را به متغیرهای قابل مشاهده و ویژه محدود می‌کند.

اکنون کار تحقیق با وضوح و تمرکز بیشتر نشان داده شده است. محقق می‌تواند با طراحی یک مطالعه آزمایشی برای مقایسه نمرات پیش و پس آزمون مفاهیم زیست‌شناختی دانش‌آموزانی که در دوره زیست‌شناسی برنامه‌ریزی شده شرکت کرده‌اند و دانش‌آموزان مشابهی که برنامه زیست‌شناسی جایگزین آن را گذرانده‌اند، کار تحقیق را به پیش برد. اکنون محقق شواهدی عینی درباره یک برنامه درسی خاص و موقعیت خاص را گردآوری کرده است که می‌تواند پرتوی بر سؤال کلی اولیه بیفکند.

بعلاوه، در تبیین مسئله، محقق باید برای ایجاد تعادل بین کلیات و جزئیات کوشش کند. اگر مسئله بیان شده بسیار کلی باشد، محقق با پهنه‌ای مبهم و بدون هیچ نشانه آشکاری از مسیر تحقیق روبرو خواهد بود. برای مثال سؤالی نظیر «تاثیر آموزش برنامه‌ریزی شده بر پیشرفت علمی در دبیرستان چیست؟»، بسیار کلی است. خیلی بهتر خواهد بود که بپرسیم «اثر استفاده از دوره‌های جبر برنامه‌ریزی شده بر پیشرفت دانش‌آموزان تیزهوش دبیرستان چیست؟» این نحوه بیان فوراً تعداد نمونه مورد نیاز، متغیرهای مرتبط، و نوع اطلاعاتی را که گردآوری خواهند شد، به نمایش می‌گذارد.

از سوی دیگر، مسئله نباید آنقدر کوچک باشد که به نظر ناچیز و بی معنا جلوه کند. محقق طالب موضوعی است که به قدر کافی جامع باشد تا بر مبنای معیارهای بحث شده با اهمیت جلوه کند، در عین حال آن قدر مشخص باشد که برای دانشجو در موقعیت خاص او انجام پذیر شود. هنگامی که مسائل تحقیقاتی بر رابطه بین دو یا چند متغیر متمرکز باشند، این بدان معنا نیست که عبارت «رابطه بین — و — چیست؟»، در بیان مسئله آمده باشد. بیان مسئله ممکن است در این قالب بیاید یا وجود رابطه را به طور ضمنی برساند و در غیراین صورت، بیان مسئله باید در قالب سؤال مطرح شود. دانشجویان اغلب در این مرحله گیج می شوند. برای مثال، (۱) بیان اخباری که «این بررسی اثر مربوط به روش خواندن الف و ب را در آموزش افراد کندآموز بررسی می کند» و (۲) بیان سؤالی که «اثر بخشی روش خواندن الف در مقایسه با روش ب در آموزش افراد کندآموز چیست؟»، رابطه بین متغیرها را پی جویی می کنند، بدون آنکه عبارت مذکور را به کار ببرند، و هر دو شیوه های قابل پذیرشی برای ارائه مسئله تحقیق هستند. همان گونه که قبلاً گفته شد، بعضی از محققان صرفاً شیوه سؤالی را ترجیح می دهند؛ چون این شیوه ساده و روشن است و از نظر روان شناختی نیز چنین به نظر می رسد که محقق را به سمت وظیفه اش، یعنی طراحی روشی برای یافتن پاسخ سؤالش، جهت می دهد. در بعضی از موارد، هم بیان اخباری و هم سؤال، توأمان مطرح می شوند: بیانه های اخباری عام تر ممکن است با یک یا چند سؤال بسیار ویژه دنبال شوند.

هنگامی که یک سؤال بالقوه صورتبندی می شود، باید روشن بودن و قابلیت انجام آن مورد ارزیابی قرار گیرد. بسط یک مسئله قابل انجام، یک فرایند تکاملی است که متضمن بسیاری از کوششها برای روشن سازی مفاهیم، تعریف شیوه عمل و بررسی راههای گردآوری داده ها است.

/// تعیین جامعه مورد بررسی و متغیرها

یک راهبرد مناسب برای شکلهی مسئله مورد نظر - با یک تصور مبهم از چیزی که فرد تمایل به بررسی آن دارد - در قالب یک مسئله قابل تحقیق، تفکر در مورد جامعه آماری تحقیق و متغیرهاست.

برای مثال، اجازه بدهید به مسئله مورد نظر خانم بورک^(۱)، یکی از مسئولان دوره ابتدایی،

توجه کنیم که آیا تدریس خصوصی توسط دانش‌آموزان سالهای بالاتر به دانش‌آموزانی که متوسط نمرهٔ قرائت آنها پایین است، اثر مثبت خواهد داشت؟

اغلب اوقات تعریف جامعهٔ آماری بررسی، آسان‌ترین کارهاست، و آن افرادی هستند که محقق مشتاق دانستن چیزهایی درباره آنهاست. جامعه مورد نظر دانش‌آموزانی هستند که خواندن آنها از حد متوسط پایین‌تر است. در این سؤال، توانایی خواندن یک متغیر نیست؛ چون همهٔ کودکان مورد نظر به عنوان دانش‌آموزان پایین‌تر از متوسط تلقی شده‌اند.

با تعیین دانش‌آموان پایین‌تر از متوسط به عنوان جامعه مورد مطالعه در بیان اولیه مسئله، خانم بورک اکنون باید از خود سؤال کند که آیا این همان جامعهٔ آماری دلخواه اوست؟ او احتمالاً باید تصمیم بگیرد که «دانش‌آموزانی که خواندن آنها پایین‌تر از سطح متوسط است» یک گروه بسیار وسیع است و لذا باید خود را به سن خاصی، یعنی دانش‌آموزان پایین‌تر از متوسط دورهٔ دوم، محدود سازد.

اکنون او آماده جستجوی متغیرهای موجود در بیان اولیه مسئله خود است. تدریس خصوصی، به دلیل تغییر دادن انواع تدریس‌های مورد استفاده یا تغییر دادن مقدار زمان تدریس و یا دریافت تدریس خصوصی از سوی بعضی از کودکان و عدم دریافت آن از طرف بقیه، می‌تواند یک متغیر باشد. خانم بورک تصمیم می‌گیرد که این مورد آن چیزی است که او واقعاً می‌خواهد بداند. بنابراین او بخش مربوط به این سؤال را به این صورت بازنویسی می‌کند که: آیا دریافت یک میزان مشخص تدریس خصوصی در مقابل هیچگونه تدریس خصوصی...؟ بنابراین، تدریس خصوصی متغیر غیر وابسته است؛ چون بر مهارت قرائت مقدم است، و نکته مهم پیش‌بینی این امر است که آیا تدریس خصوصی در مهارت خواندن که یک متغیر وابسته است، اثر خواهد داشت؟

اکنون آشکار شده است که کلمه تدریس خصوصی بسیار کلی است و بدون آنکه همهٔ آزمودنیها یکسان و به یک میزان تدریس خصوصی دریافت‌کنند، نتایج بررسی بی‌معنی خواهد بود. خانم بورک تصمیم می‌گیرد که از تمرین یادآوری سریع کلمه، به عنوان نوع ویژه‌ای از تدریس خصوصی و ۱۵ دقیقه زمان در هر روز، به عنوان مقدار زمان استفاده کند.

عبارت یک تاثیر مثبت دارد، تا وقتی که خانم بورک به آن به مثابه متغیر غیر وابسته ننگرد، کاملاً مبهم است. تمرین یادآوری سریع کلمات بر چه ... تاثیر دارد؟ او می‌داند که بر یادآوری

سریع کلمه اثر دارد، ولی او می‌خواهد اثرات آنرا بر دیگر جنبه‌های رفتاری خواندن، که ممکن است مشاهده شود، بررسی کند: خواندن بر احساس با صدای بلند، بی‌صدا خوانی، احساسات مثبت نسبت به خواندن، تعداد کتابهای قرائت شده، ادراک و غیره. در میان دانش‌آموزان کلاس دوم که خواندن آنها از سطح حد متوسط پایین‌تر است، آیا تفاوتی در نمره قرائت CAT (آزمون پیشرفت کالیفرنیا)^(۱) بین کسانی که روزانه ۱۵ دقیقه توسط دانش‌آموزان کلاسهای بالاتر به تمرین یادآوری کلمات می‌پردازند و کسانی که هیچ تمرینی در این زمینه نداشته‌اند، وجود دارد؟ او به یک متغیر وابسته فارغ از قضاوت معلم نیاز دارد و تصمیم می‌گیرد که از نمره قرائت آزمون پیشرفت کالیفرنیا به عنوان متغیر وابسته استفاده کند.

بیان مسئله تجدید نظر شده خانم بورک حالا اینگونه عنوان می‌شود: در میان دانش‌آموزان کلاس دوم که خواندن آنها از سطح حد متوسط پایین‌تر است، آیا در نمره قرائت CAT بین کسانی که روزانه ۱۵ دقیقه توسط دانش‌آموزان کلاسهای بالاتر به تمرین یادآوری سریع کلمات می‌پردازند و کسانی که هیچ تمرینی در این زمینه نداشته‌اند، تفاوتی وجود دارد؟ این سؤال می‌گوید که او در مورد چه کسانی بررسی می‌کند، برای بعضی، به نسبت بقیه، چه کار متفاوتی صورت خواهد گرفت و او انتظار چه اثری را در مقابل تیمار متفاوت خود دارد. همچنین توجه داشته باشید که قضاوت ارزشی تأثیر مثبت در این سؤال حذف شده است.

پیگیری این فرایند به شیوه متداول و مشابه با آنچه که در طراحی یک جمله به کار می‌رود، مفید است. فرد می‌تواند کار را با ترسیم یک خط عمودی و نوشتن **جامعه آماری** در سمت چپ و متغیرها در سمت راست آغاز کند. این عوامل موجود در بررسی سپس در پایین خط افقی فهرست می‌شوند. در مثال ما نمودار به صورت ذیل دنبال می‌شود.

متغیرها	جامعه آماری
- انجام تمرین هر روز به مدت ۱۵ دقیقه توسط دانش‌آموزان بالاتر از متوسط در مقابل انجام ندادن هیچ گونه تمرین - نمره خواندن در تست CAT	دانش‌آموزان کلاس دوم که خواندن آنها از سطح متوسط پایین‌تر است.

اجازه بدهید سؤال دیگری را مطرح کنیم. اثر تجربه کردن یک برنامه پیش دبستانی یا نداشتن چنین تجربه‌ای در پیشرفت خواندن در کلاس اول دبستان چیست؟

متغیرها	جامعه آماری
- تجربه کردن یا نکردن یک برنامه پیش دبستانی (متغیر غیر وابسته)	کلاس اولی‌ها
- پیشرفت خواندن (متغیر وابسته)	

این سؤال کامل است؛ زیرا یک جامعه آماری معین و دو متغیر دارد. از آنجا که برنامه پیش دبستانی بر پیشرفت خواندن کلاس اولی‌ها مقدم است، اولی می‌تواند به عنوان متغیر مستقل و دومی به عنوان متغیر وابسته تعیین شوند.

اجازه دهید مثال دیگری را بررسی کنیم: آیا آموزش رانندگی در دبیرستان فایده‌ای دارد؟ همانطور که اشاره شد، این سؤال نه جامعه آماری دارد و نه متغیر. یک محقق که با این سؤال شروع می‌کند، نخست باید تصمیم بگیرد رانندگان ۱۹ ساله‌ای را که آموزش رانندگی دوره دبیرستان را دیده‌اند، با آنهایی که این آموزش را ندیده‌اند، مقایسه کند. حالا ما یک جامعه آماری معین و یک متغیر مستقل داریم. اکنون می‌توانیم توجه خود را به انتخاب یک متغیر وابسته معطوف کنیم. آموزش رانندگی یا عدم آموزش آن در دوره دبیرستان ممکن است چه اثری بر رانندگان ۱۹ ساله داشته باشد؟ چنین تصمیم گرفته شده است که نرخ تصادف می‌تواند یک متغیر وابسته مناسب باشد. با گذاشتن این عوامل در یک نمودار، خواهیم داشت:

متغیرها	جامعه آماری
- گرفتن و یا نگرفتن تعلیم رانندگی در دوره دبیرستان (متغیر مستقل)	رانندگان ۱۹ ساله
- نرخ تصادف (متغیر وابسته)	

اکنون می‌توان یک سؤال کامل را مطرح کرد: آیا نرخ تصادف رانندگان ۱۹ ساله‌ای که در دوره دبیرستان، تعلیم رانندگی دیده‌اند، از رانندگان ۱۹ ساله‌ای که تعلیم رانندگی ندیده‌اند، کمتر است؟ این سؤال که چه رابطه‌ای بین جزم اندیشی و نگرشهای سیاسی در میان دانشجویان سال اول دانشکده وجود دارد، نکته دیگری را روشن می‌کند.

متغیرها

جامعه آماری

دانشجویان سال اول دانشکده	- جزم اندیشی
	- نگرشهای سیاسی

این سؤال با داشتن جامعه آماری و دو متغیر، کامل است. اما تا وقتی که نتوان تعیین کرد که کدامیک بر دیگری مقدم است، نخواهیم توانست متغیرها را به عنوان متغیر مستقل و وابسته برچسب بزنیم.

اگر بررسی انجام بگیرد که به جای رابطه بین متغیرها، وضعیت موجود را بسنجد، ممکن است تنها با یک متغیر کامل شود. برای مثال، محقق ممکن است عقیده دانشجویان دانشکده را در مورد قانونی شدن مصرف ماری جوآنا، مورد بررسی قرار دهد. در این مورد، جامعه آماری، دانشجویان ارشد دانشکده و تنها متغیر، عقاید آنها در مورد موضوع است.

روشهای مختلفی برای پاسخ دادن به انواع سؤالات تحقیقاتی به کار گرفته می شود. هرگاه یک متغیر مستقل بتواند توسط محقق دستکاری شود، روشهای آزمایشی مناسب هستند (به فصل ۹ مراجعه کنید). اولین مثال در این بخش مربوط است به اثر تمرین یادآوری سریع کلمه در آزمون قرائت CAT که یک تحقیق آزمایشی است. بسیاری از متغیرها در امر آموزش نمی توانند مورد دستکاری قرار گیرند، بنابراین روش آزمایشی نمی تواند در بررسی این متغیرها مورد استفاده قرار گیرد. موضوع تحقیقی مربوط به مقایسه نرخ تصادف رانندگان ۱۹ ساله ای که تعلیم رانندگی گرفته اند و یا نگرفته اند، احتمالاً روش علی - مقایسه ای را می طلبد (به فصل ۱۰ مراجعه کنید). متغیر مستقل (گرفتن تعلیم رانندگی و یا نگرفتن تعلیم رانندگی) توسط محقق مورد دستکاری قرار نگرفت.

برای بعضی از موضوعات تحقیقی، بسته به آنکه چگونه طراحی شده باشند، هم روش آزمایشی و هم علی - مقایسه ای، مناسب است. در مثال تحقیقاتی یادشده در مورد اثر برنامه های پیش دبستانی در پیشرفت خواندن، محقق بسته به آنکه دستکاری متغیر مستقل، یعنی برنامه های پیش دبستانی را برگزیند یا آزمودنی هایی را انتخاب کند که قبلاً تربیت پیش دبستانی داشته و یا نداشته اند، می تواند از هر یک از این دو روش استفاده کند.

موضوعات تحقیقی دیگر با توصیف ماهیت وقوع یک یا چند متغیر آموزشی سروکار دارند.

روش توصیفی یکی از مناسبترین روشها برای این گونه مسائل است (به فصل ۱۱ مراجعه کنید). مثال مربوط به رابطه بین جزم اندیشی و نگرشهای سیاسی بین دانشجویان کالج، یک تحقیق توصیفی است. سنجش عقاید در مورد مسائلی از قبیل قانونی شدن مصرف ماری جوآنا نیز یک تحقیق توصیفی است.

// مقاله مجلات

تعیین جامعه آماری و متغیرهای مستقل و وابسته، چهارچوبی برای تجزیه و تحلیل گزارشهای تحقیقاتی چاپ شده به دست می دهد. مارش در مقاله تحقیقات آموزشی عواملی را به نمایش می گذارد که معمولاً در گزارش هر تحقیق آزمایشی مشاهده می شود.^(۱) مقدمه این گزارش دلایل منطقی برای انجام بررسی و نیز تاریخچه کوتاهی از یافته ها و تحقیقات موجود را در مورد موضوع ارائه می کند. در بخش روشها، عوامل متعددی وجود دارد: (۱) نمونه (۱۰ کلاس از ۵ دانشگاه) که انتخاب شده است تا معرف جامعه آماری مورد نظر (دانشجویان دانشگاه) باشد؛ (۲) متغیر مستقل (امتحان در خانه، در مقابل امتحان کلاس) توصیف و به طور کاربردی تعریف شده است؛ (۳) متغیرهای وابسته (پاسخگویی دانشجویان به سؤال تحقیق در مورد اینکه آنها چقدر انگیزه دارند تا جهت آمادگی برای هر یک از انواع امتحانات و کسب نمرات خاصی از طریق یک آزمون درک و بکارگیری مطلب، که بدون اطلاع قبلی صورت می گیرد، به مطالعه و یادگیری بپردازند) توصیف و به طور کاربردی تعریف شده اند؛ (۴) شیوه ها با جزئیات کافی توصیف شده اند تا سایر محققان را قادر سازند این آزمایش را با گروه نمونه متفاوت تکرار کنند. بخش نتایج، داده های حاصله از آزمایش و نتایج آزمونهای آماری را به دست می دهد. بخش نتیجه گیری حاوی تفسیر محقق از این نتایج است.

مقایسه ای بین امتحان در خانه، در مقابل امتحان در کلاس

این بررسی تلاشی بود برای تعیین اینکه آیا برای آنکه یادگیری صورت گیرد، امتحان در خانه به اندازه امتحان سنتی در کلاس خوب است، یا از آن بهتر است. نتایج این آزمایش مهم است؛ چون کارشناسان امر آموزش به طور فزاینده ای در حال حذف امتحان در سر کلاس هستند (گی و

گالاگر^(۱)، ۱۹۷۶)، و به نظر می‌رسد هیچ پایه عملی برای این باور که امتحان انجام شده در منزل شیوه‌ای مؤثر است، وجود ندارد. بنابراین، روندهای جاری در جهت حذف امتحانهای کتبی رسمی ممکن است در واقع بر خلاف یکی از اهداف اساسی آموزش و پرورش، یعنی یاد داری مفاهیم آموخته شده از طرف دانش‌آموزان، عمل کند.

احتمالاً نگرانی اصلی از امتحانهای کلاسی این است که این امتحانات موجب اضطرابهای تضعیف‌کننده می‌شوند؛ هر چند تحقیق در مورد این ادعا نتایج متناقضی به بار آورده است. برای مثال، ثابت شده است که در شرایط امتحانی، اضطراب یک اثر منفی دارد (ساراسون، دیویدسون، لایتال، ویت، و روبوش، ۱۹۶۰)^(۲)، یک اثر مثبت دارد (کاستاندا، مک‌کندلس، و پالرمو، ۱۹۶۴)^(۳) و بی‌تاثیر است (دنی، پاترسون، و فلدهاسن، ۱۹۶۴)^(۴). فرضیه بعضی از محققان این است که این، میزان اضطراب است که در عملکرد امتحان تاثیر می‌گذارد. آناستازی^(۵) (۱۹۷۴) معتقد است که امتحان، اگر به طور درست گرفته شود، یک اثر کلی مثبت دارد. او می‌گوید که رابطه بین اضطراب و عملکرد امتحان احتمالاً خطی نیست و بنابراین مقداری اضطراب مفید است؛ در حالی که میزان زیاد آن زیانبخش است. آزمایشهای زیادی برای از میان بردن اضطراب در شرایط امتحان کلاسی طراحی شده است، اما آناستازی اظهار می‌کند که «خود روش آزمونگران و نیز انجام یک امتحان خوب سازماندهی شده و از آسان به مشکل، به همین هدف دست خواهد یافت» (آناستازی، ۱۹۷۶).

تجربیات دانشجویان سطوح مختلف در مورد اضطراب آنها قبل از امتحانات، احتمالاً حذف امتحان را به دلیل ایجاد اضطراب تایید نمی‌کند و هدف ما نباید احتراز کامل از تنش باشد؛ چرا که این امر در دنیای ما امکان‌ناپذیر است، اما آموختن چگونگی تشخیص پاسخ نوعی به تنش و سپس تلاش برای سازگار کردن شیوه مواجهه ما با آن، امری شدنی است (سلی، ۱۹۷۸)^(۶). پیشینه تحقیق در خصوص مقایسه امتحانهای انجام شده در خانه و کلاس بسیار قلیل است. شواهدی وجود دارد که امتحان شفاهی و کتبی هر دو به یک میزان تاثیر دارند و هر دوی این روشها بهتر از اصلاً امتحان نکردن است (کلھون، ۱۹۶۲)^(۷).

1- Gay and Gallagher

2- Sarason, Davidson, Lighthall, Waite, and Ruebush, 1960

3- Castaneda, McCandless, and Palermo, 1964

5- Anastasi

7- Calhoun, 1962

4- Denny, Paterson, and Feldhusen, 1964

6- Selye, 1978

همچنین، بررسیها نشان می‌دهد که پیشرفت تحصیلی دانشجویان دورهٔ لیسانس بر اساس نظام رده‌بندی قبول / رد، کمتر از دانشجویان تحت نظام رده‌بندی استاندارد است (بین، هیلز، و راند، ۱۹۷۳؛ گلد، ریلی، سیلمن، و لهر، ۱۹۷۱)^(۱).

در سال ۱۹۷۶ یک بررسی (گی و گالاگر، ۱۹۷۶) انجام شد که در آن یک کلاس دورهٔ لیسانس به طور تصادفی به سه قسمت تقسیم شده بود. یک گروه تمرینهای دوره‌ای دریافت می‌کردند تا در خانه انجام دهند، گروه دوم آزمونهای دوره‌ای دریافت می‌کردند؛ و به سومین گروه حق انتخاب یکی از این دو روش داده شد. جالب است اشاره کنیم که هیچیک از افراد گروه سوم، آزمون دوره‌ای را انتخاب نکردند. بجز این نحوهٔ برخورد متفاوت، همهٔ دانشجویان در حد ممکن یکسان رفتار کردند. در پایان دوره، یک آزمون غیر منتظره یکسان به هر سه گروه داده شد. نتایج این آزمون نشان داد نمرات گروهی که قبلاً آزمون دوره‌ای دریافت می‌کردند، به طور قابل ملاحظه‌ای از دو گروه دیگر بیشتر بود.

این بررسی از جنبه‌های مختلف با سایر تحقیقات متفاوت است. نخست آنکه، این بررسی کلاسهای متنوعی را در رشته‌های ادبی، رفتاری و کاربردی در بر می‌گرفت؛ بنابراین نتایج احتمالاً قابل تعمیم به مخاطبان وسیعتری هستند. در این بررسی تلاش بر این بود تا سطوح ویژه یادگیری بر اساس سلسله مراتب بلوم در یادگیری شناختی اندازه‌گیری شود (بلوم، انگل‌هارد، فارست، هیل، و کرث‌هول، ۱۹۵۶)^(۲). بالاخره تلاش شد که با توجه به نتایج آزمایش، علت مسئله شناسایی شود.

روشها

دانشجویان ۱۰ کلاس از ۵ دانشگاه، به عنوان شرکت کنندگان این بررسی انتخاب شدند. مجموع نمونه را ۲۵۸ نفر دانشجو از کلاسهای روان‌شناسی، آمار، تعلیم و تربیت، تحقیق و تاریخ تشکیل می‌دادند. در شروع ترم تحصیلی استاد به همراه یک پاکت محتوی دستورالعملهای دقیق، توضیح کوتاهی در مورد شیوه‌های کار به هر دانشجو می‌داد. به دانشجویان گفته می‌شد که نیمی از آنان یک امتحان در خانه دریافت خواهند کرد، که یک

1- Bain, Hales, and Rand, 1973; Gold, Reilly, Silbeman, and Lehr, 1971

2- Bloom, Engelhart, Furst, Hill, and Krathwohl, 1956

آزمایش علمی است، هیچ کس از نتیجه‌ای که به دست می‌آورد، لطمه نخواهد خورد و هر کس که از نمره خود رضایت نداشته باشد، حق انتخاب یک آزمون دیگر را خواهد داشت. انتظار می‌رفت که این شیوه در عامل اضطراب اثر خنثی سازی داشته باشد.

حدود دو هفته قبل از شروع نیمه دوم دوره، کلاسها به طور تصادفی به دو گروه مساوی الف و ب تقسیم شد. در هر کلاس، به افرادی که به طور تصادفی در گروه الف انتخاب شده بودند، یک امتحان در منزل داده شد که می‌باید یک هفته بعد برمی‌گرداندند. به اعضای گروه ب یک آزمون یکسان در سر کلاس داده شد. یک هفته بعد در سر کلاس عادی به هر دو گروه یک آزمون غیر منتظره داده شد. در آزمون اخیر همان مواد درسی گنجانده شده بود که هفته قبل نیز مورد آزمون قرار گرفته بودند. ضمن آنکه اعضای هر دو گروه تحت شرایطی یکسان و زیر نظر یک استاد آموزش می‌دیدند، گروه آزمایشی و گروه گواه در حد امکان به طور یکسان در معرض آزمایش از نوع کلاسی قرار گرفتند. تنها روش (نوع آزمون) برای هر کلاس متفاوت بود.

آزمون غیر منتظره حاوی ۵۰ سؤال چند جوابی بود، که ۲۵ سؤال اطلاعات دانشجویان و ۲۵ سؤال دیگر ترکیب فهم مطلب و کاربرد محتوای دوره را می‌سنجید. آخرین دسته‌بندیها، سه سطح اولیه یادگیری شناختی در سلسله مراتب بلوم را (بلوم و دیگران، ۱۹۵۶) شامل می‌شد. برای کمک به تعیین میزان انگیزه هر دانشجو در خواندن و یاد گرفتن درس، جهت آماده شدن برای هریک از انواع امتحان، یک پیمایش صورت گرفت. ۹۶/۱ درصد دانشجویان به این پرسشنامه پاسخ دادند. در نهایت، یک تحلیل واریانس دو عاملی مورد استفاده قرار گرفت. از آنجا که دو گروه از ۱۰ کلاس مختلف انتخاب شده بودند، طرح بررسی را یک جدول 2×10 تشکیل می‌داد (اری و جیکوبز، ۱۹۷۶).^(۱)

نتایج

نمرات گروه ب (در آزمون کلاسی) به طور قابل ملاحظه‌ای از گروه الف (آزمون در منزل) در همه طبقه بندیهای آزمون غیر منتظره، بالاتر بود.

گروه آزمایشی به طور قابل توجهی یادگیری بیشتر، دانش بیشتر، و فهم بیشتری را از مطالب و کاربرد مواد درسی به نمایش گذاشت. به این ترتیب تجزیه و تحلیل‌های آماری در مورد هر طبقه

نشان داد که متغیر مستقل (نوع آزمون) و متغیر وابسته (یادگیری) به شدت به هم مربوطند. در پیمایشی که در مورد آزمون غیر منتظره انجام شد، یکی از سؤالات این بود: «اگر شما یک آموزگار بودید و می خواستید شاگردانتان مواد درسی را بواقع یاد بگیرند، کدامیک از انواع آزمون را به آنها می دادید؟» $25/8$ درصد از پاسخگویان اظهار داشتند که آزمون قابل انجام در منزل را خواهند داد، $59/1$ درصد آزمون کلاسی را انتخاب کردند، و $15/1$ درصد قادر به تصمیم گیری نبودند. این طور به نظر می رسید که آنها به نسبت 2 به 1 این روش را ترجیح می داده اند. دو سؤال برای تعیین نگرش آزمودنیها در مورد نوع خاصی از آزمون، طراحی شده بود. در اولین سؤال از دانشجویان خواسته می شد تا مجموعه تلاش خود را برای آموختن مطالب موجود در آزمون معلم ساخته آنها که هفته پیش به عنوان آزمون غیر منتظره دریافت کرده بودند، در یک پیوستار ده درجه ای، مشخص کنند. در سؤال دوم، که آن هم شامل یک پیوستار ده درجه ای بود، از آنها خواسته شد تا پاسخ دهند چه میزان تلاش برای انجام دیگر انواع آزمون (آزمون قابل انجام در منزل یا کلاس) بایستی در نظر گرفته شود.

امتیاز گروه الف به طور قابل ملاحظه ای در سؤال اول پایین تر بود. بنابراین، دانشجویان اظهار داشتند که آنها در صورت داشتن آزمون کلاسی، تلاش بیشتری از خود نشان خواهند داد. بر عکس امتیاز گروه ب به طور قابل ملاحظه ای در سؤال دوم پایین تر بود. محقق می تواند این نتیجه را چنین تفسیر کند که دانشجویان در صورت داشتن آزمون قابل انجام در منزل کمتر درس خواهند خواند. نتیجه ای که تفاوت در یادگیری بین گروهها را توضیح می دهد، این است که آن دسته از دانشجویانی که آزمون کلاسی داشته اند، به اعتراف خود برای مطالعه تلاش بیشتری کرده اند.

نتیجه گیری

این داده ها شاهی است بر مقبولیت آزمون کلاسی، در مقابل آزمون قابل انجام در منزل، در بین دانشجویان مختلف دانشگاه، در رشته های روان شناسی، آمار، تحقیق، تعلیم و تربیت و تاریخ. در این آزمایش دو نوع آزمون به کار گرفته شد. در مجموع، یادگیری دانشجویان در سه سطح اولیه سلسله مراتب بلوم در یادگیری شناختی، یعنی دانش، فهم مطالب و کاربرد، تاثیر داشته است. به طور وضوح آزمون کلاسی شیوه بهتری بود، و به نظر می رسید که دانشجویان با آن موافق هستند. امّا، به دلایل بسیار، محقق نمی تواند اظهار کند که آزمونهای سنتی کلاسی

چون یادگیری مؤثر را افزایش می‌دهند، تنها روش معتبر هستند. با تأکید بر آرمانها و اهداف آزمونها به تحقیق بیشتری در مورد روشهای اندازه‌گیری و مقایسه سطوح یادگیری در دامنه‌ای وسیع نیاز است.

/// خلاصه

اولین وظیفه‌ای که در پیش روی محققان است، انتخاب و صورتبندی یک مسئله است. یک مسئله تحقیقی، عبارتی سؤالی یا اخباری در مورد رابطه موجود بین متغیرهاست. در تلاش برای یافتن یک مسئله قابل تحقیق، پژوهشگر می‌تواند به تجربیات شخصی خود، نظریه‌هایی که می‌توان از آنها سؤالاتی استنتاج کرد یا به متون تحقیقی جاری در حوزه‌های مورد نظر مراجعه کند. آنها باید اهمیت مسئله پیشنهادی را برحسب معیارهای ویژه ارزیابی کنند، یعنی از خود بپرسند: آیا این مسئله در پیکره دانش امروزی مؤثر خواهد بود؟ آیا این مسئله امکان حرکت به سمت تحقیق آتی را فراهم می‌آورد؟ آیا قابل آزمون است، یعنی آیا متغیرها می‌توانند مشاهده و اندازه‌گیری شوند؟ مسئله تا چه حد با علایق، تجربیات و دانش من در این زمینه متناسب است؟ آیا به داده‌هایی که برای این مسئله مورد نیاز است، دسترسی دارم و آیا وسایل اندازه‌گیری موجود است، یا می‌توان آنها را برای اندازه‌گیری متغیرها ساخت؟ سؤالات نباید مستقیماً با مسائل فلسفی درگیر و همچنین نباید آنقدر کلی باشند که تحقیق آن غیر ممکن به نظر رسد. بیان پرسش باید جامعه آماری مورد نظر و متغیرهای مورد بررسی را معین کند.

// مفاهیم اصلی

بیان مسئله	جامعه آماری
متغیر مستقل	متغیر وابسته
مسئله تحقیق	معیارهایی برای مسئله تحقیق
منابع موجود در زمینه مسئله تحقیق	

/// تمرینها

- ۱- یک گزارش تحقیقی چاپ شده در یک مجله پیدا کنید و بر اساس مطالب خوانده شده به پرسشهای زیر پاسخ دهید:

الف - چه مسئله‌ای در این بررسی مورد پژوهش قرار گرفته است؟

ب - فرضیه (ها) کدامند؟

ج - متغیرهای مستقل و وابسته کدامند؟

د - مسئله و فرضیه (های) بیان شده در گزارش را در کجا یافتید؟

ه - آیا مسئله و فرضیه (ها) آنقدر روشن بیان شده‌اند که شما بدانید چه چیزی دقیقاً باید

مورد پژوهش قرار گیرد؟

۲- در حیطه‌ای که احتمالاً به تحقیق در آن علاقه دارید، یک زمینه وسیع را انتخاب کنید. یک

جنبه خاص از این زمینه وسیع در نظر بگیرید و سپس یک مسئله تحقیق را که به آن علاقه‌مند هستید، برگزینید. این مسئله را در یک قالب تحقیقی بیان کنید.

۳- نمونه‌های زیر بیانهای نامناسبی از مسائل تحقیقی هستند. هر یک را دوباره چنان بیان

کنید که به یک سؤال خاص مناسب برای تحقیق تبدیل شوند.

الف - اثرات شیوه‌های مختلف یادگیری مفاهیم علمی

ب - اضطراب و پیشرفت تحصیلی

ج - نگرشهای کودکانی که دارای محرومیت فرهنگی هستند.

د - مشاوره و دانش‌آموزان ضعیف

ه - اثر بخشی روش کوئیزنر^(۱) در تدریس ریاضیات ابتدایی

۴- مسائل تحقیقی ذیل را ارزشیابی کنید:

الف - آیا فلسفه دکتر اسپاک در مورد تربیت کودکان سهل انگار، تأثیر نامطلوبی در آموزش و

پرورش آمریکا داشته است؟

ب - چه رابطه‌ای بین شیوه پا روی پا انداختن و هوش دانشجویان زن در کالج آمریکایی

وجود دارد؟

ج - با توجه به بررسیهای کاربردی اخیر در پیشرفت ریاضیات، آیا باید ریاضیات «جدید» در

مدارس ابتدایی منسوخ شود و روش سنتی ریاضیات جایگزین آن شود؟

د - دانش‌آموزان نقش مدیر مدرسه را در Central Middle School چگونه تعبیر می‌کنند؟

۵ - محتمل‌ترین متغیرهای مستقل و وابسته را در بررسیهای زیر بیان کنید:

- الف - لین، اس، و ج، برگن^(۱) (۱۹۸۸). اثرات متغیرهای آموزشی در توانایی یادگیری زبان در بین کودکان پیش دبستانی. *مجله تحقیقات آموزشی آمریکا*^(۲)، ۲۸۴ - ۲۷۱ و ۲۵.
- ب - لین، آر و سی هاستینگز^(۳) (۱۹۸۴). اعتبار پیش‌بینی کننده‌های موفقیت در دانشکده حقوق. *مجله اندازه‌گیری آموزشی*^(۴)، ۲۵۸ - ۲۴۵ و ۲۱.
- ج - هاتیوا، ان^(۵) (۱۹۸۸). مشق کامپیوتری و تمرین در حساب: وسیع تر شدن شکاف بین دانش‌آموزان قوی و ضعیف. *مجله تحقیقات آموزشی آمریکا*، ۳۹۷ - ۳۶۶ و ۲۵.
- د - چاپمن، ج.^(۶) (۱۹۸۸). خود پندارهٔ کودکان ضعیف در یادگیری. *مجله بررسی تحقیقات آموزشی*^(۷)، ۳۶۶ - ۳۴۷ و ۵۸.
- ه - ویر، ان و وی لی^(۸) (۱۹۸۸). تفاوت‌های جنسی در انتخاب رشته‌های علوم در دانشکده. *مجله تحقیقات آموزشی آمریکا*، ۶۱۴ - ۵۳۹ و ۲۵.
- ۶ - منابع مسائل تحقیقاتی مورد استفاده محققان را فهرست کنید و برای هر کدام یک مثال بیاورید.
- ۷ - خواندن کتابهای خوب چه تأثیری در مهارت‌های خواندن کودکان دارد؟ آیا این سؤال قابل تحقیق است؟ اگر نیست، برای آنکه بتوان آنرا به طور کاربردی مورد پژوهش قرار داد، چه تغییراتی را پیشنهاد می‌کنید؟

/// پاسخها

- ۱- پاسخها متنوع خواهند بود.
- ۲- پاسخها متنوع خواهند بود.
- ۳- الف - آیا بین نمرات پیشرفت درس علوم دانش‌آموزان کلاس ششم انجمن تحقیق علم (SRA)^(۹) که یک سال برنامهٔ آموزش علوم ابتدایی (ESS) را گذرانده‌اند، با نمرات آنهایی که دورهٔ یکساله مبتنی بر کتاب درسی علوم را دیده‌اند، تفاوتی وجود دارد؟

1- Lane, S., and J. Bergan (1988).

3- Linn, R., and C. Hastings (1984).

5- Hativia, N. (1988).

7- View of Educational Research, 58, 347 -366.

9- Science Research Associates (SRA).

2- American Educational Research Journal 25, 271 -284.

4- Journal of Educational Measurement, 21, 245- 258.

6- Chapman, J. (1988).

8- Ware, N., and V. Lee (1988).

ب - آیا بین نمرات آزمون ترکیبی استانفورد و نمرات مقیاس اضطراب آشکار^(۱) در بین دانش‌آموزان دوره دبیرستان تفاوتی وجود دارد؟

ج - محقق ابتدا باید مشخص کند که کدامیک از نگرشهای کودکان مورد بررسی قرار گرفته‌اند و سپس آنها را با نگرشهای یک گروه کنترل مناسب (شاید کودکانی که دچار محرومیت فرهنگی نیستند) مقایسه کند، یا در مورد رابطه بین انواع نگرشها و دیگر رفتارها فرضیه‌ای ارائه کند. برای مثال، آیا کودکان دارای محرومیت فرهنگی با نگرش مثبت نسبت به مدرسه، از کودکان با نگرش منفی نسبت به مدرسه، نمرات بهتری کسب می‌کنند؟

د - اثر یک برنامه مشاوره بر نگرشهای دانش‌آموزان ضعیف دوره راهنمایی در مورد مدرسه چیست؟

ه - آیا نمرات درس حساب دانش‌آموزان کلاس سوم، که یک سال از روش کوئیزر در تدریس ریاضیات ابتدایی بهره‌مند بوده‌اند، با نمرات آنهایی که روش تدریس دیگری داشته‌اند، متفاوت است؟

۴- الف - این سؤال متضمن یک قضاوت ارزشی است که بررسی کاربردی آن غیرممکن است.

ب - این سؤال پیش‌پاافتاده است، و پاسخ آن سهم ناچیزی در دانش موجود دارد.

ج - تحقیق قادر به پاسخگویی به سؤالات ارزشی نیست؛ تنها قادر به فراهم ساختن اطلاعاتی است که بر مبنای آن بتوان تصمیماتی اتخاذ کرد.

د - با اینکه سؤال می‌تواند مورد بررسی قرار گیرد، اما از جنبه تعمیم دادن به سایر موقعیتها بسیار محدود است.

۵- الف - متغیر مستقل: روشهای آموزشی

متغیر وابسته: برخی از اندازه‌های توانایی زبان

ب - متغیر مستقل: پیش‌بینی‌های متعدد (نمره‌های آزمونها و غیره)

متغیر وابسته: بعضی اندازه‌ها در عملکرد دانشکده حقوق (شاید معدل)

ج - متغیر مستقل: استفاده از تمرین کامپیوتری در ریاضیات

متغیر وابسته: پیشرفت در درس حساب برای دانش‌آموزان ضعیف و قوی.

د - متغیر مستقل: ضعیف بودن در یادگیری

متغیر وابسته: بعضی اندازه‌های خود پنداره

ه - متغیر مستقل: جنسیت

متغیر وابسته: انتخاب رشته علمی در دانشکده

۶ - تجربیات روزمره کارشناسان آموزش و پرورش

قیاس از نظریه

تحقیقات پیشین

منابع غیر آموزشی

۷ - اصطلاح کتاب خوب و مهارت خواندن نیاز به تعریف دارد. بیان مجدد آن اینگونه است

که اثر خواندن چهار متن کلاسیک مشخص، در گنجینه لغات کودکان چیست.

فصل سوم

تحقیقات پیشین

هدفهای آموزشی

دانشجو پس از مطالعه این فصل، قادر خواهد بود:

- ۱- هفت کارکرد اصلی در تحقیق را که از مرور تحقیقات پیشین حاصل می‌شود، فهرست کند.
- ۲- منابع مرجع اصلی در تعلیم و تربیت را توضیح دهد.
- ۳- نمایه‌های اریک را نام برد و تشابهات و تفاوت‌های آنها را تشریح کند.
- ۴- یک رویکرد کلی برای استفاده نمایه‌های نشریات علمی در تحقیق را خلاصه کند.
- ۵- حداقل سه نمایه غیر از اریک را نام برد و از اطلاعات زمینه‌ای در هر کدام نمونه‌ای ذکر کند.
- ۶- چگونگی استفاده از SSCI را در مروری از تحقیقات پیشین تشریح کند و عملکرد هر یک از چهار فهرست را توضیح دهد.
- ۷- کاربرد سالنامه اندازه‌گیری روانی^(۱) را به عنوان یک منبع اطلاعاتی در آزمون‌ها و ابزار اندازه‌گیری، توصیف کند.
- ۸- شیوه‌های یک جستجوی کامپیوتری از تحقیقات پیشین را توضیح دهد.
- ۹- پیشرفت نظام‌مند مرحله به مرحله در سازماندهی تحقیقات پیشین را شرح و مقصود از هر مرحله را توضیح دهد.

1- Mental Measurment Yearbooks.

این فصل با همکاری ساموئل. تی. هوآنگ، بنیانگذار کتابخانه دانشگاه کارولینای شمالی، تجدید نظر شده و اطلاعات آن به روز شده است.

وقتی که موضوع انتخاب شد، محقق طبیعتاً مشتاق عمل خواهد بود. به هر حال حمله شتابزده برای طراحی و انجام بررسی، قبل از یافتن زمینه کاملی از آنچه که پیش از این در حیطه مورد نظر انجام یافته، یک اشتباه است. محقق باید موضوع مورد نظر خود را به دانش موجود در زمینه مورد علاقه خود ارتباط دهد. قابلیت تعیین، سازماندهی و استفاده از منابع در زمینه مورد نظر، برای کارشناس تعلیم و تربیت و سایر افراد درگیر در تحقیق اهمیت یکسانی دارد.

این فصل مربوط است به (۱) نقش تحقیقات پیشین در یک پروژه تحقیقی، (۲) منابع مرجع در تعلیم و تربیت و (۳) سازماندهی تحقیقات پیشین برای ارائه در گزارش.

/// نقش تحقیقات پیشین در یک پروژه تحقیقی

قبل از شروع بررسی باید جستجو برای تحقیقات پیشین به طور عملی تکمیل شده باشد. این مرحله به اهداف متعدد مهمی خدمت می‌کند.

۱- آگاهی از تحقیقات پیشین محقق را قادر می‌سازد که حدود موضوع تحقیق خود را تعیین کند. اگر از قیاس تمثیلی استفاده کنیم، یک کاشف ممکن است بگوید: «ما می‌دانیم در دو هزار مایلی غرب و در پشت آن گیاهان یک رشته کوه قرار گرفته است، ولی نمی‌دانیم که در پشت این رشته کوه چه موقعیت و وضعیتی وجود دارد. من پیشنهاد می‌کنم که از منطقه گیاهان گذر کنیم. به بالای کوهها برویم و از آنجا به طرف غرب حرکت کنیم.» به همین نحو، یک محقق می‌گوید: «محققان الف، ب و ج سؤال مرا تا این حد پاسخ داده‌اند، محقق د نیز با تحقیقات خود، این قدر به دانش ما افزوده است. من پیشنهاد می‌کنم برای اینکه از تحقیقات د فراتر رویم، به این شیوه عمل کنیم.»

۲- جستجوی کامل در نظریه‌ها و تحقیقات مربوطه، محقق را قادر می‌سازد که سؤال خود را در یک چشم‌انداز قرار دهد. محقق باید تعیین کند که آیا تلاشهایش به شیوه‌ای معنادار احتمالاً بر دانش موجود خواهد افزود؟ دانش موجود در هر یک از زمینه‌های مورد نظر مبتنی بر انباشت نتایج بررسیهای نسلهای بشمار محققان و نیز نظریه‌های طراحی شده برای تلفیق این دانش و توضیح پدیده مورد مشاهده است. یک محقق ممکن است جستجو در تحقیقات پیشین را با هدف یافتن پیوستگی بین بررسی خود و مجموعه دانش موجود در حوزه مورد نظر آغاز کند. بررسیهایی که فاقد هر گونه پیوستگی با دانش موجود هستند، به ندرت به زمینه خاص خود

کمک معنی داری می‌کنند. این‌گونه بررسیها به اطلاعاتی مهجور یا بهره‌وری محدود تبدیل می‌شوند.

۳- جستجو در تحقیقات پیشین به محقق کمک می‌کند تا سؤال خود را محدود کند و مفاهیم بررسی را تعریف و روشن سازد. یک سؤال تحقیقی ممکن است برای انجام بررسی، بسیار گسترده باشد و یا برای آنکه بتوان به طور عینی کار را با آن شروع کرد، بسیار مبهم باشد. جستجوی دقیق در تحقیقات پیشین می‌تواند به محققان در جرح و تعدیل سؤال اولیه به نحوی که قابل تحقیق باشد، کمک کند. این کار همچنین در روشن ساختن مفاهیم موجود در بررسی و تبدیل این مفاهیم به تعاریف کاربردی مفید است. بسیاری از سازه‌های آموزشی و رفتاری از قبیل تنش، خلاقیت، ناکامی، پرخاشگری، پیشرفت، انگیزش و سازگاری، نیاز به روشن شدن و تعریف شدن به شکل کاربردی دارند. این سازه‌ها به موازات بسیاری از مفاهیم آموزشی و رفتاری دیگر، تا زمانی که نتوان آنها را کمی کرد، برای تحقیقات علمی مناسب نیستند. محقق، در جستجوی منابع پیشین، با تلاشهای گذشته برای روشن ساختن این مفاهیم و تعریف کاربردی آنها، آشنا می‌شود. جستجوهای موفق اغلب در تشکیل فرضیات مربوط به روابط بین متغیرهای موجود در بررسی نتیجه‌بخش است. بررسیهایی که در آنها فرضیاتی مورد آزمون قرار می‌گیرند، اغلب از آنهایی که فاقد فرضیه هستند، مفیدترند.

۴- یک جستجوی موشکافانه در تحقیقات پیشین اغلب به شناخت دلایل مربوط به نتایج متناقض در یک حیطه خاص منجر می‌شود. نتایج متضاد برای عموم شناخته شده نیستند. دلایل این تناقضات را می‌توان در انواع رویکردهای انتخاب شده برای حل مسئله یا انواع وسایل به کار گرفته شده، روشهای مورد استفاده و یا تجزیه و تحلیل‌های انجام شده، یافت. حل چنین تضادهایی به بحث و جدل می‌انجامد، اما آنها می‌توانند به دانش موجود در حوزه مورد نظر به طور چشمگیری کمک کنند.

۵- پژوهشگراز طریق جستجوی تحقیقات پیشین در می‌یابد که کارآیی کدام روشها به اثبات رسیده است و کدامیک اطمینان بخش به نظر نمی‌آیند. محقق با غور و تفحص در تحقیقات پیشین و تعالی بخشیدن به دانش خود ممکن است بزودی راهی را بیابد که از آن طریق بتواند مطالعات خود را به پیش برد. البته پس‌نگری^(۱) همیشه بهتر از پیش‌نگری^(۲) است؛ از این رو شاید این امر که مطالعات اولیه در حیطه‌ای خاص اغلب خام و نامؤثر به نظر می‌رسند،

اجتناب ناپذیر باشد. بسیاری از پروژه‌های تحقیقاتی به دلیل استفاده از شیوه‌ها، ابزار، طرح بررسی و یا تجزیه و تحلیل‌های آماری نامناسب، رد می‌شوند. واری‌های کامل روش‌های به کار گرفته شده در بررسی‌های پیشین اغلب هم در یافتن دلیل شکست آنها و هم در ایجاد بینش برای انتخاب روش مناسب برای بررسی مورد نظر محقق ثمر بخش است. هم موفقیت و هم عدم موفقیت کارهای گذشته متضمن بینشی برای محقق در طراحی بررسی‌اش خواهد بود. با بنیان نهادن بررسی‌های آتی بر تحقیقات پیشین، می‌توانیم پیشرفت فزاینده‌ای در دانش خود در زمینه تحقیقات آموزشی انتظار داشته باشیم.

۶- یک جستجوی کامل در تحقیقات مرتبط با موضوع، از تکرار سهوی بررسی‌های گذشته جلوگیری می‌کند. بارها اتفاق افتاده است که یک محقق ایده‌پردازشی را ارائه و سپس کشف کرده است که این ایده در تحقیقی مشابه قبلاً به کار گرفته شده است. در چنین مواردی او باید تصمیم بگیرد که آیا باید بررسی قبلی را به طور عمد تکرار کند و یا برنامه‌اش را تغییر دهد و جنبه دیگری از مسئله را بررسی کند.

۷- مطالعه پیشینه تحقیق، محقق را برای تفسیر اهمیت نتایج خود در موقعیت بهتری قرار می‌دهد. آشنایی با نظریه مرتبط با موضوع و تحقیقات گذشته، محقق را برای هماهنگ ساختن یافته‌های تحقیقاتش با پیکره دانش موجود در همان حیطه، آماده می‌سازد.

/// منابع مرجع در تعلیم و تربیت

برای هر دانشمند یا محقق آگاهی از نحوه دستیابی به کارهای پیشین در زمینه مورد نظر، امری بسیار ضروری است. برای انجام این مهم، محقق باید با (۱) با منابع کارهای پیشین، (۲) سازمانی که این اطلاعات را جمع‌آوری و آنها را در بانکهای اطلاعاتی سازماندهی می‌کند، (۳) راه دستیابی به این بانکهای اطلاعاتی و (۴) راههای مؤثر برای یافتن اطلاعات مورد نیاز، آشنا باشد. برای استفاده از این منابع، فرد باید از تسهیلات و خدمات کتابخانه‌ای قابل دسترس مطلع باشد. بسیاری از کتابخانه‌ها کتابچه راهنمایی چاپ کرده‌اند که خدمات و مقررات آنها را شرح می‌دهند یا به منظور آشنا ساختن علاقه‌مندان، بازدیدهایی ترتیب می‌دهند.

این مسئله بویژه برای آشنایی با نحوه سازماندهی کارت کاتالوگها بسیار مهم است. بسیاری از کتابخانه‌ها از طریق سیستم کامپیوتر کتابخانه، از یک فهرست کامپیوتری استفاده می‌کنند. اگر

سیستم کتابخانه‌ای شما کامپیوتری شده باشد، شاید بخواهید با چگونگی استفاده مؤثر از این سیستم آشنا شوید. در این بین شما باید دریابید که آیا کتابخانه می‌تواند کتابها و سایر مدارک نظیر رساله‌ها و مقالات نشریات علمی را از طریق سرویس امانتدهی بین کتابخانه‌ای، از سایر موسسات به امانت بگیرد یا خیر.

// راهنماهای اصلی

برای آغاز کردن بررسیهای تحقیقی، رجوع به راهنماهای اصلی تحقیقات پیشین مفید است. یکی از این راهنماهای مفید *راهنمای کتابشناسی در تحقیقات آموزشی* بری^(۱) است. این راهنما به «عنوان یک راهنمای فشرده، برای کمک به دانشجویان در رشته‌های تعلیم و تربیت و برای استفاده مؤثر آنها از منابع کتابخانه دانشکده یا دانشگاهشان، در نظر گرفته شده است» این یک فهرست مشروح بالغ بر ۷۰۰ منبع تحقیقی اصلی است که بر اساس انواع مواد، اعم از نشریات علمی، بررسیهای تحقیقی، نشریات دولتی، متون مرجع و آزمونها، مرتب شده است. متون مرجع شامل سالنامه، کتاب راهنما، منابع زندگینامه‌ای و کتابچه‌های راهنما در روش شناسی تحقیقات آموزشی، در سبک و سیاقی مناسب برای نوشتن مقالات تحقیقی است.

منابع کتابشناسی روز در تعلیم و تربیت^(۲) که به سه زبان (انگلیسی، فرانسه و اسپانیایی) منتشر می‌شود، عمدتاً بر اساس مجموعه مراجع دفتر بین‌المللی اسناد آموزشی یونسکو در پاریس پایه‌ریزی شده است. این نشریه شامل کتابشناسی‌ها، فهرستهای نشریات علمی و نامنامه تحقیقات به ترتیب کشور مربوطه، یا فهرستهای مجزا برای سازمانهای نشر بین‌المللی و منطقه‌ای است.

راهنمای منابع اطلاعات آموزش وودبری^(۳) راهنمای جامع برای یافتن اطلاعات در زمینه تعلیم و تربیت است. این راهنما، منابع چاپی، غیر چاپی و سازمانی را انتخاب و توصیف می‌کند و در بسیاری از موارد آنها را برای تحقیق در تعلیم و تربیت مورد ارزشیابی قرار می‌دهد. ترتیب آن اساساً بر مبنای نشریات، با فصولی در مورد تامین مالی و دولت، تعلیم و تربیت استثنایی، مواد آموزشی، آزمونها و وسایل سنجش است. این فهرست همچنین شامل توصیفی از مراحل یک تحقیق مؤثر و یک راهنما برای پژوهشگران است.

1- Berry, D. M. (1980). *A Bibliographic Guide to Educational Research* (2d ed.). Metuchen, NJ: Scarecrow Press.

2- Unesco (1984). *Current Bibliographical Sources in Education* (2d ed.). Paris: Unesco, International Bureau of Education.

3- Woodbury, M. (1982). *A Guide to Sources of Educational Information* (2d ed.). Arlington, VA: Information Resources

راهنمای کتابهای مرجع شی‌می^(۱) چند هزار منبع مرجع را در موضوعات مختلف، از جمله تعلیم و تربیت، به طور مختصر توصیف و ارزشیابی می‌کند. اطلاعات آن از طریق ضمايم دو سالانه به روز در می‌آیند. کتابهای مرجع سالانه آمریکا^(۲) کتابهای مرجع منتشره سال قبل (شامل کتابهای تجدید چاپ شده) را در کلیه موضوعات در بر می‌گیرد و حواشی توصیفی و ارزیابانه‌ای را فراهم می‌آورد.

// نشریات علمی در زمینه تعلیم و تربیت

نشریاتی که بررسیهای تحقیقاتی را در موضوعات خاص ادغام و خلاصه می‌کنند، محقق را قادر می‌سازند تا بررسی خود را با مرور تحقیقات پیشین، آغاز کند.

دائرة المعارف تحقیقات آموزشی میتزل^(۳) برای ارائه یک ترکیب و تحلیل انتقادی از گزارشهای تحقیقی در تعلیم و تربیت طراحی شده است و شامل مقالات مؤثر با بحثهای مستند به مدارک کتابخانه‌ای از جریانه‌ها و تحولات اخیر، به همراه موضوعات سستی است. این دائرة المعارف چهار جلدی شامل حدود ۱۶۴ موضوع است. این دائرة المعارف منبع اصلی مناسب برای تحقیق در مورد موضوعات بنیادی در زمینه تعلیم و تربیت می‌باشد. دائرة المعارف ارزشیابی آموزشی^(۴) سعی بر ارائه مفاهیم و تکنیکهای اصلی برای ارزشیابی برنامه‌های تربیت معلم دارد و مقالات آن به ترتیب الفبا تنظیم شده است. دائرة المعارف کارشناسان تعلیم و تربیت آمریکا^(۵) شامل حدود ۲۰۰۰ مدخل کوتاه برای نامها و اصطلاحاتی است که اغلب در متون تخصصی تعلیم و تربیت به آنها برمی‌خوریم. دائرة المعارف بین‌المللی تعلیم و تربیت: «پژوهشها و بررسیها»^(۶)، یک بررسی کلی از جدیدترین دستاوردهای دانش‌پژوهی در حل مسائل تعلیم و تربیتی، تجربیات، و سازمانهای تحقیقی در سراسر جهان است. این دائرة المعارف کارهای تحقیقی و کارشناسی در تعلیم و تربیت را به مفهوم وسیع در بر می‌گیرد و آخرین پیشرفته‌ها در شاخه‌های مختلف تعلیم و تربیت، قابلیت دستیابی به اطلاعات علمی

press

ادامه پاورقی صفحه قبل

- 1- Sheehy, E. P. (1986). Guide to Reference Books (10 Th ed.). Chicago: American Library Association.
- 2- Wynar, B. S. (ed.) (1970-). American Reference Books Annual. Littleton, co: Libraries Unlimited.
- 3-Mitzel, H. E. (ed.) (1982). Encyclopedia of Educational Research (5th ed., vols. 1-4). New York: Free Press.
- 4- Anderson, S. B., et al. (1975). Encyclopedia of Educational. San Francisco: Jossey - Bass.
- 5- Dejnozka, E.L, and D.E. Kapel (1982). American Educators' Encyclopedia. Westport, CT: Greenwood Press.
- 6- The International Encyclopedia of Education: Research and Studies (1985). (10vols.) Oxford: pergamon.

مطمئن و معتبر در این حیطه و انواع تحقیقات تکمیلی مورد نیاز را مورد بررسی قرار می دهد. دائرة المعارف بین المللی آموزش عالی^(۱) شامل ۲۸۲ مقاله در مورد موضوعات جدید در آموزش عالی است که توسط پښتازان تعلیم و تربیت در سراسر دنیا نوشته شده است.

یکی دیگر از منابع مفید، کتابچه راهنمای تحقیق در تدریس است.^{(۲) (۳)} از این کتابچه راهنما در فاصله ۱۰ سال سه چاپ متفاوت منتشر شده است. این کتابچه تحقیقات موجود در عرصه تدریس را فهرست، خلاصه و تحلیل انتقادی می کند. هر چاپ شامل مقالات معتبر کارشناسان در موضوعات منتخب در این زمینه است. کتابشناسی های جامع تحت موضوعات منتخب در آن گنجانده می شود. در بین موضوعات چاپ سوم می توان از «اندازه گیری تدریس»، «روشهای کمی تحقیق در زمینه تدریس»، «روشهای کیفی تحقیق در زمینه تدریس»، «مشاهده به عنوان کندوکاو و روش»، «ترکیبات تحقیق در تدریس»، «تدریس راهبردهای یادگیری»، «رفتار معلمان و پیشرفت دانش آموزان»، «تحقیق در مورد تدریس هنرها و زیبایی شناسی» و بررسیهای تحقیقی بیشمار دیگر در زمینه تدریس نام برد. دائرة المعارف تعلیم و تربیت استثنایی^(۴) اطلاعات پایه ای را برای بسیاری از شیوه ها و تخصصهای مربوط به تعلیم و تربیت کودکان استثنایی و ویژگیهای مخصوص به آنها، نیازها و مسائل آنها، فراهم می سازد. این دائرة المعارف سه جلدی شامل متجاوز از ۲۰۰۰ موضوع است. هر چند ترتیب این کتاب الفبایی است، اما بسیاری از موضوعات را می توان از نظر محتوایی در هفت زمینه اصلی دسته بندی کرد: زندگی نامه ها، آزمونهای آموزشی و روان شناختی، مداخله و ارائه خدمات، شرایط معلولیت، خدمات مربوطه، مسائل حقوقی و موضوعات متفرقه.

چهار چاپ از نشریه بررسی تعلیم و تربیت استثنایی^(۵) به همراه نشریه پیشرفتهای تعلیم و تربیت استثنایی^(۶) بررسیهای جامعی در مورد حیطه های تعلیم و تربیت استثنایی ارائه می کند و به ارزشیابی های محدودتر حیطه های خاص آن می پردازد. بررسی تحقیقات آموزشی^(۷) یک

1- Knowles, A. S. (ed.). The International Encyclopedia of Higher Education (1978). (10vols.). San Francisco: Jossey. Bass.

2- Gage, N. L. (ed.). (1963). Handbooks of Research on Teaching: Chicago: Rand McNally.

3- Travers, R. (ed) (1973). Second Handbook of Research on Teaching (2d ed.) Chicago: Rand McNally.

4- Reynolds, C.R. and L. Mann (eds) (1987). Encyclopedia of Special Education: A Reference for the Education of the Handicapped and Other Exceptional Children and Adults (3vols.). New York: John Wiley.

5- Mann, L., and D. Sabatino (eds.) (1973-1980). Review of Special Education (1st- 4th eds). Philadelphia: JSE Press.

6- Keogh, B. K. (1981). Advances in Special Education: A Research Annual, Greenwich, CT: JAI Press.

7- Review of Educational Research (1931-). Washington: American Educational Research Association.

فصلنامه از انجمن تحقیقات آموزش آمریکا (AERA) است که بررسیهای ترکیبی و تفسیر متون تحقیقات آموزشی را هم در مسائل ماهوی و هم در مسائل روش شناختی، منتشر می‌کند. از سال ۱۹۳۱ تا ۱۹۷۰ در هر شماره نشریه، به ترتیب، یکی از ۱۵ موضوع گسترده در زمینه تعلیم و تربیت مورد بررسی قرار گرفت. در ژوئن سال ۱۹۷۰ نشریه این سیاست را ترک و شروع به چاپ مقاله‌های ارائه شده کرد، از آن پس هر شماره آمیزه‌ای از موضوعات گوناگون است. در تلاشی برای تهیه چکیده‌هایی از بررسیها در زمینه موضوعات گسترده یک حیطه تخصصی، AERA مجموعه بازنگری تحقیق در تعلیم و تربیت^(۱) را که از سال ۱۹۷۳ هر ساله چاپ می‌شود، تدارک دید. تلاش این مجموعه بر آن است که مشخص کند در حیطه تعلیم و تربیت چه تحقیقاتی انجام شده است، چه تحقیقاتی در دست اجرا است و چه تحقیقاتی لازم است که صورت گیرد. برای مثال، جلد سیزدهم بازنگری تحقیق در تعلیم و تربیت^(۲) شامل مرور پیشرفتهای انجام‌شده در تحقیق و در عمل، در چهار زمینه اصلی: آموزش ابتدایی قرائت، پیشرفت مهارتهای نوشتن، نیروی تدریس و فرا تحلیل انجام آزمایشهای گوناگون در تعلیم و تربیت است.

بررسیهای روان شناختی نقش مهمی در تحقیقات آموزشی ایفا می‌کند. سالنامه بررسیهای روان شناختی، که از سال ۱۹۵۰ هر ساله چاپ می‌شود، متخصصان حیطه‌های مختلف روان شناسی را به کار می‌گیرد تا در مورد متون تحقیقاتی گزارش تهیه و آن تحقیقات را ارزیابی کنند، و روندها و پیشرفتهای جدید را در همه جنبه‌های علم روان‌شناسی شناسایی و حوزه‌های فراموش شده را مشخص کنند. کتاب راهنمای روان‌شناسی عمومی^(۳) یک منبع مطالعاتی مفید است که به حوزه‌های وسیعی از تاریخ، نظریه، روش‌شناسی، ارگانیزم انسان، ادراک، یادگیری، زبان، تفکر و هوش، انگیزش و هیجان، شخصیت و سایر حوزه‌های خاص می‌پردازد. این نشریه شامل منابع کتابشناختی از بررسیهای تحقیقی موجود است.

// نمایه نشریات دوره‌ای، مجلات خلاصه تحقیقات و فهرست نقل نامها
امروزه با پی‌ریزی شالوده‌گسترده‌ای از تحقیق، نظریه و عقاید عمومی در یک موضوع خاص،

1- Review of Research in Education (1973 -). Washington: American Educational Research Association.

2- Annual Review of Psychology (1950-). Palo Alto: Annual Reviews.

3- Wolman, B. B.(1973). Handbook of General Psychology, Engle Wood- Cliffs, NJ: Prentice - Hall.

محقق می‌تواند مدارک و مطالب اضافی را که در منابع اصلی بررسی نقل نشده‌اند، پیدا کند. نمایه فصلنامه‌ها، مجلات خلاصه تحقیقات و فهرست نقل نامها فرد را قادر می‌سازد تا این اطلاعات را بیابد. این نشریات، که در شماره‌های متوالی و در مقاطع زمانی منظم چاپ می‌شوند، در یافتن اطلاعاتی که به طور وسیع در مجلات و دیگر منابع پراکنده هستند، به عنوان راهنما عمل می‌کنند. ناشران این نمایه‌ها از کارشناسانی که منابع چاپ شده و چاپ نشده را بررسی و دسته‌بندی می‌کنند، استفاده به عمل می‌آورند. این مقالات سپس بر حسب موضوع، و گاهی با حاشیه‌نویسی، گروه‌بندی می‌شوند؛ به گونه‌ای که محققان از طریق فهرست نسبتاً جامعی از آثار موجود در حوزه تخصصی‌شان، که به طور منطقی به روز درآمده‌اند، مطلع می‌شوند.

/ نمایه‌های اریک

مرکز اطلاعاتی منابع آموزشی (ERIC) به منظور گردآوری، نگهداری و انتشار اطلاعات در زمینهٔ تعلیم و تربیت که توسط دفتر تعلیم و تربیت آمریکا (USOE) تأسیس شده است، دو نمایه مهم تهیه می‌کند.

پیش از تأسیس سیستم اریک، گزارشهایی که توسط طرفهای قرار داد و اهداکنندگان مقالات به USOE ارائه می‌شد، به صورت ابتدایی و پراکنده توزیع و همانند گزارشهای منابع دیگر، ناپدید می‌شد. هدف اریک تصحیح این شرایط آشفته، به منظور گردآوری و حفظ اسناد و مدارک غیر چاپی «ناپایدار» و مورد نظر دست اندرکاران تعلیم و تربیت و قرار دادن این بانک اطلاعاتی در اختیار مردم بود.

اکنون سیستم اریک، که بودجهٔ آن را انجمن ملی تعلیم و تربیت (NIE) تأمین می‌کند، اطلاعات را گردآوری، ارزشیابی و فهرست می‌کند تا از طریق شبکه‌ای مشتمل بر ۱۶ دفتر مرکزی واقع در دانشگاههای متعدد و سازمانهای حرفه‌ای در سراسر کشور و با بهره‌جویی از پردازش اطلاعات مرکزی، وارد نمایه‌های اریک شوند. هریک از دفاتر مرکزی، مسئول یک زمینه خاص در تعلیم و تربیت است. یک کتابچهٔ راهنما، با عنوان *راهنمای اطلاعاتی خدمات اریک*^(۱) این دفاتر مرکزی و سایر سازمانهایی را که راه دستیابی به بانک اطلاعاتی اریک و منابع

مربوط به آن را برای مراجعان و مشتریان خود فراهم می‌سازند، فهرست می‌کند.

سه نشریه علمی از مهمترین نشریاتی که توسط اریک تهیه می‌شود، نمایه روز مجلات تعلیم و تربیت^(۱) (CIJE) و منابع تعلیم و تربیت^(۲) (RIE)، و منابع آموزشی کودکان استثنایی^(۳) (ECER)، هستند. برای دستیابی به اسناد و مدارک اریک از یک سیستم فهرستبندی پیشرفته و دقیق استفاده می‌شود. این سیستم، در گنجینه توصیفی اریک^(۴) که در بیشتر کتابخانه‌ها موجود است، تعریف شده است.

/ ماهنامه نمایه روز مجلات تعلیم و تربیت

ماهنامه نمایه روز مجلات تعلیم و تربیت (CIJE) از محصول کار کارشناسان مراکز اطلاعاتی اریک ترکیب یافته است. مقالات موجود در بالغ بر ۷۰۰ مجله طبق سیستم گنجینه اریک دسته‌بندی و فهرست می‌شوند. متجاوز از ۳۵۰۰۰۰ مقاله در CIJE فهرست و هر ساله در حدود ۱۸۰۰۰ مقاله به آنها اضافه می‌شود.

CIJE به چهار بخش تقسیم می‌شود: (۱) فهرست موضوعی، (۲) فهرست مؤلف، (۳) بخش مدخل اصلی، و (۴) فهرست محتوایی مجلات. فرد می‌تواند عنوان و شماره‌های مقالات مورد علاقه خود را با اولین نگاه به فهرست موضوعی بیابد و سپس از این شماره‌ها برای پیدا کردن مدخلها در بخش مدخل اصلی استفاده کند. اگر نتوانید موضوع مورد علاقه خود را در فهرست موضوعی بیابید، در فهرست موضوعات مشابه به جستجو بپردازید. گنجینه لغات اریک برای تدارک یک سیستم فهرست بندی نظام‌مندتر عمده‌اً محدود شده است. فهرست نویسندگان متفرقه برای آن دسته از افراد مفید است که مشتاق یافتن کارهای محققان خاصی هستند. فهرست محتوایی مجلات شامل محتوای کلی مجلات و محتوای هر مقاله است.

شکل ۱-۳ نمونه‌ای از نحوه معرفی یک مقاله در بخش مدخل اصلی، به همراه تشریح قسمتهای مختلف آن به همان نحوی است که به وسیله CIJE برای روشن شدن نوع اطلاعات در فهرست محتوایی مجلات شامل محتوای کلی مجلات و محتوای هر مقاله است.

شکل ۱ - ۳ نمونه‌ای از یک مدخل *CIE*

Accession No.	EJ 123 465	RC 503 097— Clearinghouse No.
Article Title	Native American Techniques of Survival in the Country—Price, John A. <i>Indian Historian</i> ;	
Author	v11 n4 p3-11 Dec 1978 (Reprint: UMI)	
Volume No.	Descriptors: *American Indians; Fire Science	Journal Title
Issue No.	Education; *Foods Instruction; *Medicine;	Availability
Pages	*Outdoor Education; *Plant Identification;	Publication Date
Major and Minor Descriptors (major descriptors are starred)	*Safety; Trees	Descriptive Note
Major and Minor Identifiers (major identifiers are starred)	Identifiers: American Indian Education; *Survival Techniques	Annotator's Initials
	Note: Theme Issue: Survival and the American Indian.	
	Provides a review of basic outdoor survival information: (1) building a shelter; (2) making a fire; (3) finding and keeping food; (4) safety and medicine; (5) orientation to directions; and (6) aids in traveling in the country. (RTS)	

شکل ۱ - ۳ نمونه‌ای از نحوه معرفی یک مقاله در بخش مدخل اصلی، به همراه تشریح قسمتهای مختلف آن به همان نحوی است که به وسیله *CIE* برای روشن شدن نوع اطلاعات در هر دفتر، تدارک دیده شده است.

حاشیه نویسی بویژه بسیار مفید است؛ چون توضیح کوتاهی در مورد هر مقاله ارائه می‌کند. توجه کنید که در نمونه بالا عنوان آن تنها ایده مختصری از محتوای هر مقاله به دست می‌دهد. قبل از ژانویه ۱۹۷۰، زمانی که *CIE* حاشیه نویسی را ضمیمه کار خود کرد، محققان در مراجعه به مقالاتی که عناوین همراه کننده و نارسا داشتند و در بررسی بیشتر ثابت شد که با موضوع بی ارتباط هستند، با ناکامی بزرگی مواجه شدند. اگر عنوان یک مقاله بتواند گرایش اصلی آن را به طور مشخص منتقل کند، خلاصه کنندگان، حاشیه‌ای به آن ضمیمه نمی‌کنند. انتشار *CIE* از سال ۱۹۶۹ آغاز شد. این مجله به طور ششماهه و سالانه منتشر می‌شود.

شیوه درست در استفاده از *CIE* شامل شش مرحله است:

۱- کلمات کلیدی را که ممکن است مقالات مربوط به بررسی شما در زیر آنها فهرست شده

باشند، معین کنید. این کلمات کلیدی نوعاً شامل جامعه مورد بررسی و متغیرهایی است که شما باید در تبیین مسئله تحقیقی خود آنها را تعریف کنید.

۲- برای دریافتن آنکه کدامیک از کلمات کلیدی شما به عنوان توصیفگر مورد استفاده قرار خواهند گرفت، گنجینه توصیفی اریک را بازبینی کنید. شما احتمالاً برای کلمات کلیدی که فهرست کرده‌اید، نیاز به یافتن معادلها دارید.

۳- بررسی خود را از آخرین شماره‌های *CIE* شروع کنید و سپس به کارهای گذشته در نشریات پیشین بپردازید.

۴- کلیه منابع داده شده برای هر عنوان را که احتمالاً می‌تواند مفید باشد، از *CIE* رونوشت بردارید. این روش کار، یافتن مقالات اصلی را آسان می‌کند.

۵- مقالات را در مجلات مربوط به آنها جستجو کنید. اگر کتابخانه شما فاقد آن مجله بود، ممکن است بتوانید نسخه زیراکس شده آن را از کتابخانه دیگری به عاریت بگیرید.

۶- اگر یک چکیده به همراه خود مقاله ضمیمه شده بود، اول آن را بخوانید تا دریابید که آیا خواندن کامل مقاله ضروری است یا خیر. در صورت فقدان چکیده از بخش خلاصه و نتیجه‌گیری مقاله آغاز کنید.

/ منابع آموزشی کودکان استثنایی

این فصلنامه (سابقاً کتابچه تحقیقهای مربوط به تعلیم و تربیت کودکان استثنایی) شامل چکیده تحقیقهای ذخیره شده در بایگانی کامپیوتری مرکز اطلاعات شورای کودکان استثنایی به عنوان بخشی از برنامه‌آریک است. چکیده‌های مربوط به تعلیم و تربیت معلولین و کودکان و جوانان سرآمد نیز فهرست شده‌اند. ترتیب آنها شبیه *CIE* و *RIE* است. تا سال ۱۹۷۵، هر مجلد شامل فهرست تراکمی سالهای قبل بوده و از ۱۹۷۶، فهرستهای تراکمی مؤلف، عنوان و موضوع در نسخه نهایی سال گنجانده شده است. کتابها با مقالات موجود و مجلات، اسناد دولتی و غیره از موارد مشمول *ECER* هستند.

/ منابع تعلیم و تربیت

چکیده گزارشهای تحقیقی از منابعی به غیر از مجله‌ها نیز فهرست و به صورت ماهنامه به

وسیله اریک در نشریه منابع تعلیم و تربیت (RIE) منتشر می‌شود. در این زمینه یک مجموعه شش ماهه از بخش فهرستها نیز به صورت تراکمی منتشر می‌شود. هر مرکز اطلاعاتی اسناد و مدارک مربوط به حیطه تخصصی خود را جمع‌آوری و سپس آنها را وارد فهرست می‌کند و فهرستها و چکیده‌ها را تهیه می‌کند. اسناد و مدارک مادر (اصلی)، به همراه چکیده‌های تهیه شده توسط مراکز اطلاعاتی، بر طبق یک روال منظم به بخش پردازش اطلاعات، جایی که به یک بخش از بانک اطلاعات مرکزی تبدیل می‌شوند، ارسال می‌شوند، به طریقی که مراجعه‌کنندگان سیستم اریک قادر به استفاده از آن باشند. بخش پردازش مسئول جمع‌آوری داده‌های همه مراکز اطلاعاتی و چاپ مجموعه به صورت ماهانه در RIE است. منابعی که مراکز اطلاعاتی تخصصی از آنها فهرست تهیه می‌کنند، شامل گزارشهای تحقیقات آموزشی با هزینه دولت فدرال، چکیده‌ها، جزوه‌ها، کتابچه‌های راهنما، مقالات برگزیده از پیشرفتهای انجمنها و سازمانهای علمی، کتابشناسی‌ها، نمونه برگزیده متون درسی دانشگاهی، راهنمای معلمان و طرح کلی برنامه درسی، به همراه تعداد بسیار زیادی پیشنهاد تحقیق و گزارش پروژه‌ها است. با چنین پوشش گسترده‌ای، روشن است که اسناد و مدارک مجموعه اریک مورد توجه شدید دست‌اندرکاران و همچنین محققان در زمینه تعلیم و تربیت باشد. در حال حاضر متجاوز از ۳۵۰۰۰ هزار سند و مدرک در این مرکز وجود دارد، و هر سال ۱۸۰۰۰ سند و مدرک به این مجموعه در حال گسترش اضافه می‌شود.

به منظور تعیین و سفارش نسخه‌ای از اسناد و مدارک مادر، هر چکیده با شماره دسترسی خود مشخص می‌شود. در حال حاضر سه فهرست در این زمینه منتشر می‌شود: فهرست نویسنده، مؤسسه و موضوع. محقق با استفاده از فهرست مناسب، می‌تواند عناوین اسناد و مدارک و شماره‌های دسترسی به آنها را پیدا و سپس از آنها برای یافتن چکیده‌های خاص استفاده کند. روال استفاده از RIE شبیه به آن چیزی است که در CIJE به کار می‌رود. محقق نوعاً بررسی خود را با جدیدترین نسخه شروع می‌کند و برای رجوع به کارهای گذشته به جستجوی موضوع یا موضوعات مورد علاقه می‌پردازد.

/ مجموعه اسناد اریک

نسخه میکروفیش اکثر اسناد ما در (اصلی) به محض درخواست در اختیار متقاضی قرار

شکل ۲-۳ نمونه‌ای از یک مدخل RIE

ERIC Accession Number— identification number sequentially assigned to documents as they are processed	ED 654 321	CE 123 456	Clearinghouse Accession Number
Author(s)	Smith, John D.	Johnson, Jane	Sponsoring Agency— agency responsible for initiating, funding, and managing the research project
Title	Career Planning for Women	Central Univ., Chicago, IL	Report Number—assigned by originator
Organization where document originated	Spons Agency—National Inst. of Education (ED), Washington, DC	Report No.—CU-2081-5	Descriptive Note (pagination first)
Date Published	Pub Date—May 83	Contract—NIE-C-83-0001	
Contract or Grant Number	Note—129p. Paper presented at the National Conference on Career Education (3rd, Chicago, IL, May 15-17, 1983).	Available from—Campus Bookstore, 123 College Ave., Chicago, IL 60690 (\$3.25).	
Alternate source for obtaining document	Language—English. French	Pub Type—Speeches/Meeting Papers (150)	
Language of Document—documents written entirely in English are not designated, although "English" is carried in their computerized records	EDRS Price—MF01/PC06 plus postage.	Descriptors—subject terms found in the <i>Thesaurus of ERIC Descriptors</i> that characterize substantive content. Only the major terms, preceded by an asterisk, are printed in the subject index	
Publication Type—broad categories indicating the form or organization of the document, as contrasted to its subject matter. The category name is followed by the category code.	Identifiers—Consortium of States, "National Occupational Competency Testing Institute	Identifiers—additional identifying terms not found in the <i>Thesaurus</i> . Only the major terms, preceded by an asterisk, are printed in the subject index	
ERIC Document Reproduction Service (EDRS) Availability—"MF" means microfiche; "PC" means reproduced paper copy. When described as "Document Not Available from EDRS," alternate sources are cited above. Prices are subject to change; for latest price code schedule, see section "How to Order ERIC Documents" in the most recent issue of RIE	Women's opportunities for employment will be directly related to their level of skill and experience and also to the labor market demands through the remainder of the decade. The number of workers needed for all major occupational categories is expected to increase by about one-fifth between 1980 and 1990, but the growth rate will vary by occupational group. Professional and technical workers are expected to have the highest predicted rate (39 percent), followed by service workers (35 percent), clerical workers (26 percent), sales workers (24 percent), craft workers and supervisors (20 percent), managers and administrators (15 percent), and operatives (11 percent). This publication contains a brief discussion and employment information concerning occupations for professional and technical workers, managers and administrators, skilled trades, sales workers, clerical workers, and service workers. (SB)	Informative Abstract	
		Abstractor's Initials	

می‌گیرد، یا می‌توان یک قرارداد دائمی با ERIC منعقد کرد تا تمام اسناد به مجرد آماده شدن در اختیار طرف قرارداد قرار گیرد.

بسیاری از کتابخانه‌ها چنین قرار دادی با اریک دارند، بنابراین کلیه مجموعه اسناد اریک با شرایطی سهل در اختیار مراجعان آنها قرار می‌گیرند.

بیشتر اسناد به دو شکل در دسترس متقاضی قرار می‌گیرند: نسخه چاپی و میکروفیش. در بسیاری از کتابخانه‌ها سیستم میکروفیش به دلیل ارزان‌تر بودن و اشغال جای کمتر ترجیح داده می‌شوند. مجموعه کامل میکروفیش اریک، حجمی برابر یک قفسه بایگانی را اشغال می‌کند. دستگاهی که میکروفیش خوان نامیده می‌شود، برای بزرگ کردن صفحه میکروفیش بر روی پرده ثابت است و همچنین برای متقاضی امکان حرکت از تصویر یک صفحه به صفحه‌ای دیگر را فراهم می‌سازد. بیشتر کتابخانه‌ها و دانشگاه‌ها از چنین تسهیلاتی برخوردارند. قیمت هر یک از اسناد اریک در هر یک از فرمهای چاپی و میکروفیش در *RIE* فهرست می‌شود.

تمام پروژه‌هایی که به هزینه USOE انجام می‌شوند، به طور خودکار به سیستم اریک منضم می‌شوند و نسخه‌های گزارشهای کامل شده آنها از طریق اریک در دسترس مراجعان قرار می‌گیرند. اگر متن منتشر شده برای بار اول و با امتیاز حق مؤلف چاپ شده باشد، به عنوان مرجع از آن نام برده می‌شود؛ هر چند نسخه‌های آن از طریق اریک در دسترس همگان قرار نمی‌گیرد. روشهای مشابهی در مورد کتابها و سایر مواد چاپی که برای فروش آماده می‌شوند، نیز به کار می‌رود. اریک در عرصه تعلیم و تربیت ارزشمند است؛ چون مواد مورد نظر برای محققان در زمینه تعلیم و تربیت گستره وسیعی از منابع را گردآوری و خلاصه می‌کند. دستیابی به بسیاری از این منابع در گذشته تقریباً غیرممکن بوده است.

مراکز اطلاعاتی اریک در زمینه موضوعات منتخب، کتابشناسی‌هایی با حواشی فراهم می‌آورد. قبل از آنکه برای بررسی خود کتابشناسی تهیه کنید، بهتر است از طریق مکاتبه با مرکز اطلاعاتی اریک که مربوط به زمینه تحقیق شماست، تماس برقرار کنید تا چنانچه کارکنان آنجا کتابشناسی مفیدی در دست دارند، آن را در اختیار شما بگذارند.

توضیح کاملتر در مورد سیستم اریک و چگونگی استفاده از آن در کتابچه اریک تحت عنوان «این مرکز چه می‌تواند برای شما انجام دهد، چگونه از آن استفاده کنیم»^(۱)، موجود است که در بسیاری از کتابخانه‌ها در دسترس متقاضیان قرار می‌گیرد.

مراکز اطلاعاتی شبکه اریک امیدوار هستند که این مجموعه را تا حد امکان جامع تر کنند و از متخصصان تعلیم و تربیت درخواست می کنند تا مواد مربوط را به این مراکز ارائه کنند. اگر گزارشی از یک سخنرانی و یا مقاله ای تهیه کرده اید و علاقه دارید از طریق اریک در کشور منتشر شود، نسخه ای از آن را به بخش پردازش و منابع اریک به آدرس 4833 Rugby Avenue, Bethesda, MD 20814 یا مرکز اطلاعاتی مسئول برای گردآوری اسناد درحوزه موضوع مربوطه، ارسال دارید.

دستیابی به سیستم اریک از طریق کامپیوتر

یک نسل قبل، جستجوی کامل ۵۰۰ مجله و ۱۰۰۰۰۰ سند و مدرک دیگر برای یافتن مطالب مربوط به مسئله مورد نظر، یک کار عظیم بود. امروزه دسترسی کامپیوتری به سیستم اریک امکان انجام چنین جستجویی را با استفاده از کامپیوتر شخصی یا ارتباط تلفنی، فراهم آورده است.

محتوای *CIE* و *RIE* بر روی نوارهای کامپیوتری، موجود و در دسترس است و با استفاده از این نوارها می توان به سیستم اریک دست یافت. برای مثال، برای شناسایی مقالات و اسناد *RIE* که برای تدریس زبان فرانسه با استفاده از مطالب برنامه ریزی شده به کار می روند، کامپیوتر به نحوی برنامه ریزی شده است که با کاربرد توصیفگر آموزش، و نیز توصیفگر زبان فرانسه، تمام اسناد و مقالات مربوط به این موضوع را شناسایی می کند. برخی از برنامه های کامپیوتری تنها شماره های دسترسی به اسناد اریک *ED* و مجلات اریک *EJ* را چاپ کامپیوتری می کنند، اما بیشتر برنامه ها مدخل کامل *CIE* و مدخل کامل *RIE* هر مقاله و سند را که دارای هر دو توصیفگر هستند، چاپ می کنند.

متجاوز از ۹۰۰ مؤسسه نوارهای اریک را در اختیار دارند و می توانند به جستجوهای کامپیوتری بپردازند. بیشتر ادارات آموزش و پرورش ایالتی این نوارها را در اختیار دارند و برای دست اندرکاران تعلیم و تربیت در ایالت خود به جستجوهای کامپیوتری می پردازند. بیشتر این ادارات همچنین از متخصصان بازیابی اطلاعات استفاده می کنند که اطلاعات بازیابی کامپیوتری مورد نیاز مشتریان را از طریق برگه تقاضا برای گرفتن کمک از سیستم اریک، در اختیار آنها قرار می دهند. بسیاری از دانشگاهها خدمات مشابهی برای دانشجویان خود و دیگر متقاضیان فراهم ساخته اند.

برخی از کتابخانه‌ها نیز سیستم اریک و دیگر بانکهای اطلاعاتی را بر روی CD-ROM^(۱) یعنی دیسک فشرده با حافظه غیر قابل دستکاری منتقل کرده‌اند و در اختیار دارند، که یک روش جدید برای دسترسی به اریک و سایر بانکهای اطلاعاتی است. مؤسسه سیلور پلاتر^(۲) از سال ۱۹۶۶ تاکنون بانک اطلاعات اریک را به طور کامل بر روی سه CD-ROM منتشر کرده است. هر ثبت اطلاعات شامل شرح کتابشناختی، چکیده‌ها، و توصیفگران می‌شود. بر خلاف فهرستهای چاپی اریک، این دیسکهای CD-ROM را می‌توان در یک کیف دستی کوچک حمل کرد. برای استفاده از CD-ROM به یک کامپیوتر شخصی با حافظه ۵۱۲ کیلوبایت و یک دیسک‌گردان نیاز است. به غیر از بانک اطلاعات کامپیوتری اریک، حدود ۵۰۰ بانک اطلاعاتی مربوط به بسیاری از موضوعات، شامل چکیده‌های روان‌شناختی، چکیده‌های رساله‌های بین‌المللی، چکیده‌های جامعه‌شناختی، و فهرست نقل نامهای علوم اجتماعی نیز وجود دارد. بانکهای اطلاعاتی جدید تقریباً هر روز از طریق نرم افزار Vendors جهت جستجوهای کامپیوتری در حال تاسیس هستند. نمایه‌تعلیم و تربیت^(۳)، نشریه یکی از بانکهای اطلاعاتی بیشمار است که به سیستم کامپیوتری ویلسون لاین وابسته است. این خدمات نیز در بسیاری از کتابخانه‌ها قابل دسترس است.

جستجو در سیستم اریک یک مرحله مهم در بررسی پیشینه تحقیقاتی است، ولی نمی‌توان فرض کرد که وقتی این مرحله خاتمه یافت، جستجوی پیشینه تحقیقاتی نیز به پایان رسیده است. مطالب مرتبط با سؤال محقق ممکن است وارد سیستم اریک نشده باشند.

/ نمایه سایر نشریات علمی

بسیاری از نشریات دیگر نیز برای دست یافتن به اطلاعات روز در زمینه تحقیقات نظریه‌ها و عقاید در تعلیم و تربیت سودمند هستند. یکی از مراجع استاندارد برای کارهای تحقیقاتی در زمینه تعلیم و تربیت، نمایه تعلیم و تربیت است که به طور منظم از سال ۱۹۲۹ توسط مؤسسه اچ.و. ویلسون در نیویورک منتشر می‌شود. این نمایه مقالات حدود ۳۵۰ نشریه علمی، سالنامه بولتن، صورتجلسه و تک‌نگاری را فهرست می‌کند. نمایه تعلیم و تربیت بهترین منبع برای

1- Compact Disc - Read - Only Memory.

2- Silver Platter Information, Inc., 37 Walnut st., Wellesley Hills, MA 02181.

3- Education Index (1929-). New York: H. W. Wilson.

دستیابی به مقالات مندرج در مجلات، قبل از تأسیس نمایه روز مجلات تعلیم و تربیت در سال ۱۹۶۹، و نیز بهترین منبع برای دستیابی به مقالات جدید است؛ زیرا این نشریه مقاله‌ها را شش ماه قبل از *CIE* فهرست می‌کند.

علاوه بر این نمایه‌های عمومی تعدادی نمایه تخصصی مفید نیز در دسترس محققان قرار می‌گیرد. چکیده‌های روان‌شناختی،^(۱) متون روان‌شناسی و مطالب مربوط به آن را در سراسر جهان فهرست می‌کند. این فهرست حاوی کتابها، رساله‌های دکترا و مقالات علمی به همراه خلاصه‌ای از هر یک است که خواننده را قادر می‌سازد تا ارتباط مطالب را با یکدیگر دریابد. چکیده‌ها و کتابشناسی رشد کودک^(۲) مطالب موجود در حیطه‌های نوباوگی، طب بالینی، بهداشت عمومی، رشد و روان‌شناسی تطبیقی، روان‌شناسی آزمایشی (شامل پدیده‌های یادگیری، شخصیت و روان‌شناسی تعلیم و تربیت و مشاوره) را برحسب نام مؤلف و موضوع فهرست می‌کند. چکیده‌های جامعه‌شناسی تعلیم و تربیت^(۳)، که یک نشریه بین‌المللی است، مقالات مجلات علمی، کتابها، و رساله‌ها را در حوزه خود، بر اساس نام مؤلف و موضوع، به همراه خلاصه‌ای از هر کدام فهرست می‌کند. چکیده‌های مدیریت آموزش^(۴) با فهرست‌بندی مقالات نشریه‌های تخصصی در زمینه مدیریت آموزشی، راه دستیابی به مطالب آنها را فراهم می‌آورد. چکیده‌های آموزش عالی^(۵) (سابقاً چکیده‌های دانشجویان - کارکنان کالج) ترکیبی از چکیده‌های مجلات، خلاصه مذكرات کنفرانسها و گزارشهای تحقیقی مربوط به دانشجویان و خدمات دانشجویی است. موضوعات شامل مشاوره، مسکن، کمکهای مالی، و آزمون و اندازه‌گیری است. نمایه علوم اجتماعی^(۶) مقالات مجلات در زمینه علوم اجتماعی شامل مجلات آموزشی و نیز مجلاتی در زمینه روان‌شناسی، جامعه‌شناسی، علوم سیاسی، مردم‌شناسی، حقوق و اقتصاد را برحسب نام مؤلف و موضوع فهرست می‌کند.

هر رشته، مانند آموزش بازرگانی، مهارتهای صنعتی، پزشکی، و غیره، نمایه تخصصی خود را دارد. مراجعه به راهنماهای اصلی متون تحقیقاتی پیشین نام نمایه‌های تخصصی در زمینه‌های دیگر را نیز به خوانندگان ارائه می‌کند.

1- Psychological Abstracts (1927-). Lancaster, PA: American Psychological Association.

2- Child Development Abstracts and Bibliography (1927-). Lafayette, IN: Purdue University, Society For Research in Child Development.

3- Sociology of Education Abstracts (1965-). Abbingdon, England: CarFax Publishing Co.

4- Educational Administration Abstracts (1966-). Beverly Hills; Sage Publications.

5- Higher Education Abstracts.

6- Social Sciences Index (1974-). New York: H. W. Wilson.

/ فهرست رساله‌ها و پایان نامه‌ها

رساله‌های فوق لیسانس و دورهٔ دکترا از منابع اطلاعاتی مفید برای محققان است. علاوه بر نمایه‌های مجلات علمی و مجلات خلاصه تحقیق، نظیر RIE اریک و چکیده تحقیقات روان شناختی، که برخی از رساله‌ها و پایان نامه‌ها را فهرست می‌کنند، راهنماها و نمایه‌های تخصصی متعدد نیز بسیار سودمند هستند.

نمایه جامع رساله‌ها^(۱)

این نمایه تلاش دارد که کلیهٔ رساله‌های پذیرفته شده در دانشگاه‌های آمریکا را طی سالهای ۷۲-۱۹۶۱ تحت پوشش قرار دهد. این نمایه شامل ۳۷ جلد در ۱۷ طبقه بندی اصلی موضوعی به همراه شرح کامل کتابشناختی برحسب مؤلف و موضوع است. یک ضمیمه سالانه برای روزآمد کردن این نمایه چاپ می‌شود. چکیده رساله‌های بین‌المللی (DAI)^(۲) که هر ماه منتشر می‌شود، شامل چکیده رساله‌های دورهٔ دکتراست که با همکاری دانشگاه‌ها به مرکز میکرو فیلم‌های بین‌المللی دانشگاهی تحویل می‌شود (حدود ۳۷۰ عدد در سال ۱۹۸۴). DAI به سه بخش تقسیم می‌شود:

علوم انسانی و علوم اجتماعی، علوم پایه و مهندسی و دانشگاه‌های اروپا. جستجوهای کامپیوتری در DAI از طریق بسیاری از بانکهای اطلاعاتی اصلی وندورز از جمله، دیالوگ^(۳)، BRS^(۴) و کامپوسرو^(۵) در دسترس است. بیشتر کتابخانه‌ها برای همکاری با محققان، جستجوهای کامپیوتری در بانکهای اطلاعاتی را تدارک می‌بینند.

یک نشریه دیگر مشابه با DAI، رساله‌های دورهٔ دکترای آمریکاست^(۶). این نشریه رساله‌های پذیرفته شدهٔ دورهٔ دکترا در دانشگاه‌های ایالات متحده و کانادا طی هر سال تحصیلی را به همراه رساله‌هایی که به شکل میکروفیلم در میکروفیلمهای دانشگاهی موجود است، در یک فهرست مشترک گردآوری می‌کند. این نشریه حاوی تعدادی از رساله‌هاست که در DAI وجود

1- Comprehensive Dissertation Index (1861- 1972). (37vols.) Ann Arbor: University MicroFilms.

2- Dissertation Abstracts International (1938-). Ann Arbor: University MicroFilms.

3- Dialog Information Service, Inc., 3460 Hillview Avenue, Palo Alto, CA 94304.

4- BRS Information Technologies, 1200 Route 7, Latham, NY 12110.

5- Compu Serve Information Service, Inc, 5000 Arlington Center Boulevard, Columbus, OH43220.

6- American Doctoral Dissertations (1957-) Ann Arbor: University MicroFilms.

ندارد. رساله‌های دوره دکترای آمریکا برحسب طبقه‌بندی موضوعی و مؤسسه مرتب شده و شامل یک فهرست است، ولی هیچ چکیده‌ای تهیه نمی‌کند.

راهنمای فهرست پایان نامه‌های دوره فوق لیسانس مؤسسه پلک^(۱) به دو بخش اصلی تقسیم می‌شود: فهرست پایان نامه‌ها در زمینه‌های خاص، با حاشیه‌نویسی‌ها، و فهرست پایان نامه‌های مؤسسات خاص. چکیده پایان نامه‌های فوق لیسانس: چکیده پایان نامه‌های فوق لیسانس منتخب بر روی میکروفیلم^(۲) شامل چکیده یک فهرست منتخب از پایان نامه‌های دوره فوق لیسانس از دانشگاهها و دانشکده‌های ایالات متحده آمریکا است که به صورت میکروفیلم در دسترس است.

پایان نامه‌های فوق لیسانس در تعلیم و تربیت مؤسسه سیلوی^(۳)، یک فهرست سالانه از پایان نامه‌ها در زمینه تعلیم و تربیت از دانشکده‌ها و دانشگاههای ایالات متحده آمریکاست. برای هر پایان نامه فهرست شده، اطلاعات کتابشناسی داده می‌شود، اما چکیده‌ای ضمیمه آنها نیست.

/ فهرست نقل نامهای علوم اجتماعی (SSCI)

مقاله‌ای که با مراجعه به یکی از این نمایه‌ها یافته شده است، خود، مراجعی در زمینه کارهای قبلی در آن حیطه، ارائه می‌کند. زمانی که فرد مقالات مهم را در زمینه‌ای خاص می‌یابد، یک نمایه نقل نامها می‌تواند در سرعت بخشیدن بر کار او مؤثر باشد. برای مثال اگر شما یک مقاله فوق العاده مفید را که از سال ۱۹۷۶ منتشر شده است، خوانده باشید، از طریق نمایه‌های بعدی می‌توانید مقالاتی را بیابید که نام مقاله مورد نظر شما در کتابشناسی آن وجود دارد. بسیاری از این مقاله‌ها، تداوم، گسترش یا تغییراتی را که در بررسی مورد نظر اولیه رخ داده است، توصیف خواهند کرد.

فهرست نقل نامهای علوم اجتماعی (SSCI)^(۴)، که هر ساله در سه جلد توسط مؤسسه اطلاعات علمی منتشر می‌شود، می‌تواند به شما بگوید که طی سال از کدام نویسندگان در همه

1- Black, D. (ed.) (1965). Guide to Lists of Master's Theses. Chicago: American Library Association.

2- Masters Abstracts: Abstracts of Selected Masters theses on Microfilm 1962- 1985; Masters Abstracts International (1986-). Ann Arbor: University MicroFilms.

3- Slvey, H. M. (1951-). Master's Theses in Education. Cedar Falls, IA: Research Publications.

4- Social Sciences Citation Index (1973-) Philadelphia: Institute For Scientific Information.

زمینه‌های علوم اجتماعی، از جمله تعلیم و تربیت، چه چیزی ذکر شده و در زمینه‌های مختلف چه نوشته شده است، همچنین کتابشناسی اطلاعات مورد لزوم درباره مؤلفانی که از آنها از آنها ذکر شده یا از کسی ذکر کرده‌اند نیز ضمیمه است. این اطلاعات از طریق چهار نمایه قابل دستیابی است:

۱- نمایه مرجع یک فهرست الفبایی از نام همه مؤلفانی را که طی سال کتابی از آنها منتشر شده است، ارائه می‌کند، که حاوی اطلاعات کتابشناسی کاملی از جمله فهرست الفبایی منابع و مأخذی است که در کتابشناسی مولف مرجع آمده است.

۲- فهرست نقل نامها، نام نویسندگان نقل شده در نمایه مرجع را اخذ می‌کند و آنها را به ترتیب الفبایی به شیوه‌ای ارائه می‌کند که محقق دریابد یک مقاله خاص در کجا ذکر شده است. مراجع کتابشناختی برای هر مورد نقل شده ارائه می‌شود. یک نمایه از ناشر مؤلف در زمینه نقل نامها و یک نمایه از مقاله‌های بی‌نام مؤلف نیز تهیه می‌شود. این نمایه فرد را قادر می‌سازد تا ایده‌هایش را با سرعت دنبال کند.

۳- فهرست موضوعی اصطلاحات همه واژه‌های مهم را با واژه‌های مهم دیگر زیر هر عنوان زوج می‌کند. اریک عنوان هر واژه را در حکم واژه اولیه فهرست می‌کند و آن را با سایر واژه‌های معادل تلفیق می‌کند. یک فهرست الفبایی از نامهای مولفان، که عناوین کتابهای آنان حاوی آن واژه‌ها است، برای هر زوج واژه و معادل آن ارائه می‌کند. بنابراین اطلاعات کتابشناسی را می‌توان برای هر مولف در نمایه مرجع یافت.

۴- نمایه نشانی ناشر یک فهرست الفبایی از سازمانهای انتشاراتی است که مؤلفان طی آن سال به آن وابسته‌اند. زیر هر مدخل، شرکت نشر، فهرست مؤلفان با اطلاعات کتابشناسی کامل قرار دارد.

// انتشارات دولتی

دولت فدرال یک منبع اصلی اطلاعات تعلیم و تربیتی است که بیشتر از هر یک از سازمانهای ایالات متحده، متعهد به انجام تحقیقات، پیمایشها، و گردآوری آمار است. اداره آموزش و پرورش به موازات مرکز ملی آمار در زمینه آموزش و پرورش، تعداد بیشمار نشریه، شامل گزارشهای تحقیقی، پیمایشها، فعالیتهای اجرایی و برنامه‌های توصیفی به چاپ می‌رساند.

برای یافتن نشریات ویژه دولت آمریکا، کاتالوگ ماهانه نشریات دولت آمریکا^(۱) فهرستی اصلی برای مراجعه است. این نشریه شامل یک بخش اصلی است که مؤسسه انتشاراتی، اسناد منتشره و عنوان نویسنده، موضوع، و فهرست کلمات کلیدی عناوین را فهرست می‌کند. تا سال ۱۹۷۵ یک مجموعه سالانه برحسب عنوان و موضوع در نسخه چاپ دسامبر گنجانده شده بود، از سال ۱۹۷۶ این مجموعه‌ها به صورت شش ماهه منتشر می‌شود. همچنین یک نمایه تراکمی ۵ سالانه برای جستجوی سریعتر در تحقیقات سالهای گذشته وجود دارد.

نشریات وابسته به ادارات ایالتی آموزش و پرورش و دیگر مؤسسات ایالتی را می‌توان از طریق نمایه ماهانه نشریات ایالتی^(۲) پیدا کرد.

// منابع آزمونها

در انجام هر تحقیقی یک آزمون یا وسایل اندازه‌گیری مورد نیاز است. سالنامه اندازه‌گیری روانی بوروس^(۳)، عمده‌ترین مأخذی است که آزمونها را فهرست می‌کند و به طور انتقادی به بررسی آنها می‌پردازد. این کتابها بخصوص برای کمک به متقاضیان در عرصه تعلیم و تربیت، روان‌شناسی و ترغیب کردن آنها به استفاده هر چه هوشمندانه‌تر از آزمونهای استاندارد شده، طراحی شده است. هر سالنامه، با این هدف با آرایش یکسان طراحی شده است که مجلات قبلی را تکمیل کند، تا اینکه جایگزین آنها شود. آزمونها بر اساس موضوع گروه‌بندی می‌شوند و توضیح در مورد هر آزمون با جستجوها و بررسیهای مهم و منابع بررسیهایی که آزمون در آن به کار رفته است، دنبال می‌شود. هر مجلد مروری بر بررسیها، قطعات منتخب و مراجع کتابشناسی مجلدات قبلی دارد. مجلدات شامل آزمونهای استعداد و پیشرفت تحصیلی در حیطه‌های موضوعی متعدد، آزمونهای شخصیت و حرفه‌ای و آزمونهای هوش و اطلاعات کامل در مورد قیمت و نحوه استفاده از هر کدام از آزمونهای هوش است. اطلاعات کامل در مورد قیمت و نحوه استفاده از هر کدام از آزمونها نیز موجود است. آزمونهای چاپ شده شماره^(۴) به مثابه نمایه و

1- U.S. Superintendent of Documents (1895-). Monthly cata of U.S. Goverment publications. Washington: U.S.Government Printing Office.

2- U.S. Library of Congress, Exchange and Gift Division (1910-). Monthly Checklist of state publication. Washington: U.S. Goverment Printing Office.

3- Buros, O.K. (ed.) (1938-) Mental Measurment Yearbooks. Lincoln: Buros Institute of Mental Measurments, University of Nebraska.

4- Mitchell, J.V., Jr.(ed.) (1983) Tests in Print III: An Index to Tests, Test Rewiews, and the Literature on Specific Tests. Lincoln: Buros.

ضمیمه‌ای برای مجلدات گذشته سالنامه اندازه‌گیریهای روانی به کار می‌رود. مؤسسه بوروس همچنین مطالب موجود در سالنامه‌های اندازه‌گیری روانی را برای تکن‌نگاریهای ویژه در آزمونه‌های شخصیت، خواندن، هوش، مهارت شغلی و حرفه‌ای، زبانهای انگلیسی و خارجی، ریاضیات، علوم و بررسیهای اجتماعی سازماندهی می‌کند.

همچنین موسسه بوروس یک مجموعه از داده‌های کامپیوتری را از طریق مرکز تکنولوژیهای اطلاعاتی BRS ارائه می‌کند که اطلاعات آن به طور ماهانه به روز می‌شود.

نمای تمام و کمال هر سند به متقاضی امکان می‌دهد که متن کامل هر سند را پیدا کند و نسخه‌ای از آن را دریافت و یا آن را به طور کامل بررسی کند و بخشهای منتخب آن را از چاپگر کامپیوتر به دست آورد.

این سیستم در بین سایر سالنامه‌ها یک سیستم بازیابی اطلاعات عالی برای به روز کردن اطلاعات است.

یک مرجع استاندارد برای آزمونها، مرجع جامع ارزشیابی در روان‌شناسی تعلیم و تربیت و تجارت^(۱) است که حاوی متجاوز از ۳۰۰۰ آزمون موجود به زبان انگلیسی است. هر آزمون یک دسته‌بندی مقدماتی به دست می‌دهد و در یکی از بخشها به طور مشروح تعریف و احتمالاً یک منبع رجوع در یک دسته‌بندی ثانوی محسوب می‌شود.

از سال ۱۹۸۴، نقد آزمونها^(۲) تعریفها و نقدهای کلی از وسایل اندازه‌گیری را فراهم ساخته است. یک فهرست مفید از مجموعه موضوعات آزمونها، بر اساس نوع متغیرهای آنها نیز ضمیمه است. راهنماهای ویژه آزمونها شامل اندازه‌گیری نگرشهای روان‌شناسی اجتماعی^(۳)، آزمونه‌های خواندن و مرور تحقیقات پیشین^(۴)، راهنمایی بر ۶۵ آزمون تشخیص در تعلیم و تربیت استثنایی^(۵)، راهنمای بین‌المللی برای آزمونه‌های فضایی^(۶)، سنجش فراگیران ناتوان:

1- Sweetland, R.C., and D.J. Keyser (eds.) (1983). Tests: A Comprehensive Reference for Assessments in Psychology, Education and Business. Kansas City, Mo: Test Corporation of American.

2- Keyser, D.J., and R.C. Sweetland (eds.) (1984-). Test Critiques vol. 1-). Kansas City, Mo: Test Corporation of America.

3- Robinson, J.P., and P.R. Shaver (1973). Measures of Social Psychological Attitudes. Ann Arbor: Survey Research Center, Institute for Social Research.

4- Buros, O.K. (ed.) (1968- 1975). Reading Tests and Reviews (2vols.). Lincoln: Buros Institute of Mental Measurement, University of Nebraska.

5- Compton, C. (1980). A Guide to 65 Diagnostic Tests for Special Education. Belmont, CA: Fearon Education.

6- Eliot, J., and I. MacFarlane (1983). An International Directory of Spatial Tests. New York: Humanities Press.

وسایل برگزیده^(۱)، آزمونهای به صورت میکروفیش در آمده: فهرستهای حاشیه‌نویسی^(۲) و کاتالوگ مجموعه آزمونهای مرکز خدمات آموزشی^(۳) است.

اطلاعات در مورد آزمونها را می‌توان در نمایه‌های نشریات علمی، از جمله نمایه روز مجلات تعلیم و تربیت، نمایه تعلیم و تربیت، منابع تعلیم و تربیت کودکان استثنایی، و چکیده‌های روان‌شناختی، که تحت نام خاص خود در فهرست موضوعی ثبت شده‌اند، یافت. شما همچنین می‌توانید زیر عناوین موضوعی مانند «آزمونها» و «مروری بر آزمونها» به جستجو بپردازید.

// جستجوی کامپیوتری

اکثر کتابخانه‌های دانشگاهی تسهیلات جستجوی کامپیوتری را تحت عناوین جستجوی کامپیوتری، جستجوی پایگاه اطلاعاتی، سرویس رجوع کامپیوتری، و از این قبیل تدارک دیده‌اند. در بیشتر کتابخانه‌ها این خدمات از طریق پست و در برخی اوقات و به طور محدود با مراجعه بدون وقت قبلی صورت می‌گیرد. بعضی از کتابخانه‌ها ممکن است به دلیل تقاضا برای جستجوی کامپیوتری اطلاعات ویژه (تخصصی) از فرد بخواهند که در آنجا حضور یابد. این امر نه تنها به فرد فرصت می‌دهد تا نمونه‌ای از نتایج اطلاعات بازیابی شده را ملاحظه کند، بلکه این بخت را به او می‌دهد که حافظه کامپیوتر را بیشتر بکاود و راهبردهای تحقیقی خود را بهبود بخشد.

بسیاری از منابع نقل قول شده در این فصل از جمله نمایه‌های تعلیم و تربیت، نمایه‌های اریک، منابع تعلیم و تربیت کودکان استثنایی، چکیده‌های روان‌شناختی، نمایه نقل نامهای علوم اجتماعی، چکیده‌های رساله‌های بین‌المللی، و سالنامه‌های اندازه‌گیری‌های روانی، در پایگاههای اطلاعاتی کامپیوتری گنجانده شده‌اند. بعلاوه، مجموعه داده‌های دیگری مانند سندرم نارسایی اکتسابی در سیستم ایمنی بدن (ایدز) تحقیقات علوم پزشکی، منابع در تعلیم و تربیت حرفه‌ای (RIVE)، مجموع آزمون مرکز خدمات آموزشی (ETSF) و چکیده‌های جامعه‌شناختی (SOCA)، نیز وجود دارند. برای تهیه فهرستهای مشروح از بیشتر پایگاههای

1- Mauser, A.J. (1977). Assessing the Learning Disabled: Selected Instruments (2d ed.). Navato, CA: Academic Therapy Publications.

2- Tests in Microfiche: Annotated Index (1975-). Princeton: Educational Testing Service.

3- The Educational Testing Service Test Collection Catalog (1986-). Phoenix: Oryx Press.

اطلاعاتی، موجود از دو کاتالوگ پایگاه اطلاعاتی منتشره توسط سرویس خدماتی دیالوگ و تکنولوژیهای اطلاعاتی BRS مشترکاً استفاده می‌شود.^(۱) بسیاری از مؤسسات، کتابداران متخصص را که در جستجوی کامپیوتری مهارت دارند و می‌توانند در خدمات مختلف به شما کمک کنند، در استخدام دارند.

از فواید جستار کامپیوتری می‌توان به این موارد اشاره کرد: (۱) صرفه جویی در وقت: به دلیل وجود بسیاری از اصطلاحات مترادف و همانند که برای مفهوم یا مفاهیم مورد نظر مفید است، کامپیوتر در چند دقیقه می‌تواند اطلاعات را بازبایی و چاپ و ازهدر رفتن ساعات و یا روزهای محقق در استخراج دستی جلوگیری کند؛ (۲) در جریان روز بودن: پایگاههای اطلاعاتی کامپیوتری مجموعاً بیشتر از فهرستهای چاپی در جریان اطلاعات روز هستند. داده‌های بیشتر پایگاههای اطلاعاتی هر هفته و یا دو هفته یک بار به روز در می‌آیند؛ (۳) پیوند: چون هر موضوع تحقیقی به ترکیبی از دو و یا چند موضوع مربوط است، کامپیوتر می‌تواند همزمان بیش از یک عنوان را جستجو کند؛ (۴) دقت: زمانی که یک موضوع را نمی‌توان به سهولت از طریق فهرستهای چاپی شناسایی کرد، کامپیوتر می‌تواند از طریق فراهم ساختن امکان دستیابی به مواد مختلف از قبیل: عناوین، کلمات کلیدی، رمزهای موضوع، یا سر تیتراهای موضوعی ضمیمه که فهرستهای متداول قادر به ارائه آنها نیستند، این مهم را به انجام رساند. جستجوی کامپیوتری همچنین می‌تواند نتایج شما را بر اساس گزینه‌های معین، از جمله تاریخ و نوع نشریه، نام نویسنده و یا زبان محدود کند (۵) منابع منحصر به فرد: برخی از پایگاههای اطلاعاتی داده‌هایی را فراهم می‌کنند که در منابع چاپی موجود نیست و (۶) دسته بندی: جستجوهای کامپیوتری اطلاعات بازبایی شده را بر طبق تاریخ انتشار (درخواست حضوری یا غیرحضوری) یا نام نویسنده و عنوان با ترتیب حروف الفبایی دسته‌بندی می‌کند.

برای بیشتر محققان جستجوی کامپیوتری تنها روش برای انتخاب است. در یک جستجوی «دستی» فرد باید برای یک عنوان و ویژه، تعداد بیشماري از فهرستها را بررسی کند، در بین فهرستها به دنبال آنها بگردد و تعداد ناچیزی داده‌های مربوط بیابد که اصطلاحات منتخب را با دیگر زمینه‌های مورد علاقه ترکیب کرده است. هر چند با استفاده از مفاهیم منطقی به نام عمل کننده‌های بول^(۲) (از سیستم منطقی‌ای که توسط جرج بول، ریاضیدان انگلیسی، در قرن نوزدهم

1- 1988 Dialog Database Catalog (1988). Palo Alto: Dialog Information Services, Inc.; 1988 BRS Database Catalog (1988).

Latham, NY: BRS Information Technologies.

2- Boole.

تحول یافت) می‌توان همزمان برای بسیاری از موضوعات به جستجوی کامپیوتری پرداخت و آنها را با یکدیگر ترکیب کرد.

شکل ۳-۳ راه‌هایی که مفاهیم و یا اصطلاحات مختلف را می‌توان با استفاده از عمل‌کننده‌های منطقی بول به یکدیگر پیوند داد، نشان می‌دهد.

طبق این شکل، $A \text{ AND } B$ برای مواردی به جستجو می‌پردازند که هر دو اصطلاح کلیدی را دارا هستند. برای مثال، در مارس ۱۹۸۹ سیستم اریک ۱۰۴۷ مورد با اصطلاح کلیدی دانش‌آموزان بزرگسال و ۲۵۰۲ مورد با اصطلاح کلیدی آموزش ریاضیات در اختیار داشت. با استفاده از برنامه سیلور پلاتر چهار مورد با دو اصطلاح کلیدی معین شد.

در این شکل، $A \text{ OR } B$ اصطلاحاتی را جستجو می‌کند که یک یا دو (و نه منحصرأً یکی) اصطلاح را دارا باشند. در مثال مشابهی درخواست برای آموزش ریاضیات یا پیشرفت ریاضیات به منابعی دست خواهد یافت که یک یا هر دو اصطلاح کلیدی را دارا باشد.

یک جستجوی معمولی از هر دوی این راهبردها استفاده می‌کند. یک مربی بالغ که به آموزش ریاضیات علاقه‌مند است، از عبارت OR ، شامل اصطلاحات کلیدی متعددی که احتمالاً دامنه آموزش بزرگسالان را پوشش می‌دهد و یا عبارت OR ، شامل عبارتهای متعددی که دامنه آموزش ریاضیات را در می‌گیرد، به همراه یک عبارت از نوع AND برای پیوند هر دو زمینه (عوامل) استفاده می‌کند.

عوامل آموزش ریاضیات

پیشرفت ریاضیات

OR

آموزش ریاضیات

OR

مهارتهای ریاضیات

عوامل آموزش بزرگسالان

آموزش ابتدایی بزرگسالان

OR

آموزش بزرگسالان

یادگیری بزرگسالان

OR

برنامه‌های بزرگسالان

OR

دانش آموزان بزرگسال

OR

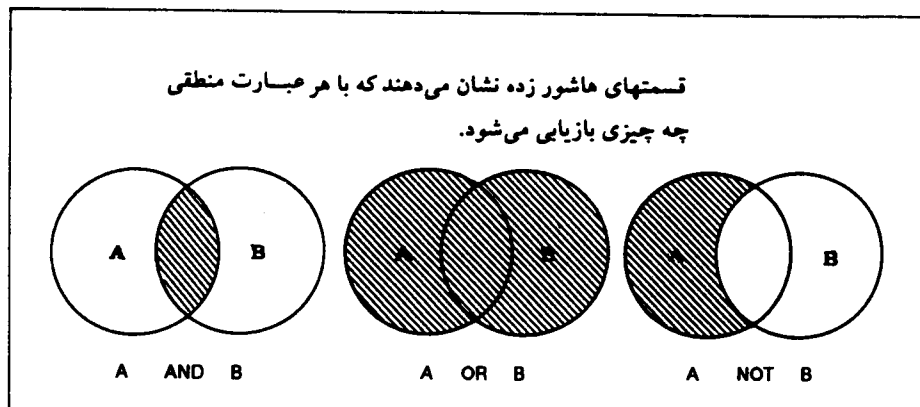
ادامه تحصیل

OR

دانش آموزان غیر سنتی

با استفاده از این عوامل، برنامه سیلور پلاتر، نشان داد که ۱۵۳۱۶ مورد از عامل اول و ۳۰۰۹ مورد از عامل دوم یک یا چند اصطلاح کلیدی داشتند. فقط ۹۴ مورد از هر دو عامل در نهایت

شکل ۳-۳ منطق بول



یک اصطلاح کلیدی داشتند که از آن میان ۷۳ مورد از اسناد اریک و ۲۱ مورد مجلات تحت پوشش CIJE بودند. (همه این اطلاعات از طریق کامپیوتر و در کمتر از چهار دقیقه به دست آمد!) در این مرحله، فرد می‌تواند تعداد موارد را با گزینش راهبردی که عوامل دیگری را ارائه می‌کند، محدود کند و موارد را به آنهایی منحصر کند که انگلیسی نیستند یا جستجو را به یک تاریخ خاص محدود سازد. همچنین فرد می‌تواند اجزای متعددی از اطلاعات را درباره هر مورد که بر صفحه کامپیوتر ظاهر می‌شود، بر اساس این تصمیم که چه چیزی به کارش مربوط می‌شود، انتخاب کند. در پایان فرد می‌تواند برای موارد باقیمانده، از این اطلاعات نسخه چاپی کامپیوتری تهیه کند. شکل ۳-۴ یک نسخه چاپ شده کامپیوتری از ۲۱ مقاله مجلات است که از طریق جستجوی کامپیوتری تعیین شده، و شامل شماره سند، نویسنده، عنوان، مقاله (حاوی عنوان، مجله، شماره بخشها، شماره، و تاریخ)، توصیفگرها و چکیده است.

مثال پیشین رایجترین موارد استفاده از منطق بول را نشان می‌دهد. عمل کننده‌های دیگر بعضی اوقات به کار می‌روند. برای نمونه، NOT می‌تواند برای موارد «آموزش بزرگسالان که دارای توصیفگر «آموزش ابتدایی بزرگسال» نیز نیستند، به جستجو بپردازند.

// واژه‌نامه‌ها

تقریباً می‌توان گفت که هر رشته علمی واژگان تخصصی خود را داراست. عملکرد یک واژه‌نامه موضوعی، تعریف مختصر کلمات، چه اصطلاحات و چه اسامی است که زبان فنی یک رشته علمی خاص را به وجود می‌آورند. واژه‌نامه‌های متعددی در زمینه تعلیم و تربیت وجود دارند. **واژه‌نامه تعلیم و تربیت گود^(۱)**، یک واژه‌نامه دانشگاهی حاوی کلمات و اصطلاحاتی است که در زمینه تعلیم و تربیت دارای معانی ویژه هستند. نام حدوداً ۲۰۰ هماهنگ کننده، دستیار و مرورگر در پایان چاپ سوم این واژه‌نامه، که حاوی بیش از ۳۳۰۰۰ مدخل است، فهرست شده است. اصطلاحات تعلیم و تربیتی مورد استفاده در کانادا، انگلستان و ولز در بخشهای جداگانه در انتهای این واژه‌نامه، ارائه شده است. **واژه‌نامه بین‌المللی تعلیم و تربیت^(۲)** حاوی بالغ بر ۱۰۰۰۰ مدخل، شامل تعاریف و اصطلاحات، سازمانهای بین‌المللی، مؤسسات و سازمانهای اصلی ملی، کارشناسان تعلیم و تربیت، و از این قبیل است.

واژه‌نامه تعلیم و تربیت^(۳)، که توسط هیل ویراستاری شده است، مقالات کوتاهی در ۱۵ زمینه تعلیم و تربیت ارائه می‌کند که براساس یک سیستم الفبایی از اصطلاحات (عمدتاً اصطلاحات انگلیسی) فهرست می‌شوند. جلد اول **دائرة المعارف بین‌المللی آموزش عالی^(۴)** نیز تعاریف کوتاهی از کلمات اختصاری و یک واژه‌نامه از اصطلاحات مهیا می‌سازد. در زمینه تعلیم و تربیت کودکان استثنایی، **واژه‌نامه ناتوانیها در خواندن و یادگیری بوش^(۵)** و **منابع راهنمای کارشناسان تعلیم و تربیت در آموزش کودکان استثنایی: اصطلاحات - قوانین - آزمونها و سازمانها تألیف دیویس^(۶)**، و **واژه‌نامه اصطلاحات تعلیم و تربیت کودکان استثنایی^(۷)** سه واژه‌نامه عمده و کاربردی هستند.

1- Good. c.v. (1973). Dictionary of Education (3d ed.). New York: McGraw - Hill.

2- Page, G.P., J.B. Thomas, and A.R. Marshall (1977). International Dictionary of Education. New York: Nicholas Publishing.

3- Hill, P.J. (ed.) (1982). A Dictionary of Education. Boston: Kegan Paul.

4- Knowles, A.S. (ED.) (1977). The International Encyclopedia of Higher Education (10 vols.). San Francisco: Jossey- Bass.

5- Bush, C.L., and R.C. Andrews (1978). Dictionary of Reading and Learning Disabilities. Los Angeles: Western Psychological Services.

6- Davis, W.E. (1980). Educator's Resource Guide to Special Education: Terms - Laws- Tests- Organizations. Boston: Allyn & Bacon.

7- Moore, B.C., W. Abraham, and C.R. Laing (1980). A Dictionary of Special Education Terms. Springfield, IL: Thomas.

شکل ۴-۳ نمونه‌ای از یک برگه جستجوی کامپیوتری ERIC

AN: EJ293655

AU: Dolan, -Lawrence

TI: Affective Characteristics of the Adult Learner: A Study of Mastery-Based Instruction.

JN: Community/Junior-College-Quarterly-of-Research-and-Practice;
v7 n4 p367-78 Jul-Sep 1983

DE: Affective-Behavior; Community-Colleges; Remedial-Mathematics; Two-Year-Colleges

DE: *Adult-Learning; *Mastery-Learning; *Mathematics-Achievement;
*Mathematics-Instruction; *Reentry-Students; *Student Characteristics

AB: Presents an investigation of the benefits of mastery learning strategies for adults with mathematics skill deficiencies. Examines which student types benefit most from mastery learning, centering on affective characteristics. Demonstrates that students with low self-concept, control, and instructional mastery skills had higher achievement in mastery than nonmastery classes. (DMM)

// قرضه بین کتابخانه‌ای

خدمات قرضه بین کتابخانه‌ای نیز می‌تواند فوق‌العاده مؤثر باشد. اگر شما منابعی برای بررسی تحقیقی در مورد نشریه‌های علمی، کتابها، یا رساله‌هایی بیابید که در کتابخانه همیشه شما موجود نباشد، از طریق خدمات قرضه بین کتابخانه‌ای امکان به امانت گرفتن آن از کتابخانه مورد نظر وجود دارد.

/// سازماندهی تحقیقات پیشین

هنگامی که محقق متقاعد شد که یک بررسی منطقاً جامع از تحقیقات پیشین در زمینه مورد

نظرش انجام یافته است، می تواند به سازماندهی آن بپردازد. یک روش سودمند، منظم کردن بررسیها بر طبق موضوع و تعیین این است که هر یک از این موضوعات چگونه به بررسی مورد نظر مربوط هستند.

تحقیقات پیشین باید به نحوی ارائه شوند که با نشان دادن آنکه چه نتایجی به دست آمده و چه مواردی برای تحقیق در زمینه مورد نظر باقی مانده است، انجام آن را از طرف محقق توجیه کنند. فرضیات، چهارچوبی برای یک چنین سازماندهی فراهم می کنند. همانند مکتشفی که قصد سفر دارد، محقق نیز از قلمروی شناخته شده نقشه برداری می کند و راه رسیدن به قلمروی ناشناخته را نشان می دهد. اگر بررسی ابعاد متعددی داشته باشد یا در مورد بیش از یک فرضیه پژوهش کند، فرایند سازماندهی برای هر کدام مجزا خواهد بود.

محقق باید از وسوسه ارائه تحقیقات پیشین به عنوان مجموعه ای از چکیده ها پرهیز کند. به معنای دقیق تر، تحقیقات پیشین باید به گونه ای ارائه شوند که شالوده ای نظام مند برای بررسی فراهم سازند.

این مسئله نیز تقریباً اجتناب ناپذیر است که تعدادی از گزارشهایی که محقق به دقت مطالعه کرده و از آنها یادداشت نیز برداشته است، در رجوع بعدی با موضوع مورد مطالعه کم ارتباط به نظر آیند. آوردن کلیه بررسیهایی که محقق در جستجوی تحقیقات پیشین به آن برخورد می کند در طرح اولیه، نه ضروری و نه مطلوب است. خوانندگان به صرف کمیت تحقیقات پیشین تحت تأثیر قرار نخواهند گرفت. ربط و سازماندهی مواد از اولویت اول برخوردار است. چنانچه محقق از ابتدا روشی نظام مند برای گردآوری منابع و مآخذ موجود اختیار نکند، دچار بی نظمی خواهد شد. پیشنهادها زیر می تواند در این زمینه چاره ساز باشد.

۱- از آخرین بررسیها در زمینه مورد نظر خود آغاز کنید و سپس به مطالعات قبلی باز گردید. یکی از نقاط قوت آشکار این رویکرد این است که شما با بررسیهایی آغاز می کنید که از تفکرات و یافته های گذشته ترکیب یافته اند. سوء تفاهمات اولیه تصحیح و از روشهای نامناسب دوری شده است. مزیت دیگر، این است که این بررسیها منابع مرجع کارهای اولیه را در بر دارند، بنابراین ما را به منابعی هدایت می کنند که در غیر این صورت، با آنها مواجه نخواهیم شد. آشکار است که در فرآیند گردآوری تحقیقات پیشین، باید محدوده ای تعیین شود. از یک سو، فراهم آوردن شالوده ای پر معنی برای تحقیق متضمن برخورداری از تمام کارهای مهم گذشته در

همان زمینه است و از سوی دیگر، صرف وقت بیش از حد در این کوشش می‌تواند خوانندگان گزارش شما را به دلیل ذکر جزئیات زاید خسته کند. مطمئن شوید که مآخذ و منابع مربوطه، به بررسی شما خدمت می‌کنند و نه حکومت.

۲- ابتدا چکیده یا خلاصه یک گزارش را مطالعه کنید تا بتوانید معین کنید که این گزارش با سؤال شما رابطه دارد یا خیر. این کار از اتلاف وقت زیاد در جریان خواندن مقالات غیر سودمند جلوگیری می‌کند.

۳- قبل از یادداشت برداری، گزارش را به طور سریع بخوانید تا بخشهایی را که با سؤال شما مرتبط است، پیدا کنید. این، راه دیگری برای جلوگیری از اتلاف وقت است.

۴- یادداشتهای خود را بر روی فیش کارت بنویسید، چون مرتب کردن و سازماندهی آنها نسبت به کاغذهای متفرقه راحت تر است. بسیاری، فیش کارتهای ۶-۴ اینچی را ترجیح می‌دهند، چرا که هم فضای قابل قبولی برای یادداشت برداری دارد و هم آنقدر کوچک است که می‌توان آنرا در پاکت و یا کیف جای داد.

۵- در جریان بررسی هرکاری کتاب شناسی کامل آنرا یادداشت کنید. اگر در آغاز کار بدانید که چه شیوه‌ای در پایان گزارش به کار خواهید برد، با یادداشت برداری بر اساس آن شیوه در زمان صرفه جویی خواهید کرد. شماره تلفن کتابخانه را نیز یادداشت کنید. این امر ممکن است در بازبایی مجدد بررسی ضروری باشد.

۶- برای تسهیل در مرتب کردن و سازماندهی کارتها، هرگز بیش از یک منبع را بر روی آنها یادداشت نکنید. مرتب کردن منابع شما به صورت الفبایی و یا با هر روش دیگری بدون در نظر گرفتن این مسئله ممکن نخواهد بود.

۷- مشخص کنید که چه بخشی از یادداشتهای سؤالات مستقیم مؤلف و کدامیک تفسیرهای شماست. عدم موفقیت در انجام این تفکیک، می‌تواند شما را سهواً به سوی دزدی تالیفات و افکار دیگران سوق دهد. همچنین عاقلانه خواهد بود که ارزیابی نویسنده از کار خودش را با ارزیابی خود از آن تفکیک کنید.

۸- اگر جستجوی کامپیوتری انجام شده باشد، شیوه جستجوی خود را در فایل کامپیوتر حفظ کنید. این اطلاعات به شما در بازبایی اطلاعات دقیق و کاهش هزینه روزآمد کردن اطلاعات مورد نیاز، کمک می‌کند.

جدول ۱-۳ منابع تحقیقات مربوط به زمینه تعلیم و تربیت

منبع	محتوا
چکیده‌ها و کتابشناسی رشد کودک	چکیده‌های مقالات مجلات در زمینه رشد کودک
نمایه روز مجلات تعلیم و تربیت	عناوین، مؤلفان و مشخصات نشریات از زمینه مقالات مربوط به تعلیم و تربیت، با حاشیه نویسی‌های مورد نیاز. طبقه‌بندی برحسب موضوع، مؤلف، و محله بدون چکیده.
چکیده رساله‌های بین‌المللی نمایه تعلیم و تربیت	چکیده رساله‌های دوره دکترا در ایالات متحده، کانادا و اروپا. عناوین، مؤلفان و مشخصات نشریات در زمینه مقالات مربوط به تعلیم و تربیت. فهرست شده بر اساس موضوع و عنوان. بدون چکیده.
دائرة المعارف تحقیق در تعلیم و تربیت مجموعه میکروفیشهای اریک	خلاصه‌ها و ارزشیابیهای تحقیقات منتشر شده در زمینه آموزش و پرورش در پایان هر دهه. اسناد کامل از تحقیقاتی که چکیده‌های آنها در منابع تعلیم و تربیت موجود است
منابع تعلیم و تربیت کودکان استثنایی	چکیده مقالات مجلات و منابع دیگر در تعلیم و تربیت
سالنامه اندازه‌گیری‌های روانی	اطلاعات و ارزشیابیهای آزمونهایی که به صورت تجاری در دسترس است.
کاتالوگ ماهانه نشریات دولتی	فهرست موضوعی نشریات دولت فدرال ایالات متحده
چکیده‌های روان شناختی	چکیده مقالات مجلات در زمینه روان شناسی
راهنمای خوانندگان در پیشینه مجلات علمی	عناوین، مؤلفان و مشخصات نشریات در زمینه مقالات دیگر نشریات مشهور، فهرست شده بر اساس موضوع و عنوان بدون چکیده.
منابع تعلیم و تربیت	چکیده گزارشهای تحقیقی و سایر اسناد گردآوری شده توسط دفاتر مرکزی اریک
بررسی تحقیقات آموزشی	بررسی موضوعات متنوع در فصلنامه‌ها
نمایه نقل نامهای علوم اجتماعی نقد آزمون	اطلاعات کتاب شناختی از نویسندگان و موضوعات نقل شده بررسی وسایل اندازه‌گیری، شامل توصیفهای کلی، کاربردهای مفید / موارد استفاده، نقدها، و جنبه‌های تکنیکی
آزمونهایی چاپ شده شماره سه	فهرست و مضمون سالنامه اندازه‌گیریهای روانی از شماره اول تا شماره هشت
گنجینه توصیفگرهای اریک	سیستم برای طبقه‌بندی و فهرست کردن اسناد اریک

/// خلاصه

اگر محقق همه راههای منتهی به اطلاعات مورد نظر را همانطور که پیشنهاد شد، بپیماید، تصویر کامل و قابل قبولی از جایگاه بررسی خود در آن حیطه تحقیقاتی خاص خواهد داشت. در جدول ۱-۳ مهمترین منابع و مضامین پیشینه تحقیقات در تعلیم و تربیت خلاصه شده است.

// مفاهیم اصلی

بانک اطلاعات

جستجوی کامپیوتری

دائرةالمعارفهای تخصصی

راهنماهای کتابشناختی

راهنماهای اسناد دولتی

سیستم اریک

فهرستهای رساله‌ها و پایان‌نامه‌ها

مروورهای دوره‌ای

منابع آزمونها

منطق بول

/// تمرینها

- ۱- سه نقش مهم تحقیقات پیشین در پروژه تحقیقی را بیان کنید.
- ۲- شایستگی CIE و فهرست تعلیم و تربیت را با هم مقایسه کنید.
- ۳- کدامیک از راهبردهای ذیل برای سازماندهی جستجو در تحقیقات پیشین قابل قبول است.

الف - سازماندهی بررسیها بر طبق موضوع

ب - با مقالات اولیه آغاز کنید و به موقع به کارهای تازه پردازید.

ج - ابتدا بخشهای چکیده یا خلاصه یک گزارش را مطالعه کنید.

د - گزارش را به طور اجمال بخوانید تا بخشهایی را که به سوالات شما مربوط است، بیابید.

ه - برای هر کار یک منبع کتابشناسی کامل بنویسید.

- ۴- در چه نقطه‌ای از روند انجام پروژه تحقیقی، محقق باید تحقیقات پیشین و بررسیهای مربوط به سوالات انتخابی و مورد نظر را مطالعه کند؟

۵ - حاشیه نویسی چیست؟

- ۶- اهمیت کتاب سالانه اندازه‌گیری ذهنی بوروس در ایجاد منابع آزمون‌ها را توضیح دهید.
- ۷- چرا اریک چنین منبع مفیدی برای مشاوره در جستجوی پیشینه تحقیقات است؟
- ۸- موارد استفاده از فهرست نقل نامهای علوم اجتماعی در جستجوی پیشینه تحقیقات را توضیح دهید.
- ۹- مراحل معین در استفاده از سیستم اریک برای مروری بر تحقیقات پیشین کدام است؟
- ۱۰- مزایای استفاده از جستجوی کامپیوتری در پیشینه تحقیقات تخصصی بانکهای اطلاعاتی چیست؟
- ۱۱- چگونه از منطق بول در جستجوی بانکهای اطلاعاتی استفاده می‌شود؟
- ۱۲- در شکل ۱-۳ کدامیک از منابع مربوط به پیشینه تحقیقات در کتابخانه‌های قابل دسترس شما موجود است؟

/// پاسخها

- ۱- شناخت تحقیقات پیشین محقق را قادر می‌سازد تا پیشتازان حیطه مورد نظر را تعیین کند، سؤال تحقیق را در یک دورنما قرار دهد، و از تکرار سهوی بررسیهای قبل اجتناب کند.
- ۲- فهرست تعلیم و تربیت راه دستیابی به مقالات منتشره را از ۱۹۲۹ تا به امروز فراهم می‌کند. CIJE بیشتر مجلات را پوشش می‌دهد و حاشیه‌نویسی را تهیه می‌کند، در حالی که فهرست تعلیم و تربیت مقالات را سریعتر از CIJE فهرست می‌کند.
- ۳- الف، ج، د و ه.
- ۴- محقق به مجرد آنکه موضوع تحقیق انتخاب و طرح شد، پیشینه تحقیقات را مورد بررسی قرار می‌دهد.
- ۵- یک حاشیه نویسی، توضیح کوتاهی از یک مقاله است و می‌تواند به محقق در انتخاب مقالات مربوط به موضوع کمک کند.
- ۶- آنها پر معناترین فهرست و تبیین از آزمونهای استاندارد شده موجود را ارائه می‌کنند. مرورهای دوره‌ای از آزمون‌ها به محقق در انتخاب یک آزمون مناسب کمک می‌کند.
- ۷- اریک در پوشش دادن بسیاری از موضوعات در زمینه تعلیم و تربیت، یک سیستم جامع محسوب می‌شود. هر یک از بخشهای تخصصی به یک زمینه موضوعی ویژه اختصاص دارند

که اطلاعات و مواد را جمع‌آوری و وارد سیستم می‌کنند. سیستم اریک همچنین محدوده وسیعی از مواد شامل: بررسیها و پروژه‌های تحقیقی، راهنماهای درسی، مقالات معتبر از انجمنها و مؤسسات شناخته شده، کتابشناسی‌ها، و موارد درسی را تحت پوشش دارد. بسیاری از این مواد چاپ نشده‌اند و امکان دستیابی به آنها در جای دیگر میسر نیست. سیستم اریک نه تنها مواد را فهرست می‌کند، بلکه از بیشتر آنها میکروفیش تهیه می‌کند.

۸- از طریق **فهرست نقل نامها** فرد می‌تواند با تعیین مقالاتی که از کارهای ویژه‌ای نام برده‌اند، یک خط تحقیقی رو به جلو را تعقیب کند. **فهرست موضوعی اصطلاحات** فرد را قادر می‌سازد تا مقالات را توسط کلمات معادل با عناوین آنها بیابد.

۹- کلمات کلیدی در بررسی خود را تعیین کنید.

الف - کلمات کلیدی یا معادلهای آنها را در دائرةالمعارف توصیفگران اریک پیدا کنید.

ب - از یک برنامه جستجوی کامپیوتری برای یافتن مقالات مجلات و آن دست از اسناد اریک، که حاوی اصطلاحات و یا توصیفگران کلیدی مورد نظر شماست، استفاده کنید، یا از طریق CIE و RIE برای یافتن شماره‌های دسترسی مقالات و اسناد با اطلاعات معادل خود، به یک جستجوی غیر کامپیوتری (دست) پردازید.

ج - چکیده‌ها یا حاشیه‌نویسی‌ها را برای حذف مقالات و اسناد بی‌ربط مطالعه کنید.

د - مقالات و اسناد باقیمانده اریک را شناسایی کنید.

ه - ابتدا چکیده، خلاصه و نتیجه‌گیری را بخوانید.

۱۰- جستجوگر کامپیوتری می‌تواند در سیستم بانکهای اطلاعاتی سریعاً به جستجو پردازد. آنها بویژه برای شناسایی مقالات و اسناد دیگر، با کمک ترکیب خاص از کلمات کلیدی، مفید هستند. آنها می‌توانند به طرق مختلف، برای مثال: عنوان، تاریخ، نویسندگان، کلمات کلیدی، یا رمزهای موضوع، به جستجو پردازند. اطلاعات بانکهای اطلاعاتی کامپیوتری از فهرستهای چاپی تازه‌تر هستند و اطلاعاتی را شامل می‌شوند که درجای دیگری قابل دستیابی نیستند.

۱۱- جستجو به طریق A و B تنها اسناد دارای هر دو توصیفگر را شناسایی می‌کند.

جستجو از طریق A یا B همه اسناد با یک توصیفگر را شناسایی می‌کند.

جستجو از طریق A و نه B بدون شناسایی اسناد با توصیفگر B همه اسناد با توصیفگر

A را شناسایی می‌کند.

۱۲- پاسخها متفاوت خواهند بود.

فصل چهارم

فرضیه

هدفهای آموزشی

دانشجو پس از مطالعه این فصل، قادر خواهد بود:

- ۱- فرضیه را تعریف کند.
- ۲- هدف از فرضیه (ها) را در یک بررسی تحقیقی بیان کند.
- ۳- بین فرضیه استقرایی و قیاسی تمایز قائل شود.
- ۴- معیارهای مورد استفاده در ارزشیابی فرضیات تحقیق را بیان کند.
- ۵- یک تعریف کاربردی را بیان و یک مثال ارائه کند.
- ۶- از مثالهای داده شده یک فرضیه قابل آزمون اقامه کند.
- ۷- یک فرضیه صفر را تعریف و هدف از آن را در یک بررسی تحقیقی تشریح کند.
- ۸- یک فرضیه تحقیقاتی و یک فرضیه صفر را برای یک بررسی تحقیقی طرح کند.
- ۹- بین فرضیات جهت دار و غیر جهت دار تمایز قائل شود.
- ۱۰- مراحل آزمون یک فرضیه را تعریف کند.
- ۱۱- هدف (هدفها) از انجام یک بررسی مقدماتی را بیان کند.
- ۱۲- عوامل مورد لزوم در یک طرح تحقیقی را فهرست کند.

فرضیه، ابزاری قدرتمند در تحقیق علمی است. این ابزار ما را قادر می‌سازد تا نظریه را به مشاهده و مشاهده را به نظریه مربوط سازیم. استفاده از فرضیه‌ها، انسان عصر حاضر را قادر می‌سازد که ایده‌های فیلسوفان استقرایی را با توجه به تاکید آنها بر مشاهده، و منطق فیلسوفان قیاسی را با توجه به تاکید آنها بر تعقل، با هم به کار گیرند. استفاده از فرضیه‌ها، تجربه و تعقل را برای خلق ابزاری قدرتمند جهت جستجوی حقیقت پیوند می‌دهد.

محقق، بعد از تعیین و بیان مسئله و بررسی متون تحقیقاتی پیشین، برای ساختن فرضیه خود آماده است. یک فرضیه می‌تواند به صورت یک گزاره^(۱) آزمایشی پیشنهادی، به عنوان راه حل یک مسئله، یا به عنوان توصیف چند پدیده، به طور دقیق تعریف شود. فرضیه، بیان ساده انتظارات محقق نسبت به رابطه بین متغیرهای موجود در مسئله است و سپس در جریان یک بررسی تحقیقاتی مورد آزمون قرار می‌گیرد. از این روست که فرضیه تنها به عنوان راه حل پیشنهادی برای یک مسئله و با درک این موضوع که نتیجه تحقیق می‌تواند باعث ابقا و یا رد آن شود، ارائه می‌شود.

برای مثال، محقق ممکن است با این سؤال آغاز کند: ادراک کودکان از خود، چه نقشی در فرآیند یادگیری خواندن آنها ایفا می‌کند؟ او سپس، باید فرض کند: بین ادراک کودکان کلاس اول از خود و پیشرفت در امر خواندن رابطه مستقیم وجود دارد. یا محقق ممکن است با سؤالی از این دست شروع کند: آموزش پیش دبستانی چه اثری بر پیشرفت تحصیلی کودکان کلاس اول که محرومیت فرهنگی دارند، می‌گذارد؟ فرضیه را می‌توان این گونه عنوان کرد: در کلاس اول، پیشرفت تحصیلی کودکان دارای محرومیت فرهنگی که آموزش پیش دبستانی را گذرانده‌اند، نسبت به کودکانی که محرومیت فرهنگی دارند، اما آموزش پیش دبستانی را ندیده‌اند، بیشتر است. در هر دو مثال، می‌توان مشاهده کرد که فرضیه، گزاره‌ای است که دو متغیر را به هم ربط می‌دهد. در مثال اول، متغیرها، ادراک خویشتن و پیشرفت در امر خواندن هستند، و در مثال بعدی متغیرها، آموزش پیش دبستانی و پیشرفت تحصیلی در کلاس اول هستند.

اگرچه صاحب‌نظران استفاده از فرضیه را توصیه می‌کنند، و همان طور که خواهیم دید، فرضیه در خدمت مقاصد متعدد و مهمی است، اما وجود آن در همه مطالعات تحقیقی کاملاً ضروری نیست. فرضیات در فرایند تحقیق یک ابزار هستند، اما به خودی خود هدف نیستند. بررسیها

اغلب در حیطه‌ای صورت می‌گیرند که مجموعه‌ی کوچکی از اطلاعات پس زمینه‌ای در آن مورد وجود داشته باشد. اگر محقق بینشی نسبت به حوزه مسئله مورد نظر، یا متغیرهای اصلی که در یک پدیده مؤثر هستند، یا محیطی که متغیرها در آن بروز می‌کنند، نداشته باشد، آن وقت است که تبیین یک فرضیه با معنا، برای او بسیار دشوار خواهد بود. برای مثال، پیمایشهایی که سعی در توضیح ویژگیهای پدیده‌ای خاص دارند، یا در پی تعیین نگرشها و عقاید گروهها هستند، اغلب کار تحقیق را بدون فرضیه دنبال می‌کنند. در این مرحله، داده‌های تجربی که بر طبق آن فرضیاتی را بتوان بیان کرد، هنوز در دسترس محقق نیست. در چنین بررسیهایی بیان مسئله تحقیق کافی است. در تحقیقاتی که رابطه علی بین متغیرها مورد بررسی قرار می‌گیرد، فرضیه‌ها اهمیت بسیار پیدامی‌کنند.

فرضیه باید قبل از مرحله گردآوری داده‌ها ساخته شود. این امر به دو دلیل ضروری است: (۱) فرضیه با شالوده درست نشان می‌دهد که محقق در حیطه تحقیق خود دارای دانش معتبری است. (۲) فرضیه به گردآوری و تفسیر داده‌ها جهت می‌دهد، فرضیه به ما می‌گوید چه شیوه‌ای را در پیش بگیریم و چه داده‌هایی را گردآوری کنیم و بنابراین می‌تواند به مقدار زیادی از اتلاف وقت و تلاش محقق جلوگیری کند. باید تاکید شود که این امر، نه تنها برای تحقیق آزمایشی، بلکه برای همه انواع تحقیق درست است.

هدفهای کاربرد فرضیه را می‌توان اینگونه خلاصه کرد:

۱- فرضیه‌ها توصیفهایی آزمایشی از پدیده‌ها فراهم می‌آورند و میزان گسترش دانش ما را در یک حیطه خاص افزایش می‌دهند. انسان، برای رسیدن به دانش معتبر در زمینه مسائل تعلیم و تربیت باید از گردآوری صرف واقعیتهای مجزا از هم فرا تر رود تا به تعمیم‌ها و روابط متقابل موجود در این واقعیتهای دست یابد. این روابط متقابل و تعمیم‌ها طرحی را فراهم می‌آورند که برای درک مسئله مهم است. چنین طرحی، احتمالاً تا زمانی که گردآوری داده‌ها جهت مشخصی نداشته باشد، آشکار نخواهد شد. فرضیاتی که خوب طراحی شده‌اند، جهت را مشخص و توصیفها را پیشنهاد می‌کنند. از آنجا که فرضیه‌ها می‌توانند با بررسی علمی، مورد آزمون قرار گیرند و اعتبار یابی شوند، امکان بسط دانش ما را فراهم می‌آورند.

۲- فرضیه، محقق را با یک بیان رابطه‌ای مجهز می‌کند که به طور مستقیم در جریان مطالعه علمی قابل آزمون است. سوالات نمی‌توانند به طور مستقیم مورد آزمون قرار گیرند. تحقیق با

یک سؤال آغاز می‌شود، ولی تنها روابط پیشنهادی بین متغیرها می‌توانند مورد آزمون قرار گیرند. برای مثال، محقق نمی‌تواند این سؤال را مورد آزمون قرار دهد: آیا اظهار نظرهای معلم در مورد نوشته‌های دانش‌آموزان باعث پیشرفت قابل توجهی در کار آنها می‌شود؟ در عوض، محقق فرضیه‌ای را که این سؤال بر آن دلالت دارد، آزمون می‌کند: اظهار نظر معلم در مورد نوشته‌های دانش‌آموزان موجب پیشرفت قابل توجهی در عملکرد آنها می‌شود. یا به طور مشخص‌تر، نمره عملکرد دانش‌آموزانی که از اظهار نظر معلم خود درباره نوشته‌های قبلی خویش برخوردار شده‌اند، از نمره عملکرد دانش‌آموزانی که چنین امکانی نداشته‌اند، بیشتر خواهد بود. سپس محقق به‌سوی بررسی رابطه بین این دو متغیر، یعنی نظرات معلم و عملکرد دانش‌آموز حرکت خواهد کرد.

۳- **فرضیه‌ها تحقیق را هدایت می‌کنند.** فرضیه هدف مشخصی را معرفی می‌کند و به این ترتیب ماهیت داده‌های ضروری را برای آزمون رابطه پیشنهادی تعیین می‌کند. به زبان ساده‌تر، فرضیه به محقق می‌گوید که چه کند. واقعیتها باید انتخاب شوند و مشاهدات در ارتباط با سؤال مشخص صورت گیرند و این فرضیه است که ارتباط بین این واقعیتها را تعیین می‌کند. فرضیه‌ها بنیانی برای انتخاب نمونه و روشهای مورد استفاده در تحقیق فراهم می‌آورند و تجزیه و تحلیل آماری مورد نیاز و رابطه‌هایی که باید مورد آزمون قرار گیرند، نیز از طریق فرضیه بیان می‌شوند. بعلاوه، فرضیه بررسی را در چهارچوب لازم نگه می‌دارد و از گستردگی بیش از حد آن جلوگیری می‌کند. برای مثال، فرضیه مربوط به آموزش پیش دبستانی کودکان دارای محرومیت فرهنگی و پیشرفت تحصیلی آنها را در کلاس اول در نظر بگیرید. این فرضیه، روش تحقیق ضروری و نمونه مورد استفاده را نشان می‌دهد و حتی محقق را در بکارگیری آزمونهای آماری مورد لزوم در تجزیه و تحلیل داده‌ها، راهنمایی می‌کند. از نحوه بیان فرضیه روشن است که محقق یک بررسی علی - مقایسه‌ای انجام خواهد داد، یعنی کودکان کلاس اول را که دارای محرومیت فرهنگی هستند و از آموزش پیش دبستانی برخوردار بوده‌اند با گروه مشابه، که آموزش پیش دبستانی نداشته‌اند، مقایسه می‌کند. هر تفاوتی در میانگین پیشرفت تحصیلی دو گروه را می‌توان با روشهای معنی‌دار آماری نظیر آزمون t و تحلیل واریانس مورد تجزیه و تحلیل قرار داد (این روشها در فصل ۶ مورد بحث قرار خواهند گرفت).

۴- **فرضیه‌ها چهارچوبی برای گزارش نتایج بررسی فراهم می‌آورند.** چنانچه محقق، هر فرضیه را جداگانه در نظر گیرد و نتیجه‌گیریهای مربوط به هر یک را تبیین کند، کار خود را بسیار

آسان خواهد یافت. به این ترتیب، او می‌تواند این بخش از گزارش‌نویسی را بر حول تدارک پاسخهایی برای فرضیات اصلی سازمان دهد. این امر موجب می‌شود که گزارش پرمعنا تر و خواندنی‌تر شود.

/// پیشنهادهایی برای استنتاج فرضیه‌ها

محقق چگونه یک فرضیه را استنتاج می‌کند؟ همچنان که در فصل دوم توضیح داده شد، یک بررسی می‌تواند از مسائل عملی، از موقعیتهای رفتاری مشاهده شده که نیازمند توصیف است، یا حتی بهتر از آن، از نظریه‌های آموزشی، روان‌شناختی یا جامعه‌شناختی نشئت گرفته باشد. بنابراین، فرضیه‌ها به طور استقرایی از مشاهده رفتار و یا به طور قیاسی از نظریه یا یافته‌های تحقیقات پیشین استنتاج می‌شوند.

// فرضیه‌های استقرایی

روش استقراء عبارت از صورتبندی یک فرضیه، از طریق تعمیم رابطه‌های مشاهده شده است؛ به این معنا که، محقق مشاهداتی در مورد رفتار صورت می‌دهد، به روندها یا رابطه‌های احتمالی توجه می‌کند و سپس برای توصیف این رفتار مشاهده شده فرضیه‌ای را ارائه می‌کند. البته، این فرآیند استدلالی، برای تعیین یافته‌های محققان دیگر باید با بررسی تحقیقات قبلی همراه شود. روش استقرایی منبع سودمندی از فرضیه‌ها برای معلم است. او هر روز رفتار دانش‌آموزان خود را مشاهده و سعی می‌کند این رفتار را با رفتار خود، رفتار سایر دانش‌آموزان، با روش تدریس خود و با تغییرات محیط مدرسه و غیره مربوط کند. این معلم می‌تواند بر مبنای چنین تجربه‌ای و بر اساس دانش خود در زمینه رفتار در محیط مدرسه، به طور استقرایی تعمیمی را صورتبندی کند که رابطه مشاهده شده را توصیف می‌کند. به هر حال اعتبار این توصیف برای تبدیل شدن به فرضیه، باید از طریق یک بررسی علمی تعیین شود.

شاید آموزگاری مشاهده کند که آزمونهای کلاسی او موجب بروز اضطراب شدیدی در دانش‌آموزان می‌شود و فکر کند که این اضطراب بر عملکرد آنها تأثیر منفی می‌گذارد. بعلاوه به این مسئله نیز توجه کرده باشد که وقتی به دانش‌آموزان امکان نوشتن نظراتشان را در مورد سؤالات می‌دهد، به نظر می‌رسد که عملکرد آزمون آنها بهتر می‌شود. معلم استدلال می‌کند که آزادی ابراز

نظرات می‌باید به هر حال باعث کاهش اضطراب در دانش‌آموزان شده و در نتیجه نمرات آنها را افزایش داده باشد. این مشاهده یک فرضیه را صورت‌بندی می‌کند: نمره دانش‌آموزانی که به نوشتن نظرات خود در مورد سؤالات آزمون بر روی پاسخنامه ترغیب شده‌اند، از نمره دانش‌آموزانی که این امکان را نداشته‌اند، بیشتر خواهد شد. سپس او می‌تواند برای آزمون اعتبار این فرضیه آزمایشی را طراحی کند. توجه کنید که فرضیه، اعتقاد معلم را در مورد رابطه بین دو متغیر نشان می‌دهد: ابراز نظر در مورد پرسشهای آزمون و عملکرد در آزمون. همچنین توجه کنید **اضطراب**، که حلقه‌ای از زنجیره قیاس برای ساخت فرضیه به شمار می‌آید، بخشی از فرضیه نهایی را به خود اختصاص نداده است. بنابراین، نتایج تحقیق تنها اطلاعاتی را در مورد رابطه بین ابراز نظر و عملکرد در آزمون فراهم خواهد آورد. رابطه بین اضطراب و نظرات و اضطراب و عملکرد آزمون می‌توانند موضوعاتی برای فرضیه‌های آتی جهت تحقیق باشند. محقق اغلب در خواهد یافت که ایده اصلی او مجموعه‌ای از روابط را دربردارد که به طور مستقیم قابل مشاهده نیستند. او سؤالش را برای تحقیق در مورد روابطی که قابلیت مشاهده مستقیم دارند، دوباره صورت‌بندی می‌کند.

در ادامه، چند مثال ساده دیگر از فرضیاتی که می‌توانند حاصل مشاهدات یک معلم باشند، ارائه می‌شوند: بین اضطراب در مورد ریاضی و اضطراب در مورد کامپیوتر در بین دانش‌آموزان دوره دبیرستان، رابطه مثبت وجود دارد؛ یادگیری کامپیوتر از طرف دانش‌آموزان در دوره اول دبیرستان، رشد تفکر منطقی را در آنها افزایش می‌دهد؛ استفاده از دفترچه یادداشت‌های الکترونیکی پیشرفته^(۱) یادگیری دانش‌آموزان دبیرستانی را از آموزش به کمک کامپیوتر در درس شیمی افزایش می‌دهد؛ آموزش خاصیت جابجایی، سرعت یادگیری ترکیبهای عمل جمع را افزایش می‌دهد؛ کودکانی که در گروههای کوچک آموزش می‌بینند، در آزمون نهایی خواندن در کلاس اول نمره بیشتری از آن دسته از دانش‌آموزان که در گروههای بزرگتر آموزش دیده‌اند، کسب می‌کنند؛ و رشد شناختی و احساسی کودکان کلاس اول، تحت تأثیر مجموعه تجربیات قبلی آنها در کودکی قرار دارد. محقق در فرآیند استقراء در وهله اول مشاهده می‌کند، سپس در مورد مسئله فکر می‌کند و در مرحله بعدی برای یافتن سرنخ به مآخذ موجود مراجعه می‌کند، بعد مشاهدات بیشتری انجام می‌دهد و پس از آن فرضیه‌ای را که سعی در تبیین علتی برای رفتار

مشاهده شده دارد، صورتبندی می‌کند. در مرحله بعدی فرضیه تحت شرایط کنترل شده آزمون می‌شود تا فرضهای معلم درباره رابطه بین متغیرها به طور علمی اعتباریابی شوند.

تحقیق در مورد فرضیه‌های استقرایی که از مشکلات روزمره استنتاج شده‌اند، اغلب می‌تواند در تعیین راه حل یک چنین مسائلی مفید باشد. به این دلیل که این گونه فرضیات از مسائل مشخص عملی سرچشمه می‌گیرند، به هر حال نتایج حاصل از آنها می‌تواند مجموعه‌ای از یافته‌های پرارزش، ولی ناپایسته را فراهم آورد که از نظر توان توصیفی، محدودیت دارند.

// فرضیه‌های قیاسی

بر خلاف فرضیاتی که با تعمیم از رابطه‌های مشاهده شده صورتبندی می‌شوند، فرضیات قیاسی از طریق قیاس از نظریه ساخته می‌شوند. اینگونه فرضیات دارای این مزیت هستند که می‌توانند به سیستم کلی تری از دانش، به مثابه چهارچوبی برای تلفیق معنادار آنها در پیکره معرفتی که از پیش وجود دارد، منتج شوند. اگر هر بررسی در حد یک کوشش مجزا باقی بماند، علم نمی‌تواند به طور مؤثری پیشرفت کند. یک علم با اتکاء به پیکره موجود واقعیتهای و نظریه‌ها می‌تواند به انباشتگی برسد. یک فرضیه استنتاج شده از نظریه به عنوان یک فرضیه قیاسی شناخته می‌شود.

شاید تشخیص بین نظریه و فرضیه مفید باشد. یک نظریه شامل مجموعه‌ای از مفاهیم، به همراه احکامی است که به چگونگی و چرایی روابط درونی موجود بین این مفاهیم می‌پردازد. نظریه‌ها به صورت تبیینهای کلی ارائه می‌شوند که در دامنه وسیعی از پدیده‌ها به کار می‌روند. در یک نظریه بیان روابط اغلب چیزی بیشتر از فرضیه‌هایی نیست که درست انگاشته شده‌اند؛ فرضیاتی که می‌توانند در بررسیهای بعد مورد آزمون قرار گیرند. فرد می‌تواند بر مبنای روابط درونی پیشنهادی در یک نظریه، پیامدهای خاص را که به طور منطقی می‌توان فرض کرد از آن روابط حادث می‌شوند، بیان کند. این پیامدهای فرض شده پایه‌هایی برای فرضیات هستند. یک نظریه علمی باید متضمن نتایجی باشد که بتواند در جریان یک تحقیق تجربی تأیید شوند، یعنی فرد باید بر اساس نظریه قادر به پیش‌بینی رویدادهای خاصی باشد که قابل مشاهده یا غیر قابل مشاهده هستند. این پیامدهای قیاسی تبدیل به فرضیه‌هایی می‌شوند که موضوع بررسیهای تجربی هستند.

هنگامی که فرضیات مشتق شده از یک نظریه در تحقیقی مورد تأیید قرار می‌گیرند، به همان ترتیب نظریه نیز مورد تأیید قرار می‌گیرد. بنابراین، فرضیات شواهدی فراهم می‌کنند که نظریه‌ها را تأیید می‌کنند، بسط می‌دهند، نقض می‌کنند و یا به تجدید نظر در نظریه‌هایی منجر می‌شوند که از آن نشئت گرفته‌اند. باید بر این امر تأکید کرد که تحقیق هر چند می‌تواند نظریه‌ای را رد کند، ولی هرگز قادر به اثبات درستی یک نظریه نیست. دلیل این امر آن است که نظریه‌ها تعمیم‌هایی هستند که برای تمام موارد ممکن پدیده‌ای که سعی در تبیین آن دارند، به کار می‌روند و این امکان وجود ندارد که نظریه را در مورد تمام احتمالات، مورد آزمون قرار داد. اما، یک نظریه هر قدر تأیید بیشتری در تحقیقات مختلف به دست آورد، ما را بیشتر مطمئن می‌کند که معتبر و مفید است.

نظریه‌ها از نظر محتوا نسبت به فرضیه کلی‌تر هستند. یک نظریه ممکن است منشأ تعدادی فرضیه برای تحقیق در بررسیهای متعدد و مجزا باشد. یک محقق ممکن است مطالعه‌ای را با انتخاب یکی از نظریه‌های موجود در حیطه مورد علاقه‌اش آغاز کند. البته انتخاب نظریه طبق هدف تحقیق و سهمی که نظریه می‌تواند در فهم مسئله ایفا کند، صورت می‌گیرد. هنگامی که نظریه انتخاب شد، محقق برای آنکه به پیامدهای منطقی نظریه برسد، با استفاده از استدلال قیاسی، از آن نظریه، فرضیه استخراج می‌کند. این استدلال قیاسی سپس تبدیل به فرضیات بررسی می‌شوند. محقق که به بررسی انگیزه پیشرفت در بین کودکان علاقه‌مند است، باید نظریه کلاسیک مک کلند در مورد انگیزه پیشرفت را مطالعه کند.

یکی از فرضهای نظریه مک کلند در باب انگیزه این است که شدت انگیزه پیشرفت تابعی است که به طور مستقیم با آموزش استقلال و خودکفایی تناسب دارد.^(۱) محقق از این نظریه می‌تواند پیش‌بینی کند کودکانی که والدینشان استقلال و خودکفایی را در آنها ترغیب می‌کنند، در کلاس انگیزه بیشتری برای پیشرفت از خود بروز می‌دهند. محقق سپس برای رسیدن به پیامدهای منطقی نظریه که می‌تواند به طور تجربی به اثبات برسد، از استدلال قیاسی استفاده می‌کند. در بررسی تحقیقی فرضی زیر، این استنتاج قیاسی، فرضیه محسوب می‌شود. مسئله و فرضیه می‌توانند به شکل زیر بیان شوند:

مسئله: چه رابطه‌ای بین شیوه‌های پرورش کودکان در منزل و انگیزه پیشرفت در بین دانش‌آموزان دوره دبستان وجود دارد؟

فرضیه: کودکانی که والدینشان بشدت آزادی آنها را محدود می‌کنند، در مقایسه با کودکان دیگری که والدینشان آزادی آنها را محدود نمی‌کنند، در یک تکلیف درسی، که مقدار کار انجام شده برای آن تابعی از انگیزش آنهاست، نمره‌های پایین‌تری کسب می‌کنند. محقق سپس برای آزمون این فرضیه به طراحی یک بررسی، و به احتمال قوی از نوع علی - مقایسه‌ای، اقدام می‌کند. اگر نتایج پیش‌بینی شده مشاهده شوند، نظریه مک کللند در مورد انگیزهٔ پیشرفت مورد تأیید قرار می‌گیرد. اگر هیچ تفاوتی در دو گروه از کودکان مشاهده نشود، محقق باید نظریه یا حداقل کاربرد آن در این شرایط را مورد سؤال قرار دهد.

محققی که به رشد تفکر منطقی در بین کودکان علاقه‌مند است، احتمالاً به نظریه پیاژه مراجعه می‌کند. پیاژه عقیده داشت که کودکان در جریان رشد ذهنی خود از مراحل مختلفی می‌گذرند. یکی از مراحل، مرحله عملیات عینی است که از ۷ یا ۸ سالگی آغاز می‌شود و مشخصه آن انتقال از مرحله وابستگی به ادراک، به مرحله بکارگیری برخی عملیات منطقی است. این عملیات در یک سطح عینی قرار دارند، ولی استدلال نمادین را نیز دربردارند.^(۱) با استفاده از این نظریه به عنوان یک نقطهٔ شروع، محقق می‌تواند این فرضیه را ارائه کند: نسبت کودکان ۹ ساله‌ای که قادر به پاسخگویی درست به این مسئله از نوع استدلال انتقالی هستند که «فرانک بلندتر از جورج است؛ جورج کوتاه‌تر از رابرت است؛ چه کسی از همه بلندتر است؟» از نسبت کودکان ۶ ساله‌ای که قادر به پاسخگویی درست به این سؤال هستند، بیشتر است.

در مطالعه‌ای که برای آزمون قیاس از یک نظریه طراحی شده است، وارسی هر یک از فوایل منطقی حایل بین نظریه و فرضیه از اهمیت فوق‌العاده‌ای برخوردار است. محقق باید از خود سؤال کند که آیا فرضیه به طور منطقی از نظریه پیروی می‌کند؟ اگر فرضیه به طور حقیقی از نظریه پیروی نکند، محقق قادر به دستیابی به نتایج معتبر در مورد شایستگی نظریه نخواهد بود. اگر فرضیه مورد تأیید قرار گرفته باشد، ولی به دقت از نظریه قیاس نشده باشد، محقق نمی‌تواند بگوید یافته او اعتباری برای نظریه فراهم آورده است. بر عکس، اگر داده‌ها فرضیه را تأیید نکردند، نظریه‌ای که این فرضیه از آن نشئت گرفته است، لزوماً نامعتبر نخواهد بود. این یک حقیقت است که بسیاری از فرضیات که می‌توان آنها را از نظریه‌های شناخته شده تر قیاس کرد،

تاکنون مورد بررسی قرار گرفته‌اند، ولی بسیاری از چنین قیاس‌هایی هنوز ساخته و آزمون نشده‌اند. همچنین یک قیاس که در گذشته مورد تحقیق قرار گرفته است، می‌تواند برای ایجاد فرضیه‌هایی در شرایط وسیعتر و متنوعتر و برای وسعت بخشیدن به کاربرد نظریه مورد استفاده قرار گیرد.

/// ویژگی‌های فرضیه‌های قابل استفاده

بعد از آنکه فرضیه به طور آزمایشی صورتبندی شد و قبل از هرگونه تلاشی برای آزمون تجربی آن، توانایی فرضیه به عنوان ابزار تحقیق باید سنجیده شود. یک فرضیه باید معیارهای ویژه‌ای را دارا باشد. ارزش نهایی یک فرضیه نمی‌تواند قبل از آزمون تجربی آن مورد قضاوت قرار گیرد، اما معیارهای ویژه‌ای وجود دارند که فرضیه‌های ارزشمند را مشخص می‌سازند و محقق باید برای قضاوت در مورد شایستگی فرضیه پیشنهادی از آنها استفاده کند.

// یک فرضیه باید از قدرت تبیین برخوردار باشد

فرضیه باید امکان تبیین آنچه را که سعی در توصیف آن دارد، داشته باشد. این یک معیار واضح، ولی مهم است. برای روشن شدن مسئله، فرض کنید که شما سعی دارید اتومبیل خود را روشن کنید، ولی موفق نمی‌شوید. این فرضیه که اتومبیل روشن نمی‌شود چون شما شیر آب حمام را باز گذاشته‌اید، قابل قبول نیست، اما فرضیه‌ای که می‌گوید باطری اتومبیل تمام شده است، یک تبیین قابل قبول است و ارزش آزمون شدن را دارد.

// یک فرضیه باید رابطه مورد انتظار بین متغیرها را بیان کند

یک فرضیه باید رابطه مورد انتظار بین دو یا چند متغیر را گمانه زند. در مثال قبل، بیان اینکه «اتومبیل روشن نخواهد شد و یک سیستم سیم‌پیچی دارد»، به این دلیل که هیچ رابطه‌ای بین متغیرها مشخص نشده است و از اینرو هیچ رابطه‌ای پیشنهادی برای آزمون وجود ندارد، مفید نخواهد بود. یک فرضیه سودمند این خواهد بود: اتومبیل روشن نخواهد شد، به این دلیل که در سیستم سیم‌پیچی آن اشکالی وجود دارد. این معیار به طور آشکار روشن است، ولی تبیین زیر را ملاحظه کنید: اگر کودکان در خودانگاره با یکدیگر متفاوت باشند، از نظر پیشرفت در درس علوم

اجتماعی نیز با یکدیگر متفاوت خواهند بود. از نگاه اول به نظر می‌رسد که این جمله یک فرضیه باشد، اما با کمی دقت روشن می‌شود که در آن هیچ رابطه مورد انتظاری بیان نشده است. یک رابطه مورد انتظار می‌توانست این گونه بیان شود: خودانگاره والتر احتمالاً پیش‌بینی برای موفقیت در درس علوم اجتماعی است. این فرضیه را می‌توان به این نحو نیز بیان کرد: بین خودانگاره و موفقیت در درس علوم اجتماعی یک رابطه مثبت وجود دارد. اگر خلاف آن پیش‌بینی شده باشد، یعنی خودانگاره والتر به موفقیت کمتر در درس علوم اجتماعی منجر شود، فرضیه این خواهد بود: بین خودانگاره و موفقیت در درس علوم اجتماعی یک رابطه منفی وجود دارد. هر دو بیان دومین معیار ما را برآورده می‌سازند.

// یک فرضیه باید قابل آزمون باشد

گفته شده است که مهمترین ویژگی یک فرضیه «خوب»، داشتن قابلیت آزمون است. این بدان معنی است که فرضیه قابل اثبات است. بر این اساس است که قیاسها، نتیجه‌گیریها یا معانی ضمنی را به نحوی فراهم می‌آورند که مشاهدات تجربی، فرضیه را رد و یا تأیید کنند. اگر فرضیه درست باشد، پس باید نتایج قابل پیش‌بینی متجلی شوند. یک فرضیه قابل آزمون محقق را قادر می‌سازد تا از طریق مشاهده تعیین کند که پیامدهایی که به طور قیاسی استدلال شده است، در عمل رخ خواهند داد یا خیر. در غیر این صورت تأیید و یا رد فرضیه غیر ممکن خواهد بود. در مثال مایین فرضیه که «روشن نشدن ماشین، کیفری است برای گناهان من»، در دنیای ما به طور آشکار قابل آزمون نیست.

بسیاری از فرضیات و یا گزاره‌هایی که ممکن است به طور بنیادین نیز بیان شده باشند، اساساً قابل آزمون نیستند. برای مثال فرضیه «تجربیات پیش دبستانی، سازگاری همه جانبه کودکان را در دوره پیش از دبستان ارتقاء می‌بخشد»، به دلیل مشکل بودن تعریف و اندازه‌گیری مفهوم سازگاری همه جانبه، بسیار مشکل خواهد بود. مثال دیگر در مورد یک فرضیه غیر قابل آزمون این است: شیوه تقلید از آثار مشهور در درس هنر مدرسه، باعث خاموشی خلاقیت هنری دانش‌آموزان می‌شود. در این مورد یکی از مشکلات، تعریف و اندازه‌گیری خلاقیت هنری و همچنین تعیین معیار برای تشخیص وقوع خاموشی در خلاقیت است.

یک فرضیه قابل آزمون، باید متغیرهایی را به هم ربط دهد که قابلیت اندازه‌گیری داشته

باشند. اگر هیچ وسیله‌ای برای اندازه‌گیری متغیرها موجود نباشد، گردآوری داده‌های ضروری برای آزمون اعتبار فرضیه، غیر ممکن خواهد بود. در این مورد نمی‌توان تاکید بسیار کرد. بدون امکان تعریف دقیق شاخصهای هر متغیر و در پی آن اندازه‌گیری این متغیرها، فرضیه قابل آزمون نخواهد بود.

از شاخصهای متغیرها به عنوان تعریف کاربردی یاد می‌شود. تعریف کاربردی، همانطور که پیش از این توضیح داده شد، یک متغیر را با تبیین «عملیات» یا روشهای لازم برای اندازه‌گیری آن، مشخص می‌کند. برای مثال، به این فرضیه توجه کنید: بین عزت نفس^(۱) در کودک و پیشرفت در امر خواندن در کلاس اول دبستان رابطه مثبت وجود دارد. برای آنکه این فرضیه از معیارهای قابلیت پذیرش بهره‌مند شود، تعریف کاربردی متغیرها ضروری است. عزت نفس را می‌توان به عنوان نمره‌های به دست آمده از مقیاس عزت نفس کوپر سمیت^(۲) و پیشرفت در امر خواندن را به عنوان «نمره‌های به دست آمده از آزمون خواندن کالیفرنیا» و یا رتبه‌بندی معلم کلاس اول در پیشرفت خواندن تعریف کرد.

یکی از ملاحظات اساسی در صورتبندی یک فرضیه، اطمینان یافتن از این امر است که می‌توان متغیرها را به طور کاربردی تعریف کرد. از بکارگیری سازه‌ها در تعریف متغیرها بپرهیزید؛ چراکه یافتن اندازه‌های مناسب برای آنها مشکل و یا غیرممکن خواهد بود. توافق بر روی تعریفی کاربردی از سازه‌هایی مانند خلاقیت، استبدادگرایی، دموکراسی و امثالهم، از آنجا که این مفاهیم معانی متعددی به خود می‌گیرند، اگر غیرممکن نباشد، مشکل است. به خاطر داشته باشید که اصطلاحات باید به رفتار عینی قابل شناسایی اطلاق شوند.

پرهیز از تبیینات ارزشی در قالب یک فرضیه، امر مهمی است. عبارتی مانند «یک برنامه مشاوره‌ای در مدارس ابتدایی مطلوب است»، نمی‌تواند در یک بررسی تحقیقی مورد مطالعه قرار گیرد؛ هر چند که محقق می‌تواند این فرضیه را مورد آزمایش قرار دهد: دانش‌آموزان کلاس اول که از مشاوره برخوردار شده‌اند، به طور کلامی نسبت به دانش‌آموزانی که فاقد این امکان بوده‌اند، رضایت بیشتری از مدرسه ابراز خواهند کرد. محقق می‌تواند تعبیرات کلامی ناشی از رضایت را اندازه‌گیری کند، اما اینکه آیا آنها مطلوب هستند یا نه، یک قضاوت شخصی است.

1- Self- Esteem

2- Coopersmith, S. (1981). The Antecedents of Self- Esteem. San Francisco: Consulting Psychology.

// یک فرضیه باید با پیکره دانش موجود سازگار باشد

فرضیه‌های جدید نباید ناقض فرضیه‌ها، نظریه‌ها و قوانین تثبیت شده قبلی باشند. فرضیه‌ای نظیر «ماشین من روشن نخواهد شد، چون آب باطری آن تبدیل به طلا شده است»، اگر چه معیار اول را داراست، ولی چنان با آگاهی از طبیعت مواد تناقض دارد که ارزش پیگیری ندارد. فرضیه «ماشین روشن نخواهد شد، چون آب باطری آن تا پایین‌ترین سطح تبخیر شده است»، با دانش موجود سازگار است و بنابراین ارزش پیگیری را دارد. فرض اینکه بین خود انگارهٔ پسران و دختران نوجوان و میزان رشد بدنی آنها رابطه‌ای وجود ندارد، بی‌فایده است؛ چرا که انبوه شواهد چنین رابطه‌ای را تأیید می‌کنند. در تاریخ علم، مردانی مانند انیشتین، نیوتن، داروین و کپرنیک را مشاهده می‌کنیم که حقیقتاً فرضیاتی انقلابی را ارائه کردند که با دانش آن عصر مغایرت داشته است، ولی باید به یاد داشت که کار چنین پیشتازانی تکذیب کامل دانش گذشته نبوده، بلکه سازماندهی مجدد آن، در نظریه‌های اقناع‌کننده‌تری بوده است. در بیشتر موارد و بخصوص برای محقق تازه کار، بهترین پیشنهاد این است که فرضیه، با آخرین آگاهی‌های حاصله در زمینه مورد نظر منطبق باشد. دوباره بر ضرورت بررسی کامل مآخذ و منابع موجود تأکید می‌کنیم، چون فرضیه‌ها بر اساس نتایج تحقیقات گزارش شده در زمینه مورد نظر صورتبندی خواهند شد.

// فرضیه‌ها باید تا حد امکان ساده و موجز بیان شوند

بیان فرضیه به روشی ساده نه تنها آزمون آنرا بسیار آسان‌تر می‌کند، بلکه به هنگام نتیجه‌گیری نیز مبنایی برای ارائه گزارشی روشن و قابل درک فراهم می‌آورد. گاهی ممکن است لازم باشد که یک فرضیه کلی و گسترده را برای آنکه قابل آزمون و واضح باشد، به چند فرضیه کوچک مشخص تقسیم کنیم. برای مثال توما و لیوسن این فرضیه‌های کلی را مورد توجه قرار دادند: وضعیت اقتصادی - اجتماعی خانواده در تعیین میزان سازگاری نوجوان در موقعیتهای مختلف اجتماعی نقش دارد، یا اجزای گوناگون این وضعیت اجتماعی از نظر میزان تأثیر بر نگرش نوجوان نسبت به اقتدار تفاوت می‌کنند. توما و لیوسن برای آنکه این فرضیه‌های کلی را واضح و قابل آزمون سازند، آنها را به چند فرضیه مشخص تقسیم کردند: (۱) بین نگرش یک پسر نوجوان نسبت به اقتداری که در خانه تجربه می‌کند و وضعیت اقتصادی - اجتماعی خانواده‌اش، یک رابطه معنی‌دار منفی وجود دارد؛ (۲) بین نگرش یک پسر نوجوان نسبت به اقتداری که در

مدرسه تجربه می‌کند و موقعیت اقتصادی - اجتماعی خانواده‌اش یک رابطه معنی‌دار منفی وجود دارد، (۳) بین نگرش یک پسر نوجوان نسبت به اقتداری که در بین همتایانش تجربه می‌کند و وضعیت اقتصادی - اجتماعی خانواده‌اش یک رابطه معنی‌دار منفی وجود دارد، (۴) بین نگرش پسران نوجوان نسبت به اقتدار و سطح تحصیلات پدران‌شان یک رابطه معنی‌دار منفی وجود دارد؛ (۵) بین نگرش پسران نوجوان نسبت به اقتدار و سطح تحصیلات مادران‌شان یک رابطه منفی معنی‌دار وجود دارد و غیره.^(۱) برای انعکاس نگرشهای دختران نوجوان ضروری است که این فرضیه‌ها از نو بیان شوند. بیان فرضیات در این شکل بسیار مشخص و جزئی، گزارش یافته‌ها و نتایج را نیز تسهیل می‌کند. فرد آزمایشگر می‌تواند هر فرضیه را به طور جداگانه در نظر بگیرد و یافته‌ها و نتایج مربوط به هر یک را ارائه کند.

بنابراین، چنین به نظر می‌رسد که ممکن است در یک مطالعه نیاز به فرضیه‌های متعددی باشد. عموماً توصیه می‌شود که محقق برای هر یک از وجوه فرعی مسئله و یا هر ابزار گردآوری داده‌ها که به کار می‌برد، یک فرضیه بیان کند. برای مثال، یک محقق ممکن است این فرضیه را بیان کند: دانش‌آموزانی که ریاضی را از طریق PLATO (یک نرم افزار کامپیوتری آموزشی) آموخته‌اند، در مفاهیم ریاضی، یادگیری و یادداری بیشتری را نسبت به دانش‌آموزانی نشان خواهند داد که با کتابهای سنتی و متداول آموزش می‌بینند. از آنجا که گزارش نتایج، هم برای یادگیری و هم برای یادداری ضروری خواهد بود، فرضیه کلی باید به مثابه دو فرضیه مجزا مورد آزمون قرار گیرد. این فرضیات را می‌توان اینگونه بیان کرد: (۱) دانش‌آموزانی که ریاضیات را از طریق PLATO یاد گرفته‌اند، در مفاهیم ریاضی یادگیری بیشتری را نسبت به دانش‌آموزانی نشان خواهند داد که از طریق کتابهای سنتی و متداول آموزش دیده‌اند، و (۲) دانش‌آموزانی که ریاضیات را از طریق PLATO یاد گرفته‌اند، یاد داری بیشتری را نسبت به دانش‌آموزانی نشان خواهند داد که از طریق کتابهای سنتی و متداول آموزش دیده‌اند. به این ترتیب، امکان نشان دادن این مسئله که داده‌های به دست آمده در هر یک از موارد مجزا سؤال کلی را تأیید می‌کنند، وجود دارد. ممکن است داده‌ها تأثیر آموزش کامپیوتری را در یادگیری اولیه نمایان کنند، اما در مورد یادداری آنها چیزی نشان ندهند. محقق نباید از اطناب کلامی حاصل از بیان چندین فرضیه هراس داشته باشد. به یاد داشته باشید که دو اصل قابلیت آزمون و وضوح، با کمک فرضیه‌های

مشخص تر به نحو بهتری تأمین می شود.

همچنین توصیه می شود که در تبیین فرضیه از ساده ترین واژه ها و اصطلاحات استفاده شود تا مفهوم مورد نظر انتقال یابد، از بکارگیری سازه های مبهم باید پرهیز شود. اصطلاحات را به شیوه ای به کار برید که به طور کلی برای اشاره به پدیده مورد نظر، قابل قبول باشند. وقتی دو فرضیه دارای قدرت توصیف یکسانی باشند، آنکه ساده تر است، باید مرجح باشد؛ زیرا تبیین ضروری را با مفروضه ها و متغیرهای کمتری که نیاز به تعریف داشته باشند، ارائه می کند. به یاد داشته باشید که اصل ایجاز در ارزشیابی فرضیه ها مهم است.

بسیاری از فرضیات که صورتبندی شده اند، بعد از آزمون تجربی رد خواهند شد. آنها پیش بینی هایی بوده اند که داده ها آنها را تأیید نکرده اند. در تاریخ تحقیق علمی، تعداد فرضیاتی که مورد تأیید قرار نگرفته اند، بیشتر از آنهایی بوده است که تأیید شده اند. محققان با تجربه عقیده دارند که فرضیات تأیید نشده، یک بخش مفید و مورد نظر در تجربیات علمی هستند. آنها قادرند تحقیق را به سمت بازبینی مجدد نظریه سوق دهند و اغلب ما را به توصیفی دقیق از وضعیت امور نزدیکتر کنند. داروین گفته است: «من دائماً تلاش کرده ام تا مغزم را از هر گونه فرضیه رها کنم، هر چند بسیار به آن علاقه مند بوده ام (و نمی توانم در مقابل ساختن آن در مورد هر موضوعی مقاومت کنم)، اما واقعیات سرعت خلاف این را نشان دادند. براستی هیچ چاره ای جز حرکت در این راستا نداشتم. به استثنای فرضیه صخره های مرجانی، من هیچ فرضیه جدید مجرد را به یاد نمی آورم که بعد از مدتی یا کاملاً کنار گذاشته نشود یا کاملاً تغییر نیابد.»^(۱) حتی یک فرضیه تأیید نشده نیز می تواند مفید باشد؛ زیرا ضرورت توجه به سایر جنبه های یک مسئله را مسجل می کند. از این روی می تواند محقق را در رسیدن به توصیفی قابل قبول، یک گام جلوتربرد. در صورتبندی فرضیه باید از ابهام پرهیز شود.

به جز موارد استقرای کامل، حتی اگر محقق تأییدی برای فرضیه خود بیابد، به معنای اثبات درستی آن نخواهد بود. یک فرضیه هرگز اثبات و یا رد شده نیست، فرضیه تنها تأیید می شود و یا تأیید نمی شود. فرضیات در ماهیت خود اساساً احتمالی هستند، شواهد تجربی می توانند محقق را به سوی این نتیجه گیری هدایت کنند که تبیین او احتمالاً حقیقت دارد یا پذیرفتن فرضیه منطقی است، ولی هرگز فرضیه را ثابت نمی کنند. نمونه ای از یک بررسی که فرضیه آن از نظریه قیاس شده است، در اینجا ارائه می شود.

تأثیر ساعت آموزش طی روز در پیشرفت تحصیلی^(۱)

مسئله تحقیق

این که آموزش در چه ساعتی از روز صورت گرفته است، در پیشرفت درس زبان انگلیسی و ریاضیات دانش آموزان کلاس هشتم چه تأثیری دارد؟

نظریه

نظریه‌ای در مورد حافظه کوتاه مدت (STM)^(۲) و حافظه بلند مدت (LTM)^(۳) می‌گوید که تکالیفی که عمدتاً به استفاده از حافظه کوتاه مدت مربوط می‌شوند، در صبح بهتر یاد گرفته و انجام می‌شوند و تکالیفی که عمدتاً به استفاده از حافظه بلند مدت مربوط می‌شوند، در بعدازظهر بهتر یاد گرفته و انجام می‌شوند. دو عامل روان شناختی دلیل تفاوت‌های یادگیری را در اوقات روز بیان می‌کنند: (الف) انگیزتگی پایه از سطح نازل در صبح به بالاترین حد آن در عصر می‌رسد و استفاده از حافظه کوتاه مدت و بلندمدت به طور قابل توجهی با آن همبسته هستند؛ (ب) به همین ترتیب ریتم‌های روزانه در سطوح پلازما و هورمون، در شیوه‌ای که افراد اطلاعات را رمزگشایی، بایگانی و بازیابی می‌کنند، مؤثر است و با پردازش بهتر حافظه بلند مدت در اواخر روز و پردازش بهتر حافظه کوتاه مدت در ساعت اولیه روز همبسته است.

فرضیه

به دلیل آنکه مهارت خواندن، اطلاعات چاپی را با آگاهی و تجربه قبلی مرتبط می‌کند، این فرضیه صورتبندی می‌شود: دانش آموزان هنگامی در خواندن پیشرفت خواهند کرد که آموزش آنها در بعدازظهر انجام شود. و چون آموزش مهارت‌های ریاضی نیازمند دستکاری بیشتر داده‌ها در ذهن و استفاده از حافظه کوتاه مدت است، فرضیه اینگونه صورتبندی می‌شود: دانش آموزان هنگامی در ریاضی پیشرفت بیشتری خواهند کرد که آموزش آنها به جای عصر، در صبح صورت گیرد.

1- Davis. Z.T.(1988). The Effect of Time- of- day of Instruction on Eight-Grade Student' English and Mathematics Achievement. The High School Journal, 71, 78- 80.

2- Short- Term Memory

3- Long- Term Memory

روشها

هشتاد دانش آموز کلاس هشتم به طور تصادفی در گروههای تیمار ذیل مشخص شدند: (۱) دوره اول انگلیسی، (۲) آخرین دوره انگلیسی، (۳) دوره اول ریاضی، (۴) آخرین دوره ریاضی. یک معلم انگلیسی واحد و یک معلم ریاضی واحد، بخشهای دوره اول و دوره آخر درسها را با استفاده از روشها و مواد درسی یکسان آموزش دادند. یک آزمون جامع مهارتهای پایه‌ای (CTBS)^(۱) در شروع (پیش آزمون) و پایان (پس آزمون) سال تحصیلی برای اندازه‌گیری میزان پیشرفت به اجرا در آمد. آزمونهای مربوط به هر محتوا در یک روز واحد داده شد و به موازات آن در ساعات میانه روز برای کنترل تأثیر بالقوه ساعات روز در میزان بازدهی، به همه گروههای دیگر نیز این آزمون داده شد.

تحلیل

مجموعه نمرات آزمونهای مربوط به انگلیسی و ریاضی هر کدام به وسیله تحلیل کوواریانس^(۲) و با استفاده از نمرات پیش آزمون به عنوان متغیر کمکی، تحلیل شد. تحلیل کوواریانس تغییرات موجود بین نمرات پیش آزمون دانش آموزان را در نظر می‌گیرد و نمرات پس آزمون را در رابطه با آن تطبیق می‌دهد، به گونه‌ای که می‌توان مقایسه‌های معتبری صورت داد.

نتایج

گروه انگلیسی بعد از ظهر به نسبت گروه انگلیسی صبح پیشرفت معنی دارتی داشتند. هیچ تفاوت معنی داری بین پیشرفت گروههای ریاضی صبح و بعد از ظهر وجود نداشت.

نتیجه‌گیری

نتیجه آن که دانش آموزان در درس انگلیسی کلاس هشتم، از آموزش بعد از ظهر، بیش از آموزش صبح بهره می‌گیرند. ساعت آموزش روزانه هیچ تفاوتی در پیشرفت درس ریاضی دانش آموزان کلاس هشتم نداشت. محقق نتیجه گرفت که فرضیه مربوط به درس ریاضی باید مورد بررسی بیشتری قرار گیرد. شاید کارکرد حافظه‌های کوتاه مدت و بلند مدت در تحقیق قبلی، به طور کامل معرف آن چیزی نبود که در استدلال ریاضی مورد نیاز است، یا شاید استدلال ریاضی استفاده نسبتاً متعادل از حافظه‌های کوتاه مدت و بلند مدت را در بردارد.

/// انواع فرضیه

// فرضیه تحقیق

فرضیاتی را که پیش از این مورد بحث قرار دادیم، در نهایت فرضیه تحقیق و یا اساساً فرضیات می‌نامیم. آنها فرضیاتی هستند که از طریق مشاهده، پیشینه تحقیق و یا نظریه‌ای که در تحقیق توصیف شده است، ساخته می‌شوند. فرضیه تحقیق گزاره‌ای است درباره رابطه‌ای که محقق انتظار دارد آن را به عنوان نتیجه تحقیق پیدا کند. فرضیه می‌تواند بیانی درباره رابطه یا تفاوت مورد انتظار بین متغیرهای موجود در بررسی باشد. یک فرضیه درباره هوشبهر و اضطراب کودکان در کلاس درس را می‌توان اینگونه بیان کرد: یک رابطه مثبت بین هوشبهر و اضطراب کودکان در دوره اول دبستان وجود دارد، یا کودکان با هوشبهر بالا نسبت به کودکان با هوشبهر پایین در کلاس درس اضطراب بیشتری از خود نشان می‌دهند. فرضیات تحقیق را می‌توان به شکل جهت‌دار یا غیر جهت‌دار بیان کرد. یک فرضیه جهت‌دار ماهیت رابطه یا تفاوت مورد انتظار را مشخص می‌کند. دو نمونه فرضیه در مورد هوشبهر و اضطراب، که در بالا ذکر شد، از نوع فرضیات جهت‌دار هستند. از سوی دیگر، یک فرضیه غیر جهت‌دار، وجود یک رابطه و یا تفاوت را بیان می‌کند، ولی ماهیت یافته‌ی مورد انتظار را مشخص نمی‌کند. برای مثال، در کودکان بین هوشبهر و اضطراب رابطه وجود دارد.

// فرضیه صفر

فرضیات تحقیق نمی‌توانند از طریق شیوه‌های آماری موجود مستقیماً مورد آزمون قرار گیرند. در بررسیهای تجربی که در آنها آزمونهای آماری باید مورد استفاده قرار گیرند، فرضیه تحقیق باید به نوع دیگری از فرضیه که به فرضیه صفر معروف است (با H_0 نشان داده می‌شود)، تبدیل شود. این فرضیه را از اینرو فرضیه صفر می‌نامند که بیان‌کننده «هیچ تفاوت»، «هیچ اثر» یا «هیچ رابطه‌ای» می‌باشد. یک فرضیه صفر، نقیض آنچه را بیان می‌کند که آزمایشگر انتظار دارد و یا پیش‌بینی می‌کند. یک آزمایشگر ممکن است مشتاق نشان دادن این مسئله باشد که بعد از یک تیمار آزمایشی، دو جامعه آماری میانگین‌های متفاوتی خواهند داشت، اما فرضیه صفر می‌گوید که میانگین‌های جامعه‌های آماری با یکدیگر تفاوت نخواهند داشت.

یک فرضیه صفر به این دلیل به کار می‌رود که محققان را قادر می‌سازد تا یافته‌های خود را با

کمک روشهای آماری در مقابل انتظارات اتفاقی مقایسه کنند. فرضیه صفر فرض می‌کند از آنجایی که تفاوت‌های مشاهده شده صرفاً از روی اتفاق رخ می‌دهند، اصلاً نشان‌دهنده تفاوت‌های واقعی نیستند. آزمونه‌های آماری برای تعیین احتمال صحیح بودن فرضیه صفر به کار می‌روند. اگر آزمونها نشان دهند که احتمال وقوع تفاوت‌های مشاهده شده به طور اتفاقی بسیار کم است، فرضیه صفر به یک فرض بسیار غیر محتمل تبدیل می‌شود. بنابراین، می‌توان آن را به نفع فرضیه مقابل رد کرد. شواهد برای این نتیجه‌گیری غیر قطعی که تفاوت واقعی است، معتبر است. اگر تفاوت‌های مشاهده شده بسادگی ناشی از اتفاق باشند، شواهد معتبر نیست و فرضیه صفر حفظ می‌شود. آزمون یک فرضیه صفر شبیه کار دادستان در یک دادگاه جنایی است. دادستان به منظور اثبات جرم (در سیستم حقوقی ایالات متحده) باید شواهد معتبری فراهم کند تا هیأت منصفه بتواند از و رای تردیدهای منطقی، بی‌گناهی مجرم را رد کنند. برای دادستان این امکان وجود ندارد که به طور قطع جرم را ثابت کند، و به همین ترتیب محقق نیز نمی‌تواند برای یک فرضیه تأیید بی‌چون و چرا به دست آورد. وکیل مدافع، چنین فرض می‌کند که تا وقتی شواهد معتبری مبنی بر مجرم نبودن متهم وجود دارد، وی بی‌گناه است، و تا زمانی که شواهد معتبر خلاف آن را ثابت کند، فرضیه صفر صحیح فرض می‌شود.

اجازه بدهید که به تحقیق فرضی سطوح اضطراب بین کودکان با هوشبهر بالا و پایین برگردیم. فرضیه تحقیق را می‌توان اینگونه بیان کرد: کودکان با هوشبهر بالا نسبت به کودکان با هوشبهر پایین اضطراب بیشتری از خود نشان خواهند داد. فرضیه صفر می‌گوید: سطح اضطراب کودکانی که هوشبهر بالا دارند با سطح اضطراب کودکان با هوشبهر پایین متفاوت نیست. فرضیه صفر از طریق روشهای آماری متعددی مورد آزمون قرار می‌گیرد که در فصل ششم تشریح شده است. اگر به هنگام تکمیل بررسی، داده‌های تجربی نشان دهند که تفاوت بین گروه‌های نمونه بیشتر از آن است که صرفاً ناشی از اتفاق باشد، پس می‌توان فرضیه صفر را رد کرد. رد فرضیه صفر نشان می‌دهد که برابری سطح اضطراب دو گروه از کودکان محتمل نیست. محقق سپس یک فرضیه مقابل (با H_1 نشان داده می‌شود) را می‌پذیرد که اینگونه بیان می‌شود: میانگین نمرات اضطراب در جامعه آماری برابر نیست. گفته می‌شود که این فرضیه مقابل غیر جهت‌دار است، چون نشان نمی‌دهد که متوسط نمرات کدام گروه بیشتر خواهد بود. احتمالاً بیشتر تحقیقات در زمینه تعلیم و تربیت با فرضیات غیر جهت‌دار آغاز می‌شوند. هر چند در

بعضی از موقعیتهای، ممکن است محقق برای بیان یک فرضیه جهت‌دار (همانطور که از نام آن پیداست، تعیین جهت یافته‌های مورد انتظار) یک دلیل صریح داشته باشد. یک نمونه از فرضیه جهت‌دار این است: میانگین نمرات اضطراب کودکان با هوشبهر بالا به نسبت کودکان با هوشبهر پایین بیشتر خواهد بود.

/// آزمون فرضیه

یک بررسی با یک فرضیه تحقیقی آغاز می‌شود، که باید بیانی ساده و روشن، از رابطه مورد انتظار بین متغیرها باشد. پیش از این توضیح دادیم که فرضیه باید قابل آزمون باشد؛ به این معنا که در عمل قابل اثبات باشد. زمانی که محققان از آزمون یک فرضیه صحبت می‌کنند، در هر حال منظورشان فرضیه صفر است. متنها فرضیه صفر را می‌توان از طریق روشهای آماری، مستقیماً مورد آزمون قرارداد. آزمون یک فرضیه شامل این مراحل است:

- ۱- در صورت درستی فرضیه تحقیق، روابطی را که باید مورد مشاهده قرار گیرند، در قالب اصطلاحات کاربردی، بیان کنید.
- ۲- فرضیه صفر را بیان کنید.
- ۳- روش تحقیقی انتخاب کنید که امکان مشاهده و یا انجام آزمایش لازم برای نشان دادن وجود یا عدم این روابط را برای شما فراهم آورد.
- ۴- داده‌های تجربی را گردآوری و تحلیل کنید.
- ۵- معین کنید که آیا شواهد برای رد فرضیه صفر کافی هستند یا خیر.

// یک مثال از آزمون فرضیه

یک مثال ممکن است در روشن‌تر شدن فرآیند آزمون تجربی یک فرضیه کمک کند. فرض کنید که یک محقق، مشتاق بررسی نظریه تقویت در کلاس درس است. معلم بر مبنای درک خود از نظریه تقویت، ممکن است فرض کند که تحسین یا تشویق باعث افزایش انگیزه در بین پاره‌ای از کودکان می‌شود. اگر این فرضیه درست باشد، منطقی خواهد بود اگر فرض کنیم که پیامد نظرات تشویق آمیز معلم بر روی ورقه‌های آزمون شاگردان باید موجب بهبود عملکرد دانش‌آموزان شود. این فرض ایجاب می‌کند که افزایش انگیزه با بهبود نمرات آزمون دانش‌آموزان نمایان شود.

(مرحله اول) این معنای قیاسی را می‌توان اینگونه بیان کرد: اظهار نظر معلم بر ورقه‌ها موجب بهبود عملکرد دانش‌آموزان در آزمون‌ها می‌شود. بین دو متغیر، نظرات معلم و عملکرد دانش‌آموزان، یک رابطه وجود دارد که باید مورد آزمون قرار گیرد.

(مرحله دوم) برای آزمون آماری، فرضیه تحقیقی بالا باید به یک فرضیه صفر تبدیل شود. فرضیه صفر اینگونه بیان می‌شود: نظرات معلم بر روی ورقه‌های دانش‌آموزان به بهبود عملکرد دانش‌آموزان در آزمون‌ها منجر نخواهد شد.

(مرحله سوم) این نوع فرضیه را می‌توان با یک آزمایش، مورد آزمون قرار داد. محقق می‌تواند به طور تصادفی تعدادی از کلاسها را برای بررسی انتخاب کند. در هر کلاس دانش‌آموزان باید به طور تصادفی به دو گروه تقسیم شوند: برای دانش‌آموزانی که در گروه الف قرار دارند، معلم درباره نحوه انجام آزمون نظرات تشویق‌آمیزی را می‌نویسد. این نظرات می‌تواند شامل کلمات تشویق‌آمیزی چون «عالی است»، «کار خود را به همین خوبی ادامه بده» و یا «کار تو پیشرفت داشته است»، باشد. این نظرات نباید با محتوای ورقه دانش‌آموزان یا تصحیح اشتباهات ویژه او مرتبط باشند. در غیر این صورت، بهبود عملکرد دانش‌آموزان را می‌توان به سود بخشی آموزشی این نظرات نسبت داد، نه به افزایش انگیزه. دانش‌آموزانی که در گروه ب قرار دارند، در ورقه آزمون خود هیچ اظهار نظری دریافت نمی‌کنند.

(مرحله چهارم) آموزگار، یک آزمون عینی که بخش معینی از محتوا را در بر می‌گیرد، به اجرا در می‌آورد. ورقه‌ها نمره‌گذاری می‌شوند و تیمار آزمایشی، به شیوه‌ای که در بالا توصیف شد، ارائه می‌شود. سپس معلم آزمون دوم را که از نظر سطح دشواری همانند آزمون اول و شامل محتوایی است که بعد از آزمون اول و تیمار آزمایشی تدریس شده است، به اجرا در می‌آورد. تغییر نمرات هر دانش‌آموز از آزمون اول به آزمون دوم باید مشخص شود و میانگین این تغییر برای گروه نیز به دست آید. سپس تعیین اینکه آیا متوسط افزایش نمره‌ها در آزمون دوم با تیمار آزمایشی (نظرات معلم بر روی ورقه‌ها رابطه دارد، از طریق تجزیه و تحلیل داده‌ها امکان‌پذیر خواهد بود. (مرحله پنجم) اگر این نتیجه به دست آید که دانش‌آموزانی که نظر معلم را دریافت کرده‌اند (گروه الف)، مشخصاً نمرات بالاتری از گروهی که این نظرات را دریافت نکرده‌اند (گروه ب)، به دست آورده‌اند، آن وقت است که می‌توان فرضیه صفر را رد کرد. رد فرضیه صفر به این معنا خواهد بود که اثرات دو روش همسان نیستند. محقق می‌تواند به طور غیرقطعی نتیجه بگیرد که نظرات معلم بر ورقه دانش‌آموزان منجر به بهبود عملکرد دانش‌آموزان در آزمون‌ها می‌شود.

// بررسی مقدماتی^(۱)

قبل از آماده‌سازی طرح تحقیق، آزمایش کردن روشهای پیشنهادی بر روی تعداد محدودی از آزمودنیها می‌تواند مفید واقع شود. این اجرای آزمایشی یا بررسی مقدماتی بیش از همه به محقق کمک می‌کند تا در مورد عملی بودن و ارزش ادامه بررسی تصمیم‌گیری کند، این آزمایش امکان ارزیابی متناسب و عملی بودن وسایل گردآوری اطلاعات را فراهم می‌آورد. این آزمایش همچنین امکان آزمون مقدماتی فرضیهی را فراهم می‌کند و دلایلی جهت تثبیت و یا اصلاحات بعدی مورد نیاز فرضیه ارائه می‌کند.

بررسی مقدماتی، همچنین کفایت روشهای تحقیقی و اندازه‌های انتخاب شده برای متغیرها را نشان می‌دهد. مسائل پیش‌بینی نشده می‌توانند در این مرحله حل شوند، در نتیجه از صرف وقت و انرژی در مراحل بعدی جلوگیری به عمل می‌آید. یک بررسی مقدماتی فرصت مقتنمی است که بخصوص برای محقق تازه کار توصیه می‌شود.

/// برنامه تحقیق

بعد از آنکه سؤال و فرضیه صورتبندی شدند، محقق برای تکمیل برنامه تحقیق آزمایشی خود آماده است. این نیاز وجود دارد که هدف تحقیق و چگونگی برنامه‌ریزی برای رسیدن به آن با جزئیات نوشته شود. تهیه برنامه تحقیق امری اساسی است و محقق را وادار می‌سازد که ایده‌های خود را به شیوه‌ای عینی بنویسد. بسیاری از ایده‌های اولیه تا وقتی که فرد آنها را بر روی کاغذ نیاورده است، نویدبخش هستند، اما بعد از آن که بر روی کاغذ آمدند، اشکالات و نارساییهای آنها آشکار می‌شود.

برنامه تحقیق مدون را می‌توان برای دریافت نظرات و انتقادات به دیگران ارائه کرد. یافتن نقایص و خطاها در یک برنامه مدون بسیار ساده‌تر از زمانی است که بخواهیم آنرا از طریق ارتباط کلامی بیان کنیم.

در این مرحله برنامه تحقیق نوعاً تنها یک طرح مقدماتی است و تا نگارش طرح نهایی و رسمی احتمالاً نیاز به تغییرات زیادی خواهد داشت. به هر حال بهتر است به خاطر داشته باشیم که هر چه طرح پیشنهادی اولیه کاملتر و با جزئیات بیشتر بیان شود، برای محقق مفیدتر است و از اتلاف وقت او نیز جلوگیری می‌کند. برنامه تحقیق شامل قسمتهای ذیل است.

// مسئله

برنامه با یک بیان روشن از سؤالی که محقق قصد پاسخگویی به آنرا دارد، آغاز می‌شود. بیان مسئله در مورد روابط بین متغیرها سؤال و جامعه آماری مورد نظر را مشخص می‌کند. یک شرح کوتاه از زمینه نظری مسئله و پیشینه تحقیقات نیز باید در این بخش گنجانده شود.

// فرضیه

سؤال با تبیینی موجز از فرضیه و یا فرضیاتی که باید مورد آزمون قرار گیرند، دنبال می‌شود. فرضیه جهت تحقیق را روشن می‌سازد. همه طرحهای آتی برای پروژه تحقیقی به بیان فرضیه بستگی دارند. همچنین بیان فرضیه و دلایل تأیید آن با روشنی هر چه تمامتر از طرف محقق ضروری است. این بخش از برنامه باید شامل تعاریف کاربردی از متغیرهای بررسی باشد.

// طرح تحقیق

مرحله بعدی برنامه، شامل توصیفی در مورد طرح تحقیق - یعنی توصیفی از روشهایی است که در آزمون فرضیه دنبال می‌شوند. انتخاب یک روش مناسب برای آزمون فرضیه بسیار مهم است. یک سؤال آزمایشی را نمی‌توان از طریق روشهای توصیفی پاسخ گفت و همچنین عکس آن نیز صادق است. این بخش همچنین باید شامل فهرستی از اندازه‌ها یا وسایل مورد استفاده در گردآوری داده‌ها باشد. محققان باید آزمونها، مقیاسها و سایر وسایل موردنیاز برای اندازه‌گیری متغیرها را تعیین کنند و اعتبار و پایایی این ابزار را بسنجند. هدف، انتخاب اندازه‌هایی است که تا حد امکان عینی و پایا باشند، بدون آنکه شایستگی آنها، یعنی متناسب بودن آنها با مفاهیمی که قرار است معرف آن باشند، قربانی شود.

// نمونه

برنامه باید شامل توصیفی از جامعه آماری مورد نظر، یعنی نوع آزمودنیهای مورد مطالعه باشد. محقق باید توجه خود را به چگونگی دسترسی به این آزمودنیها نشان داده باشد. همچنین ضروری است که روشهای نمونه‌گیری خود را توصیف کند. جامعه آماری که از آن نمونه‌گیری می‌شود و علاوه بر آن شیوه‌هایی که در انتخاب نمونه دنبال خواهد شد و اندازه پیشنهادی نمونه نیز باید مشخص شود.

// تجزیه و تحلیل آماری

این بخش شامل خط مشی محقق برای تجزیه و تحلیل آماری داده‌هاست. پیش از آنکه محقق شروع به گردآوری داده‌ها کند، باید روشهای آماری را که امکان پاسخگویی به سؤال تحقیق و یا آزمون فرضیه را فراهم می‌آورند، تعیین کند. در مرحله نخست محقق نیازمند توصیف یا خلاصه کردن داده‌های گردآوری شده از نمونه مورد بررسی است. سپس باید بتواند پایایی یا دقت استنباطها و تعمیم‌هایی را که از یافته‌های مربوط به گروه نمونه به دست آمده و به کل جامعه آماری عمومیت داده شده است، تخمین بزند. روشهای آماری به هر دوی این ویژه کارها کمک می‌کند. کار خلاصه کردن داده‌های به دست آمده از طریق آمار توصیفی صورت می‌گیرد. آمار استنباطی محقق را قادر می‌سازد تا با کمک داده‌های گروه نمونه، دست به استنتاج بزند.

بسیاری از محققان با تجربه، و همچنین افرادی که به تازگی فرایند تحقیق را آموخته‌اند، مشورت با یک کارشناس آمار را قبل از تکمیل طرح بررسی خود ضروری می‌یابند. در فصل ششم بحث کوتاهی در مورد نقش تجزیه و تحلیل آماری در آزمون فرضیه، ارائه خواهد شد.

/// خلاصه

برای رسیدن به مراحل منطبق با بررسی تحقیقی، داشتن فرضیه‌هایی که بروشنی بیان شده‌اند، دارای اهمیت است. فرضیه، پیش‌بینی محقق در مورد نتیجه بررسی است. فرضیه‌ها به شیوه استقرایی از مشاهده و یا به شیوه قیاسی از یک نظریه شناخته شده استنتاج می‌شوند. تجربه، دانش و آشنایی با تحقیقات قبلی در حیطه موردنظر در صورتبندی یک فرضیه رضایت‌بخش، عوامل مهمی به شمار می‌روند.

فرضیه در تحقیق، به عنوان عملکردی چند منظوره ایفای نقش می‌کند؛ زیرا توصیفی ارائه می‌کند که به طور تجربی قابل آزمون است و به گسترش دانش خدمت می‌کند. فرضیه با مشخص کردن روش تحقیق و نوع داده‌های مرتبط با حل مسئله، به تلاشهای محقق جهت می‌دهد. همچنین فرضیه چهارچوبی برای تفسیر و تبیین نتایج بررسی مهیا می‌سازد.

یک فرضیه خوب، باید این معیارهای مشخص را دارا باشد: (۱) باید قدرت توصیف داشته باشد، (۲) باید قابل آزمون باشد؛ به عبارتی باید متغیرهایی را با هم مرتبط کند که بتوان آنها را

اندازه‌گیری کرد، (۳) باید با داده‌های مسلط موجود سازگار باشد، و (۴) باید بروشنی و ایجاز هر چه تمام‌تر بیان شود.

بعد از صورتبندی یک فرضیه، مرحله بعدی تدوین یک برنامه برای انجام بررسی، شامل بیان مسئله و فرضیه و توصیف طرح تحقیق، گروه نمونه و تجزیه و تحلیل آماری مورد استفاده است. این برنامه مدون اولیه، امکانی برای محقق و دیگران فراهم می‌کند تا تعیین کنند آیا یک برنامه عملی برای آزمون فرضیه را می‌توان به اجرا درآورد؟

پس از صورتبندی و ارزشیابی فرضیه براساس معیارهای گفته شده، محقق آماده است تا فرضیه تحقیق را مورد آزمون تجربی قرار دهد. فرضیه صفر، نقیض آنچه که مورد انتظار محقق است، بیان می‌شود. مهم است که به یاد داشته باشیم که یک فرضیه، اثبات شدنی و یا اثبات ناشدنی نیست و تنها تأیید و یا رد می‌شود. یک فرضیه حتی اگر تأیید هم نشود، هنوز می‌توان مفید باشد، چون محقق را به سوی ارزشیابی مجدد استدلالها و روشها و توجه به رویکردهای دیگر برای حل مسئله هدایت می‌کند.

// مفاهیم اصلی

آزمون فرضیه

فرضیه استقرایی

فرضیه جهت‌دار

فرضیه غیرجهت‌دار

فرضیه قیاسی

فرضیه صفر

کارکردهای فرضیه

معیار برای ارزشیابی فرضیه

نظریه

/// تمرینها

۱- مقصود از ارائه فرضیه چیست؟

۲- چه تفاوتی بین یک فرضیه استقرایی و قیاسی وجود دارد؟

۳- بر مبنای هر یک از سؤالات تحقیقی که در ذیل فهرست شده است، یک فرضیه بیان کنید؟

الف - اثر استفاده از روش کوئیزر در تدریس حساب ابتدایی چیست؟

ب - آیا بین جنسیت معلم و افزایش مهارت خواندن در بین دانش‌آموزان دختر سیاهپوست

در مدارس ابتدایی رابطهای وجود دارد؟

ج - آیا زندگی در محلات چندنژادی در نگرش فرد نسبت به مردم نژادهای دیگر تأثیر می‌گذارد؟

د - آیا بین نوع تشویق (مادی یا غیرمادی) و میزان یادگیری کودکان دارای محرومیت اقتصادی رابطهای وجود دارد؟

ه - آیا آموزش پیش‌دبستانی شکاف آموزشی بین کودکان دارای محرومیت اقتصادی و کودکان بدون محرومیت اقتصادی را کاهش می‌دهد؟
و - آیا انتظارات معلم از شاگردان برای داشتن عملکردی خردمندانه تأثیری در عملکرد واقعی آنها دارد؟

۴ - فرضیه ذیل را در قالب فرضیه صفر بازنویسی کنید: کودکانی که مهارت خواندن آنها از سطح کلاس پایین‌تر است، نسبت به کودکانی که مهارت خواندن آنها از سطح کلاس بالاتر است، رضایت کمتری از مدرسه نشان می‌دهند.

۵ - شایستگی هر یک از فرضیات ذیل را ارزیابی کنید. اگر یک فرضیه شایسته است، دلیل شایستگی آن را بیان کنید و یک فرضیه شایسته بنویسید.

الف - سزاوار است که معلمان بیش از مسئولان مدارس حقوق دریافت کنند.

ب - دانش‌آموزان دوره اول دبستان که درس علوم اجتماعی را می‌گذرانند، نسبت به کسانی که این درس را نمی‌گذرانند، توان قضاوت روشن بینانه‌تری را در مورد امور سیاسی محلی خواهند داشت.

ج - برای تدریس ترکیبات ضرب به دانش‌آموزان کندآموز، دادن تکالیف و تجربیات کامپیوتری بهتر از استفاده از کارتهای کمک آموزشی است.

د - اگر دانش‌آموزان از نظر وضعیت اجتماعی - اقتصادی متفاوت باشند، از نظر نمرات درس انگلیسی نیز متفاوت خواهند بود.

ه - کودکانی که هیجان زاید الوصفی نشان می‌دهند، براساس مقیاس اندازه‌گیری اضطراب آشکار، اضطراب بیشتری را نیز نشان خواهند داد.

و - تقویت مثبت پاسخهای دانش‌آموز از طرف معلم در کلاس، احتمال پاسخهای آتی را کاهش می‌دهد.

- ۶- یک فرضیه غیرجهت‌دار و یک فرضیه جهت‌دار مبتنی بر این سؤال تحقیقی بنویسید که چه رابطه‌ای بین سرعت بلوغ پسران نوجوان و خودانگاره آنها وجود دارد.
- ۷- چرا یک فرضیه باید پیش از شروع تحقیق بروشنی بیان شده باشد؟
- ۸- فرضیات ذیل را به عنوان فرضیات تحقیقی و یا فرضیه صفر نام‌گذاری کنید:
 - الف - دانش‌آموزانی در آزمونهای پیشرفت تحصیلی نمرات پایین خواهند گرفت که به سطوح بالای طبقه‌بندی بلوم رسیده باشند.
 - ب - هیچ تفاوتی در عملکرد دانش‌آموزانی که ریاضی را با روش الف آموخته‌اند با دانش‌آموزانی که ریاضیات را با روش ب آموخته‌اند، وجود ندارد.
 - ج - هیچ تفاوتی بین میانگین نمرات یادداری کودکانی که داروی آزمایشی را خورده‌اند با میانگین نمرات کودکانی که این دارو را نخورده‌اند، وجود ندارد.
 - د - دانش‌آموزانی که تحت آموزش معلمان آزادمنش قرار داشته‌اند، نسبت به دانش‌آموزانی که تحت آموزش معلمان مقتدر بوده‌اند، مهارتهای بیشتری در حل مسئله نشان خواهند داد.
- ۹- یک بررسی تحقیقی، که فرضیه‌ای را بیان می‌کند، پیدا کنید و سپس سعی کنید نظریه‌ای را که فرضیه از آن منتج شده است، شناسایی کنید.
- ۱۰- فرضیه ذیل را نقد کنید: برای آنکه بعدها شهروندان بهتری داشته باشیم، در کلاس علوم اجتماعی دوم دبیرستان استفاده از روش بحث، سودمندتر از روش سخنرانی است.

/// پاسخها

- ۱- هدف فرضیه تهیه یک طرح پیشنهادی موقت (غیرقطعی) به عنوان راه‌حلی برای یک مسئله و یا تبیین برخی پدیده‌هاست.
- ۲- با یک فرضیه استقرایی، محقق در مورد روابط دست به مشاهده می‌زند و سپس تبیینی برای رفتارهای مشاهده شده فرض می‌کند. از طریق یک فرضیه قیاسی، محقق فرضیه‌ای مبتنی بر نظریه‌ای شناخته شده به همراه دلیل منطقی برای فرضیه خود صورتبندی می‌کند.
- ۳- الف - دانش‌آموزان ابتدایی که از طریق روش کوئیزر حساب را آموخته‌اند، نسبت به آنهایی که حساب را از این طریق نیاموخته‌اند، در آزمون ریاضی نمرات بیشتری خواهند گرفت.
- ب - دانش‌آموزان پسر سیاهپوست دوره ابتدایی که از طریق معلم مرد آموزش دیده‌اند،

نسبت به دانش‌آموزان پس‌سیاهپوست دوره ابتدایی که از طریق معلم زن آموزش دیده‌اند، نمرات بهتری در خواندن می‌گیرند.

ج - افرادی که در محلات چندنژادی زندگی می‌کنند، نسبت به افرادی که در محلات تفکیک شده زندگی می‌کنند، در برخورد با افراد دیگر نژادها نگرش مطلوبتری از خود نشان می‌دهند.

د - دانش‌آموزان ضعیفی که از طریق پاداشهای مادی تقویت شده‌اند، نسبت به دانش‌آموزان ضعیفی که با پاداشهای غیرمادی تقویت شده‌اند، در خواندن پیشرفت بیشتری را نشان می‌دهند.

ه - کودکان محروم و غیرمحرومی که در سنین کودکستان از آموزش پیش‌دبستانی بهره‌مند بوده‌اند، نسبت به کودکان محروم و غیرمحرومی که در سنین کودکستان از آموزش پیش‌دبستانی بی‌بهره بوده‌اند، از نظر آموزشی با فاصله کمتری از یکدیگر متمایز می‌شوند. و- کودکانی که معلمشان از آنها انتظار زیادی برای عملکرد عاقلانه دارد، نسبت به کودکانی که معلمشان از آنها انتظار کمی برای عملکرد عاقلانه دارد، رفتار سطح بالاتری را نشان می‌دهند.

۴- بین کودکانی که مهارت خواندن آنها از سطح کلاس پایین‌تر است و کودکانی که مهارت خواندن آنها از سطح کلاس بالاتر است، هیچ تفاوتی در ابراز رضایت از مدرسه وجود ندارد.

۵- الف - فرضیه نامناسب است؛ چون یک تبیین ارزشی است و نمی‌تواند در یک بررسی تحقیقاتی مورد پژوهش قرار گیرد. یک فرضیه معقول این است: معلمانی که دستمزد بیشتری از مسئولان مدرسه دریافت می‌کنند، نسبت به معلمانی که دستمزد کمتری دریافت می‌کنند، رضایت شغلی بیشتری نشان می‌دهند.

ب - فرضیه نامناسب است؛ چون قضاوت‌های روشن‌بینانه یک اصطلاح ارزشی است. یک فرضیه قابل قبول می‌تواند این باشد: دانش‌آموزانی که درس علوم اجتماعی دوره اول دبیرستانی را می‌گذرانند، نسبت به دانش‌آموزانی که این درس را نگذرانده‌اند، دانش بیشتری نسبت به امور سیاسی محلی از خود نشان می‌دهند.

ج - فرضیه نامناسب است؛ چون یک تبیین ارزشی و فاقد تعاریف کاربردی روشن و دقیق است. یک فرضیه قابل آزمون می‌تواند این باشد: دانش‌آموزانی که در ریاضی عملکردی

ضعیف دارند و ترکیبات ضرب را از طریق تکالیف کامپیوتری تمرین می‌کنند، در مجموع نسبت به دانش‌آموزانی که در ریاضی عملکردی ضعیف دارند و با صرف وقت یکسان ترکیبات ضرب را از طریق کارتهای کمک آموزشی تمرین می‌کنند، در یک آزمون ملاکی، به نسبت بیشتری، پاسخ درست می‌دهند.

د - فرضیه نامناسب است؛ چون در آن بیان یک رابطه مورد انتظار بین متغیرها وجود ندارد.

ه - کودکانی که انگیزه پیشرفت زیادی از خود نشان می‌دهند؛ بیش از کودکانی که انگیزه پیشرفت کمی دارند، در مقیاس اضطراب آشکار کودکان، امتیاز کسب خواهند کرد.

و - فرضیه نامناسب است؛ چون با دانش موجود در مورد تقویت مثبت و اثرات آن در پاسخهای دانش‌آموز، متضاد است.

۶ - فرضیه جهت‌دار: پسران زودرس، خودانگاره مثبت‌تری از پسران دیررس نشان می‌دهند. فرضیه غیرجهت‌دار: بین خودانگاره نوجوان پسر زودرس و دیررس تفاوت وجود دارد.

۷ - فرضیه به گردآوری و تفسیر داده‌ها جهت می‌دهد. بیان روشن فرضیه می‌تواند نقطه ضعفهایی را نشان دهد که هنگام تکوین ایده گنگ بررسی در ذهن، آشکار نیست.

۸ - الف - تحقیق

ب - صفر

ج - صفر

د - تحقیق

۹ - پاسخها می‌توانند متعدد باشند.

۱۰ - فرضیه در یک مدت زمان منطقی قابل آزمون نیست. بعلاوه، تعریف و اندازه‌گیری اصطلاح شهروند بزرگسال بهتر، مشکل خواهد بود.

بخش سوم

تحلیل آماری

فصل پنجم

آمار توصیفی

هدفهای آموزشی

دانشجو بعد از مطالعه این فصل، قادر خواهد بود:

- ۱- ویژگیها و محدودیتهای چهار نوع مقیاس اندازه گیری: اسمی، رتبه ای، فاصله ای و نسبی را تشخیص دهد.
- ۲- داده های تحقیق را به صورت توزیع فراوانی طبقه بندی کند و آنها را به شکل نمودار چند ضلعی فراوانی و نمودار ستونی ارائه کند.
- ۳- تفاوت اندازه های گرایش مرکزی را تشخیص دهد و موقعیتهای قابل استفاده از هریک را بازشناسد و برای هر داده مفروض میانگین، میانه و نما را محاسبه کند.
- ۴- کاربردهای مناسب اندازه های تغییر پذیری را توصیف کند و برای هر مجموعه از داده های مفروض واریانس، انحراف معیار، انحراف چارکی و دامنه را محاسبه کند.
- ۵- نمره های خام را به نمره های استاندارد تبدیل و جایگاه نسبی نمره z را در یک منحنی نرمال تعیین کند.
- ۶- کاربردهای مناسب شاخصهای مختلف همبستگی را برای توصیف رابطه بین متغیرها بشناسد.

- ۷- ضریب همبستگی بین دو مجموعه از داده‌های فاصله‌ای یا نسبی را محاسبه کند.
- ۸- ضریب همبستگی بین دو مجموعه از داده‌ها را که به صورت رتبه‌ای منظم شده‌اند، محاسبه کند.
- ۹- معنای ضریب تعیین و کاربرد آن را در تفسیر ضریب همبستگی توصیف کند.
- ۱۰- معنا، ویژگی‌ها و کاربردهای فرا تحلیل^(۱) را بشناسد.

شیوه‌های آماری، اساساً روشهایی هستند که اطلاعات کمی را به گونه‌ای به کار می‌گیرند که آن اطلاعات معنادار می‌شوند. این شیوه‌ها دو مزیت اصلی دارند. نخست آنکه این شیوه‌ها ما را قادر می‌سازند که مشاهده‌های خود را طبقه‌بندی، توصیف و خلاصه کنیم. دوم، این شیوه‌ها به ما کمک می‌کنند تا دریابیم که با چه اعتباری می‌توان نتیجه‌گیری کرد که پدیده مشاهده شده در یک گروه محدود، یعنی نمونه‌ای در جامعه بزرگ مورد نظر که نمونه از آن انتخاب شده است، به وقوع خواهد پیوست - به سخن دیگر، ما با چه دقتی می‌توانیم از طریق استدلال استقرایی، نتیجه بگیریم که آنچه در جزء مشاهده کرده‌ایم، در کل نیز مشاهده خواهیم کرد. ما برای مسائلی که چنین ماهیتی دارند، نیازمند به کارگیری آمار استنباطی هستیم.

شناخت برخی از شیوه‌های آماری برای کسانی که به کار پژوهش می‌پردازند، از آنرو اساسی است که آنها می‌توانند داده‌های خود را تحلیل و تفسیر کنند و از طریق یافته‌های خود با دیگران ارتباط برقرار نمایند. بعلاوه، شایسته است که کارشناسان آموزشی، که نیازمند آگاهی از امر پژوهش و استفاده از یافته‌های پژوهشی هستند، به منظور درک و ارزشیابی مطالعات پژوهشی دیگران، با شیوه‌های آماری آشنا باشند. همچنین آگاهی از شیوه‌های آماری برای اجرا و تفسیر مناسب آزمونهایی که در مدارس ما معمول است، ضرورت دارد. آموزگاران که با این شیوه‌ها ناآشنا هستند، ممکن است در ارزشیابی تواناییها و پیشرفت دانش‌آموزان خود با اشکال مواجه شوند. همچنین، آنها بررسی پژوهشهای حیطه تخصصی خود و کسب اطلاعات جدید را نیز دشوار خواهند یافت.

/// مقیاسهای اندازه‌گیری

یک گام اساسی در انجام تحقیق، اندازه‌گیری است؛ فرآیندی که از طریق آن مشاهده‌ها به اعداد برگردانده می‌شوند. اس. اس. استیونس گفته است، «اندازه‌گیری در گسترده‌ترین معنای آن، تخصیص اعداد به اشیاء یا رویدادها، بنابر قانونمندیهاست.»^(۱) محققان با متغیرها آغاز می‌کنند و سپس برای تعیین اینکه این متغیرها چگونه به شکل عددی در می‌آیند، از قوانین استفاده می‌کنند. متغیر رحجان مذهبی را می‌توان بر مبنای شمارش پاسخهای دانشجویان به پرسشی که

از آنها می‌خواهد تا از میان چهار گزینه (۱) کاتولیک، (۲) یهودی، (۳) پروتستان، یا (۴) دیگر مذاهب، یکی را انتخاب کنند، اندازه‌گیری کرد. متغیر وزن را می‌توان با مشاهده ارقامی که با گام نهادن آزمودنیها بر روی ترازو به دست می‌آید، اندازه‌گیری کرد. متغیر بلوغ اجتماعی را می‌توان بر مبنای نمره‌هایی که از مقیاس بلوغ اجتماعی واینلند^(۱) به دست می‌آید، اندازه‌گیری کرد. ماهیت فرآیند اندازه‌گیری که اعداد را خلق می‌کند، تفسیری را که می‌توان از آنها به عمل آورد و شیوه‌های آماری را که می‌توان به طور معنی‌دار در مورد آنها به کار برد، تعیین می‌کند. رایج‌ترین طبقه‌بندی شیوه‌های اندازه‌گیری، مقیاسهای اندازه‌گیری استیونس است، که اندازه‌گیری را به انواع اسمی، رتبه‌ای، فاصله‌ای و نسبی دسته‌بندی کرده است.

// مقیاس اسمی

ابتدایی‌ترین مقیاس اندازه‌گیری، مقیاس اسمی است. اندازه‌گیری اسمی مستلزم جای‌دادن اشیاء یا افراد در طبقاتی است که نه از نظر کمی، بلکه از نظر کیفی با هم متفاوت هستند. اندازه‌گیری این سطح، تنها نیازمند آن است که هر فرد بتواند دو یا چند طبقه مرتبط را از هم تشخیص دهد و معیارهای لازم را برای دادن افراد یا اشیاء در این یا آن طبقه بشناسد. سطح فعالیت تجربی ضروری، مستلزم تشخیص این است که آیا یک شیء یا فرد فرضی به طبقه نامتداخل (مانعةالجمع) تعلق دارد یا ندارد. تنها رابطه‌ی مابین طبقات، این است که آنها با یکدیگر متفاوت هستند، و هیچ دلالتی وجود ندارد که این طبقات گویای «بیشتر» یا «کمتر» بودن ویژگی اندازه‌گیری شده باشند. دسته‌بندی کردن دانشجویان برحسب جنس، یک اندازه‌گیری اسمی خواهد بود.

در سطح اندازه‌گیری اسمی، اعداد، اغلب مورد استفاده قرار می‌گیرند، اما این امر تنها برای مشخص کردن طبقات صورت می‌گیرد. اعداد، که برحسب سلیقه به طبقات اختصاص می‌یابند، صرفاً به عنوان یک برچسب عمل می‌کنند. به همه اعضای یک طبقه‌بندی عدد یکسانی تخصیص داده می‌شود و هرگز به دو طبقه‌بندی، عددی یکسان تخصیص نمی‌یابد. برای مثال، در آماده سازی داده‌ها برای کامپیوتر، می‌توان رقم صفر را برای نمایانندن مردان و رقم ۱ را برای نمایانندن زنان مورد استفاده قرار داد. مابین اعدادی که برای اندازه‌گیری اسمی به کار می‌روند،

هیچ رابطه تجربی‌ای که با روابط ریاضی بین اعداد مطابقت کند، وجود ندارد. عدد ۱ گویای آن نیست که چیزی بیشتر از عدد صفر دارد. اعداد می‌توانند با هم جابجا شوند، بدون اینکه برچیزی به جز طرح نامگذاری مورد استفاده، تأثیر بگذارند.

اعداد بکار رفته در یک مقیاس اسمی، معرف مقدار مطلق یا نسبی هیچ ویژگی‌ای نیستند. آنها صرفاً برای شناساندن اعضای یک طبقه‌بندی فرضی عمل می‌کنند. برای مثال، اعداد تخصیص یافته به بازیکنان فوتبال، یک مقیاس اسمی تشکیل می‌دهند. ما نمی‌خواهیم بگوییم بازیکنی که شماره ۴۸ را روی پیراهن خود دارد، لزوماً بهتر از بازیکن شماره ۳۶ است. همچنین نمی‌توانیم بگوییم که بین بازیکنان شماره ۴۰ و ۴۸ همان تفاوتی در توانایی بازی وجود دارد که بین بازیکنان شماره ۵۰ و ۵۸.

البته، اعداد شناسایی در یک مقیاس اسمی هرگز نمی‌توانند از نظر محاسباتی، از طریق جمع، منها، ضرب، یا تقسیم دستکاری شوند. تنها می‌توان شیوه‌های آماری مبتنی بر شمارش صرف، نظیر گزارش تعداد مشاهده‌ها در هر طبقه‌بندی را مورد استفاده قرار داد.

// مقیاس رتبه‌ای

دومین مقیاس مهم اندازه‌گیری، مقیاس رتبه‌ای است، که در آن با در نظر گرفتن یک ویژگی، موقعیت نسبی اشیاء یا افراد، بدون نشان دادن فاصله بین آن موقعیتها، مشخص می‌شود. برای اندازه‌گیری در این سطح، یک معیار تجربی برای نظم و ترتیب دادن به اشیاء یا رویدادها، با توجه به ویژگی مورد نظر، ضرورتی اساسی دارد، یعنی به شیوه‌ای نیاز است که ترتیب افراد یا اشیاء را بر حسب دارا بودن ویژگی مورد نظر تعیین کند. برای مثال، هنگامی که آموزگاران دانش‌آموزان خود را بر مبنای برخی ویژگیها، نظیر بلوغ اجتماعی، تواناییهای رهبری، حس همکاری و غیره رتبه‌بندی می‌کنند، یک اندازه‌گیری رتبه‌ای صورت پذیرفته است. دانش‌آموزان اغلب برحسب پیشرفت تحصیلی خود یا عملکردشان در یک مسابقه ورزشی یا رقابت در زمینه موسیقی رتبه‌بندی می‌شوند.

در اندازه‌گیری رتبه‌ای، شیوه تجربی مورد استفاده برای مرتب کردن اشیاء باید معیاری به نام اصل رابطه غیر مستقیم^(۱) را داشته باشد. این اصل چنین نوشته می‌شود: اگر $(a > b)$ و $(b > c)$ ،

پس ($a > c$)، و معنای آن این است که رابطه باید به گونه‌ای باشد که اگر شیء a بزرگتر از شیء b است، و شیء b بزرگتر از شیء c است، پس شیء a بزرگتر از شیء c خواهد بود. البته، ممکن است کلمات دیگری جانشین بزرگتر از شوند، این کلمات می‌توانند شامل قوی‌تر از، جلوتر از، از یک ویژگی مقدار بیشتری دارد، و غیره باشند.

عملکرد تجربی در اندازه‌گیری رتبه‌ای تنها شامل مقایسه مستقیم اشیاء یا افراد برحسب مقدار ویژگی مورد نظر در آنهاست. بنابراین، هنگامی که اعداد به اشیاء تخصیص می‌یابند، تنها اطلاعی که مورد توجه قرار می‌گیرد، ترتیب اشیاء است. از اینرو، تنها ویژگی اعداد، که دارای معناست، ترتیب آنهاست. اعداد تخصیص یافته در اندازه‌گیری رتبه‌ای تنها مرتبه جایگاه را نشان می‌دهند و دیگر هیچ. نه تفاوت بین اعداد و نه نسبت آنها، معنایی ندارند. بکارگیری اعداد ۱ و ۲ و ۳ در اندازه‌گیری رتبه‌ای به این معنا نیست که رتبه ۱ همان قدر بالاتر از رتبه ۲ است که رتبه ۲ از رتبه ۳ و الی آخر. فاصله بین کودک رتبه ۱ و کودک رتبه ۲ ممکن است با فاصله دو کودک رتبه ۲ و ۳، یکی باشد، کمتر باشد و یا بیشتر باشد. بسادگی باید گفت که هیچ مبنایی برای تفسیر اندازه تفاوت بین اعداد یا نسبت اعداد وجود ندارد. در یک مسابقه دو زمان‌گیری نشده، ما می‌دانیم که چه کسانی اول، دوم، سوم و... به خط پایان رسیده‌اند، اما نخواهیم دانست که یک دوندۀ چقدر از دوندۀ دیگر سریع‌تر دویده است. تفاوت بین اول و دوم لزوماً با تفاوت بین دوم و سوم و یا سوم و چهارم یکسان نخواهد بود. همچنین نمی‌توانیم بگوییم که دوندۀ ای که به عنوان نفر دوم به خط پایان رسیده، دو بار سریع‌تر از دوندۀ چهارم بوده است.

مقیاس سختی مواد معدنی (کانیها) یک مثال خوب برای مقیاس رتبه‌ای است. مواد معدنی برحسب توانایی آنها در خراشیدن یکدیگر رده‌بندی می‌شوند. اگر ماده معدنی A بتواند بر ماده معدنی B خراش بیندازد، گفته می‌شود که ماده معدنی A سخت‌تر از ماده معدنی B است. بر این مبنا، الماس به عنوان سخت‌ترین ماده معدنی رده‌بندی می‌شود؛ زیرا الماس می‌تواند همه مواد معدنی شناخته شده دیگر را بخراشد، اما هیچ ماده معدنی نمی‌تواند بر آن خراش بیندازد. مجموعه‌ای از ده ماده معدنی که دامنه آنها از نرم‌ترین تا سخت‌ترین است، به عنوان استاندارد (معیار - ملاک - استاندارد) انتخاب شده و اعداد ۱ تا ۱۰ به آنها تخصیص یافته است، که ۱ نشانه نرم‌ترین و ۱۰ نشانه سخت‌ترین ماده معدنی است. به مواد معدنی دیگر نیز بر اساس آزمایش خراش‌پذیری، اعدادی تخصیص یافته است. بنابراین، ما ترتیب سختی مواد معدنی را می‌دانیم،

ولی نمی‌دانیم یک ماده معدنی چقدر از دیگری سخت‌تر است. ما نمی‌توانیم فرض کنیم که یک ماده معدنی که ارزش ۴ به آن تخصیص یافته است، دو بار سخت‌تر از ماده‌ای با ارزش ۲ است، یا تفاوت سختی بین مواد معدنی ۲ و ۴ به اندازه تفاوت سختی بین موادی با ارزشهای ۳ و ۴ است. آمارهای مناسب برای یک مقیاس رتبه‌ای محدود هستند. از آنجا که اندازه فاصله بین طبقات نامشخص است، نمی‌توانیم از هیچ شیوه آماری‌ای که فرض آن فواصل مساوی است، استفاده کنیم. آماره‌هایی که نشان‌دهنده نقاطی هستند که درصدهای معینی از موارد در زیر آن نقاط قرار می‌گیرند، برای یک مقیاس رتبه‌ای مناسب هستند.

// مقیاس فاصله‌ای

مقیاس فاصله‌ای مقیاسی است که فواصل مساوی از یک مبدا اختیاری را تأمین می‌کند. یک مقیاس فاصله‌ای نه تنها اشیاء یا رویدادها را برحسب مقدار ویژگی‌ای که دارا هستند، رتبه‌بندی می‌کند، بلکه همچنین، بین واحدهای اندازه‌گیری، فواصل مساوی بنا می‌نهد. تفاوت‌های مساوی در اعداد، معرف تفاوت‌های مساوی در ویژگی اندازه‌گیری شده است. دماسنج‌های فارنهایت و سانتیگراد نمونه‌هایی از مقیاس فاصله‌ای هستند.

در یک مقیاس فاصله‌ای، هم رابطه‌های ترتیبی و هم رابطه‌های فاصله‌ای بین اعداد، دارای معنا هستند. ما می‌توانیم بگوییم که تفاوت بین ۵۰ و ۵۱ درجه سانتیگراد با تفاوت بین ۳۰ و ۳۱ درجه سانتیگراد برابر است، اما به هر روی نمی‌توانیم بگوییم که ۵۰ درجه دو بار گرم‌تر از ۲۵ درجه است. این امر به دلیل آن است که در مقیاس فاصله‌ای نقطه صفر واقعی وجود ندارد. نقطه صفر قراردادی است، مثل مقیاس سانتیگراد، که ارزش صفر درجه را برای نقطه انجماد آب تعیین کرده است.

همچنین، در یک آزمون روان‌شناختی یا آموزشی، نقطه صفر اختیاری است. برای مثال، هوش صفر وجود ندارد. در آزمون‌های هوش استاندارد شده ما هیچ راهی برای شناسایی فرد دارای هوش صفر وجود ندارد. گاهی یک دانشجو ممکن است در آزمون آمار نمره صفر بگیرد، اما این به معنای آن نیست که دانش آماری او صفر است. اگر ما سه دانشجو داشته باشیم که در یک آزمون آمار نمره‌های ۱۵، ۳۰ و ۴۵ گرفته باشند، نمی‌توانیم بگوییم که نمره ۳۰ معرف دو برابر دانش آماری نمره ۱۵ یا نمره ۴۵ معرف سه برابر دانش آماری نمره ۱۵ است. برای درک این

مطلب، اجازه دهید فرض کنیم ۱۵ پرسش بسیار ساده به آزمون اضافه می‌شود؛ به گونه‌ای که هر سه دانشجو قادرند آنها را بدرستی پاسخ گویند. اکنون سه نمره قبلی به ۳۰، ۴۵ و ۶۰ تبدیل خواهد شد. اگر تلاش کنیم در این مقیاس از نوع فاصله‌ای، نسبت‌هایی بین نمره‌ها بسازیم، به اشتباه گزارش خواهیم کرد دانشجویی که نمره ۶۰ گرفته، دو برابر دانشجویی که نمره ۳۰ گرفته است، دانش آماری دارد، در حالی که در نسبت قبلی ما به خطا فرض کردیم که همان دانشجو سه برابر دانشجوی دیگر آمار می‌داند.

بنابراین، از آنجا که صفر اختیاری است، ضرب و تقسیم اعداد درست نیست و همان‌گونه که دیدیم نسبت‌های بین اعداد در یک مقیاس فاصله‌ای بی‌معنا هستند. به هر حال، در یک مقیاس فاصله‌ای می‌توان تفاوت بین جایگاه‌ها (موقعیتها)^(۱) را گزارش داد یا اعداد را با هم جمع کرد. همه شیوه‌های آماری مبتنی بر عمل جمع و شیوه‌های آماری مناسب برای مقیاس‌های سطح پایین‌تر را می‌توان در این مقیاس مورد استفاده قرار داد. این امر اکثر شیوه‌های رایج آماری را در بر می‌گیرد.

نکته حائز اهمیت آن است که در موارد بسیاری که از مقیاس‌های فاصله‌ای استفاده می‌کنیم، فواصل مقیاس از لحاظ خود ابزار اندازه‌گیری، مساوی هستند، اما لزوماً از لحاظ توانایی مورد اندازه‌گیری، مساوی نیستند. برای روشن شدن موضوع، یک آزمون هجی کردن را با این کلمات در نظر بگیرید: گربه، بشقاب، توپ، پرحرف، اسکیزوفرنی و سینه پهلوی. در اینجا فاصله بین یک پاسخ درست و سه پاسخ درست با فاصله بین سه پاسخ درست و پنج پاسخ درست همسان است. به هر حال، هنگامی که توانایی هجی کردن را در نظر بگیریم، تفاوت بین ۳ و ۵ پاسخ درست در مقایسه با ۱ و ۳ پاسخ درست گویای توانایی بیشتری است؛ مگر اینکه کسی بتواند بگوید که در آزمون هجی کردن، فاصله بین ۳ و ۵ و فاصله ۱ و ۳ توانایی هجی کردن را به یک میزان نشان می‌دهند، که در این صورت این نمره‌ها رتبه‌بندی دانشجویان را بیان می‌کنند.

با این همه، با دقت و تلاش زیاد می‌توان آزمونی تهیه کرد که فواصل بین نمرات آن یک تخمین منطقی از فاصله بین توانایی‌ها باشد. آزمون‌های بهترهوش، شاهد این گفته هستند. تفاوت در توانایی بین هوشبهر (IQ) ۹۰ و هوشبهر (IQ) ۹۵، شاید دقیقاً همان تفاوت بین هوشبهر (IQ) ۱۰۵ و ۱۱۰ نباشد، اما اگر هم فرض کنیم که این دو تفاوت تقریباً یکسان هستند، خیلی به خطا نرفته‌ایم.

// مقیاس نسبی

یک مقیاس نسبی، یعنی بالاترین سطح اندازه‌گیری، مقیاسی است که هم نقطه صفر واقعی و هم فواصل مساوی را تأمین می‌کند. در این مقیاس، بین هر دو مقدار فرضی می‌توان نسبتها را تشکیل داد. خط‌کشی که برای اندازه‌گیری طول در واحد اینچ یا فوت به کار می‌رود، یک مقیاس نسبی است؛ زیرا مبدأ مقیاس یک صفر مطلق و مطابق با طول صفر است. بنابراین، می‌توانیم بگوییم که طول یک تکه چوب ۶ فوتی دو برابر تکه چوب ۳ فوتی است. در یک مقیاس نسبی، این امکان وجود دارد که هر یک از مقادیر را در یک عدد معین ضرب یا تقسیم کنیم، بدون آنکه خصوصیات مقیاس تغییر کند. برای مثال، می‌توانیم ۲ پوند را در ۱۶ ضرب کنیم، برای آنکه واحد اندازه‌گیری را به ۳۲ اونس تبدیل کنیم یا می‌توانیم ۶ فوت را در ۱۲ ضرب کنیم، برای آنکه واحد را به اینچ تبدیل کنیم. ما می‌توانیم عمل ضرب را انجام دهیم و همان نسبتهای قبل از ضرب را حفظ کنیم. برای مثال، ما می‌توانیم ۴ کوارت^(۱) شیر و ۲ کوارت شیر را در ۲ ضرب و واحد اندازه‌گیری را به پینت^(۲) تبدیل کنیم. درواحدپینت هنوز ۸ پینت شیر دو برابر ۴ پینت است.

در تعلیم و تربیت تنها چند متغیر مورد نظر، ماهیت نسبی دارند. این متغیرها بیشتر به عملکرد حرکتی و دیگر اندازه‌های فیزیولوژیکی منحصر هستند. با آنکه ما می‌توانیم بگوییم که قد شخصی با ۶ فوت دو برابر قد شخص دیگری با ۳ فوت است، زیرا ارتفاع، یک داده نسبی است، اما نمی‌توانیم بگوییم که شخصی با هوشبهر (IQ) ۱۵۰ دو برابر شخص دیگری با هوشبهر (IQ) ۷۵، باهوشتر است؛ زیرا نمره‌های هوشبهر (IQ) داده‌های فاصله‌ای هستند. همه انواع شیوه‌های آماری برای یک مقیاس نسبی مناسب هستند.

/// سازماندهی داده‌های پژوهش

توصیف داده‌هایی که به نوعی مرتب شده باشند، اگر ناممکن نباشد، بسیار مشکل است. بنابراین، سازمان دادن داده‌های پژوهش در آمار توصیفی یک گام اساسی است. دو روشی که اغلب در سازمان دادن داده‌ها مورد استفاده قرار می‌گیرند، عبارتند از (۱) مرتب کردن اندازه‌ها به صورت توزیع فراوانی و (۲) ارائه آنها به شکل نگاره‌ای.

// توزیع فراوانی

آرایش نظام‌مند اندازه‌های فردی، از کمترین تا بالاترین آنها، توزیع فراوانی نامیده می‌شود. استفاده از این روش صرفاً مستلزم ساخت فهرستی از اندازه‌های فردی به صورت ستونی است که بالاترین اندازه در بالای ستون و پایین‌ترین اندازه در انتهای ستون قرار گرفته باشد. اغلب دیده می‌شود که نمرات یکسان متعددی در یک توزیع وجود دارد. به جای آنکه این نمرات جداگانه فهرست شوند، مرسوم است که ستون دومی اضافه شود و در آن فراوانی هر یک از اندازه‌ها ثبت شود. در جدول ۵-۱، نمرات ۱۰۵ دانش‌آموز در آزمون آمار نشان داده شده است. در قسمت الف این جدول، نمرات به شکل سازمان نیافته فهرست شده‌اند. در قسمت ب همان نمرات بعد از مرتب شدن در یک توزیع فراوانی ارائه شده‌اند.

شکل کلی یک توزیع را می‌توان از توزیع فراوانی آن دریافت. با سازمان دادن نمرات در یک توزیع فراوانی، می‌توان گستردگی آنها را تعیین کرد که آیا این نمرات به طور متعادل توزیع شده‌اند یا گرایش به تجمع دارند و این تجمع‌ها در چه نقاطی از توزیع صورت می‌پذیرند. برای مثال، با نگاهی به توزیع فراوانی نمرات جدول ۵-۱، بسادگی روشن می‌شود که محدوده این نمرات از ۲۱ تا ۳۵ است، و نمره ۲۹ بیشتر از دیگر نمره‌ها تکرار شده است، و این نمرات گرایش دارند که در بالای توزیع تجمع یابند، تا در پایین آن. اگر نمره‌ها سازمان داده نشده بودند، هیچ یک از این موارد آشکار نمی‌شد. سازماندهی داده‌ها در توزیعهای فراوانی، همچنین، محاسبه آماره‌های مفید مختلف را تسهیل می‌کند.

// ارائه نگاره‌ای داده‌ها

اغلب ارائه داده‌های پژوهش به شکل نگاره‌ای مفید و مناسب است. در میان انواع مختلف نگاره‌ها، نمودار ستونی و نمودار چندضلعی فراوانی، بیشترین مورد استفاده را دارند. مراحل اساسی ساخت نمودار ستونی و نمودار چند ضلعی فراوانی یکسان است:

۱- بر روی یک محور افقی (محور طولی)، از چپ به راست، از پایین‌ترین مقدار تا بالاترین مقدار، نقاطی را برای نمره‌ها مشخص کنید. در دو انتهای توزیع، فضای کافی برای یک نمره اضافی در نظر بگیرید.

۲- بر روی محور عمودی، فراوانی نمره‌ها یا فاصله‌ها را مشخص کنید. از صفر شروع کنید و رو به بالا ادامه دهید.

۳- بر بالای مرکز هر نمره در سطح فراوانی آن یک نقطه بگذارید.

از این قسمت به بعد می‌توان هر دو نمودار ستونی یا چند ضلعی را ساخت.
 برای ساخت نمودار ستونی، همانطور که در شکل ۵-۱ نشان داده شده است، باید بر هر نقطه، خطی افقی به پهنای هر نمره کشید.
 برای ساخت نمودار چند ضلعی، نقاط مجاور به هم متصل می‌شوند و دو انتهای شکل به دست آمده نیز در نقاطی که یکی کمتر از پایین‌ترین نمره و یکی بیشتر از بالاترین نمره باشد، به محور افقی (خط صفر)، همانطور که در شکل ۵-۲ نشان داده شده است، وصل می‌شوند.

/// اندازه‌های گرایش مرکزی

یکی از راههای ساده خلاصه کردن داده‌ها، یافتن یک شاخص منفرد است که بتواند معرف مجموعه‌ای از اندازه‌ها باشد. برای مثال، یافتن یک نمره منفرد، که بتواند نشانه‌ای از عملکرد یک گروه ۳۰۰ نفری دانش‌آموزان را در آزمون IQ به دست دهد، برای مقاصد مقایسه‌ای مفید خواهد

جدول شماره ۱-۵ نمرات ۱۰۵ دانش‌آموز در آزمون آمار

الف. نمرات سازمان نیافته

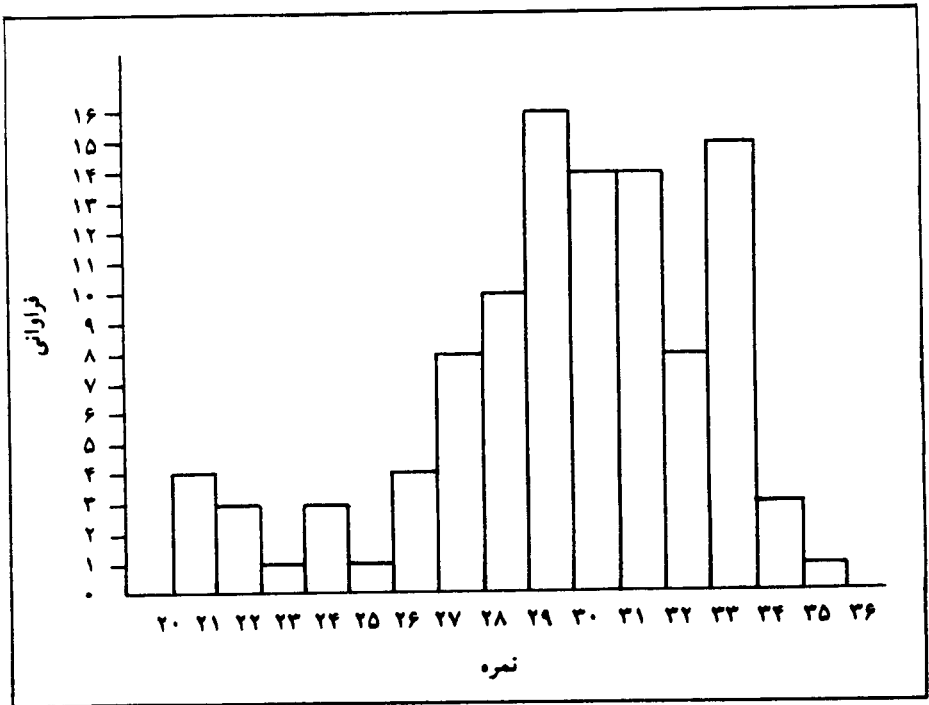
۲۳	۲۹	۳۱	۳۴	۲۴	۲۸	۳۲	۳۳	۲۹	۳۳	۳۰	۲۹	۳۳
۳۱	۳۲	۳۲	۲۹	۲۸	۳۰	۲۶	۳۳	۳۳	۲۶	۲۹	۲۵	۲۸
۲۹	۲۷	۳۱	۳۵	۳۲	۳۰	۳۱	۲۹	۳۳	۳۱	۳۴	۲۸	۲۸
۳۳	۳۴	۳۱	۳۰	۲۷	۲۸	۲۸	۲۸	۲۹	۲۸	۳۳	۲۶	۲۶
۳۱	۲۹	۳۱	۲۸	۳۳	۳۱	۳۱	۲۱	۳۰	۲۴	۲۹	۲۷	۳۰
۳۰	۳۰	۳۳	۲۲	۳۰	۲۹	۳۱	۳۲	۳۲	۲۹	۳۳	۳۱	۳۱
۳۳	۲۷	۲۸	۲۹	۳۰	۲۱	۳۳	۲۷	۲۹	۲۸	۳۳	۳۲	۳۲

ب. توزیع فراوانی

فراوانی (f)	خط نشانها	نمرات (x)
۱	I	۳۵
۳	III	۳۴
۱۵	IIII IIII	۳۳
۸	IIII III	۳۲
۱۴	IIII IIII II	۳۱
۱۴	IIII IIII II	۳۰
۱۶	IIII IIII II	۲۹
۱۰	IIII III	۲۸
۸	IIII III	۲۷
۴	IIII	۲۶
۱	I	۲۵
۳	III	۲۴
۱	I	۲۳
۳	III	۲۲
۴	IIII	۲۱
N = ۱۰۵		

بود. در آمار، سه شاخص برای چنین منظوری در دسترس است. این شاخصها اندازه‌های گرایش مرکزی یا متوسطها خوانده می‌شوند. برای بسیاری از مردم عادی، معنای واژه متوسط، جمع نمره‌ها و تقسیم آن بر تعداد نمره‌هاست. برای یک آمارشناس، واژه متوسط همین معنا را دارد، که آن را میانگین، یا یکی از دو اندازه دیگر گرایش مرکزی، یعنی نما یا میانه می‌نامند. هر یک از این سه می‌توانند به عنوان شاخصی که نماینده یک گروه به صورت یک کل باشند، عمل کنند.

شکل ۱-۵ نمودار ستونی ۱۰۵ نمره درس آمار از جدول ۵-۱



// نما

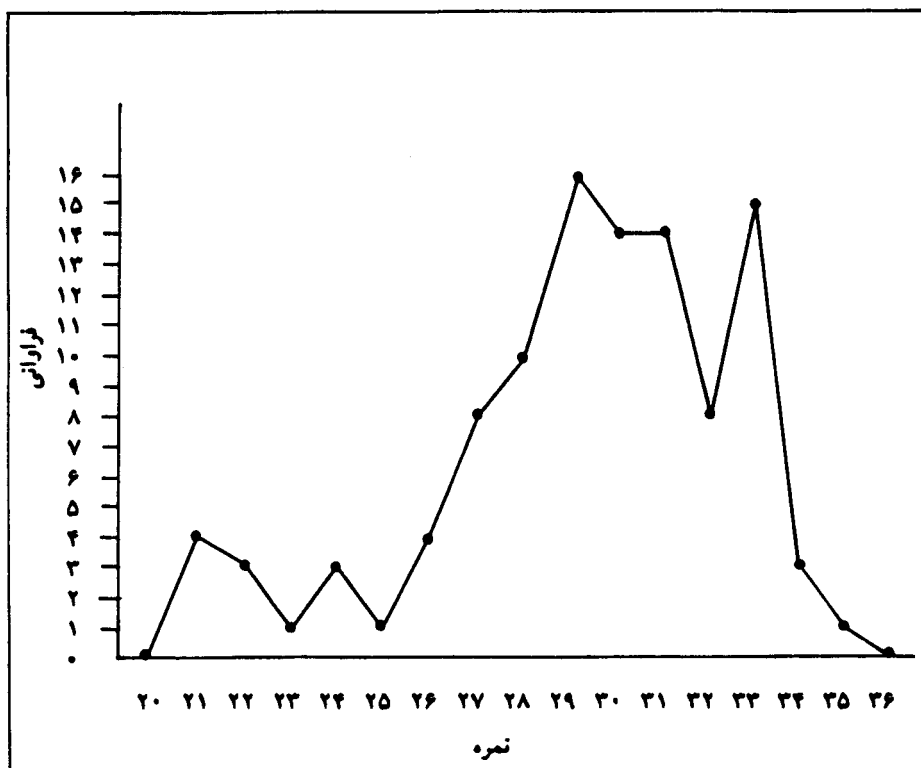
در یک توزیع، نما مقداری است که بیشترین تکرار را داراست. از میان سه اندازه گرایش مرکزی، یافتن نما ساده‌تر از دو اندازه دیگر است، زیرا استفاده از آن تنها به منظور بازبینی صورت می‌گیرد. نما را می‌توان برای هر یک از مقیاسهای اندازه‌گیری گزارش کرد، اما تنها اندازه گرایش مرکزی است که کاربرد آن در مقیاس اسمی درست است.

// میانه

میانه در یک توزیع به نقطه‌ای گفته می‌شود که ۵۰ درصد موارد در زیر آن قرار گیرند. (به این معنا که ۵۰ درصد دیگر در بالای آن نقطه قرار می‌گیرند). برای مثال، اگر توزیع نمره‌های زیر را داشته باشیم:

۱۴ ۱۶ ۱۶ ۱۷ ۱۸ ۱۹ ۱۹ ۱۹ ۲۱ ۲۲

شکل ۲-۵ نمودار چند ضلعی ۱۰۵ نمره درس آمار از جدول ۱-۵

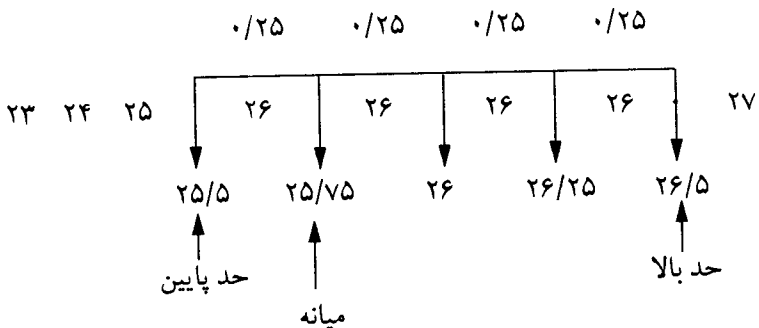


نقطه‌ای که ۵۰ درصد از موارد در زیر آن قرار می‌گیرند، در وسط ۱۸ و ۱۹ واقع شده است. بنابراین، میانه این توزیع ۱۸/۵ است. توجه کنید که برای یافتن این مقدار، ده نمره این توزیع ابتدا از کم به زیاد مرتب شده‌اند و سپس نقطه‌ای که نیمی از نمره‌ها زیر آن قرار می‌گیرند، پیدا شده است. این نقطه (۱۸/۵)، که دو مقدار ۱۸ و ۱۹ را دقیقاً جدا می‌کند، در الفبای آماری، حد بالای نمره ۱۸ و حد پایین نمره ۱۹ خوانده می‌شود. در محاسبه میانه، چنین تصور می‌شود که هر نمره، معرف دامنه یا فاصله‌ای است که از نیمه بین آن نمره و نمره پایین‌تر از خود تا نیمه بین آن نمره و نمره بالاتر از خود امتداد دارد. بنابراین، در این مثال، تصور بر آن است که ۱۸ معرف فاصله از ۱۷/۵ تا ۱۸/۵ است، در حالی که ۱۹ معرف فاصله از ۱۸/۵ تا ۱۹/۵ است.

نکته مهم آنکه میانه همیشه در خط مرزی بین دو مقدار قرار نمی‌گیرد. در واقع، میانه، اغلب در مکانی بین حد بالا و حد پایین یک فاصله جای می‌گیرد. به منظور تعیین میانه، ما باید یک نمره معرف دامنه حد پایین و بالای آن تصور کنیم، نه به صورت یک نقطه منفرد. مثال زیر را در نظر بگیرید:

۲۷ ۲۶ ۲۶ ۲۶ ۲۶ ۲۵ ۲۴ ۲۳

در این توزیع، میانه، بین دو نمره ۲۶ قرار می‌گیرد. به فراوانی نمره ۲۶ توجه کنید. در توزیع، چهار نمره ۲۶ وجود دارد، که یکی از آنها زیر نقطه میانی و سه عدد دیگر بالای نقطه میانی جای می‌گیرند. در این حالت برای یافتن میانه ما باید فاصله ۲۶/۵ و ۲۵/۵ را به چهار بخش تقسیم کنیم. پس فاصله بین هر یک از این چهار نمره، ۰/۲۵ خواهد بود. بنابراین تصور می‌شود که هر یک از این نمره‌ها معرف دامنه‌ای است که ۰/۲۵ از فاصله بین ۲۵/۵ و ۲۶/۵ را که حد پایین و بالای نمره ۲۶ هستند، می‌پوشانند. این مفهوم در زیر به تصویر درآمده است و موقعیت و مقدار میانه را نشان می‌دهد.



استفاده از یک فرمول برای یافتن میانه، ساده تر است تا اینکه هر بار چنین فرآیندی را طی کنیم.

$$Md = L + \left(\frac{\frac{N}{2} - cfb}{fw} \right) i \quad (5-1)$$

Md = میانه

L = حد پایین فاصله‌ای که میانه در آن قرار می‌گیرد

N = تعداد موارد موجود در توزیع

cfb = کل فراوانی در همه فاصله‌هایی که زیر فاصله حاوی میانه قرار دارند

fw = فراوانی موارد در داخل فاصله‌ای که حاوی میانه است

i = اندازه فاصله

در مثال قبل مقدار L برابر ۲۵/۵، N برابر ۸، cfb برابر ۳، fw برابر ۴ و i برابر ۱ می‌باشد.

$$\begin{array}{ccccccc} \frac{23}{cfb = 3} & & \frac{24}{25} & & \frac{26}{fw = 4} & & \frac{26}{27} \\ & & \downarrow & & & & . \end{array}$$

$$L = 25/50$$

این مقادیر را در فرمول می‌گذاریم و میانه را به دست می‌آوریم.

$$Md = 25/50 + \left(\frac{\frac{8}{2} - 3}{4} \right) 1 = 25/75$$

مقدار ۱ تنها زمانی ضرورت پیدا می‌کند که اندازه فاصله، ۱ نباشد، برای مثال، اگر با آزمون پیشرفت تحصیلی، که نمره‌های آن با واحدهای ده دهی نظیر ۳/۴ و ۵/۹ ارائه می‌شوند، سروکار داشته باشیم، اندازه فاصله ۰/۱ خواهد بود.

برای محاسبه ساده تر میانه می‌توان از توزیع فراوانی استفاده کرد. جدول ۵-۲ همان داده‌های جدول ۵-۱ را نشان می‌دهد و شامل یک ستون اضافی، به نام فراوانی تجمعی (cf) است که از آن

برای یافتن میانه استفاده می‌شود. فراوانیهای تجمعی، فراوانی مقادیر را در یک توزیع، تا هر فاصله به طوری که شامل همان فاصله هم باشد، معین می‌کنند. برای مثال، در جدول ۵-۲، فراوانی نمره‌ها، تا فاصله‌ای که شامل نمره ۳۰ است، همان‌طور که در ستون c_f (فراوانی تجمعی) مشخص شده، ۶۴ است. این ستون ما را قادر می‌سازد تا فاصله‌ای را که میانه در آن جای دارد، مشخص کنیم. از آنجا که در یک توزیع، میانه نقطه‌ای است که ۵۰ درصد از موردها در زیر آن قرار می‌گیرند، فاصله‌ای که میانه را در خود دارد، دارای فراوانی تجمعی برابر با $\frac{N}{2}$ خواهد بود. به عبارت دیگر، N توزیع را بر ۲ تقسیم می‌کنیم تا فاصله‌ای را که میانه در آن قرار دارد، پیدا کنیم.

در توزیع نمره‌هایی که در جدول ۵-۲ نشان داده شده است، تعداد موردها، یعنی N برابر ۱۰۵ است، بنابراین مقدار $\frac{N}{2}$ برابر ۵۲/۵ می‌باشد. با نگاه کردن به ستون متوجه می‌شویم که فراوانی تجمعی برای نمره ۲۹، برابر ۵۰ و برای نمره ۳۰، برابر ۶۴ است، بنابراین، میانه در فاصله‌ای که نمره ۳۰ معرف آن است، قرار دارد. اکنون، امکان به کار بردن فرمول ۵-۱ و یافتن میانه توزیع وجود دارد. در این مثال مقدار L ، ۲۹/۵ است؛ زیرا فاصله‌ای که میانه را در خود دارد، معرف دامنه‌ای است که حد پایین آن ۲۹/۵ و حد بالای آن ۳۰/۵ است. مقدار c_{fb} (یعنی، فراوانی مقادیر زیر فاصله‌ای که میانه در آن است) ۵۰ است. فراوانی مقادیر موجود در فاصله مربوط به نمره ۳۰ (یعنی، مقدار fw) ۱۴ است. اندازه فاصله، یا مقدار i در این توزیع، ۱ است؛ زیرا هر نمره، گستره فاصله را ۱ نشان می‌دهد. با استفاده از فرمول ۵-۱ در می‌یابیم که میانه توزیع، ۲۹/۶۷۸ است.

توجه داشته باشید که اندازه، نمره‌های فردی را به حساب نمی‌آورد. ما برای یافتن میانه، داده‌های خود را از کم به زیاد مرتب می‌کنیم و نقطه‌ای را که توزیع را به دو نیمه مساوی تقسیم می‌کند، پیدا می‌کنیم. میانه، یک آماره رتبه‌ای است؛ زیرا بنای آن بر رتبه‌بندی است. ما می‌توانیم از داده‌های فاصله‌ای یا نسبی نیز میانه را محاسبه کنیم، اما در چنین مواردی از خصیصه فاصله‌ای بودن داده‌ها استفاده نمی‌شود.

یکی از مواقعی که میانه اندازه مطلوب گرایش مرکزی است، هنگامی است که چند نمره انتهایی در توزیع وجود داشته باشد. در چنین حالتی استفاده از یک اندازه گرایش مرکزی که مقدار هر نمره را به حساب می‌آورد، منتج به برآورد بالا یا برآورد پایین نمره معرف می‌شود.

میانه به لحاظ عدم حساسیت به نمره‌های انتهایی مناسبترین شاخص برای یافتن نمره معرف است. برای روشن شدن مسئله، توزیع زیر را در نظر بگیرید:

۴۹ ۵۰ ۵۱ ۵۳ ۵۴ ۵۵ ۵۶ ۷۰ ۸۹

جدول ۲- ۵ محاسبه میانه با توزیع فراوانی نمرات ۱۰۵ دانش آموز در

آزمون آمار

نمره (X)	فراوانی (f)	فراوانی تجمعی (cf)
۳۵	۱	۱۰۵
۳۴	۳	۱۰۴
۳۳	۱۵	۱۰۱
۳۲	۸	۸۶
۳۱	۱۴	۷۸
۳۰	۱۴	۶۴
۲۹	۱۶	۵۰
۲۸	۱۰	۳۴
۲۷	۸	۲۴
۲۶	۴	۱۶
۲۵	۱	۱۲
۲۴	۳	۱۱
۲۳	۱	۸
۲۲	۳	۷
۲۱	۴	۴

$c f b = ۵۰$

$$Md = L_{۲۹/۵۰} + \left(\frac{\frac{۱۰۵}{۲} - ۵۰}{۱۴} \right) ۱ = ۲۹/۶۷۸$$

نمره ۵۴ که میانه این توزیع است، بهترین معرف است. شاخصی که نمره‌های منفرد ۷۰ و ۸۹ را در نظر بگیرد، حتماً برآورد بالایی از نمره معرف به دست خواهد داد.

// میانگین

اندازه گرایش مرکزی که بیشترین مورد استفاده را دارد، میانگین است که معمولاً به نام

متوسط یا متوسط حسابی شناخته می‌شود. میانگین، عبارت است از مجموع تمام مقادیر در یک توزیع، تقسیم بر تعداد موردها. میانگین به زبان فرمول عبارت است از:

$$\bar{X} = \frac{X_1 + X_2 + X_3 + \dots + X_n}{N} \quad (5-2)$$

که معمولاً به این صورت نوشته می‌شود:

$$\bar{X} = \frac{\sum X}{N} \quad (5-3)$$

$\bar{X}$ = میانگین

Σ = مجموع

X = نمره خام

N = تعداد موردها

با کاربرد فرمول ۵-۲ در مورد نمره‌های هوشبهر (IQ) زیر، در می‌یابیم که میانگین آنها ۱۱۱ است.

۱۰۰ ۱۰۹ ۱۱۹ ۱۰۱ ۱۱۵ ۱۲۱ ۱۱۲ :نمره IQ

$$\bar{X} = \frac{112 + 121 + 115 + 101 + 119 + 109 + 100}{7} = \frac{777}{7} = 111$$

توجه داشته باشید که در این محاسبه نمره‌ها مرتب نشدند. برای یافتن میانگین یک مجموعه از اندازه‌ها، نیازی به مرتب کردن نمره‌ها نیست.

از آنجا که میانگین یک متوسط حسابی است، آن را یک آماره فاصله‌ای به شمار می‌آورند. استفاده از آن برای داده‌های فاصله‌ای یا نسبی مناسب، اما برای داده‌های اسمی یا رتبه‌ای نامناسب است.

محاسبه میانگین از یک توزیع فراوانی

اگر داده‌ها به صورت یک توزیع فراوانی مرتب شوند، مجموع نمره‌ها را می‌توان با ضرب کردن در فراوانی آن، جمع کردن حاصلضربها و سپس تقسیم نتیجه بر تعداد موردها به دست آورد. فرمول محاسبه میانگین از یک توزیع فراوانی عبارت است از:

$$\bar{X} = \frac{\sum fX}{N} \quad (۴-۵)$$

جدول ۳-۵ محاسبه میانگین را برای داده‌های ارائه شده در جدول ۱-۵ نشان می‌دهد. آشکار است که کاربرد فرمول ۴-۵ برای آن توزیع فراوانی، کار محاسبه میانگین آن توزیع را آسان کرده است.

// مقایسه سه شاخص گرایش مرکزی

از آنجا که میانگین، یک آماره فاصله‌ای یا نسبی است، به طور کلی از میانه (آماره رتبه‌ای) یا نما (آماره اسمی) اندازه‌ای دقیق‌تر است. میانگین، مقدار هر نمره را به حساب می‌آورد. همچنین،

جدول ۳-۵ محاسبه میانگین برای توزیع فراوانی نمره‌های ۱۰۵ دانش‌آموز در

آزمون آمار

x	f	fx
۳۵	۱	۳۵
۳۴	۳	۱۰۲
۳۳	۱۵	۴۹۵
۳۲	۸	۲۵۶
۳۱	۱۴	۴۳۴
۳۰	۱۴	۴۲۰
۲۹	۱۶	۴۶۴
۲۸	۱۰	۲۸۰
۲۷	۸	۲۱۶
۲۶	۴	۱۰۴
۲۵	۱	۲۵
۲۴	۳	۷۲
۲۳	۱	۲۳
۲۲	۳	۶۶
۲۱	۴	۸۴
	$N=105$	$\sum fX=3076$

$$-X = \frac{3076}{105} = 29/295$$

همچنین، میانگین در بین سه اندازه گرایش مرکزی، ثابت ترین آنهاست؛ زیرا اگر از یک جامعه آماری، به طور تصادفی، چند نمونه انتخاب شود، میانگین های این نمونه ها، کمتر از میانه ها و نماهای آنها با هم تفاوت خواهند داشت. به همین دلیل میانگین بیشتر از دو شاخص دیگر، در پژوهشها مورد استفاده قرار می گیرد.

میانگین، بهترین شاخص عملکرد ترکیبی یک گروه یکپارچه است. با این همه، میانه بهترین شاخص عملکرد نوعی است. برای مثال، هیأت مدیره یک دبستان را در نظر بگیرید که در آمد سالانه آنها ۷۰۰۰۰، ۳۰۰۰۰، ۲۵۰۰۰، ۲۰۰۰۰ و ۲۰۰۰۰ دلار است. میانگین ۳۳۰۰۰ دلار، کل درآمد گروه را نسبت به تعداد اعضا نشان می دهد، اما این میانگین از درآمد همه اعضای هیأت مدیره بجز یک نفر، بیشتر است. در واقع میانه، ۲۵۰۰۰ دلار، تصویر بهتری از درآمد نوعی گروه را به دست می دهد.

/ شکل توزیعها

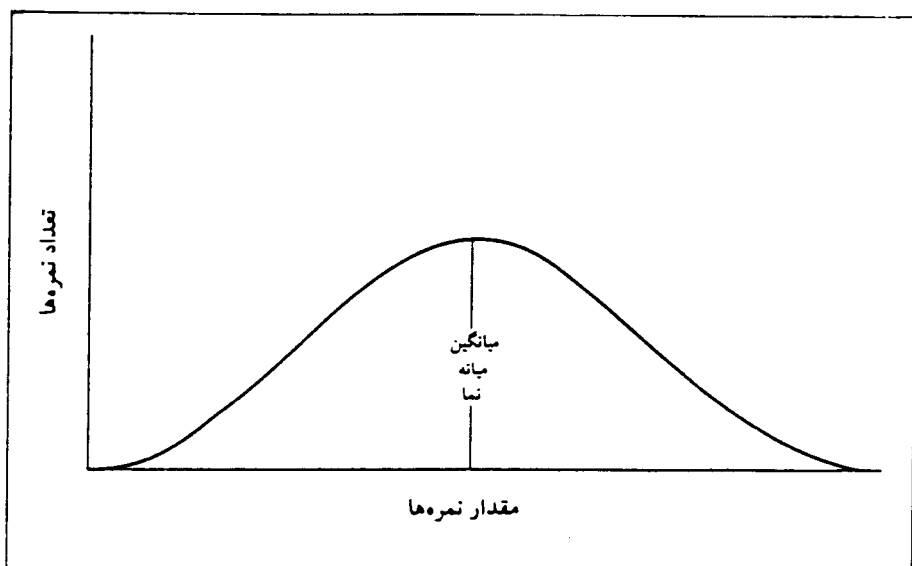
توزیعهای فراوانی می توانند شکلهای گوناگون داشته باشند. یک توزیع، هنگامی متقارن است که دو نیمه آن، تصویر انعکاسی (تصویر آینه ای)^(۱) یکدیگر باشند. در یک توزیع متقارن، مقادیر میانگین و میانه بر روی هم منطبق می شوند. اگر چنین توزیعی، به جای دو یا چند نما، تنها یک نما داشته باشد، سه شاخص گرایش مرکزی، همان طور که در شکل ۳ - ۵ نشان داده شده است، بر هم منطبق خواهند شد.

اگر توزیعی متقارن نباشد، آنرا توزیع دارای چولگی^(۲) توصیف می کنند. در توزیعهای دارای چولگی مقادیر اندازه های گرایش مرکزی با هم تفاوت دارند. در چنین توزیعی، مقدار میانگین، به دلیل آنکه از نمره های انتهایی تأثیر می پذیرد، همان طور که در شکل ۴ - ۵ و ۵ - ۵

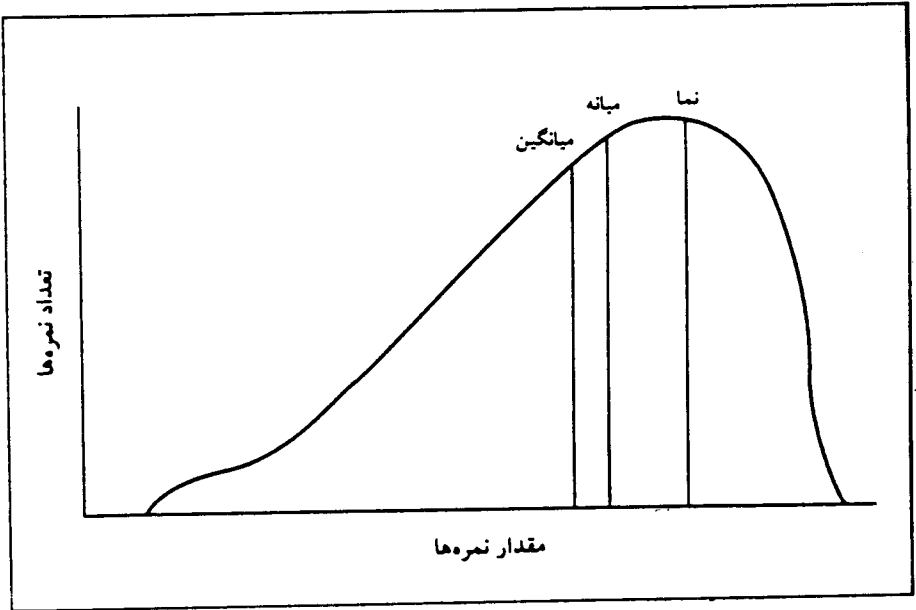
نشان داده شده است، به سمت انتهای توزیع، که نمره‌های افراطی در آن جای دارند، کشیده می‌شود. تأثیر نمره‌های افراطی بر میانه کمتر است؛ زیرا این شاخص از مقدار نمره تأثیر نمی‌پذیرد، بلکه از جایگاه نمره‌ها تأثیر می‌گیرد.

مقادیر افراطی هیچ تأثیری بر نما ندارند، زیرا این شاخص، هیچ رابطه‌ای با دو انتهای توزیع ندارد. زمانی که چولگی توزیع به سمت انتهای پایین‌تر باشد، یا چولگی منفی باشد، همیشه میانگین، کوچکتر از میانه است و میانه معمولاً کوچکتر از نما است (شکل ۴ - ۵). زمانی که چولگی توزیع به سمت انتهای بالاتر باشد، یا چولگی مثبت باشد، میانگین همیشه بیشتر از میانه و میانه معمولاً بیشتر از نما است (شکل ۵ - ۵). چولگی یک توزیع را می‌توان از مقایسه میانگین و میانه تشخیص داد، بدون آنکه لازم باشد که نمودار چند ضلعی یا ستونی آن را ترسیم کرد.

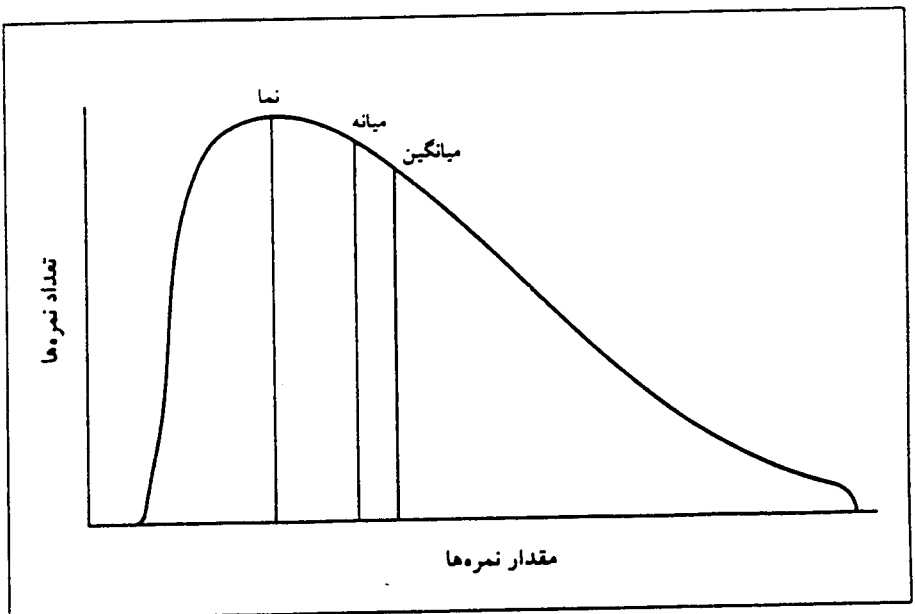
شکل ۳-۵ یک توزیع متقارن



شکل ۴-۵ یک توزیع دارای چولگی منفی



شکل ۵-۵ یک توزیع دارای چولگی مثبت



/// اندازه‌های تغییرپذیری

اگر چه شاخصهای گرایش مرکزی در توصیف داده‌ها به بیان مقدار متوسط یا اندازه معرف کمک می‌کنند، اما تصویر کاملی از توزیع را ارائه نمی‌کنند. امکان دارد میانگینهای دو توزیع یکسان باشند، در حالی که درجه پراکنش^(۱) یا تغییرپذیری نمره‌های آنها کاملاً با هم فرق کنند. امکان دارد در یک توزیع، نمره‌ها حول یک مقدار مرکزی گرد آیند، و در توزیع دیگر پراکنده باشند. برای روشن شدن موضوع، توزیع این نمره‌ها را در نظر بگیرید:

$$\bar{X} = \frac{175}{V} = 25 \quad (الف) \quad 24, 24, 25, 25, 25, 26, 26$$

$$\bar{X} = \frac{175}{V} = 25 \quad (ب) \quad 16, 19, 22, 25, 28, 30, 35$$

مقدار میانگین در هر دو توزیع ۲۵ است، اما درجه پراکندگی نمره‌ها به طور قابل ملاحظه‌ای با هم تفاوت دارد. نمره‌های توزیع الف، آشکارا، بسیار همگن‌تر از نمره‌های توزیع ب هستند. روشن است که ضرورت شاخصی که توزیع‌ها را بر مبنای تغییرپذیری نمره‌ها توصیف کند، وجود دارد. در آمار شاخص‌های متعددی برای این منظور وجود دارند. چهار شاخصی که بیشتر از بقیه استفاده می‌شوند، عبارتند از: دامنه، انحراف چارکی، واریانس و انحراف معیار.

// دامنه

ساده‌ترین شاخص تغییرپذیری، دامنه است. دامنه، تفاوت بین بالاترین و پایین‌ترین نمره در یک توزیع است و با کسر کوچکترین مقدار از بیشترین، به دست می‌آید. دامنه، به شکل فرمول عبارت است از:

$$R = X_b - X_i \quad (۵-۵)$$

دامنه = R

X_b = بیشترین مقدار در یک توزیع

X_i = کمترین مقدار در یک توزیع

برای مثال، در توزیعهای زیر، دامنه به ترتیب ۱۴ و ۶ است.

۲ ۱۰ ۱۱ ۱۲ ۱۳ ۱۴ ۱۶

۹ ۱۰ ۱۱ ۱۲ ۱۳ ۱۴ ۱۵

دامنه، شاخص تغییرپذیری ناپایداری است؛ زیرا اساس آن تنها بر دو مقدار، بیشترین و کمترین، بنا شده است. همانطور که بسادگی در دو توزیع بالا دیده می‌شود، دامنه، نشانگر استواری برای ماهیت پراکندگی اندازه‌ها، حول یک مقدار مرکزی نیست. به همین دلیل، استفاده از دامنه بیشتر به مقاصد بازیبانه محدود شده است. در برخی از گزارشهای تحقیقاتی، دامنه توزیع به عنوان یک آماره مورد استفاده قرار می‌گیرد، اما معمولاً استفاده از آن به همراه سایر اندازه‌های تغییرپذیری، نظیر انحراف چارکی و انحراف معیار صورت می‌پذیرد.

// انحراف چارکی

انحراف چارکی عبارت از نصف تفاوت بین چارکهای بالا و پایین در یک توزیع است. چارک بالا (Q_3) نقطه‌ای است که ۷۵ درصد از موارد در زیر آن قرار می‌گیرند. چارک پایین (Q_1) نقطه‌ای است که ۲۵ درصد از موردها در زیر آن قرار می‌گیرند. همچنین، چارک بالا را صدک هفتاد و پنجم و چارک پایین را صدک بیست و پنجم به شمار می‌آورند.

شیوه محاسبه Q_3 و Q_1 مشابه روشی است که برای محاسبه میانه به کار می‌رود. میانه، در واقع، چارک دوم است. در مورد Q_3 ، فرمول محاسبه عبارت است از:

$$Q_3 = L + \left(\frac{\frac{3N}{4} - cfb}{fw} \right) i \quad (5-6)$$

و برای Q_1 فرمول محاسبه عبارت است از:

$$Q_1 = L + \left(\frac{\frac{N}{4} - cfb}{fw} \right) i \quad (5-7)$$

Q_3 = چارک بالا

Q_1 = چارک پایین

L = حد پایین فاصله‌ای که چارک در آن قرار می‌گیرد

N = تعداد موردها در توزیع

cfb = فراوانی تراکمی، در زیر فاصله‌ای که حاوی چارک است

fw = فراوانی موارد، در داخل فاصله‌ای که حاوی چارک است

i = اندازه فاصله

هنگامی که مقادیر چارکهای اول و سوم پیدا شد، انحراف چارکی به این صورت محاسبه

می‌شود:

$$QD = \frac{Q_3 - Q_1}{2} \quad (8-5)$$

برای مثال، اگر در توزیع نمره‌ها، چارکهای بالا و پایین، به ترتیب، ۳۵ و ۱۵ باشند، مقدار انحراف چارکی، ۱۰ خواهد بود.

$$QD = \frac{35 - 15}{2} = 10$$

اندازه‌ای که انحراف چارکی به دست می‌دهد، عبارت است از نصف دامنه نمره‌هایی که در آن ۵۰ درصد میانی موردها قرار دارند. اگر گستردگی نمره‌ها زیاد باشد، مقدار انحراف چارکی بیشتر از زمانی خواهد بود که گستردگی نمره‌ها کوچک باشد. هر چند انحراف چارکی، نظیر دامنه، مبتنی بر دو مقدار یک توزیع (Q_3 و Q_1) است، اما اندازه تغییرپذیری مفیدتری از دامنه است. این امر به دلیل آن است که Q_3 و Q_1 بسیار پایدارتر از بالاترین و پایین‌ترین مقادیر یک توزیع هستند، که دامنه مبتنی بر آنهاست.

انحراف چارکی به خانواده آماری میانه تعلق دارد؛ زیرا یک آمارهٔ رتبه‌ای است. از این آماره غالباً به همراه میانه استفاده می‌شود. آن را همچنین دامنه نیمه چارکی نیز می‌نامند. انحراف چارکی، نظیر میانه، بویژه هنگامی مفید است که ما به یک اندازه نیاز داشته باشیم که تحت تأثیر چند نمره انتهایی قرار نگیرد. اگر انحراف چارکی برای درآمد خانوار در دهکده A، ۱۰۰۰۰ دلار و در دهکده B، ۸۰۰۰ دلار باشد، در می‌یابیم که درآمدهای خانوار در دهکده A بر اساس اندازه‌گیری پراکندگی در وسط توزیع، ناهمگن‌تر هستند. برنده شدن یک خانوار در یک

بخت آزمایی در دهکده B، تغییر بسیار اندکی در میانه ایجاد خواهد کرد و یا اصلاً تغییری ایجاد نخواهد کرد؛ هر چند این امر افزایش قابل ملاحظه‌ای در میانگین و انحراف معیار به وجود می‌آورد.

// واریانس و انحراف معیار

واریانس و انحراف معیار مفیدترین اندازه‌های تغییر پذیری هستند. هر دوی آنها بر پایهٔ نمره انحراف از میانگین بنا شده‌اند - نمره‌هایی که تفاوت بین یک نمره خام و میانگین را نشان می‌دهند. فرمول یک نمره انحراف از میانگین، عبارت است از $x = X - \bar{X}$. نمره‌های خام زیر میانگین، نمره انحراف از میانگین منفی و نمره‌های خام بالای میانگین، نمرهٔ انحراف از میانگین مثبت خواهند داشت. بنا به تعریف، مجموع نمره‌های انحراف از میانگین در یک توزیع، همیشه صفر است. بنابراین اگر کسی بخواهد از نمره‌های انحراف از میانگین برای محاسبه اندازه‌های تغییر پذیری استفاده کند، باید برای این واقعیت که $\sum x = 0$ است، چاره‌ای بیندیشد. شیوهٔ مورد استفاده، مربع کردن همهٔ نمره‌های انحراف از میانگین است تا همه آنها به اعداد مثبت تبدیل شوند. اگر ما پس از آن، مربع نمره‌های انحراف از میانگین را با هم جمع کنیم و به تعداد نمره‌ها تقسیم کنیم، میانگین مربع نمره‌های انحراف از میانگین، یا واریانس را به دست خواهیم آورد. واریانس به زبان ریاضی عبارت است از:

$$\sigma^2 = \frac{\sum x^2}{N} \quad (9-5)$$

σ^2 = واریانس

Σ = مجموع

انحراف هر نمره از میانگین، $(X - \bar{X})$ یا به عبارت دیگر نمره انحراف از میانگین x

N = تعداد موردها در توزیع

از آنجا که هر نمره انحراف از میانگین به توان دوم می‌رسد، لزوماً واریانس با واحدی بیان می‌شود که مربع واحد اصلی اندازه است. برای مثال، امکان دارد که دریابیم که واریانس قد شاگردان یک کلاس ۹ اینچ مربع است. این واریانس به ما می‌گوید که این کلاس از نظر قد ناهمگن تر از کلاسی است که واریانس آن ۴ اینچ مربع است و همگن تر از کلاسی است که واریانس آن ۱۶ اینچ مربع است.

در بسیاری از موارد کارشناسان آموزش و پرورش شاخصی را ترجیح می‌دهند که داده‌ها را به

همان واحد اندازه‌گیری داده‌های اصلی خلاصه کند. انحراف معیار (σ)، ریشه دوم (جذر) واریانس، چنین شاخصی را به دست می‌دهد. بنا به تعریف، انحراف معیار، ریشه دوم (جذر)، میانگین نمره‌های انحراف از میانگین است. اگر این تعریف را به زبان نمادین ریاضی بنویسیم، خواهیم داشت:

$$\sigma = \sqrt{\frac{\sum x^2}{N}} \quad (۱۰-۵)$$

برای وضوح بیشتر، جدول ۴-۵ را در نظر بگیرید. ستون اول این جدول توزیع نمره‌های آزمون ده دانش‌آموز را نشان می‌دهد. میانگین این توزیع ۶ است. ستون دوم، انحراف از میانگین هر نمره را ارائه می‌کند. برای مثال، انحراف نمره ۸ از میانگین، ۲+ است، انحراف نمره ۵ از میانگین، ۱- است، و الی آخر. ستون سوم، مربع هر یک از نمره‌های انحراف از میانگین را نشان می‌دهد. مجموع مربع نمره‌های انحراف از میانگین ۹۰ است. با جایگذاری این مقدار در فرمول و تقسیم آن بر ۱۰ (تعداد موردها) به عدد ۹ می‌رسیم، که میانگین مربع نمره‌های انحراف از میانگین است. ریشه دوم (جذر) این عدد ۳ است، که انحراف معیار این توزیع می‌باشد.

جدول ۴-۵ محاسبه انحراف معیار

(۱) X	(۲) $x = X - \bar{X}$	(۳) $x^2 = (X - \bar{X})^2$
۱۰	+۴	+۱۶
۹	+۳	+۹
۹	+۳	+۹
۸	+۲	+۴
۷	+۱	+۱
۶	۰	۰
۵	-۱	+۱
۳	-۳	+۹
۲	-۴	+۱۶
۱	-۵	+۲۵
$\Sigma X = ۶۰$		$\Sigma x^2 = ۹۰$

$$\bar{X} = \frac{۶۰}{۱۰} = ۱۰$$

$$\sigma = \sqrt{\frac{۹۰}{۱۰}} = \sqrt{۹} = ۳$$

روش بالا، وقتی مناسب است که میانگین توزیع، یک عدد سر راست (گرد شده) باشد، اما در بیشتر موارد چنین نیست. بنابراین برای اجتناب از زحمت کارکردن با اعداد کسری (اعشاری) از بسط فرمول اصلی، فرمول زیر به دست آمده است. فرمول زیر همان نتیجه را با کار کمتر به دست می‌دهد.

$$\sigma = \sqrt{\frac{\sum X^2 = \frac{(\sum X)^2}{N}}{N}} \quad (۱۱-۵)$$

σ = انحراف معیار

$\sum X^2$ = مجموع مربعات هر نمره (یعنی، ابتدا

نمره‌ها به توان دوم می‌رسد و سپس

با هم جمع می‌شوند)

$(\sum X)^2$ = مربع مجموع نمره‌ها (نمره‌ها ابتدا

با هم جمع می‌شوند و سپس جمع

کل به توان دوم می‌رسد)

N = تعداد موردها

جدول ۵-۵ محاسبه انحراف معیار را بر اساس داده‌های جدول ۴-۵ با این فرمول نشان می‌دهد. در این جدول، ستون اول، نمره‌ها و مجموع آنها را نشان می‌دهد. ستون دوم، مربع هر نمره و جمع مربع نمره‌ها را نشان می‌دهد. بقیه جدول، کاربرد فرمول ۱۱-۵ را برای یافتن مقدار انحراف معیار نشان می‌دهد. توجه داشته باشید که مقدار حاصله، همان مقداری است که با استفاده از فرمول ۱۰-۵ در مورد همان داده‌ها به دست آمده است.

انحراف معیار به خانواده آماری میانگین تعلق دارد، یعنی نظیر میانگین، یک آماره فاصله‌ای یا نسبی است و محاسبه آن بر اندازه نمره‌های فردی در توزیع بنا شده است. انحراف معیار، یکی از اندازه‌های تغییر پذیری است که بیشترین مورد استفاده را دارد و به همراه میانگین به کار می‌رود.

جدول ۵-۵ محاسبه انحراف معیار با استفاده از فرمول ۵-۱۱

X	X^2	
۱۰	۱۰۰	$\sigma = \sqrt{\frac{450 - \frac{(60)^2}{10}}{10}}$ $= \sqrt{\frac{450 - 360}{10}}$ $= \sqrt{9}$ $= 3$
۹	۸۱	
۹	۸۱	
۸	۶۴	
۷	۴۹	
۶	۳۶	
۵	۲۵	
۳	۹	
۲	۴	
۱	۱	
$\Sigma X = 60$	$\Sigma X^2 = 450$	

/// نمره‌های استاندارد

ما اغلب دوست داریم مقایسه‌ای بین موقعیت‌های نسبی یک فرد در دو آزمون متفاوت صورت دهیم. مقایسه درست دو نمره تنها در صورتی امکان دارد که دو آزمون، میانگین و انحراف معیار یکسان داشته باشند، اما این امر در عمل کمتر اتفاق می‌افتد. برای غلبه بر این مشکل، می‌توانیم اندازه‌ها (نمره‌های متفاوت) را به نمره‌های استاندارد برگردانیم. یک نمره استاندارد، که موارد استفاده گسترده‌ای دارد و در تحلیل‌های آماری نقش مهمی بازی می‌کند، نمره Z است، و تعریف آن عبارت است از فاصله یک نمره از میانگین که بر اساس واحدهای انحراف معیار اندازه‌گیری شده باشد. فرمول محاسبه نمره Z عبارت است از:

$$z = \frac{X - \bar{X}}{\sigma} = \frac{x}{\sigma} \quad (5-12)$$

X = نمره خام

$\bar{X}$ = میانگین توزیع

σ = انحراف معیار توزیع

x = نمره انحراف از میانگین

با کاربرد این فرمول، نمره‌ای که دقیقاً یک انحراف معیار بالاتر از میانگین باشد، +۱، و نمره‌ای که دقیقاً یک انحراف معیار زیر میانگین باشد، -۱ خواهد شد. نمره z برای نمره‌ای که مقدار عددی آن برابر میانگین باشد، صفر خواهد بود. برای روشن شدن مطلب، فرض کنید نمره دانشجویی در آزمون روان‌شناسی ۷۲ باشد، در حالی که میانگین توزیع آن ۷۸ و انحراف معیار آن ۱۲ است. همچنین، فرض کنید همان دانشجو در آزمون آمار ۴۸ گرفته باشد، در حالی که میانگین و انحراف معیار توزیع نمرات آزمون آمار ۵۱ و ۶ است. اگر اعداد مربوط را در فرمول ۱۲ - بگذاریم، برای هر آزمون یک نمره z خواهیم داشت.

$$\begin{array}{ll} \text{آمار} & \text{روان‌شناسی} \\ z_2 = \frac{48 - 51}{6} = -0.5 & z_1 = \frac{72 - 78}{12} = -0.5 \end{array}$$

هر دو این نمره‌های استاندارد به توزیع z تعلق دارند، که بنا به تعریف، میانگین و انحراف معیار آن صفر و یک است، و بنابراین به طور مستقیم قابل مقایسه‌اند. در این مثال روشن است که نمره ۷۲ در آزمون روان‌شناسی و نمره ۴۸ در آزمون آمار با هم برابر هستند، یعنی هر دوی این نمره‌ها نمایانگر سطح نسبی عملکرد همسان می‌باشند. به عبارت دیگر، موقعیت دانشجویی که این نمره‌ها را کسب کرده است، در مقایسه با عملکرد همکلاسانش، در هر دو آزمون همسان است. بدون استفاده از روش نمره z ، انجام چنین مقایسه‌ای بسیار دشوار خواهد بود.

اکنون مثال دیگری می‌زنیم: فرض کنید دانشجوی دیگری که همان آزمون‌ها را گذرانده است، در آزمون روان‌شناسی نمره ۸۱ و در آزمون آمار نمره ۵۳ کسب کرده است. مانند مثال قبل، مقایسه این نمره‌های خام برای نشان دادن اینکه این دانشجو در کدام آزمون بهتر عمل کرده، مشکل است. تبدیل این نمره‌ها به نمره z مقایسه را آسان می‌سازد. با استفاده از فرمول ۱۲ - ۵، مقادیر z_1 و z_2 را برای این مورد، پیدا می‌کنیم.

$$\begin{array}{ll} \text{آمار} & \text{روان‌شناسی} \\ z_2 = \frac{53 - 51}{6} = +0.33 & z_1 = \frac{81 - 78}{12} = +0.25 \end{array}$$

این نتیجه شگفت‌آور نشان می‌دهد که نمره ۵۳ در آزمون آمار، نمایانگر عملکرد نسبی اندک بهتری از نمره ۸۱ در آزمون روان‌شناسی است. این دانشجوی، در مقایسه با همتایانش، در آمار کار بهتری نسبت به روان‌شناسی داشته است.

یکی از اشکالات نمره Z آن است که می‌باید با مقادیر منفی و کسره‌های اعشاری سروکار داشت. ما می‌توانیم به هنگام مقایسه بین نمره‌هایی از توزیع‌های مختلف، برای غلبه بر این دشواریها نمره Z را به مقیاس نمره استاندارد دیگری تغییر دهیم که اعداد منفی و اعشاری در بر ندارد.

یکی از رایج‌ترین نمره‌های استاندارد، توزیع نمره Z است که دارای میانگین ۵۰ و انحراف معیار ۱۰ است. برای تبدیل نمره Z به نمره Z مقدار Z را در ۱۰ ضرب می‌کنیم و با ۵۰ جمع می‌کنیم.

فرمول نمره Z چنین است:

$$Z = 10z + 50 = 10 \left(\frac{X - \bar{X}}{\sigma} \right) + 50 \quad (5-13)$$

فرض کنید نمره دانشجویی در آزمون زبان اسپانیولی ۲۱ است. اگر میانگین نمره‌ها در این آزمون ۲۷ و انحراف معیار ۶ باشد، نمره Z عبارت خواهد بود از $(21 - 27) / 6$ ، که می‌توان آن را مستقیماً در فرمول Z قرار داد:

$$Z = 10 = \left(\frac{21 - 27}{6} \right) + 50 = 40$$

تبدیل نمره Z به نمره Z نه تنها ما را قادر می‌سازد که با اعداد صحیح کار کنیم، بلکه همچنین، سبب می‌شود که از نکات منفی روان‌شناختی که در توصیف عملکرد آزمودنیها با اعداد منفی وجود دارد، اجتناب کنیم.

آموزگارانی که می‌خواهند موقعیت شاگردان خود را در آزمونهای متوالی مقایسه کنند، یا می‌خواهند تمام نمراتی را که شاگردان در آزمونهای مختلف یک دوره درسی کسب کرده‌اند، با هم جمع و یک توزیع کلی فراهم کنند، می‌توانند نمرات خام دانشجویان را به نمره Z یا نمره Z

برگرداندند تا به هر مجموعه از نمره‌ها وزن یکسانی داده باشند. اکنون باید روشن شده باشد که جمع کردن و میانگین گرفتن از نمره‌هایی که به توزیعهای مختلفی تعلق دارند و دارای میانگین و انحراف معیار متفاوتی هستند، بدون تبدیل (برگرداندن) آنها به نوعی از نمره‌های استاندارد، از نظر آماری تأیید نمی‌شود.

علاوه بر Z ، توزیعهای نمره استاندارد تبدیل یافته دیگری نیز وجود دارند. برای تبدیل یک توزیع نمره‌ها به یک توزیع استاندارد شده جدید، تنها لازم است که نمره Z را در انحراف معیار دلخواه ضرب و با میانگین دلخواه جمع کنیم. فرمول کلی عبارت است از:

$$A = \mu_A + \sigma_A (Z) \quad (13a - 5)$$

A = نمره استاندارد در مقیاس جدید

μ_A = میانگین مقیاس استاندارد جدید

σ_A = انحراف معیار برای مقیاس استاندارد جدید

برای مثال، هر قسمت از امتحان ورودی بعد از لیسانس (GRE)^(۱) یک میانگین ۵۰۰ و یک انحراف معیار ۱۰۰ برای توزیع تبدیل یافته خود دارد. اگر شما در بخش کلامی GRE، ۱/۵ انحراف معیار بالای میانگین ($z = 1/5$) قرار گرفته باشید، نمره شما ۶۵۰ [$500 + 100(1/5) = 650$] گزارش خواهد شد. اگر نمره کمی شما ۵۰۰ باشد، شما نمره‌ای به دست آورده‌اید که دقیقاً برابر میانگین است.

مقیاس هوش استانفورد - بینه یک مقیاس نمره استاندارد با میانگین ۱۰۰ و انحراف معیار ۱۶ است. بنابراین برای تبدیل یک نمره خام آزمون استانفورد - بینه به مقیاس استاندارد، ابتدا باید Z آن را تعیین و سپس آن را در ۱۶ ضرب و با ۱۰۰ جمع کرد.

تبدیل مجموعه‌ای از نمره‌ها به نمره‌های استاندارد، شکل توزیع اصلی را تغییر نمی‌دهد. اگر توزیع نمره‌ها چولگی داشته باشد، نمره‌های استاندارد حاصله نیز یک توزیع دارای چولگی تولید خواهند کرد. تنها اگر توزیع اصلی نرمال باشد، نمره‌های استاندارد یک توزیع نرمال تولید خواهند کرد.

/// منحنی نرمال

دیده شده است که توزیع بسیاری از اندازه‌های فیزیکی و روانی، وقتی که به صورت چند ضلعی فراوانی ترسیم شوند، شکل زنگوله را به خود می‌گیرند. برای مثال، اگر قد پسران آمریکایی را در دهمین روز تولد اندازه بگیریم، در خواهیم یافت که قد اکثر کودکان نزدیک میانگین است، قد تعداد کمتری نیز اندکی بالاتر و پایین تر از میانگین می‌باشد. البته هر چقدر از میانگین دور شویم، در هر سطح بلندی قد، پسران کمتری یافت خواهند شد. یک نمودار چند ضلعی که این توزیع را نشان بدهد، کاملاً شبیه چندضلعی نظری است که منحنی نرمال خوانده می‌شود. جداول عرض و سطوح منحنی نرمال پیش از این محاسبه شده‌اند. سطوح منحنی نرمال در جدول ۱ - A ضمیمه، نشان داده شده‌اند. این نمودار چند ضلعی فرضی، فراوانیهای مورد انتظار (یعنی نظری) همهٔ نمره‌های ممکن z را نشان می‌دهد. این چند ضلعی نشان می‌دهد که انتظار به وقوع پیوستن نمره‌های z نزدیک به صفر، بیشتر از مقادیر دیگر نمره‌های z است و هر چقدر یک نمره z دورتر از صفر باشد، انتظار کمتری برای به وقوع پیوستن آن وجود دارد.

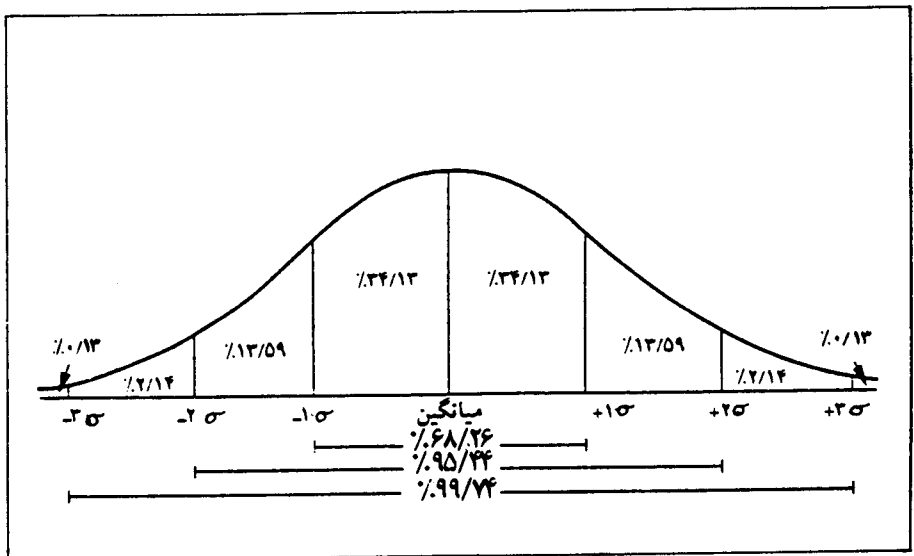
از آنجا که بسیاری از توزیعهای رخدادهای طبیعی به منحنی نرمال شباهت دارند، ثابت شده که این مدل نظری، بسیار مفید است. هر گاه دریابیم یا بر این باور باشیم که داده‌های واقعی به صورت توزیع شبیه منحنی نرمال هستند، می‌توانیم با استنتاج قیاسی برآوردهای مفیدی از خواص نظری منحنی نرمال به عمل آوریم. منحنی نرمال، توزیع مقارنی از اندازه‌ها است که در فواصل مشخص آن، در زیر میانگین و بالای میانگین، تعداد یکسانی جای گرفته‌اند. میانگین نقطه‌ای است که در زیر آن ۵۰ درصد از موردها و بالای آن ۵۰ درصد دیگر از موردها جای دارند. در چنین توزیعی میانه و نما مقادیر یکسانی هستند و بر میانگین منطبق می‌شوند. در یک منحنی نرمال، اکثر موردها نزدیک میانگین متمرکز می‌شوند. همچنانکه از میانگین در دو جهت دور می‌شویم، مقدار موردها کاهش می‌یابد. در منحنی نرمال تقریباً ۳۴ درصد از موردها بین میانگین و یک انحراف معیار بالای میانگین یا یک انحراف معیار زیر میانگین قرار می‌گیرند. سطح بین یک انحراف معیار از میانگین و دو انحراف معیار از میانگین، در هر دو سوی توزیع، شامل حدود ۱۴ درصد از موردها است. تنها حدود ۲ درصد از موردها بین دو انحراف معیار و سه انحراف معیار از میانگین و تنها حدود ۰/۱ درصد از موردها بالا یا پایین سه انحراف معیار از میانگین قرار دارند. این امر در شکل ۵-۶ به تصویر در آمده است.

می‌توان با مراجعه به جدول ۱ - A در ضمیمه، که سطوح منحنی نرمال را ارائه می‌کند، درصد موردهای پایین و بالای هر نمره z را در یک منحنی نرمال تعیین کرد. در این جدول، ستون اول شامل مقادیر مختلف z است.

در ستون دوم سطح زیر منحنی بین میانگین و هر مقدار z داده شده است. سرانجام، ستون سوم سطح باقیمانده از هر نمره z تا انتهای منحنی را نشان می‌دهد. بنابراین حاصل جمع سطوحی که در ستون دوم و سوم آمده است، به $۰/۵۰۰۰$ می‌رسد. برای مثال، نمره z برابر با $۰/۷۰$ را در نظر بگیرید. سطح بین این مقدار z و میانگین را می‌توان در ستون دوم یافت، این سطح $۰/۲۵۸۰$ است. این رقم می‌گوید که حدود ۲۶ درصد از موردها بین این مقدار z و میانگین توزیع قرار می‌گیرند.

چون میانگین منحنی نرمال بر میانه منطبق می‌شود، ۵۰ درصد از موردها زیر میانگین قرار می‌گیرند. $۰/۵$ را به $۰/۲۵۸۰$ اضافه می‌کنیم و نتیجه به ما می‌گوید می‌توانیم انتظار داشته باشیم که $۷۵/۸$ درصد از موردها زیر نمره z $۰/۷۰$ قرار بگیرند. ستون سوم نشان می‌دهد که $۲۴/۲$ درصد از موردها بالای نمره z $۰/۷۰$ قرار می‌گیرند.

شکل ۶-۵ درصد موردهای بین انحراف معیارهای متوالی در یک منحنی نرمال



زمانی که مقدار z منفی باشد، این روش بر عکس می‌شود، یعنی به جای آنکه درصد موردهای بین میانگین و مقدار z داده شده را با ۵۰ درصد جمع کنیم، آن را از ۵۰ درصد کم می‌کنیم. فرض کنید می‌خواهیم درصد موردهای زیر مقدار $z = ۰/۷۰ -$ را پیدا کنیم. سطح بین میانگین و نمره $z = ۰/۷۰ -$ ، یا به زبان درصد، ۲۵/۸ درصد از موردها است. با کم کردن ۲۵/۸ از ۵۰، عدد ۲۴/۲ را به دست خواهیم آورد.

این نتیجه نشان می‌دهد که در یک منحنی نرمال تنها در حدود ۲۴ درصد از نمره‌ها زیر نمره $z = ۰/۷۰ -$ قرار می‌گیرند. این مقدار را همچنین می‌توان در ستون سوم جدول سطوح زیر منحنی نرمال یافت. در این ستون مقدار داده شده برای یک نمره $z = ۰/۷۰$ برابر با ۰/۲۴۲۰ است. سطوح زیر منحنی نرمال بین دو نمره z دقیقاً تنها برای یک توزیع نرمال کاربرد دارد. بهره‌گیری از ویژگیهای یک منحنی نرمال در مورد توزیعهای دارای چولگی، در بهترین حالت، فقط یک تخمین به بار می‌آورد. بنابراین، ضرورت دارد که قبل از بکارگیری این ویژگیها، شکل توزیع بررسی شود.

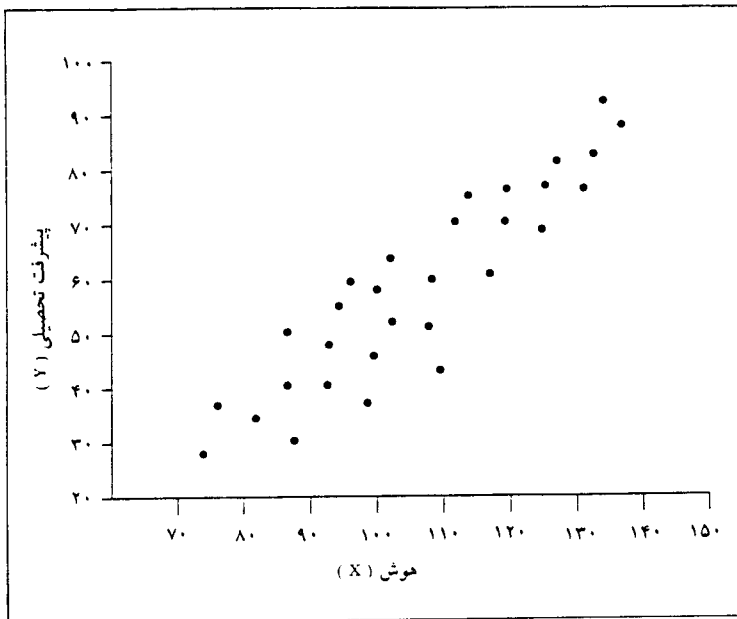
/// همبستگی

بحث ما در مورد روشهای آماری، تاکنون، درباره توصیف نمره‌های یک توزیع منفرد بوده است. اکنون می‌خواهیم درباره روشی که رابطه بین زوجهایی از نمره‌ها را نشان می‌دهد، بحث کنیم. روشهای آماری برای تعیین رابطه بین زوجهایی از نمره‌ها، شیوه‌های همبستگی نامیده می‌شوند. به طور نوعی، اندازه‌گیری دو متغیر برای هر عضو یک گروه معمول است و می‌توان وجود رابطه بین این اندازه‌گیریهای زوج شده را تعیین کرد. شیوه‌های همبستگی نشان می‌دهند که تغییر در یک متغیر، تا چه اندازه با تغییر در متغیر دیگر پیوستگی دارد. برای مثال، می‌دانیم که پیشرفت تحصیلی و هوش با هم رابطه دارند و بنابراین انتظار داریم دانشجویانی که از IQ بالایی برخوردارند، در آزمونهای پیشرفت تحصیلی نمرات بالاتر از میانگین کسب کنند. یک راه ساده نمایش این رابطه، ترسیم نمرات آزمون هوش و نمرات آزمون پیشرفت تحصیلی عده‌ای از افراد در یک جدول دو بعدی است، که آن را نمودار پراکندگی می‌نامند. یک مجموعه از نمرات بر محور افقی، به ترتیب از کمترین به بیشترین و از چپ به راست، رسم می‌شود. مجموعه دیگر نمرات بر محور عمودی به ترتیب از پایین به بالا رسم می‌شود. سپس موقعیت هر فرد در دو

آزمون با یک نقطه منفرد بر روی نمودار پراکندگی نشان داده می‌شود. در شکل ۷-۵ نمرات آزمون پیشرفت تحصیلی ۳۰ دانش‌آموز کلاس دهم در برابر نمرات آزمون هوش آنها ترسیم شده است. بررسی این شکل آشکار می‌کند که با بالا رفتن نمره هوش، نمره آزمون پیشرفت تحصیلی نیز به سمت بالا گرایش پیدا می‌کند.

ترسیم یک نمودار پراکندگی ما را قادر می‌سازد که هم جهت و هم شدت رابطه را مشاهده کنیم. جهت رابطه به مثبت یا منفی بودن رابطه اطلاق می‌شود. در شکل ۷-۵ نقطه‌ها طرحی را تشکیل می‌دهند که از سمت چپ پایین آغاز و به سمت راست بالا ختم می‌شود، به گونه‌ای که نمره‌های کم در یک متغیر (هوش) با نمره‌های کم در متغیر دیگر (پیشرفت تحصیلی) و نمره‌های زیاد در یک متغیر با نمره‌های زیاد در متغیر دیگر پیوستگی دارند (بنا به قرار داد، نمره‌های متغیر مستقل $[X]$ بر روی محور افقی و نمره‌های متغیر وابسته $[Y]$ بر روی محور عمودی ترسیم می‌شوند). یک چنین رابطه‌ای بین متغیرها مثبت نامیده می‌شود؛ زیرا نمره‌های زیاد با نمره‌های زیاد و نمره‌های کم با نمره‌های کم پیوستگی دارند.

شکل ۷-۵ یک نمودار پراکندگی نمایانگر رابطه بین هوش و پیشرفت تحصیلی



رابطه بین دو متغیر همیشه مثبت نیست. برخی از متغیرها با هم رابطه منفی دارند. برای مثال، روشن شده است که نرخ تولد و پایگاه اجتماعی - اقتصادی با هم رابطه منفی دارند، یعنی نرخ تولد با بالا رفتن پایگاه اجتماعی - اقتصادی کاهش می‌یابد. در یک رابطه منفی نمره‌های زیاد در یک متغیر با نمره‌های کم در متغیر دیگر پیوستگی می‌یابند و نقطه‌های موجود در نمودار پراکندگی از سمت چپ بالا به سمت راست پایین امتداد پیدا می‌کند. همچنین، یک نمودار پراکندگی نمره‌های z شدت رابطه بین متغیرها را آشکار می‌کند. اگر نقطه‌های موجود در نمودار پراکندگی، یک نوار باریک را تشکیل دهند؛ به گونه‌ای که اگر یک خط مستقیم در طول نوار رسم شود، نقطه‌ها نزدیک خط قرار گیرند، یک رابطه قوی بین متغیرها وجود دارد. به هر حال اگر نقطه‌های نمودار پراکندگی نمره‌های z کاملاً متفرق باشند، رابطه بین متغیرها نسبتاً ضعیف است. نمودار پراکندگی در شکل ۸-۵ رابطه‌های مثبت و منفی و قوی و ضعیف گوناگونی را نشان می‌دهد.

/// ضریبهای همبستگی

دسته‌ای از شاخصهای آماری برای این تهیه شده‌اند که هم، جهت (منفی یا مثبت) و هم، شدت رابطه بین متغیرها را نشان دهند. این شاخصها ضرایب همبستگی نامیده می‌شوند. محاسبه ضریب همبستگی بین دو متغیر، مقداری را به دست می‌دهد که دامنه آن ۱- تا ۱+ است. ضریب همبستگی ۱- رابطه منفی کامل را نشان می‌دهد، مقدار ۱+ متضمن یک رابطه مثبت کامل است و نقطه وسط این دامنه، صفر، نمایانگر آن است که هیچ رابطه‌ای بین دو متغیر وجود ندارد. یک ضریب همبستگی مثبت کامل، هنگامی به دست می‌آید که نمره z هر فرد در یک متغیر، از نظر اندازه و علامت با نمره z او در متغیر دیگر یکسان باشد. از سوی دیگر، یک ضریب همبستگی منفی کامل، هنگامی به دست می‌آید که نمره‌های z هر فرد از نظر اندازه یکسان و از نظر علامت متضاد باشند. یک همبستگی صفر، هنگامی به دست می‌آید که هیچ یک از این دو روند وجود نداشته باشند، یعنی جایگاه فرد در یک متغیر با جایگاه او در متغیر دیگر همبسته نباشد. ضریب همبستگی نزدیک به یک، چه ۱- و چه ۱+، رابطه بسیار بالا را نشان می‌دهد. چنین رابطه‌های بالایی انجام پیش‌بینی‌های دقیق را درباره یک متغیر، بر مبنای اطلاعاتی از یک متغیر دیگر، ممکن می‌سازند. از ضریب همبستگی منفی، همانند ضریب همبستگی مثبت، می‌توان برای پیش‌بینی استفاده کرد. در شکل ۸-۵ ضرایب همبستگی برای مجموعه‌هایی از داده‌ها به صورت نمره z آمده است. توجه کنید که در حالت همبستگی کامل

تمام نمره‌ها بر روی یک خط مستقیم قرار می‌گیرند. هر چه ضریب همبستگی به صفر نزدیکتر باشد، انحراف نمره‌ها از خط مستقیم بیشتر است. در حالت همبستگی صفر (و)، نمره‌ها بر سطح نگاره پراکنده هستند و در هیچ جهتی شکل معینی به خود نمی‌گیرند. ضرایب همبستگی در اندازه‌های روان‌شناختی و آموزش، به دلیل پیچیدگی این پدیده‌ها، ندرتاً به بیشترین حد، یعنی +۱ یا -۱ می‌رسند. در مورد این اندازه‌ها، هر ضریبی که بیشتر از به اضافه یا منهای ۰/۹ باشد، معمولاً همبستگی بسیار بالا محسوب می‌شود.

// همبستگی گشتاوری

ضریب همبستگی گشتاوری، که آماردان انگلیسی کارل پیرسون آن را ارائه کرده است و r پیرسون نامیده می‌شود، یک شاخص همبستگی است که بیشترین مورد استفاده را دارد. این ضریب هنگامی که مقیاس اندازه‌گیری فاصله‌ای یا نسبی باشد، به کار می‌رود. این شاخص را میانگین حاصل ضرب نمرات Z تعریف می‌کنند، یعنی نمره Z هر فرد برای متغیر (X) در نمره Z همان فرد برای متغیر دیگر (Y) ضرب می‌شود. حاصلضرب این نمره‌های Z زوج شده با هم جمع و به تعداد آنها تقسیم می‌شود. تعریف r پیرسون به زبان فرمول عبارت است از:

$$r = \frac{\sum Z_x Z_y}{N} \quad (۱۴-۵)$$

ضریب همبستگی گشتاوری پیرسون $r =$

مجموع حاصلضرب نمره‌های Z $\sum Z_x Z_y =$

تعداد نمره‌های زوج شده $N =$

به دلیل نحوه تعریف نمره‌های Z به زبان ریاضی، هنگامی که نمره Z همه افراد در دو متغیر X و Y هسان باشد، مجموع حاصلضربهای $Z_x Z_y$ با تعداد زوجهای نمره‌های Z برابر خواهد بود و میانگین حاصلضرب نمره‌های Z (همبستگی گشتاوری پیرسون) مساوی +۱ خواهد شد. اگر همخوانی کامل مثبت بین نمره‌های Z وجود داشته باشد، همبستگی گشتاوری +۱ خواهد شد. اگر نمره‌های Z از نظر عددی یکسان، اما از نظر علامت متضاد باشند، حاصلضرب منفی می‌شود و همبستگی گشتاوری پیرسون -۱ خواهد شد. در هر دو حالت وقتی که همه نمره‌ها بر روی یک نمودار پراکنده‌گی ترسیم شوند، بر روی یک خط مستقیم قرار می‌گیرند. اکنون بیایید با استفاده از

فرمول ۱۴-۵ همبستگی بین نمره‌های ۱۴ آزمودنی را در دو آزمون X (آمار توصیفی) و Y (آمار استنباطی)، آن گونه که در جدول ۶-۵ نشان داده شده است، محاسبه کنیم.

ستونهای (۲) و (۳)، به ترتیب، نمره‌های خام آزمودنیها (X) و نمره‌های انحراف از میانگین (x) را در آزمون آمار توصیفی نشان می‌دهند. ستونهای (۵) و (۶) نمره‌های خام آزمودنیها (Y) و نمره‌های انحراف از میانگین (y) را در آزمون آمار استنباطی ارائه می‌کنند. ستونهای (۴) و (۷) مربع نمره‌های انحراف از میانگین را که برای محاسبه انحراف معیار مورد استفاده قرار گرفته‌اند، نشان می‌دهند. ستونهای (۸) و (۹) نمره‌های Z مربوط به نمره‌های X و Y را که با استفاده از فرمول ۱۲-۵ محاسبه شده‌اند، معرفی می‌کنند. ستون (۱۰) حاصلضربهای $z_x z_y$ را نشان می‌دهد. مجموع این حاصلضربها ۱۱/۵۰ است. اکنون امکان یافتن ضریب همبستگی گشتاوری پیرسون، بین دو مجموعه از نمره‌ها، با کاربرد فرمول ۱۴-۵ فراهم آمده است.

$$r = \frac{11/50}{14} = +0/82$$

جدول ۶-۵ محاسبه r پیرسون بین دو مجموعه از نمره‌ها (X و Y)

(۱)	(۲)	(۳)	(۴)	(۵)	(۶)	(۷)	(۸)	(۹)	(۱۰)
آزمودنیها	X	نمره‌های x	x^2	Y	نمره‌های y	y^2	z_x	z_y	$z_x z_y$
۱	۱۸	+۳	۹	۲۸	۴	۱۶	+۱/۵	+۱	+۱/۵
۲	۱۸	+۳	۹	۳۰	۶	۳۶	+۱/۵	+۱/۵	+۲/۲۵
۳	۱۷	+۲	۴	۳۰	۶	۳۶	+۱	+۱/۵	+۱/۵
۴	۱۷	+۲	۴	۲۶	۲	۴	+۱	+۱/۵	+۱/۵
۵	۱۶	+۱	۱	۲۸	۴	۱۶	+۱/۵	+۱	+۱/۵
۶	۱۶	+۱	۱	۲۴	.	.	+۱/۵	.	.
۷	۱۵	.	.	۲۲	-۲	۴	.	-۱/۵	.
۸	۱۵	.	.	۲۰	-۴	۱۶	.	-۱	.
۹	۱۴	-۱	۱	۲۶	۲	۴	-۱/۵	+۱/۵	-۱/۲۵
۱۰	۱۴	-۱	۱	۲۲	-۲	۴	-۱/۵	-۱/۵	+۱/۲۵
۱۱	۱۳	-۲	۴	۲۴	.	.	-۱	.	.
۱۲	۱۳	-۲	۴	۱۸	-۶	۳۶	-۱	-۱/۵	+۱/۵
۱۳	۱۲	-۳	۹	۲۰	-۴	۱۶	-۱/۵	-۱	+۱/۵
۱۴	۱۲	-۳	۹	۱۸	-۶	۳۶	-۱/۵	-۱/۵	+۲/۲۵
	۲۱۰		۵۶	۳۳۶		۲۲۴			۱۱/۵۰

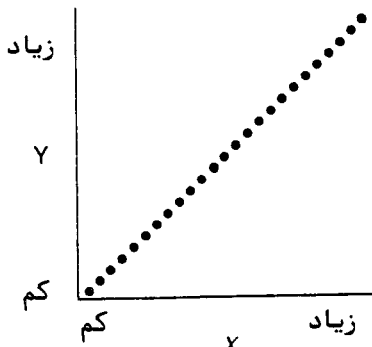
$$\bar{X} = \frac{\sum X}{N} = \frac{210}{14} = 15$$

$$\bar{Y} = \frac{\sum Y}{N} = \frac{336}{14} = 24$$

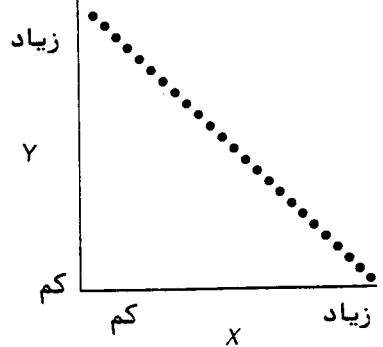
$$\sigma_x = \sqrt{\frac{\sum x^2}{N}} = \sqrt{\frac{56}{14}} = 2$$

$$\sigma_y = \sqrt{\frac{\sum y^2}{N}} = \sqrt{\frac{224}{14}} = 4$$

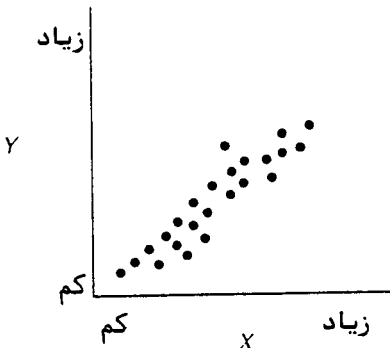
شکل ۸-۵ نمودار پراکندگی مقادیر انتخاب شده r



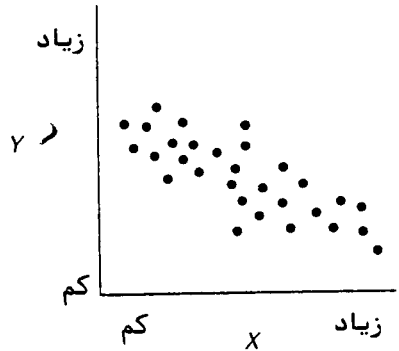
الف - همبستگی مثبت کامل (۱ +)



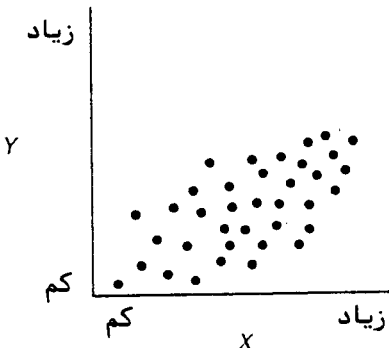
ب - همبستگی منفی کامل (۱ -)



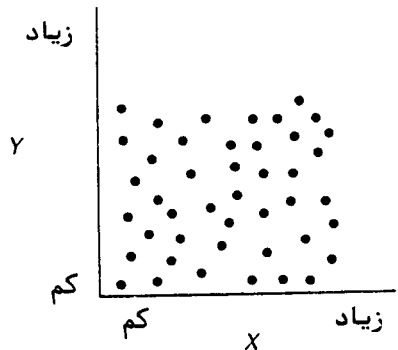
ج - همبستگی مثبت بالا (۰/۹۳ +)



د - همبستگی منفی متوسط (۰/۷۶ -)



ه - همبستگی مثبت متوسط (۰/۶۷ +)



و - همبستگی صفر

هنگامی که با تعداد زیادی آزمودنی سروکار داشته باشیم، فرآیند برگرداندن نمره‌های خام به نمره‌های Z کاری خسته کننده خواهد بود. این امکان وجود دارد که این مرحله را حذف کنیم و با استفاده از فرمول محاسباتی زیر، که از نظر ریاضی هم ارز فرمول ۱۴ - ۵ است، مستقیماً با نمره‌های خام عمل کنیم:

$$r = \frac{\Sigma XY - \frac{(\Sigma X)(\Sigma Y)}{N}}{\sqrt{\left[\Sigma X^2 - \frac{(\Sigma X)^2}{N} \right] \left[\Sigma Y^2 - \frac{(\Sigma Y)^2}{N} \right]}} \quad (۱۵ - ۵)$$

r پیرسون

ΣX = حاصل جمع نمره‌ها در توزیع X

ΣY = حاصل جمع نمره‌ها در توزیع Y

ΣXY = مجموع حاصلضرب نمره‌های X و Y زوج شده

ΣX^2 = مجموع نمره‌های مربع در توزیع X

ΣY^2 = مجموع نمره‌های مربع در توزیع Y

N = تعداد نمره‌های X و Y زوج شده (آزمودنیها)

اکنون بیاید با استفاده از همان نمره‌های خام قبلی، فرمول ۱۵ - ۵ را برای محاسبه ضریب همبستگی گشتاوری پیرسون به کار ببریم. اعداد و محاسبه‌های لازم در جدول ۵-۷ داده شده است. با گذاشتن مقادیر این جدول در فرمول ۱۵ - ۵، می‌توانیم r پیرسون را محاسبه کنیم: توجه کنید که با کاربرد این فرمول نه تنها برگردان نمره‌ها به مقادیر Z ضرورتی ندارد، بلکه همچنین محاسبه میانگینها و انحراف معیارهای دو توزیع نیز حذف شده است.

ضریب همبستگی گشتاوری به خانواده آماری میانگینها تعلق دارد. در محاسبه آن، اندازه هر یک از نمره‌ها در هر دو توزیع X و Y ، به حساب می‌آید، این ضریب، نظیر میانگین و انحراف معیار، یک آماره فاصله‌ای است که می‌توان آن را با داده‌های نسبی نیز مورد استفاده قرار داد. مفروضه‌ای که در پس ضریب همبستگی گشتاوری وجود دارد، این است که رابطه بین دو متغیر (X و Y) یک رابطه خطی است، یعنی یک خط مستقیم، توضیحی منطقی از رابطه یک

متغیر با متغیر دیگر به دست می‌دهد. اگر برای توضیح این رابطه به یک خط منحنی نیاز باشد، گفته می‌شود که یک رابطه منحنی الخط وجود دارد.

یک راه عملی برای دریافتن اینکه آیا رابطه بین دو متغیر خطی است یا منحنی الخط، بررسی نمودار پراکندگی داده‌ها است. شکل ۹ - ۵ دو نمودار را نشان می‌دهد، یکی از آنها (A) یک رابطه خطی را می‌نمایاند و دیگری (B) معرف رابطه منحنی الخط است.

اگر رابطه بین دو متغیر منحنی الخط باشد، محاسبه r پیرسون منتج به برآورد پائینی از رابطه بین دو متغیر خواهد شد که گمراه کننده است. در چنین حالتی باید شاخص دیگری، نظیر نسبت همبستگی (η) به کار رود. بحث دربارهٔ نسبت همبستگی را می‌توان در انتهای این فصل یافت.

جدول ۷ - ۵ محاسبه r پیرسون با استفاده از فرمول نمره‌های خام

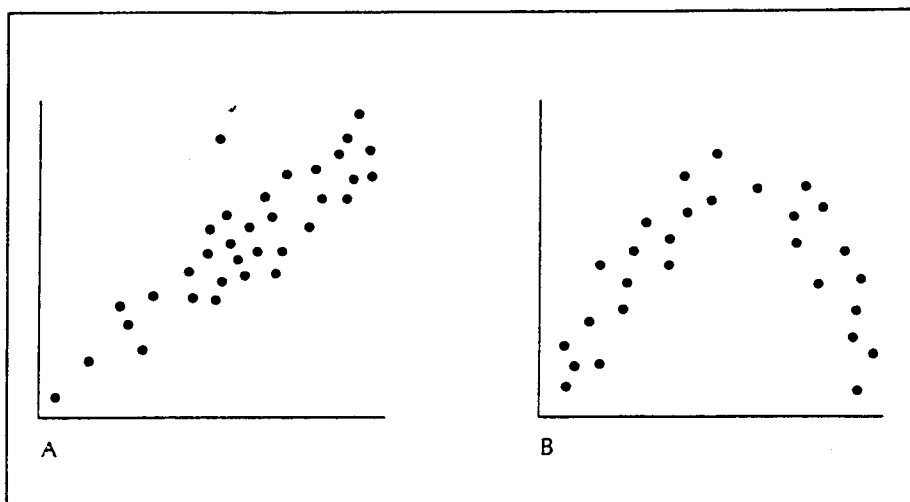
(۱)	(۲)	(۳)	(۴)	(۵)	(۶)
آزمودنها	X	Y	X^2	Y^2	XY
۱	۱۸	۲۸	۳۲۴	۷۸۴	۵۰۴
۲	۱۸	۳۰	۳۲۴	۹۰۰	۵۴۰
۳	۱۷	۳۰	۲۸۹	۹۰۰	۵۱۰
۴	۱۷	۲۶	۲۸۹	۶۷۶	۴۴۲
۵	۱۶	۲۸	۲۵۶	۷۸۴	۴۴۸
۶	۱۶	۲۴	۲۵۶	۵۷۶	۳۸۴
۷	۱۵	۲۲	۲۲۵	۴۸۴	۳۳۰
۸	۱۵	۲۰	۲۲۵	۴۰۰	۳۰۰
۹	۱۴	۲۶	۱۹۶	۶۷۶	۳۶۴
۱۰	۱۴	۲۲	۱۹۶	۴۸۴	۳۰۸
۱۱	۱۳	۲۴	۱۶۹	۵۷۶	۳۱۲
۱۲	۱۳	۱۸	۱۶۹	۳۲۴	۲۳۴
۱۳	۱۲	۲۰	۱۴۴	۴۰۰	۲۴۰
۱۴	۱۲	۱۸	۱۴۴	۳۲۴	۲۱۶
$N = ۱۴$	$\Sigma X = ۲۱۰$	$\Sigma Y = ۳۳۶$	$\Sigma X^2 = ۳۲۰۶$	$\Sigma Y^2 = ۸۲۸۸$	$\Sigma XY = ۵۱۳۲$

$$r = \frac{5132 - \frac{(210)(336)}{14}}{\sqrt{\left[3206 - \frac{(210)^2}{14}\right] \left[8288 - \frac{(336)^2}{14}\right]}} = +0.82$$

// تفسیر r پیرسون

هنگامی که دو متغیر در جهت مثبت رابطه بالایی با هم داشته باشند، همبستگی بین آنها به سوی $+1$ میل می‌کند. زمانی که این دو متغیر در جهت منفی با هم رابطه بالایی داشته باشند، همبستگی آنها به سوی -1 میل می‌کند. هنگامی که رابطه کمی بین متغیرها وجود داشته باشد، همبستگی به سمت صفر میل می‌کند. بنابراین r پیرسون شاخص بامعنایی برای نمایاندن رابطه بین متغیرها فراهم می‌آورد، که علامت ضریب همبستگی جهت رابطه را نشان می‌دهد و تفاوت بین ضریب همبستگی و صفر، درجه رابطه را مشخص می‌کند.

شکل ۹-۵ همبستگی خطی (A) و منحنی الخط (B)



به هر روی، در تفسیر ضریب همبستگی باید نکات زیر را به خاطر داشت:

- ۱- همبستگی لزوماً علیت را نشان نمی‌دهد. هنگامی که در می‌یابیم دو متغیر با هم همبستگی دارند، این امر گویای آن است که جایگاه نسبی در یک متغیر با جایگاه نسبی در متغیر

دیگر همبسته است. این موضوع لزوماً به معنای آن نیست که تغییر در یک متغیر باعث تغییر در متغیر دیگر می‌شود. ما در مثال خود همبستگی $+0.82$ را بین نمره‌های آزمون آمار توصیفی و نمره‌های آمار استنباطی یافتیم. این ضریب همبستگی به ما می‌گوید فردی که در یکی از این دو آزمون نمره بالای متوسط می‌گیرد، احتمالاً در آزمون دیگر نیز نمره بالای متوسط کسب خواهد کرد. ما نمی‌توانیم بگوییم که عملکرد عالی در یک آزمون موجب عملکرد عالی در آزمون دیگر می‌شود. نمره‌های هر دو آزمون ممکن است نتیجه دلایل دیگری، نظیر استعداد ریاضی افراد شرکت کننده در آزمون باشد.

برای روشن شدن موضوع از مثال دیگری استفاده می‌کنیم، فرض کنید که ما همبستگی مثبت بالایی بین ثروت خانوادگی و هوش کودکان آن خانواده‌ها به دست آورده‌ایم. چنین همبستگی بالایی به هیچ وجه نمایانگر یک رابطه علت و معلولی بین این دو متغیر نیست. برای این افراد، نه ثروت لزوماً هوش زیاد خلق می‌کند و نه هوش لزوماً ایجاد ثروت می‌کند.

۲- اندازه همبستگی، تا حدی، تابعی از تغییر پذیری دو توزیعی است که می‌خواهیم همبستگی آنها را تعیین کنیم. محدود کردن دامنه نمره‌هایی که قرار است با هم همبسته شوند، درجه رابطه مشاهده شده بین دو متغیر را کاهش می‌دهد. برای مثال، مشاهده شده است که موفقیت در بازی بسکتبال، با قد رابطه دارد، یعنی می‌توانیم بگوییم، هر اندازه قد کسی بلندتر باشد، احتمال بیشتری وجود دارد که این ورزش را بهتر انجام دهد. این حکم در مورد یک جمعیت بزرگ، که در آن تنوع قد وجود دارد، صادق است؛ هر چند، در یک تیم بسکتبال که همه اعضای آن بلند قد هستند، ممکن است هیچ رابطه‌ای بین قد و موفقیت نباشد، یا رابطه کمی وجود داشته باشد؛ زیرا دامنه قد در اعضای آن تیم محدود شده است.

در دانشکده‌ای که دانشجویان خود را بر مبنای دامنه گسترده نمره‌های آزمون استعداد تحصیلی^(۱) می‌پذیرد، انتظار خواهیم داشت بین نمره‌های این آزمون و نمره‌های دروس دانشگاهی رابطه وجود داشته باشد. اما در دانشکده‌ای که تنها با نمره‌های بسیار بالای آزمون استعداد تحصیلی دانشجو می‌پذیرد، انتظار همبستگی بسیار کمی بین نمره‌های آزمون و نمره‌های دانشگاهی خواهیم داشت؛ زیرا در این مورد دامنه نمره‌های آزمون محدود شده است.

۳- در تفسیر ضریب همبستگی نباید آن را درصدی از همبستگی کامل محسوب کرد. از

آنجا که ضرایب همبستگی به شکل اعداد اعشاری بیان می‌شوند، افرادی که تبحری در علم آمار ندارند، گاهی ضرایب همبستگی را به عنوان درصدی از همبستگی کامل تفسیر می‌کنند. یک ضریب همبستگی $0/8$ نمایانگر 80 درصد رابطه کامل بین دو متغیر نیست. این تفسیر نادرست است؛ زیرا برای مثال، یک ضریب همبستگی $0/8$ رابطه‌ای را که دو برابر یک ضریب همبستگی $0/4$ است، بیان نمی‌کند. یک شیوه برای تشخیص درجه‌ای که می‌توان با آن یک متغیر را از روی متغیر دیگر پیش‌بینی کرد، محاسبه شاخصی به نام ضریب تعیین^(۱) است. ضریب تعیین مربع ضریب همبستگی است. احتمالاً بهترین شیوه برای معنابخشی به اندازه ضریب همبستگی ترسیم درجه پراکندگی اندازه‌های متفاوت همبستگی (همان طور که در شکل ۸ - ۵ نشان داده شده است) و آشنا شدن با اندازه همبستگی‌هایی است که عموماً بین متغیرهای مورد نظر مشاهده می‌شود.

۴- در تفسیر ضریب همبستگی از مطلق نگری اجتناب کنید. در تفسیر درجه همبستگی، هدف مورد نظر را به خاطر داشته باشید. برای مثال، ضریب همبستگی $0/5+$ ممکن است برای پیش‌بینی عملکرد آتی یک گروه از افراد کافی باشد، اما امکان دارد که استفاده از همین ضریب همبستگی برای پیش‌بینی عملکرد یک فرد در آینده، کار عاقلانه‌ای نباشد. پس ضریب $0/5+$ یک مقدار مطلق با مفهوم یکسان در هر دو حالت نیست.

// ضریب همبستگی رتبه‌ای

ضریب همبستگی گشتاوری پیرسون یک شاخص آماری است که برای یافتن رابطه بین دو مجموعه داده فاصله‌ای که به صورت خطی توزیع شده باشند، به کار می‌رود. ما گاهی در تحقیق می‌خواهیم ضریب همبستگی بین دو مجموعه از اندازه‌ها را که به صورت رتبه‌ای منظم شده‌اند، پیدا کنیم، یعنی به جای داده‌های فاصله‌ای با داده‌های رتبه‌ای سروکار داریم.

برای مثال، فرض کنید ما می‌خواهیم همبستگی رتبه‌هایی را که دو آموزگار در زمینه خلاقیت به یک گروه از دانش‌آموزان تخصیص داده‌اند، محاسبه کنیم. شاخصی که در چنین مواردی به کار می‌رود، ضریب همبستگی رو (رتبه) اسپیرمن^(۲) (ρ) است که با فرمول زیر محاسبه می‌شود.

$$\rho = 1 - \frac{6 \sum D^2}{N(N^2 - 1)} \quad (۵-۱۶)$$

ρ ضریب همبستگی رو اسپیرمن =

$\sum D^2$ = مجموع مربعات تفاوت‌های بین رتبه‌ها

N = تعداد موردها

برای روشن شدن مطلب، جدول ۵-۸ را که در آن رتبه‌بندی ۱۱ دانش‌آموز توسط دو آموزگار نشان داده شده است، در نظر بگیرید. ستونهای (۲) و (۳) جدول ۵-۸ به ترتیب نشان دهنده رتبه‌بندی آموزگار ۱ و آموزگار ۲ است ستون (۴) تفاوت، بین این رتبه‌بندی‌ها را نشان می‌دهد. برای مثال، تفاوت بین رتبه‌بندی آموزگاران در مورد دانش‌آموز A، ۳-، در مورد دانش‌آموز B، ۱- و الی آخر است. مجموع مقادیر این ستون همیشه صفر است. در ستون (۵) مربع این تفاوتها داده شده است. مجموع مقادیر D^2 ، ۲۶ و تعداد موردها ۱۱ است. هنگامی که این مقادیر در فرمول ۱۶-۵ جایگزین شوند، ضریب همبستگی رتبه‌ای ۰/۸۸+ به دست می‌آید.

جدول ۵-۸ محاسبه ضریب همبستگی بین دو مجموعه از رتبه‌ها

(۱) دانش‌آموز	(۲) رتبه‌بندی آموزگار اول R_1	(۳) رتبه‌بندی آموزگار دوم R_2	(۴) تفاوت D	(۵) D^2
A	۱	۴	-۳	۹
B	۲	۳	-۱	۱
C	۳	۱	+۲	۴
D	۴	۲	+۲	۴
E	۵	۵	۰	۰
F	۶	۶	۰	۰
G	۷	۸	-۱	۱
H	۸	۹	-۱	۱
I	۹	۷	+۲	۴
J	۱۰	۱۱	-۱	۱
K	۱۱	۱۰	+۱	۱
			۰	۲۶

$$\rho = 1 - \frac{(6)(26)}{11(121 - 1)} = +0/88$$

به هنگام رتبه‌بندی افراد یا اشیاء، به منظور یافتن همبستگی بین دو مجموعه رتبه‌بندی شده، احتمال دارد که دو یا چند فرد یا شیء رتبه‌ای همسان را به خود تخصیص دهند. برای مثال، زمانی که دو نفر در رتبه سوم قرار می‌گیرند، در واقع، آنها رتبه‌های سوم و چهارم آن مجموعه هستند و از آنجا که ناممکن است که بگوییم کدام یک باید در رتبه سوم و کدامیک در رتبه چهارم قرار گیرند، ضروری است که به هر دوی آنها، متوسط این جایگاه، در این مورد $3/5$ ، اختصاص داده شود. به نفر بعد، رتبه 5 اختصاص خواهد یافت.

گاهی از اوقات، ما می‌خواهیم رابطه بین مجموعه‌ای از رتبه‌ها و مجموعه‌ای از اندازه‌های فاصله‌ای، نظیر دسته‌ای از نمرات امتحانی را پیدا کنیم. این امکان وجود ندارد که رتبه‌ها را از داده‌های رتبه‌ای به داده‌های فاصله‌ای ارتقاء دهیم، بنابراین، ابتدا باید نمره‌ها را به رتبه تبدیل کنیم و سپس فرمول رواسپیرمن را به کار ببریم. به عنوان مثال، جدول ۹ - ۵ را در نظر بگیرید.

جدول ۹-۵ محاسبه همبستگی رتبه‌ای بین یک مجموعه از داده‌های رتبه‌ای و

یک مجموعه از داده‌های فاصله‌ای

(۱)	(۲)	(۳)	(۴)	(۵)	(۶)
دانش‌آموزان	رتبه‌بندی ۱	نمره‌ها	رتبه‌بندی ۲	D	D ^۲
جک	۱	۱۹	۱	۰	۰
لیندا	۲	۱۷	۳/۵	-۱/۵	۲/۲۵
لوسی	۳	۱۸	۲	۰.۱	۱
دیگ	۴	۱۷	۳/۵	-۰/۵	۰/۲۵
تام	۵	۱۵	۶	-۱	۱
مارشا	۶	۱۴	۸	-۲	۴
جان	۷	۱۵	۶	۰.۱	۱
دیوید	۸	۱۵	۶	۰.۲	۴
جوآن	۹	۱۲	۱۰	-۱	۱
آنا	۱۰	۱۳	۹	۰.۱	۱
جورج	۱۱	۸	۱۱	۰	۰
سو	۱۲	۵	۱۲	۰	۰
				۰	۱۵/۵

$$\rho = 1 - \frac{(6)(15/5)}{12(144 - 1)} = +0.95$$

ستون (۲) جدول، پیش‌بینی یک آموزگار را در مورد رتبه‌های گروهی از دانش‌آموزان در یک امتحان، پیش از انجام آن، نشان می‌دهد. ستون (۳) نمره‌های واقعی این دانش‌آموزان را پس از امتحان نشان می‌دهد. این آموزگار برای تعیین رابطه بین رتبه‌های پیش‌بینی شده و رتبه‌های واقعی، باید نمره‌های دانش‌آموزان را به رتبه تبدیل کند. ستون (۴) رتبه‌بندی دانش‌آموزان را بر مبنای نمره‌های امتحانی آنها نشان می‌دهد. توجه داشته باشید که هم لیندا و هم دیک نمره ۱۷ را به دست آورده‌اند و بنابراین در جایگاههای سوم و چهارم قرار می‌گیرند، از این رو، به هر دوی آنها رتبه یکسان ۳/۵، که متوسط رتبه‌های ۳ و ۴ است، داده شده است. همین موقعیت برای تام، جان و دیوید پیش آمده است، که هر سه نفر نمره ۱۵ گرفته‌اند و در جایگاههای پنجم، ششم و هفتم سهم شده‌اند. رتبه ۶ که متوسط رتبه‌های ۵، ۶ و ۷ است، به هر یک از این سه نفر تخصیص داده شده است. طریقه یافتن D و D^2 دقیقاً نظیر همان شیوه‌ای است که در جدول ۸-۵ نشان داده شد. ضریب همبستگی بین این دو مجموعه داده‌ها، $+0/95$ است.

همبستگی رتبه‌ای اسپیرمن حالت خاصی از همبستگی گشتاوری پیرسون است و فرمول آن از فرمول r پیرسون مشتق شده است. بنابراین، امکان استفاده از فرمول ۱۵-۵ برای محاسبه همبستگی بین دو مجموعه از رتبه‌ها وجود دارد. برای مثال، اگر ما فرمول ۱۵-۵ را در مورد رتبه‌های جدول ۸-۵ به کار ببریم، یک ضریب همبستگی همانند، یعنی $+0/88$ ، به دست خواهیم آورد، اما وجود هم رتبه‌ها یک استثنا ایجاد می‌کند. با حضور هم رتبه‌ها، نظیر آنچه که در جدول ۹-۵ دیدیم، استفاده از فرمول r پیرسون، دیگر یک ضریب همبستگی همانند با آنچه که از بکارگیری فرمول رتبه‌ای اسپیرمن به دست می‌آید، حاصل نخواهد کرد.

ضریب همبستگی رتبه‌ای رواسپیرمن، عضوی از خانواده آماری میانه است. این همبستگی یک آماره رتبه‌ای است که به منظور استفاده با داده‌های رتبه‌ای طراحی شده است. نظیر ضریب همبستگی گشتاوری، دامنه آن از ۱- تا ۱+ تغییر می‌کند. هنگامی که هر فرد، رتبه یکسانی در هر دو متغیر داشته باشد، همبستگی رتبه‌ای ۱+ خواهد شد و هنگامی که رتبه‌های آنان در یک متغیر، دقیقاً مخالف رتبه‌هایشان در متغیر دیگر باشد، ضریب رو ۱- خواهد شد. اگر اصلاً هیچ رابطه‌ای بین رتبه بندیها وجود نداشته باشد، ضریب همبستگی رتبه‌ای صفر خواهد شد. رواسپیرمن همسان r پیرسون تفسیر می‌شود.

// سایر شاخصهای همبستگی

علاوه بر شاخصهای پیرسون و اسپیرمن، شاخصهای همبستگی متعدد دیگری وجود دارند که برای یافتن شدت رابطه بین انواع مختلف متغیرها مناسب هستند. بخش بعدی، برخی از این شاخصها را بدون آنکه جزئیات محاسباتی آنان را مطرح سازد، معرفی می‌کند. به دانشجویان علاقه‌مند توصیه می‌شود برای آگاهی از شیوه‌های محاسباتی به دیگر کتابهای آماری مراجعه کنند.^(۱)

// همبستگی دو رشته‌ای و همبستگی دو رشته‌ای نقطه‌ای

روشهای همبستگی دو رشته‌ای و دو رشته‌ای نقطه‌ای برای مواقعی تهیه شده‌اند که می‌خواهیم رابطه بین یک متغیر فاصله‌ای یا نسبی پیوسته و یک متغیر اسمی دو مقوله‌ای را پیدا کنیم. استفاده از همبستگی دو رشته‌ای مستلزم این مفروضه است که متغیر دو مقوله‌ای ما، پیوسته و نرمال باشد. به سخن دیگر، دو مقولگی به طور تصنعی از یک متغیر پیوسته خلق شده است. برای مثال، می‌توان افراد را از نظر خلاقیت، بر مبنای نمره‌های یک آزمون خلاقیت، به زیر متوسط و بالای متوسط طبقه‌بندی کرد. اگر نمره‌های آزمون هوش (متغیر پیوسته) را با خلاقیت، که به شکل زیر متوسط و بالای متوسط اندازه‌گیری شده است (دو مقولگی مصنوعی)، همبسته کنیم، ضریب همبستگی دو رشته‌ای، اندازه مناسبی برای تعیین رابطه آن دو خواهد بود.

همبستگی دو رشته‌ای نقطه‌ای هنگامی به کار می‌رود که دو مقولگی، امری واقعی باشد. برای مثال، هنگامی که رابطه بین جنس و نمره را در یک آزمون استدلال مطالعه می‌کنیم، به دو مقوله، مقادیر عددی ۱ یا ۰ تخصیص می‌یابد و این مقادیر با مقادیر مربوط به متغیر پیوسته همبسته می‌شوند. دیگر متغیرهای دو مقوله‌ای واقعی عبارتند از: تبعه / غیر تبعه ایالات متحده، چپ دست / راست دست، فارغ‌التحصیل دانشگاه / دانشجو، سیگاری / غیرسیگاری و متغیرهایی نظیر آنها. همبستگی دو رشته‌ای نقطه‌ای، ساده‌سازی ریاضی r پیرسون است.

۱ - برای بحث بیشتر و جزئیات محاسباتی مراجعه کنید به:

G.A. Ferguson (1981). *Statistical Analysis in Psychology and Education* (5Th ed.). New York: McGraw- Hill. (2) A.L. Edwards (1976). *An Introduction to Linear Regression and Correlation*. San Francisco. (3) J.D. Wynne (1982). *Learning Statistics*. New York: Macmillan.

// همبستگی چهارخانه‌ای^(۱) و ضریب فی^(۲)

همبستگی چهارخانه‌ای و ضریب فی برای یافتن رابطه بین متغیرها، در مواردی که هر دو متغیر دو مقوله‌ای هستند، به کار می‌روند. ماهیت دو مقولگی، شاخصی را که باید مورد استفاده قرار گیرد، تعیین می‌کند. اگر هر دو متغیر، به طور واقعی پیوسته باشند، اما به طور مصنوعی به دو مقوله تبدیل شوند، شاخص مناسب، ضریب همبستگی چهارخانه‌ای است. اگر کسی بخواهد همبستگی بین دو متغیر دو مقوله‌ای واقعی را پیدا کند، ضریب فی، شاخصی است که باید مورد استفاده قرار گیرد. برای مثال، همبستگی چهارخانه‌ای برای یافتن رابطه بین خلایق و هوش، هنگامی که هر دو متغیر به دو مقوله زیر متوسط و بالای متوسط طبقه‌بندی شده‌اند، به کار خواهد رفت. این شاخص همبستگی، به طور گسترده مورد استفاده قرار نمی‌گیرد. با تبدیل متغیرهای پیوسته به متغیرهای دو مقوله‌ای، اطلاعات زیادی از دست می‌رود. ضریب فی برای توصیف رابطه بین جنس یک گروه از دانش‌آموزان دبیرستان و اینکه آیا آنها کمک هزینه تحصیلی دانشکده را به دست خواهند آورد یا نه، به کار خواهد رفت. جنس، به صورت مرد - زن و کسب کمک هزینه تحصیلی، به صورت بله - خیر، به دو مقوله تبدیل می‌شوند که هر دو، مقوله‌ای واقعی هستند. مقادیر عددی (۱ و ۰) به مقوله‌ها تخصیص می‌یابند و ضریب فی محاسبه می‌شود. ضریب فی نیز ساده‌سازی ریاضی χ^2 پیرسون است. جدول ۱۰ - ۵ ضرایب همبستگی را که با انواع مختلف مقیاسها به کار می‌روند، خلاصه می‌کند.

// نسبت همبستگی^(۳)

خطی بودن رابطه بین دو متغیر، مفروضه‌ای است که پشتوانه ضریب همبستگی گشتاوری می‌باشد. به هر روی، گاهی رابطه بین متغیرها منحنی الخط است، یعنی این رابطه به جای خط راست، با یک خط منحنی توصیف می‌شود. در چنین مواردی، کاربرد فرمول χ^2 پیرسون مناسب نیست؛ زیرا برآورد پایینی از درجه رابطه را به دست می‌دهد. هنگامی که رابطه بین X و Y منحنی الخط باشد، شاخص دیگری به نام نسبت همبستگی، اتا^(۴)، مورد استفاده قرار می‌گیرد. این شاخص، شامل تصحیحی برای غیرخطی بودن است و بنابراین برآورد دقیق‌تری از مقدار

وابستگی بین متغیرها ارائه می‌کند.

برای مثال، فرض کنید کسی می‌خواهد همبستگی بین سن و قدرت بدنی را پیدا کند. ما می‌دانیم که قدرت بدنی تا نقطه معینی با سن رابطه مثبت دارد. از آن پس، افزایش سن با کاهش قدرت بدنی همراه می‌شود. نسبت همبستگی برای استفاده در چنین مواردی شاخص مناسبی خواهد بود.

// همبستگی سهمی و همبستگی چند متغیری

روشهای همبستگی که تاکنون مورد بحث قرار گرفتند، تنها برای استفاده با دو متغیر مناسب هستند. به هر روی، مواردی پیش می‌آیند که ما با بیش از دو متغیر سروکار داریم. همبستگی سهمی روشی است که با استفاده از آن تعیین می‌کنیم وقتی اثر متغیر سوم حذف شود، چه همبستگی‌ای بین دو متغیر دیگر باقی می‌ماند. همبستگی می‌تواند بین دو متغیر واقع شود؛ زیرا هر دوی آنها با متغیر سوم همبستگی دارند. همبستگی سهمی این متغیر سوم را کنترل می‌کند. برای مثال، فرض کنید کسی به همبستگی بین سن عقلی و مهارت‌های روانی - حرکتی علاقه‌مند

جدول ۱۰ - ۵ انواع ضرایب همبستگی و انواع مقیاسهای متناسب با آنها

نوع مقیاس	ضریب همبستگی
۱. هر دو متغیر در مقیاس فاصله‌ای یا نسبی	۱. گشتاوری پیرسون
۲. هر دو متغیر در مقیاس اسمی	۲. رتبه‌ای اسپیرمن
۳. یک متغیر در مقیاس فاصله‌ای، دیگری یک متغیر دو مقوله‌ای واقعی در مقیاس اسمی	۳. دو رشته‌ای نقطه‌ای
۴. یک متغیر در مقیاس فاصله‌ای یا نسبی، متغیر دیگر مقوله‌ای مصنوعی (۱)	۴. دو رشته‌ای
۵. دو مقولگی مصنوعی (مقیاس اسمی) مورد استفاده در هر دو متغیر، هر دو دارای توزیع پیوسته زیربنایی	۵. چهارخانه‌ای
۶. هر دو متغیر دو مقوله‌ای واقعی (مقیاس اسمی)	۶. ضریب فی

۱ - یک دو مقولگی مصنوعی، متغیر پیوسته را با استفاده از یک نقطه انقطاع، به طور اختیاری، به دو طبقه تقسیم می‌کند - برای مثال، نمره‌های آزمونی که به دو طبقه قبول و رد تقسیم شده باشند. مثالهای دو مقولگی واقعی عبارتند از مرد - زن و زنده - مرده.

شده است. هر دوی این متغیرها به متغیر سوم، یعنی سن زمانی، وابسته‌اند. برای مثال، کودکان ۱۲ ساله، به طور کلی، دارای بلوغ عقلی بیشتری از کودکان ۸ ساله‌اند و همچنین آنها دارای مهارت‌های بدنی رشد یافته‌تری هستند. نمره‌های یک آزمون سن عقلی و یک آزمون روانی - حرکتی، با یکدیگر همبستگی خواهند داشت؛ زیرا هر دوی آنها با سن زمانی همبستگی دارند. در کار با چنین داده‌هایی، برای به دست آوردن یک اندازه‌ای از همبستگی که در آن اثر سن حذف شده است، از همبستگی سهمی استفاده می‌شود. همبستگی باقیمانده بین دو متغیر، پس از حذف همبستگی آن دو با متغیر سوم، یک همبستگی سهمی مرتبه اول نامیده می‌شود. از همبستگی سهمی می‌توان برای حذف اثر بیش از یک متغیر استفاده کرد؛ هر چند که از این روش، به دلیل تفسیر دشوار آن، اغلب استفاده نمی‌شود.

همبستگی چند متغیری نیز بیش از دو متغیر را در بردارد. این روش، محقق را قادر می‌سازد که بهترین شکل وزن دادن به دو یا چند متغیر مستقل را بیابد تا همبستگی بیشینه (ماکزیمم) را با یک متغیر وابسته به دست آورد. برای مثال، محققان برای پیش‌بینی معدل دوره دانشگاه از رتبه صدکی دبیرستان (HSPR)^(۱)، نمره‌های آزمون استعداد تحصیلی (SAT) و نمره‌های آزمون شخصیت استفاده کرده‌اند. نتایج حاصله نشان داده‌اند که (۱) HSPR بهترین پیش بین منفرد است، (۲) ترکیب وزنی مناسبی از HSPR و SAT، از هر یک از این دو متغیر مستقل، پیش بین بهتری برای معدل دوره دانشگاه است، (۳) نمره‌های آزمون شخصیت چیز مفیدی به اعتبار پیش بینی ترکیب SAT-HSPR اضافه نمی‌کند.^(۴) همبستگی چند متغیری با جزییات بیشتر در فصل ۱۱ مورد بحث قرار خواهد گرفت.

// محاسبه گر‌ها و برنامه‌های کامپیوتری برای آمار

بسیاری از ماشین حساب‌های جیبی حاوی برنامه‌های از پیش آماده شده برای روش‌های آماری هستند که در این فصل و فصل بعد توصیف شده‌اند. توصیف نرم‌افزارهای از پیش آماده شده برای کامپیوترها و ریز کامپیوترها در فصل ۱۵ آمده است.

1- High School Percentile Rank

۲- یک بحث کامل درباره همبستگی چند متغیری را می‌توان در این کتاب یافت:

E.J. Pedhazur (1982), Multiple Regression in Behavioral Research (2d ed.), New York: Holt, Rinehart and Winston.

/// فرا تحلیل

فرا تحلیل روشی برای ترکیب نظام‌مند داده‌های کمی حاصل از چند بررسی است که بر پرسشی یکسان متمرکز شده و از متغیرهای مشابهی استفاده کرده‌اند. گلاس، این روش را تحلیل تحلیلها تعریف کرده است.^(۱) با استفاده از آماره‌های بررسیهای اصلی، که از مقایسه گروههای آزمایشی دریافت‌کننده تیمار و گروههای کنترل به دست می‌آید، کمیت مشترکی استخراج می‌شود که آن را برای بررسی مورد نظر، یک بیان کلی نتایج، تفسیر می‌کنند.

برای هر بررسی تفاوت بین میانگین گروههای کنترل و آزمایشی را با تقسیم آن به انحراف معیار گروه کنترل، به واحدهای انحراف معیار تبدیل می‌کنند. این نسبت در واحدهای انحراف معیار، اندازه تأثیر^(۲) نامیده می‌شود.

$$ES = \frac{\bar{X}_E - \bar{X}_C}{\sigma_C} \quad (5-17)$$

اندازه تأثیر برآورد شده = ES

میانگین گروه آزمایشی = $\bar{X}_E$

میانگین گروه کنترل = $\bar{X}_C$

انحراف معیار گروه کنترل = σ_C

اگر فرض کنیم که توزیع مقادیر اندازه تأثیر نرمال است، پس می‌توان هر مقدار را به عنوان یک نمره z با میانگین صفر و انحراف معیار ۱ تلقی کرد، که نقطه مرجع، یا نقطه صفر آن، $\bar{X}_C$ است. متوسط مقادیر اندازه تأثیر همه بررسیها، تأثیر کلی برای متغیرهای آزمایشی را توصیف می‌کند.

برای مثال، اسمیت و گلاس از فرا تحلیل برای بررسی این پرسش رایج که آیا روان‌درمانی تغییری ایجاد می‌کند، استفاده کردند. مرور پیشینه تحقیقاتی روشن ساخت که در این زمینه ۱۰۰۰ آزمایش وجود دارد. بررسیهایی که برای یک تحلیل کامل مناسب بود، ۸۳۳ اندازه تأثیر را

1- Glass, G.V., B. McGaw, and M.L. Smith (1981) *Meta-analysis in Social Research*. Beverly Hills: Sage Publications.

2- Effect - Size

به آنان ارائه کرد. بررسیهای انتخابی شامل استراتژیهای درمانی خود^(۱)، پویا^(۲)، رفتاری و انسانگرایانه^(۳) بود که به طور آزمایشی به متغیرهای مرتبط با نتیجه نظیر عزت نفس^(۴)، سازگاری^(۵)، هراس - اضطراب^(۶) و عملکرد درسی^(۷) وابسته بودند. متوسط اندازه تأثیر ۰/۶۸ بود، یعنی میانگین پس - از - درمان برای آزمودنیهایی که درمان دریافت کردند، مساوی بود با نمره ۰/۶۸ انحراف معیار بالای میانگین آزمودنیهایی که درمانی دریافت نکرده بودند. اسمیت و گلاس نتیجه گرفتند که نتیجه روان درمانی عبارت است از کسب بهبودی در متغیر وابسته که میزان آن، پیشرفت از میانگین به صدک هفتاد و پنجم گروه کنترل است.

فرا تحلیل را نه تنها می توان در مورد بررسیهایی که میانگینها را مقایسه کرده اند، به کار برد، بلکه از آن برای بررسیهای همبستگی، نسبت و دیگر اندازه ها نیز می توان استفاده کرد. گلاس، مک گا^(۸) و اسمیت تبیین کاملی از روشهای فرا تحلیل و کاربردهای آن ارائه کرده اند.^(۹)

/// خلاصه

آمار توصیفی در خدمت توصیف و خلاصه کردن مشاهدات است. انتخاب روش توصیفی بسته به این است که آماره مورد نظر در خدمت چه هدفی است و از چه مقیاس اندازه گیری برای ثبت داده ها استفاده شده است.

مقیاسهای اندازه گیری ابزار کمی کردن مشاهدات و چهار نوع هستند: مقیاسهای اسمی، مشاهدات را به طبقات نامتداخل دسته بندی می کنند، مقیاسهای رتبه ای اشیاء یا طبقات اشیاء را بر مبنای جایگاه نسبی آنها مرتب می کنند، مقیاسهای فاصله ای برای اندازه گیری از فواصل مساوی استفاده می کنند و نشان می دهند یک شخص یا شیء تا چه اندازه از کیفیت خاصی برخوردار است، مقیاسهای نسبی فواصل مساوی را برای اندازه گیری به کار می برند و نقطه صفر مطلق را مبنای اندازه گیری قرار می دهند.

هنگامی که مشاهدات کمی شدند، می توان داده ها را چه خام و چه دسته بندی شده

1- Ego

2- Dynamic

3- Humanistic

4- Self- Esteem

5- Adjustment

6- Fear- Anxiety

7- School Performance

8- McGaw

9- Smith, M. L., and G. V. Glass (1977). Meta-analysis of psychotherapy outcome studies. *American Psychologist*, 32,

به صورت توزیع فراوانی مرتب کرد و به صورت نگاره‌ای با نمودار ستونی یا نمودار چندضلعی نشان داد.

اندازه‌های گرایش مرکزی - نما، میانه و میانگین - شاخص منفردی را برای نمایاندن مقدار متوسط یک مجموعه کامل از اندازه‌ها در دسترس قرار می‌دهند. نما، که یک آماره اسمی است، ناپایدارترین و ناسودمندترین اندازه در تحقیقات آموزشی است. میانه یک آماره رتبه‌ای است که رتبه نمره‌ها را در داخل یک توزیع محسوب می‌دارد، اما مقدار نمره‌های فردی را در نظر نمی‌گیرد. میانگین، که یک آماره فاصله‌ای (یا نسبی) است، پایدارترین و رایج‌ترین شاخص گرایش مرکزی می‌باشد.

شیوه دیگر توصیف مشاهدات، نشان دادن تغییر پذیری، یا پراکندگی، مقادیر یک توزیع است. دامنه، انحراف چارکی و انحراف معیار سه شاخصی هستند که برای این منظور به کار می‌روند. دامنه، فاصله بین بیشترین و کمترین مقادیر یک توزیع را نشان می‌دهد و یک آماره اسمی است.

انحراف چارکی نیمه فاصله بین چارکهای بالاتر و پایین‌تر را به ما می‌دهد. انحراف چارکی یک آماره رتبه‌ای است. انحراف معیار جذر میانگین مربع مقادیر انحراف از میانگین است. انحراف معیار یک آماره فاصله‌ای (یا نسبی) و رایج‌ترین شاخص تغییرپذیری است.

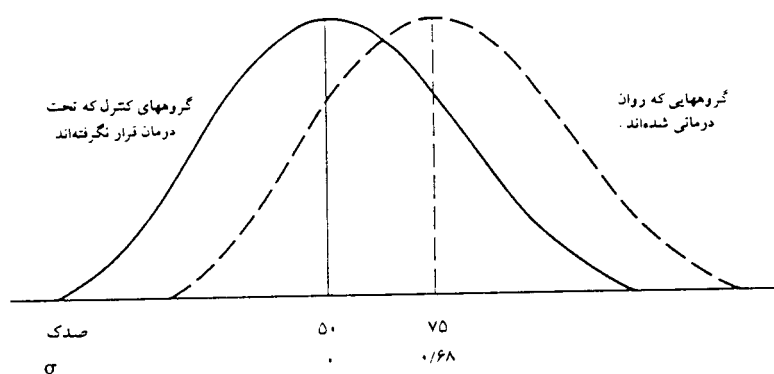
نمره‌های استاندارد برای نشان دادن مکان یک نمره منفرد در یک توزیع مورد استفاده قرار می‌گیرند. رایج‌ترین نمره استاندارد، نمره Z است، که مقادیر مختلف را به واحدهای انحراف معیار تبدیل می‌کند.

با استفاده از ویژگیها و سطوح زیر منحنی نرمال، ما می‌توانیم درصد موارد پایین و بالای هر نمره Z را در یک توزیع نرمال برآورد کنیم.

روشهای همبستگی ما را قادر می‌سازد که رابطه بین دو مجموعه از اندازه‌ها را توصیف کنیم. همبستگی گشتاوری (r پیرسون) و همبستگی رتبه‌ای (رواسپیرمن) دو شاخص رایج رابطه هستند. r پیرسون با داده‌های فاصله‌ای یا نسبی مورد استفاده قرار می‌گیرد، و برای داده‌های رتبه‌ای، رواسپیرمن برای یافتن رابطه بین دو مجموعه از رتبه‌ها به کار می‌رود. جدول ۱۱ - ۵ این آماره‌ها را خلاصه می‌کند.

شکل ۱۰-۵

شکل ۱۰-۵- متوسط برآورد شده مقدار تأثیر درمان با در نظر گرفتن گروههایی که تحت درمان قرار نگرفته‌اند (داده‌ها مبتنی بر ۴۰/۰۰۰ آزمودنی است که یا درمان شده‌اند و یا درمان نشده‌اند زمان متوسط درمان زیر ۲۰ ساعت و میزان متوسط تجربه درمانگران ۲/۵ سال بوده است)



جدول ۱۱-۵ خلاصه آماره‌های توصیفی ارائه شده در این فصل

فاصله‌ای	رتبه‌ای	اسمی	
میانگین	میان	نما	شاخصهای گرایش مرکزی
واریانس و انحراف معیار	انحراف چارکی	دامنه	شاخصهای تغییر پذیری
نمره Z، نمره Z و سایر	رتبه صدکی	برچسب یا طبقه‌بندی	شاخصهای مکان
نمره‌های استاندارد			
Γ پیرسون	رواسبیرمن	فی	شاخصهای همبستگی

// مفاهیم اصلی

آمار استنباطی

آمار توصیفی

انحراف چارکی

(دامنه نیمه چارکی)

توزیع فراوانی

تغییر پذیری

رابطه منحنی الخط

رابطه خطی

دامنه

ضریب همبستگی	ضریب گشتاوری پیرسون	ضریب فی
گرایش مرکزی	فراوانی تراکمی	فرا تحلیل
مقیاس فاصله‌ای	مقیاس رتبه‌ای	مقیاس اسمی
منحنی باچولگی منفی	منحنی باچولگی مثبت	مقیاس نسبی
میان	میانگین	منحنی نرمال
نمره استاندارد	نما	نسبت همبستگی (اتا)
نمودار پراکندگی	نمره Z	نمره Z
واریانس	نمودار ستونی	نمودار چند ضلعی فراوانی
همبستگی دورشته‌ای	همبستگی چهار خانه‌ای	همبستگی
همبستگی مثبت	همبستگی رتبه‌ای اسپیرمن	همبستگی دورشته‌ای نقطه‌ای
		همبستگی منفی

/// تمرینها

- ۱- در مورد هر یک از جملات زیر، نوع مقیاس اندازه‌گیری - اسمی، رتبه‌ای، فاصله‌ای، یا نسبی - را مشخص کنید:

الف - جان، آزمون ریاضی را در ۳۵ دقیقه تمام کرد، در حالی که جک، همان آزمون را در ۲۵ دقیقه به پایان رساند.

ب - جک به زبان فرانسه صحبت می‌کند، اما جان صحبت نمی‌کند.

ج - جک بلندتر از جان است.

د - بلندی قد جان ۶ فوت و ۲ اینچ است.

ه - IQ جان ۱۲۰ است، در حالی که IQ جک ۱۱۰ است.
- ۲- یک نمودار ستونی و یک نمودار چند ضلعی برای توزیع فراوانی زیر رسم کنید:

x	f	x	f	x	f	x	f
۸۰	۱	۷۶	۶	۷۳	۲۰	۷۰	۷
۷۹	۲	۷۵	۱۵	۷۲	۱۷	۶۹	۳
۷۸	۳	۷۴	۲۲	۷۱	۹	۶۸	۱
۷۷	۱۰						

۳- با در نظر گرفتن توزیع ۱۵، ۱۴، ۱۴، ۱۳، ۱۱، ۱۰، ۱۰، ۱۰، ۸، ۵:

الف - میانگین را حساب کنید.

ب - مقدار میانه را تعیین کنید.

ج - مقدار نما را تعیین کنید.

۴- به طور خلاصه رابطه بین چولگی یک توزیع نمره‌ها و مقادیر حاصله میانگین، میانه، و نما را شرح دهید.

۵- مناسب‌ترین اندازه‌ها - نما، میانگین یا میانه - را برای هر یک از انواع مقیاسها تعیین کنید:

الف - رتبه‌ای ب - اسمی ج - فاصله‌ای

۶- تعیین کنید که هر یک از عبارات زیر، کدام اندازه - نما، میانگین یا میانه - را تعریف می‌کند:

الف - نمره وسط

ب - متوسط حسابی

ج - نمره‌ای که بیشتر از همه تکرار شده است.

۷- نمره‌های زیر نتایج آزمون واژگان ۲۰ دانش‌آموز کلاس هفتم را نشان می‌دهد. دامنه،

انحراف معیار، و انحراف چارکی را محاسبه کنید و دربارهٔ مزایا و نقایص هر یک، به عنوان یک اندازه تغییر پذیری برای این نمره‌ها بحث کنید.

x	f	fx	x^2	fx^2
۱۶	۱	۱۶	۲۵۶	۲۵۶
۱۵	۰	۰	۲۲۵	۰
۱۴	۰	۰	۱۹۶	۰
۱۳	۰	۰	۱۶۹	۰
۱۲	۲	۲۴	۱۴۴	۲۸۸
۱۱	۰	۰	۱۲۱	۰
۱۰	۲	۲۰	۱۰۰	۲۰۰
۹	۱	۹	۸۱	۸۱
۸	۱	۸	۶۴	۶۴
۷	۱	۷	۴۹	۴۹
۶	۴	۲۴	۳۶	۱۴۴
۵	۲	۱۰	۲۵	۵۰
۴	۱	۴	۱۶	۱۶
۳	۱	۳	۹	۹
۲	۴	۸	۴	۱۶

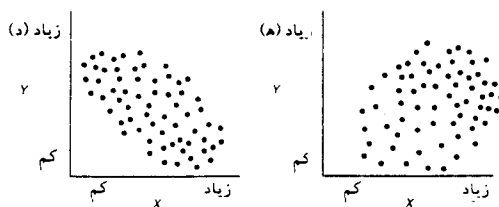
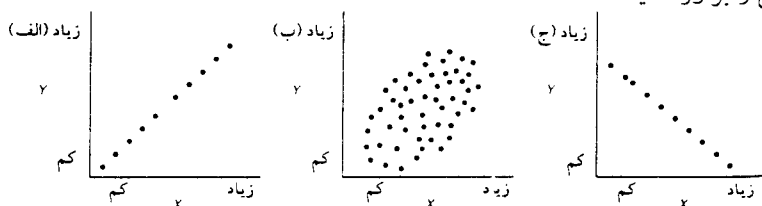
۸- برای به حداقل رساندن اثر یک نمره انتهایی، باید انحراف چارکی را به عنوان شاخص تغییرپذیری انتخاب کرد یا انحراف معیار را؟ چرا؟

۹- در یک آزمون، نمره میانگین ۴۰ و انحراف معیار ۴ است. هر یک از نمره‌های خام زیر را به یک نمره Z تبدیل کنید:

الف - ۴۱ ب - ۳۰ ج - ۴۸ د - ۳۶ ه - ۴۶

۱۰- در یک توزیع نرمال چه درصدی از نمره‌ها زیر نمره Z ۱- قرار خواهند گرفت؟ چه درصدی زیر Z صفر؟ چه درصدی زیر نمره Z ۰/۶۷+؟

۱۱- رابطه نشان داده شده به وسیله این نمودارهای پراکنندگی را توصیف کنید. سپس ضریب همبستگی را برآورد کنید.



۱۲- داده‌های زیر نمره‌های ده دانشجو در یک آزمون استدلال انتزاعی و نیز معدل آنها در رشته فلسفه است. r پیرسون را برای این داده‌ها محاسبه کنید.

معدل	استدلال انتزاعی	دانشجو
۱/۵	۱۵	الف
۲/۵	۲۰	ب
۳	۳۰	ج
۲	۳۵	د

ه	۲۵	۳
و	۴۰	۳/۵
ز	۳۵	۴
ح	۵	۱
ط	۱۲	۲
ی	۱۰	۲/۵

۱۳- محققى به مبحث اضطراب و تأثیر آن بر عملکرد افراد در آزمونهای هوش علاقه‌مند شده است. او با کمک یک روان‌شناس بالینی اضطراب آزمودنیها را سنجیده و آنها را از ۱ تا ۲۰ رتبه‌بندی کرده و سپس از آزمودنیها یک آزمون هوش استاندارد شده گرفته و نمره IQ هر ۲۰ نفر را به رتبه تبدیل کرده است. این محقق برای داده‌های خود چه ضریب همبستگی‌ای را باید محاسبه کند؟ پاسخ خود را تبیین کنید.

۱۴- محققى ضریب همبستگی $+0/6$ را بین درجه‌بندی مدیران مدارس در مورد نحوه لباس پوشیدن معلمان و عملکرد درسی دانش‌آموزان ۱۵۰ دبستان در ایالت خود نشان داد. او نتیجه‌گیری کرد که تشویق معلمان به آراستگی، عملکرد درسی را افزایش خواهد داد. نظر خود را درباره نتیجه‌گیری او بیان کنید.

۱۵- شاخص همبستگی مناسب را برای استفاده در حل هر یک از مسئله‌های زیر تعیین کنید:
الف - ما می‌خواهیم ضریب همبستگی بین هوش و خلاقیت را، هنگامی که نمره‌های یک گروه از آزمودنی‌ها را در هر دو متغیر داریم، پیدا کنیم.

ب - ما می‌خواهیم ضریب همبستگی بین عملکردهای یک گروه از آزمودنیها را در دو آزمون، هنگامی که نمره‌های آزمودنیها را در یک آزمون و جایگاه آنها را به صورت بالا یا زیر میانگین در آزمون دیگر داریم، پیدا کنیم.

ج - ما می‌خواهیم همبستگی بین جنس و شرایط لازم برای یک شغل را، هنگامی که اطلاعات رد یا قبول در آزمون صلاحیت آن شغل را برای ۶۰ مرد و ۶۰ زن داریم، پیدا کنیم.

د - ما می‌خواهیم همبستگی بین پاسخهای ۱۸۰ دانشجو را به پرسشهای دو آزمون، هنگامی که پاسخهای درست و نادرست این دانشجویان را به پرسشهای هر دو آزمون در دست داریم، پیدا کنیم.

هـ- ما می‌خواهیم رابطه بین هوش و انگیزه پیشرفت را، هنگامی که نمره‌های آزمون هوش ۲۰۰ آزمودنی را در آزمون هوش استانفورد - بین و جایگاه آنان را در آزمون انگیزه پیشرفت به صورت بالا یا زیر متوسط در دست داریم، پیدا کنیم.

و - ما می‌خواهیم رابطه بین جنس و توانایی کار با ابزار را، هنگامی که نمره‌های ۶۰ پسر و ۶۰ دختر را در یک آزمون توانایی کار با ابزار داریم، پیدا کنیم.

ز - ما می‌خواهیم رابطه بین نمره‌های گروه سنجی یک گروه از دانش‌آموزان و رتبه درسی آنها را در یک کلاس دبیرستانی پیدا کنیم.

۱۶- کدام روش آماری، تحلیل تحلیلی نامیده شده است؟

/// پاسخها

۱- الف - نسبی ب - اسمی ج - رتبه‌ای د - نسبی ه - فاصله‌ای

۲- پاسخها متفاوت است.

۳- الف - میانگین = ۱۱ ب - میانه = ۱۰/۵ ج - نما = ۱۰

۴- این سه اندازه در یک توزیع دارای چولگی مساوی نیستند. میانگین به سمت چولگی کشیده می‌شود. بنابراین، در یک توزیع دارای چولگی مثبت، میانگین همیشه بیشتر از میانه است و نما معمولاً کمترین مقدار را دارد. در یک توزیع دارای چولگی منفی، میانگین همیشه کمتر از میانه است و میانه و نما معمولاً بیشترین مقدار را دارند.

۵- الف - میانه ب - نما ج - میانگین

۶- الف - میانه ب - میانگین ج - نما

۷- دامنه ۱۴ = ۲ - ۱۶

$$\sigma = \sqrt{\frac{1173 - \frac{(133)^2}{20}}{20}} = \sqrt{14/4275} = 3/798$$

$$QD = \frac{Q_2 - Q_1}{2} = \frac{9/5 - 3/5}{2} = \frac{6}{2} = 3$$

۸- انحراف چارکی؛ زیرا نمره‌های انتهایی بر انحراف چارکی تأثیر نمی‌گذارند.

۹- الف - ۰/۲۵ ب - ۲/۵ ج - ۲ د - ۱ ه - ۱/۵

۱۰- ۱۶٪، ۵۰٪، ۷۵٪

۱۱- الف - مثبت کامل، +۱ ب - مثبت، ۰/۷۵ ج - منفی کامل، -۱

د - منفی، ۰/۷۵ ه - بدون همبستگی، ۰

۱۲-

X	Y	XY	X^2	Y^2
۱۵	۱/۵	۲۲/۵	۲۲۵	۲/۲۵
۲۰	۲/۵	۵۰	۴۰۰	۶/۲۵
۳۰	۳	۹۰	۹۰۰	۹
۳۵	۲	۷۰	۱۲۲۵	۴
۲۵	۳	۷۵	۶۲۵	۹
۴۰	۳/۵	۱۴۰	۱۶۰۰	۱۲/۲۵
۳۵	۴	۱۴۰	۱۲۲۵	۱۶
۵	۱	۵	۲۵	۱
۱۲	۲	۲۴	۱۴۴	۴
۱۰	۲/۵	۲۵	۱۰۰	۶/۲۵
Σ ۲۲۷	۲۵	۶۴۱/۵	۶۴۶۹	۷۰

$$r = \frac{\Sigma XY - \frac{(\Sigma X)(\Sigma Y)}{N}}{\sqrt{\left[\Sigma X^2 - \frac{(\Sigma X)^2}{N} \right] \left[\Sigma Y^2 - \frac{(\Sigma Y)^2}{N} \right]}}$$

$$= \frac{641/5 - \frac{(227)(25)}{10}}{\sqrt{\left[6469 - \frac{(227)^2}{10} \right] \left[70 - \frac{(25)^2}{10} \right]}} = + 0/74$$

۱۳- چون محقق داده‌های رتبه‌ای دارد، باید ضریب همبستگی رتبه‌ای اسپیرمن را محاسبه کند.

۱۴- محقق هیچ توجیهی برای استنباط یک رابطه علی، صرفاً بر مبنای شواهد همبستگی

ندارد. درجه‌بندی مدیران مدارس از نحوه لباس پوشیدن معلمان و عملکرد درسی دانش‌آموزان می‌تواند تابعی از متغیرهای دیگر باشد.

۱۵- الف - ضریب گشتاوری پیرسون ب - دو رشته‌ای ج - چهارخانه‌ای

د- ضریب فی ه- دو رشته‌ای و- دو رشته‌ای نقطه‌ای ز- رواسپیرمن

۱۶- فرا تحلیل

فصل ششم

نمونه گیری و آمار استنباطی

هدفهای آموزشی

دانشجو، بعد از مطالعه این فصل، قادر خواهد بود:

- ۱- معنا، منطق و مراحل نمونه گیری را توصیف کند و نمونه گیری احتمالاتی را از نمونه گیری غیر احتمالاتی تشخیص دهد.
- ۲- ویژگیها، کاربردها و محدودیتهای هر یک از انواع نمونه گیری احتمالاتی و نمونه گیری غیر احتمالاتی را فهرست کند.
- ۳- معنای خطای نمونه گیری و رابطه آن را با استنباطهای آماری توضیح دهد.
- ۴- معناداری آماری را توضیح دهد.
- ۵- معنای فرضیه صفر و کاربردهای آن را در تحقیق علمی توضیح دهد.
- ۶- خطاهای نوع I و نوع II را توصیف کند.
- ۷- تفاوت بین آزمونهای معناداری جهت دار و غیر جهت دار و کاربرد مناسب هر یک از این آزمونها را توضیح دهد.
- ۸- آزمون t را برای یافتن معناداری تفاوت بین میانگینهای همبسته و غیر همبسته به کار برد.

۹- آزمون F را برای یافتن معناداری تفاوت بین گروهها در تحلیل واریانسهای یکسویه و دوسویه به کار برد.

۱۰- آزمون χ^2 را برای یافتن معناداری تفاوت بین نسبتها در طبقه‌بندیهای یکسویه و دوسویه به کار برد.

۱۱- استفاده از آزمون t را در مورد T پیرسون توصیف و تعیین کند که کدام ضریب همبستگی معنادار است.

۱۲- شیوه استنباط آماری مناسب برای استفاده در آزمون یک فرضیه تحقیقاتی را انتخاب کند.

۱۳- واژه‌های فنی آماری مورد استفاده در گزارش نتایج تحقیق را درک کند.

مباحث آماری که در فصل گذشته به ذکر آن پرداختیم، برای طبقه‌بندی، خلاصه کردن و توصیف داده‌ها مورد استفاده قرار می‌گیرد، اما در تحقیق، ما اغلب نیازمند آن هستیم که از توصیف داده‌ها فراتر رویم. بعد از انجام مشاهده‌های لازم در مورد گروه نمونه، برای آنکه یافته‌های خود را به کل جامعه آماری، که گروه نمونه از آن انتخاب شده است، تعمیم دهیم، از استقراء یا استنباط استفاده می‌کنیم. برای انجام این امر، به تکنیک‌هایی نیازمندیم که ما را به استنباط‌های معتبر از گروه نمونه به کل جامعه آماری قادر سازد.

/// نمونه‌گیری

یک خصیصه مهم آمار استنباطی، فرآیند از جزء به کل رسیدن است. برای مثال، ممکن است ما یک گروه ۵۰۰ نفری از دانشجویان یک دانشگاه را به منظور انجام تعمیم‌هایی درباره کل پیکره دانشجویی آن دانشگاه، مطالعه کنیم.

گروه کوچکی که مورد مشاهده قرار می‌گیرد، نمونه و گروه بزرگتر که در مورد آن تعمیم صورت می‌گیرد، جامعه آماری نامیده می‌شود. یک جامعه آماری را این گونه تعریف می‌کنند: همه اعضای یک طبقه کاملاً تعریف شده از افراد، رویدادها یا اشیاء. برای مثال، در مطالعه‌ای که نوجوانان آمریکایی جامعه آمار آن را تشکیل می‌دهند، می‌توان این جامعه آماری را به عنوان همهٔ پسران و دختران آمریکایی که در دامنه سنی ۱۲-۲۱ سال قرار دارند، تعریف کرد. یک نمونه، بخشی از یک جامعه آماری است. برای مثال، دانش‌آموزان دبیرستان واشنگتن در ایندیانا پولیس، یک نمونه از نوجوانان آمریکایی هستند. آنها بخشی از جامعه آماری نوجوانان آمریکایی‌اند؛ زیرا هم شهروند آمریکایی هستند و هم در دامنه سنی ۱۲-۲۱ سال قرار دارند.

استنباط آماری روشی است که از طریق آن پارامترها، یعنی خصیصه‌های جامعه آماری را از آماره‌ها، یعنی خصیصه‌های نمونه، برآورد می‌کنیم. چنین برآوردهایی بر قوانین احتمالات بنا شده‌اند و بیش از آنکه واقعیت مطلق باشند، بهترین برآورد ما از این واقعیت‌ها هستند. در هر استنباطی درجه معینی از خطا نهفته است. همان گونه که بعداً خواهیم دید، آمار استنباطی، همچنین، برای آزمون فرضیه‌های مربوط به جامعه آماری به کار می‌رود.

// منطق نمونه گیری

استدلال استقرایی، یک بخش اساسی از رویکرد علمی است. روش استقرا مستلزم انجام مشاهدات و سپس استخراج نتایجی از این مشاهدات است. اگر کسی بتواند همه افراد یک جامعه آماری را مورد مشاهده قرار دهد، می تواند با اطمینان پایه نتیجه گیریهای خود را درباره جامعه آماری بر این مشاهدات بنا نهد (استقرای کامل). از سوی دیگر، اگر او تنها برخی از افراد جامعه آماری را مورد مشاهده قرار دهد، نمی تواند چیزی بیش از استنباط این امر که این مشاهدات در مورد کل جامعه آماری صدق می کنند، انجام دهد (استقرای ناقص). این مفهوم نمونه گیری است، که شامل گرفتن بخشی از جامعه آماری، انجام مشاهدات بر روی این گروه کوچکتر و تعمیم یافته ها به جامعه مرجع، یعنی جامعه بزرگتری است که نمونه از آن گرفته شده است.

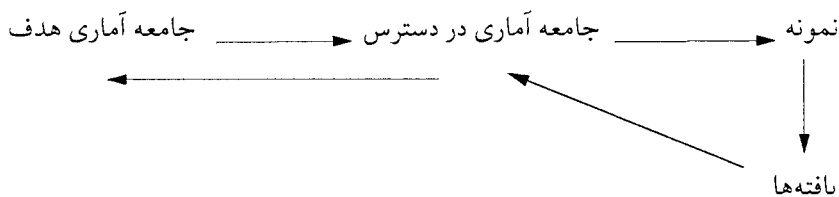
برای محقق، نمونه گیری یک ضرورت است. معمولاً زمان، هزینه و تلاشی که برای انجام یک تحقیق لازم است، به محقق اجازه نمی دهد که همه افراد ممکن یک جامعه آماری را مورد مطالعه قرار دهد. بعلاوه برای درک پدیده مورد بررسی، به طور کلی نیازی به مطالعه همه موارد ممکن نیست. نمونه گیری ابزاری است که به یاری محقق می آید و به او امکان می دهد که به جای مطالعه تمام جامعه آماری، تنها بخشی از آن را مورد مطالعه قرار دهد.

از آنجا که منظور از پیرون کشیدن یک نمونه از یک جامعه آماری کسب اطلاعاتی در مورد آن جامعه آماری است، کاملاً اهمیت دارد افرادی که در یک نمونه قرار می گیرند، برش عرضی از آن جامعه آماری را تشکیل دهند که معرف همه افراد آن باشند؛ به عبارتی برای اینکه بتوان با اطمینان، از نمونه به جامعه آماری تعمیم داد، نمونه ها باید معرف باشند. برای مثال، ممکن است محقق فرض کند که دانش آموزان دبیرستان واشنگتن معرف جوانان آمریکایی هستند. به هر ترتیب، اگر افرادی که در نمونه گنجانده شده اند، خصلتهایی متفاوت با جامعه آماری مرجع داشته باشند، ممکن است این نمونه، معرف جامعه آماری نباشد. مکان مدرسه آنها، زمینه اقتصادی - اجتماعی آنها، موقعیت خانوادگی آنها، تجارب قبلی آنها و بسیاری دیگر از خصلتهای این گروه، ممکن است از آنها گروهی غیر معرف برای جوانان آمریکایی ساخته باشد. این نوع از نمونه ها را نمونه اریب نام می گذارند. یافته های حاصل از یک نمونه اریب، قابل تعمیم به جامعه آماری آن نیست.

/ مراحل نمونه‌گیری

اصل نخستین در نمونه‌گیری، شناسایی جامعه آماری است که در بررسی معرفی می‌شود. اگر پژوهشگری بخواهد دربارهٔ آموزگاران مدارس سنت لوئیز اطلاعاتی کسب کند، تمام کسانی که در آن مدارس تدریس می‌کنند، جامعه آماری هدف^(۱) او را تشکیل می‌دهند در یک بررسی که موضوع آن نگرشها و ارزشهای جوانان آمریکایی باشد، جامعه آماری هدف، همه پسران و دختران آمریکایی در محدودهٔ سنی ۱۲-۲۱ خواهند بود؛ با این فرض که پژوهشگر، تعریف عملی جوانی را دوره سنی بین ۱۲ تا ۲۱ سال دانسته باشد.

به هر حال، از آنجا که معمولاً بررسی تمام جامعه آماری هدف امکان‌ناپذیر است، پژوهشگر باید آن بخش از جامعه آماری را که قابل دسترس است، تعیین کند. این بخش را جامعه آماری در دسترس می‌نامند و پژوهشگر نمونه‌ای را که مورد بررسی قرار می‌دهد، از این بخش انتخاب می‌کند. ماهیت جامعه آماری در دسترس از زمان و منابعی که پژوهشگر در اختیار دارد، تأثیر می‌پذیرد. در یک بررسی نگرش نوعی، برای مثال، ممکن است پژوهشگر تمام پسران و دختران کالیفرنیا را به عنوان جامعه آماری در دسترس تعیین کند یا فقط پسران و دختران سان فرانسیسکو را مورد نظر قرار دهد. پژوهشگر، از جامعه آماری در دسترس، نمونه‌ای انتخاب می‌کند که معرف آن جامعه آماری باشد. برای مثال، اگر جوانان کالیفرنیا به عنوان جامعه آماری در دسترس تعیین شده باشد، پژوهشگر از بین تمام جوانان ایالت کالیفرنیا نمونه انتخاب خواهد کرد، یا اگر جوانانی که در سان فرانسیسکو زندگی می‌کنند، به عنوان جامعه آماری در دسترس تعیین شوند، نمونه از بین این گروه خاص انتخاب خواهد شد.



یک پژوهشگر با چه اطمینانی می‌تواند یافته‌های حاصل از یک نمونه را به جامعه آماری

هدف تعمیم دهد؟ اگر نمونه انتخاب شده به راستی معرف جامعه آماری در دسترس باشد، پس برداشتن اولین گام در فرآیند تعمیم با دشواری کمی روبرو خواهد شد. اصل کلی این است: اگر یک نمونه به گونه‌ای انتخاب شده است که معرف جامعه آماری در دسترس است، یافته‌های حاصل از نمونه می‌تواند به آن جامعه آماری تعمیم یابد. برای مثال، اگر پژوهشگری یک نمونه معرف از جوانان کالیفرنیا انتخاب کرده بود، بنابراین، می‌توانست تعمیم‌هایی در مورد نگرشها و ارزشهای تمام پسران و دختران جوان کالیفرنیا انجام دهد.

به هر ترتیب، تعمیم دادن از جامعه آماری در دسترس به جامعه آماری هدف، به طور کلی، خطر^(۱) بیشتری در بر دارد. در این گام، اطمینانی که پژوهشگر می‌تواند نسبت به تعمیم نتایج داشته باشد، به تشابه جامعه آماری در دسترس و جامعه آماری هدف بستگی دارد. در مثال بالا، اگر پژوهشگر به جای محدود کردن جامعه آماری در دسترس به ایالت کالیفرنیا، ایالت‌های مختلفی را در سراسر کشور به عنوان جامعه آماری در دسترس تعیین کرده بود، با اطمینان بیشتری می‌توانست نتایج حاصله را به جوانان آمریکایی تعمیم دهد. به این ترتیب تمام قسمتهای ایالات متحده در جامعه آماری جای می‌گرفت و نمونه‌گیری جامع‌تری از نگرشهای جوانان آمریکایی ممکن می‌شد.

این بیان درست است که به هنگام استفاده از مشاهده‌های یک نمونه برای برآورد خصلتهای یک جامعه آماری، محقق ناچار است که در «اعتماد» به نتایج حاصله، به «جهشی» استنباطی دست زند. احتمال درست بودن چنین استنباطهایی، تا اندازه زیادی تابع روش نمونه‌گیری است. دو شیوه اصلی نمونه‌گیری در اختیار محقق است: نمونه‌گیری احتمالاتی و نمونه‌گیری غیر احتمالاتی. نمونه‌گیری احتمالاتی نمونه‌گیری‌ای است که عناصر آن به شیوه اتفاقی برگزیده می‌شوند. ویژگی اصلی نمونه‌گیری احتمالاتی این است که هر عضو یا عنصر جامعه آماری احتمال مشخصی برای قرار گرفتن در نمونه دارد. نمونه‌گیری غیر احتمالاتی شامل روشهایی است که در آنها اعضای نمونه با شیوه‌های اتفاقی برگزیده نمی‌شوند. در این نوع نمونه‌گیری هیچ راهی وجود ندارد که بتوان احتمال قرار گرفتن یک عنصر را در نمونه مشخص کرد. موفقیت آن به دانش، تخصص و قضاوت محقق بستگی دارد. نمونه‌گیری غیر احتمالاتی هنگامی مورد استفاده قرار می‌گیرد که کاربرد نمونه‌گیری احتمالاتی امکان‌پذیر نباشد. مزایای آن، سهولت و صرفه‌جویی است.

// نمونه‌گیری احتمالاتی

نمونه‌گیری احتمالاتی نوعی نمونه‌گیری است که در آن احتمال انتخاب هر عنصر جامعه آماری، بالاتر از صفر است.^(۱) در این نوع نمونه‌گیری، امکان قرار گرفتن هر عنصر جامعه آماری در نمونه، برحسب اتفاق است و این امر از طریق انتخاب تصادفی صورت می‌گیرد. هنگامی که از نمونه‌گیری احتمالاتی استفاده می‌شود، آمار استنباطی محققان را قادر می‌سازد که تفاوت احتمالی یافته‌های مبتنی بر نمونه را با آنچه که از مطالعه تمام جامعه آماری به دست خواهد آمد، برآورد کنند. چهار نوع نمونه‌گیری احتمالاتی رایج در تحقیقات آموزشی عبارتند از: نمونه‌گیری تصادفی ساده، نمونه‌گیری طبقه‌ای، نمونه‌گیری خوشه‌ای، و نمونه‌گیری نظام‌مند (سیستماتیک).

/ نمونه‌گیری تصادفی ساده

یکی از شناخته‌شده‌ترین روشهای نمونه‌گیری، روش نمونه‌گیری تصادفی ساده است. خصلت اساسی نمونه‌گیری تصادفی ساده آن است که همه افراد جامعه آماری بخت مساوی و مستقل برای انتخاب شدن دارند، یعنی برای هر زوج X و Y ، بخت X برای انتخاب شدن مساوی بخت Y است، و انتخاب X به هیچ ترتیبی بر احتمال انتخاب Y تأثیر ندارد. مراحل نمونه‌گیری تصادفی ساده از این قرارند:

۱- تعریف جامعه آماری

۲- فهرست کردن همه اعضای جامعه آماری

۳- انتخاب نمونه با اعمال روشی که اعضای فهرست تنها بر حسب تصادف، در نمونه قرار بگیرند.

نظام‌مندترین روش برای انتخاب یک نمونه تصادفی، مراجعه به جدول اعداد تصادفی است. این جدول حاوی ستونهایی از اعداد است که برای اطمینان از ترتیب تصادفی آنها، به طور ماشینی و معمولاً با استفاده از کامپیوتر، تعیین شده‌اند. جدول ۶-۱ در ضمیمه، نمونه‌ای از جدول اعداد تصادفی است. نخستین مرحله در انتخاب یک نمونه تصادفی از یک جامعه آماری، تخصیص یک شماره شناسایی مشخص به هر عضو جامعه آماری است. سپس جدول اعداد

تصادفی برای انتخاب شماره شناسایی آزمودنیهایی که باید در نمونه قرار بگیرند، به کار می آید. اجازه دهید چگونگی استفاده از این جدول را با گرفتن نمونه‌ای از جامعه آماری دانش آموزان دبیرستان واشنگتن، روشن کنیم. ابتدا لازم است تمام افرادی را که در جامعه آماری قرار می گیرند، معین کنیم. دفتر مدیر مدرسه می تواند فهرستی از تمام دانش آموزانی که در مدرسه ثبت نام کرده اند، فراهم کند. سپس به منظور شناسایی هر فرد جامعه آماری شماره‌ای به هر یک تخصیص داده می شود. با فرض اینکه مدرسه ۸۰۰ دانش آموز داشته باشد، محقق می تواند از شماره‌های ۰۰۰، ۰۰۱، ۰۰۲، ۰۰۳، ۰۰۴، الی ۷۷۹ برای این منظور استفاده کند. سپس محقق به جدول اعداد تصادفی مراجعه می کند تا اعداد سه رقمی را به دست آورد، و تنها از اعدادی استفاده می کند که کمتر از یا مساوی با ۷۹۹ باشد. به ازای هر عدد انتخابی، فردی که در جامعه آماری با آن عدد همخوان است، در نمونه قرار می گیرد. محقق این فرآیند را تا آنجا ادامه می دهد که افراد گروه نمونه به تعداد مطلوب انتخاب شوند. مرسوم است که به هنگام استفاده از جدول اعداد تصادفی، نقطه شروع را به طور اتفاقی، انتخاب کنند. یکی از راهها آن است که چشمها را ببندند و انگشت را بر صفحه کاغذ جدول اعداد تصادفی بگذارند و از همان عدد شروع کنند.

معنای عام کلمه تصادفی^(۱) عبارت است از «بدون منظور یا به طور اتفاقی». به هر ترتیب نمونه گیری تصادفی روشی هدفمند و مبتنی بر اسلوب معین است. بدیهی است که نمونه‌ای که به طور تصادفی انتخاب شود، از سوگیریهای پژوهشگر مبرا است. هنگامی که پژوهشگری این روش را به کار می برد، خود را مکلف می سازد که نمونه را به گونه‌ای انتخاب کند که جهت گیریهای او مجال بروز پیدا نکند. او تعهد می کند که از انتخاب آگاهانه آزمودنیهایی که فرضیه اش را تأیید خواهند کرد، خودداری کند. او اجازه می دهد که تنها تصادف تعیین کند که چه عناصری از جامعه آماری در نمونه قرار خواهند گرفت.

پژوهشگر انتظار خواهد داشت که یک نمونه تصادفی معرف جامعه آماری مرجعی باشد که نمونه از آن گرفته شده است، به هر حال، انتخاب تصادفی، بویژه در مورد نمونه‌های کوچک، تضمینی برای این نیست که نمونه، معرف جامعه آماری باشد. انتخاب تصادفی ضمانت می کند که هر تفاوتی بین نمونه و جامعه آماری مرجع، تنها تابعی از تصادف باشد و از جهت گیری محقق منتج نباشد. تفاوتی بین نمونه‌های تصادفی و جامعه آماری مرجع، نظام مند نیستند.

برای مثال، امکان دارد میانگین پیشرفت خواندن یک نمونه تصادفی از دانش‌آموزان کلاس ششم بیشتر از میانگین پیشرفت خواندن جامعه آماری مرجع باشد، اما همان قدر نیز محتمل است که میانگین نمونه، کمتر از میانگین جامعه آماری مرجع باشد، به سخن دیگر، در یک نمونه تصادفی، به یک میزان احتمال دارد که خطای نمونه‌گیری منفی یا مثبت باشد.

بعلاوه، نظریه پردازان علم آمار، با استدلال قیاسی نشان داده‌اند که تا چه حد می‌توان انتظار داشت که مشاهدات حاصل از یک نمونه تصادفی با آنچه که در جامعه آماری مشاهده خواهد شد، تفاوت داشته باشد. تمام روشهای آماری که در این فصل توصیف شده‌اند، بر این مقصود مبتنی هستند. شما به خاطر خواهید آورد که احتمال اینکه در یک نمونه کوچک خصلتهای مشاهده شده با خصلتهای جامعه آماری تفاوت کند، بیشتر از زمانی است که یک نمونه بزرگ به کار ببریم. هنگامی که پژوهشگر از نمونه‌گیری تصادفی استفاده می‌کند، می‌تواند با روشهای آمار استنباطی، احتمال تفاوت جامعه آماری را با نمونه خود، برآورد کند. آمار استنباطی که در این فصل آمده است، تماماً مبتنی بر نمونه‌گیری تصادفی است و تنها در مواردی به کار می‌رود که در آنها انتخاب تصادفی رعایت شده باشد.

متأسفانه انجام نمونه‌گیری تصادفی ساده نیازمند تهیه فهرستی از تمام افراد موجود در یک جامعه آماری محدود است. این امر اغلب مانعی بزرگ در راه استفاده از این روش است.

/ نمونه‌گیری طبقه‌ای (نمونه‌گیری لایه‌ای)

هنگامی که جامعه آماری شامل تعدادی زیر گروه یا طبقه باشد که احتمالاً در خصلتهای مورد مطالعه با هم فرق داشته باشند، اغلب استفاده از نوعی نمونه‌گیری احتمالاتی، که نمونه‌گیری طبقه‌ای خوانده می‌شود، مطلوب است. برای مثال، اگر کسی یک نظرخواهی برای سنجش افکار در مورد یک مسئله خاص سیاسی انجام دهد، تقسیم جامعه آماری به گروههای مختلف بر مبنای سن یا شغل خردمندانه خواهد بود. دلیل این امر آن است که انتظار می‌رود در میان گروههای سنی یا شغلی، به طور نظام‌مند، نظرات متفاوتی وجود داشته باشد. در نمونه‌گیری طبقه‌ای، ابتدا طبقات مورد نظر را تعیین و سپس تعداد مشخصی از آزمودنیها را به طور تصادفی از هر طبقه انتخاب می‌کنند. امکان دارد مبنای طبقه‌بندی جغرافیایی یا خصلتهای جامعه آماری، نظیر درآمد، شغل، جنس، سن، سال تحصیلی، دانشکده یا سطح تدریس باشد.

برای مثال، در بررسی جوانان، ممکن است پژوهشگر نه تنها به پیمایش نگرشهای جوانان نسبت به پدیده خاصی علاقه‌مند باشد، بلکه بخواهد نگرش جوانانی را که در شهرهای کوچک ساکن هستند، با آنها که در شهرهای متوسط و بزرگ زندگی می‌کنند، مقایسه کند. در چنین موردی، پژوهشگر، جامعه جوانان را بر مبنای بزرگی شهری که در آن اقامت دارند، به سه گروه تقسیم می‌کند و سپس به طور تصادفی نمونه‌های مستقلی از هر طبقه (گروه) را انتخاب می‌کند.

یک مزیت نمونه‌گیری طبقه‌ای آن است که معرف بودن نمونه را افزایش می‌دهد و محقق را قادر می‌سازد که تفاوت‌هایی را که ممکن است بین زیرگروه‌های یک جامعه آماری وجود داشته باشند، مطالعه کند. در این نوع نمونه‌گیری محقق می‌تواند هم از هر طبقه به تعداد مساوی و هم از هر طبقه به نسبت بزرگی طبقه در جامعه آماری انتخاب کند. روش دوم، نمونه‌گیری طبقه‌ای متناسب خوانده می‌شود، که هنگامی به کار می‌رود که ویژگی‌های تمام جامعه آماری مسئله اصلی بررسی باشد. هر طبقه به نسبت دقیق فراوانی آن در جامعه آماری، در نمونه نمایانده می‌شود. برای مثال، اگر ۱۰ درصد از رأی‌دهندگان دانشجو باشند، پس ۱۰ درصد از نمونه‌ای که محقق برای نظرخواهی از رأی‌دهندگان انتخاب می‌کند، از این طبقه گرفته خواهد شد.

با این همه، در برخی از مطالعات تحقیقاتی، مسئله اصلی، تفاوت‌های بین طبقات مختلف است. در این حالت، محقق نمونه‌ای انتخاب می‌کند که از هر طبقه به اندازه مساوی در آن باشد. برای مثال، اگر کسی در مورد تفاوت عادات مطالعه دانشجویان دوره فوق لیسانس و لیسانس بررسی می‌کند، از هر دو گروه به تعداد مساوی در نمونه خود می‌گنجاند و سپس تفاوت‌های موجود بین آنان را مطالعه می‌کند. شیوه مورد استفاده برحسب ماهیت سؤال تحقیق انتخاب خواهد شد. اگر تأکید بر انواع تفاوت‌های بین طبقات باشد، محقق از هر طبقه به تعداد مساوی انتخاب می‌کند. اگر ویژگی‌های تمام جامعه آماری مسئله اصلی باشد، نمونه‌گیری متناسب مطلوب‌تر است.

نمونه‌گیری طبقه‌ای هر جا که کاربرد داشته باشد، احتمالاً به ما نمونه‌ای معرفتر از نمونه‌گیری تصادفی ساده ارائه می‌کند. در نمونه‌گیری تصادفی ساده، ممکن است طبقات خاصی، بر حسب اتفاق، وزن بیشتر یا کمتر در نمونه داشته باشند. برای مثال، در نمونه تصادفی ساده از دانش‌آموزان دبیرستان، به طور نظری ممکن است (اگر چه بسیار نامحتمل) که تنها دانش‌آموزان دختر انتخاب شوند. به هر حال، اگر پسران و دختران به طور جداگانه فهرست و یک نمونه

تصادفی از هر گروه انتخاب شود، چنین امری نمی‌تواند اتفاق افتد.

مزیت بزرگ نمونه‌گیری طبقه‌ای آن است که حضور گروه‌های تعریف شده جامعه آماری را در نمونه تضمین می‌کند.

نمونه‌گیری خوشه‌ای

همان طور که قبلاً توضیح داده شد، فهرست کردن همه افراد یک جامعه آماری موردنظر و انتخاب نمونه‌ای از میان آنها، اگر ناممکن نباشد، بسیار دشوار است. جامعه آماری دانش‌آموزان مدارس آمریکا، برای مثال، به قدری بزرگ است که کسی نمی‌تواند همه افراد آن را به قصد انتخاب نمونه‌ای از میان آنها، فهرست کند. علاوه بر آن، مطالعه نمونه‌ای که افراد آن در سراسر آمریکا پراکنده‌اند، کار بسیار پرهزینه‌ای است. در چنین حالتی، مطالعه آزمودنی‌ها در گروه‌هایی که به طور طبیعی تشکیل می‌شوند، یا خوشه‌ها، ساده‌تر خواهد بود، یعنی پژوهشگر از فهرست مدارس، تعدادی را به طور تصادفی انتخاب خواهد کرد و سپس همه دانش‌آموزان آن مدارس را در نمونه خود می‌گنجاند. این نوع نمونه‌گیری احتمالاتی، به عنوان نمونه‌گیری خوشه‌ای شناخته شده است؛ زیرا واحد انتخابی فرد نیست، بلکه گروهی از افرادند که به طور طبیعی گرد هم آمده‌اند. این افراد تا جایی که از نظر خصلت‌های مربوط به متغیر بررسی به هم شباهت دارند، یک خوشه را تشکیل می‌دهند. اکنون فرض کنیم یک نظرخواهی همگانی در آتلانتا انجام می‌شود. احتمالاً محقق به فهرست همه افراد بالغ دسترسی نخواهد داشت، بنابراین انتخاب یک نمونه تصادفی ساده برای او ناممکن خواهد بود. روش عملی‌تر، انتخاب یک نمونه تصادفی مرکب از ۵۰ منطقه از روی نقشه شهر و سپس نظرسنجی از همه افراد بالغی است که در آن مناطق زندگی می‌کنند. هر منطقه نمایانگر یک خوشه از آزمودنی‌ها است که در خصلت‌های ویژه‌ای که با زندگی در همسایگی همبسته است، شباهت دارند.

این یک امر اساسی است که خوشه‌هایی که واقعاً در مطالعه گنجانده می‌شوند، به طور تصادفی از جامعه آماری خوشه‌ها انتخاب شوند. اگر تنها یک خوشه منفرد به کار رفته بود - برای مثال، یک مدرسه ابتدایی در یک شهر بزرگ - پژوهشگر نمی‌توانست نتایج را به جامعه آماری تعمیم دهد. یک ضرورت دیگر آن است که زمانی که یک خوشه انتخاب می‌شود، همه افراد آن خوشه باید در نمونه گنجانده شوند. خطای نمونه‌گیری در یک نمونه خوشه‌ای، بسیار بزرگتر از خطای نمونه‌گیری تصادفی واقعی است.

/ نمونه‌گیری نظام‌مند (سیستماتیک)

شکل دیگری از نمونه‌گیری احتمالاتی را نمونه‌گیری سیستماتیک می‌نامند. این روش عبارت است از گرفتن یک نمونه با انتخاب هر مورد K ام از فهرست جامعه آماری. ابتدا پژوهشگر در مورد تعداد آزمودنیهای نمونه (n) تصمیم می‌گیرد.

از آنجا که او تعداد کل اعضای جامعه آماری (N) را می‌داند، به آسانی N را بر n تقسیم و فاصله نمونه‌گیری K را تعیین می‌کند. اولین فرد نمونه به طور تصادفی از میان اولین K نفر فهرست انتخاب می‌شود و سپس هر فرد K ام جامعه آماری برای نمونه انتخاب می‌شود. برای مثال، فرض کنیم که یک جامعه آماری مرکب از ۵۰۰ آزمودنی داریم و می‌خواهیم نمونه‌ای با ۵۰ عضو انتخاب کنیم، بنابراین، $K = N/n = 500/50 = 10$.

پژوهشگر از ابتدای فهرست شروع می‌کند؛ به طوری که مورد اول از میان ده مورد اول انتخاب شود، سپس با شمارش، موردهای ۱۰ ام فهرست را در نمونه می‌گنجاند. فرض کنیم سومین نام یا شماره فهرست، اولین کسی است که انتخاب می‌شود. سپس فاصله نمونه‌گیری K یا ۱۰، به ۳ اضافه می‌شود و نفر سیزدهم در نمونه قرار می‌گیرد، همچنین نفر بیست و سوم و به همین نحو تا آخر فهرست ادامه می‌یابد.

نمونه‌گیری سیستماتیک از این نظر با نمونه‌گیری تصادفی ساده تفاوت دارد که در آن انتخابهای مختلف، مستقل نیستند، هنگامی که مورد اول انتخاب می‌شود، تمام موردهای بعدی که باید در نمونه گنجانده شوند به طور خودکار (خود بخود) تعیین می‌شوند.

اگر فهرست اصلی جامعه آماری دارای ترتیب تصادفی بود، نمونه‌گیری سیستماتیک نمونه‌ای را حاصل می‌کرد که از نظر آماری می‌توانست یک جانشین منطقی برای نمونه تصادفی به شمار آید. به هر حال، اگر فهرست مثلاً الفبایی باشد، امکان دارد هر نفر K ام جامعه آماری از خصیلت‌های خاصی برخوردار باشد که بر متغیر وابسته بررسی تأثیر بگذارد و بنابراین یک نمونه اریب به دست دهد.

نمونه‌گیری سیستماتیک از روی یک فهرست الفبایی، احتمالاً یک نمونه معرف از گروه‌های مختلف ملی به بار نخواهد آورد؛ زیرا برخی از گروه‌های ملی تمایل دارند که در زیر حروف خاصی گردآیند و فاصله‌نمونه‌گیری می‌تواند آنها را به طور کامل حذف کند یا حداقل به میزان

کافی در نمونه نگنجانند.

باید توجه داشت که روشهای مختلف نمونه‌گیری احتمالاتی که مورد بحث قرار گرفت، نا متداخل (مانعة الجمع) نیستند. ترکیبهای گوناگونی ممکن است مورد استفاده قرار گیرد. برای مثال، اگر پژوهشگری یک جامعه آماری بسیار بزرگ و خیلی پراکنده را مطالعه می‌کند، می‌تواند از نمونه‌گیری خوشه‌ای استفاده کند.

در همان زمان، ممکن است پژوهشگری بخواهد نمونه را طبقه‌بندی کند تا به سؤالهایی دربارهٔ طبقات مختلف پاسخ گوید. در این مورد، پژوهشگر باید جامعه آماری را طبق معیارهای معین طبقه‌بندی کند و سپس به طور تصادفی خوشه‌هایی از آزمودنیها را از میان هر طبقه انتخاب کند.

// نمونه‌گیری غیر احتمالاتی

در بسیاری از موقعیتهای تحقیقاتی، سرشماری عناصر جامعه آماری، که یک ضرورت اساسی برای نمونه‌گیری احتمالاتی است، اگر ناممکن نباشد، دشوار است. در چنین مواردی، محقق از نمونه‌گیری غیر احتمالاتی استفاده می‌کند. گونه‌های اصلی نمونه‌گیری غیر احتمالاتی عبارتند از نمونه‌گیری اتفاقی^(۱)، نمونه‌گیری هدفمند^(۲) و نمونه‌گیری سهمیه‌ای^(۳).

/ نمونه‌گیری اتفاقی

نمونه‌گیری اتفاقی، که ضعیف‌ترین شیوه نمونه‌گیری تلقی می‌شود، عبارت است از استفاده از موردیهای قابل دسترسی برای انجام یک مطالعه. مصاحبه با اولین فردی که در محوطه دانشگاه با او روبرو می‌شویم، استفاده از دانشجویان کلاس خود به عنوان افراد نمونه، یا گرفتن داوطلب برای انجام مصاحبه در تحقیق پیمایشی، مثالهای مختلفی از نمونه‌گیری اتفاقی هستند. هیچ راهی (بجز تکرار مطالعه با استفاده از نمونه‌گیری احتمالاتی) برای برآورد خطایی که از طریق شیوه‌های نمونه‌گیری اتفاقی در نتایج تحقیق وارد می‌شود، وجود ندارد. اگر کسی از نمونه‌گیری اتفاقی استفاده می‌کند، در تفسیر یافته‌های خود باید بی‌نهایت محتاط باشد.

/ نمونه گیری هدفمند

در نمونه گیری هدفمند، که از آن با عنوان نمونه گیری قضاوتی نیز نام برده می شود، آن دسته از عناصر جامعه آماری که حدس زده می شود معرف آن باشند، انتخاب می شوند. بنیان این نوع نمونه گیری بر این مفروضه قرار دارد که قضاوت های نادرست در انتخاب عناصر جامعه آماری، یکدیگر را تعدیل می کنند.

نمونه گیری هدفمند اغلب برای پیش بینی نتایج انتخابات سراسری به کار می رود. در هر ایالت، نواحی کوچکی که نتایج آنها در انتخابات قبلی معرف کل ایالت بوده است، انتخاب می شوند. سپس با تمام رأی دهندگان واجد شرایط در این نواحی مصاحبه می شود و نتایج این مصاحبه ها برای پیش بینی الگوی رأی دهی آن ایالت مورد استفاده قرار می گیرد. کارشناسان نظرسنجی با بکارگیری همین شیوه در تمام ایالتها، نتایج سراسری را پیش بینی می کنند.

در نمونه گیری هدفمند سؤال حیاتی این است که برای دستیابی به یک نمونه معرف، تا چه میزان می توان به قضاوتها اتکا کرد. این مفروضه که خطای قضاوتها لزوماً یکدیگر را تعدیل خواهند کرد، همیشه معتبر نیست. بعلاوه، هیچ دلیلی برای این فرض وجود ندارد که واحدهایی که معرف جامعه آماری بوده اند، در یک دوره زمانی همچنان معرف جامعه آماری باقی بمانند. بنابراین، استفاده از نمونه گیری هدفمند می تواند نتایج گمراه کننده ای به بار آورد.

نمونه گیری هدفمند، به دلیل سهولت کار و هزینه کم، در پیمایشهای نگرش سنجی و نظرسنجی سودمند بوده است؛ هر چند که محقق باید از محدودیتهای این روش آگاه باشد و از آن با نهایت احتیاط استفاده کند.

/ نمونه گیری سهمیه ای

نمونه گیری سهمیه ای انتخاب موردهای معرف از طبقات مختلف یک جامعه آماری است. سهمیه ها بر مبنای ویژگیهای آشکار جامعه آماری، که محقق می خواهد نتایج را به آن تعمیم دهد، تعیین می شوند. اجزای نمونه به گونه ای انتخاب می شوند که نمونه به دست آمده، با توجه به ویژگیهای انتخاب یک تقریب مینیاتوری از جامعه آماری باشد. برای مثال، اگر نتایج سرشماری نشان بدهد که ۲۵ درصد جامعه آماری یک ناحیه شهری در حومه آن زندگی می کنند، پس ۲۵ درصد نمونه باید از حومه شهر انتخاب شوند.

برای نمونه‌گیری سهمیه‌ای باید این مراحل را طی کنید:

۱- تعدادی از متغیرها را که کاملاً با سؤال مورد بررسی وابسته‌اند، تعیین کنید تا مبنایی برای طبقه‌بندی به دست آید. برای این کار، معمولاً از متغیرهایی نظیر جنس، سن، سطح تحصیلات و طبقه اجتماعی استفاده می‌شود.

۲- با استفاده از نتایج سرشماری یا سایر داده‌های موجود، اندازه هر بخش جامعه آماری را مشخص کنید.

۳- سهمیه هر بخش جامعه آماری را متناسب با اندازه آن بخش محاسبه کنید.

۴- مردهای معرف را از هر بخش یا طبقه جامعه آماری انتخاب کنید تا سهمیه‌ها پر شوند. نقص اصلی نمونه‌گیری سهمیه‌ای، در مرحله ۴، یعنی انتخاب افراد از هر طبقه است. محقق به آسانی تشخیص نمی‌دهد که آیا افراد انتخاب شده معرف طبقه مورد نظر هستند یا خیر. انتخاب افراد احتمالاً با توجه به قابلیت دسترسی و سهولت کار انجام می‌شود. اگر محقق برای یک بررسی پیمایشی، ۲۵ درصد از خانوارهای شهرنشین را انتخاب کند، بیشتر محتمل است که به جای خانه‌های کهنه و زهوار در رفته به خانه‌های زیبا برود، به خانه‌هایی برود که بیشتر در دسترس هستند، به خانه‌هایی برود که اعضای آن در طول روز در خانه هستند و الی آخر. چنین روشهایی خود بخود اریبی نظام‌مند در نمونه ایجاد می‌کنند؛ زیرا برخی از اجزای جامعه آماری کنار گذاشته می‌شوند. بعلاوه، هیچ مبنایی برای محاسبه خطای موجود در نمونه‌گیری سهمیه‌ای در دست نیست. با وجود این کمبودها، نمونه‌گیری سهمیه‌ای در بسیاری از پروژه‌های تحقیقاتی مورد استفاده قرار گرفته است که احتمالاً بدون استفاده از این نمونه‌گیری، انجام آنها امکانپذیر نمی‌بود. بسیاری احساس می‌کنند که سرعت گردآوری داده‌ها در این نوع نمونه‌گیری، بر نواقص آن می‌چربد. بعلاوه، سالها تجربه‌اندوزی با نمونه‌های سهمیه‌ای، شناسایی برخی از خطرات آن و نحوه اجتناب از این خطرات را ممکن کرده است.^(۱)

// اندازه نمونه

یکی از اولین سؤالاتی که برای محقق مطرح می‌شود، تعداد آزمودنیهایی است که باید در

۱- برای بحث بیشتر درباره اندازه نمونه مراجعه کنید به:

W.C. Cochran (1985), Sampling Techniques (2d ed.), New York: Wiley; L.Kish (1965), Survey Sampling, New York: Wiley, and S Sudman (1976), Applied Sampling, New York: Academic Press.

نمونه قرار بگیرند. از نظر تکنیکی، اندازه نمونه بستگی دارد به دقتی که در برآورد پارامتر جامعه آماری در یک سطح اطمینان خاص، مورد نظر محقق است. هیچ قانون یگانه‌ای وجود ندارد که بتوان با آن اندازه نمونه را تعیین کرد. اگر پژوهشگری به طور دقیق واریانس جامعه آماری، تفاوت مورد انتظار و احتمالات مطلوب خطاهای نوع اول و دوم (به صفحه ۲۴۹ مراجعه کنید) را تعریف کند، می‌تواند اندازه نمونه لازم را از نظر جبری محاسبه کند. شماری از کتب درسی آمار، این شیوه را توصیف کرده‌اند.

بهترین پاسخ به پرسش اندازه نمونه، این است که تا آنجا که امکان دارد، نمونه بزرگتری به کار ببریم. نمونه بزرگ بسیار محتمل‌تر است که معرف جامعه آماری باشد. علاوه بر آن، با یک نمونه بزرگ، احتمالاً داده‌ها درست‌تر و دقیق‌تر خواهند بود. به بیان دیگر، هر چه نمونه بزرگتر باشد، خطای معیار کوچکتر می‌شود. به طور کلی، خطای معیار میانگین یک نمونه با جذر n نسبت عکس دارد. بنابراین، برای دو برابر کردن دقت یک برآورد، اندازه نمونه باید چهار برابر بشود.

سوی این حکم کلی، اندازه نمونه به همگنی جامعه آماری بستگی دارد. اگر جامعه آماری مورد مطالعه همگن باشد، یک نمونه کوچک معرف آن خواهد بود، اما با افزایش تنوع جامعه آماری، یک نمونه بزرگ ضرورت خواهد داشت. برای مثال، اگر دو جامعه آماری دانش‌آموزان کلاس اول و دانش‌آموزان دوره ابتدایی را در نظر بگیریم، برای اولی یک نمونه کوچکتر و برای دومی یک نمونه بزرگتر مورد نیاز خواهد بود.

بسیاری از مؤلفان توصیه می‌کنند که حداقل ۳۰ آزمودنی در نمونه گنجانده شود؛ زیرا این تعداد، اجازه استفاده از آماره‌های نمونه بزرگ را می‌دهد. بنابراین، در تحقیق آزمایشی باید نمونه‌ای انتخاب کنیم که به ما اجازه دهد تا در هر گروه ۳۰ نفر قرار دهیم. با این همه، در تحقیق توصیفی استفاده از نمونه بزرگ مطلوب است، بویژه هنگامی که جامعه آماری مورد نظر ناهمگن باشد.^(۱)

به هر حال، باید تأکید شود که اندازه نمونه به تنهایی، دقت را تضمین نمی‌کند. معرف بودن، مهمترین جنبه در انتخاب یک نمونه است. ممکن است یک نمونه بزرگ و هنوز دارای اریبی

۱ - برای بحث بیشتر درباره روشهای نمونه‌گیری مراجعه کنید به:

باشد. چنین وضعیتی در نظر سنجی سال ۱۹۳۶ نشریه Literary Digest که شکست پرزیدنت روزولت را پیش‌بینی کرد، بخوبی روشن شده است. اگر چه نمونه این نظر سنجی شامل دو و نیم میلیون پاسخگو بود، اما این نمونه معرف رای دهندگان نبود و بنابراین به نتیجه‌گیری اشتباهی منجر شد. اریبی نمونه، حاصل انتخاب پاسخگویان این نظر سنجی از دفاتر ثبت اتومبیلها، دفاتر راهنمای تلفن و فهرست آبنومان مجلات بود. این آزمودنیها به یقین معرف کل جامعه آماری رای دهندگان در سال ۱۹۳۶ نبوده‌اند. همچنین، از آنجا که نظر سنجی به وسیله پست انجام گرفت، تفاوت‌های بین کسانی که پاسخ را ارسال کردند و آنهایی که پاسخ ندادند، در نتایج اریبی ایجاد کرد. بنابراین پژوهشگر باید آگاه باشد که اندازه نمونه، اریبی حاصله از روشهای غلط نمونه‌گیری را جبران نخواهد کرد. در انتخاب نمونه، معرف بودن باید به عنوان هدف اصلی در نظر باشد.

// مفهوم خطای نمونه‌گیری

هنگامی که از یک نمونه، دربارهٔ جامعه آماری استنباطی صورت می‌گیرد، مقدار معینی خطا در آن استنباط وجود دارد؛ زیرا می‌توان انتظار داشت که بین نمونه‌های تصادفی از یک جامعه آماری نیز تفاوت‌هایی موجود باشد. امکان دارد که میانگین‌های نمره هوش دو نمونه تصادفی دانش‌آموزان کلاس چهارم، که از یک جامعه آماری گرفته شده‌اند، با هم متفاوت باشند. چنین تفاوت‌هایی، که خطای نمونه‌گیری نامیده می‌شوند، از این واقعیت ناشی می‌شود که محقق کل جامعه آماری را مشاهده نکرده است، بلکه یک نمونه را مبنای نتیجه‌گیری خود قرار داده است. خطای نمونه‌گیری را «تفاوت بین پارامتر یک جامعه آماری و آماره یک نمونه» تعریف می‌کنند. بنابراین، اگر کسی از میانگین تمام جامعه آماری (μ) و نیز میانگین یک نمونه تصادفی از همان جامعه آماری ($\bar{X}$) آگاه باشد، تفاوت بین این دو، $\bar{X} - \mu$ ، خطای نمونه‌گیری (e) را می‌نماید. بنابراین $e = \bar{X} - \mu$. برای مثال، اگر ما بدانیم که میانگین نمره هوش یک جامعه آماری ۱۰۰۰ نفری از دانش‌آموزان کلاس چهارم $\mu = ۱۰۰$ است و یک نمونه تصادفی ۲۰۰ نفری، میانگین $\bar{X} = ۹۹$ را دارد، پس خطای نمونه‌گیری $\bar{X} - \mu = ۹۹ - ۱۰۰ = -۱$ خواهد بود. از آنجا که معمولاً برای برآورد پارامتر جامعه آماری به آماره نمونه تکیه می‌کنیم، این نکته که انتظار می‌رود نمونه‌ها با جامعه آماری چقدر تفاوت داشته باشند، یک رکن اساسی در آمار

استنباطی است. به هر حال، رویکرد آمار استنباطی این است که به جای تلاش در جهت تعیین تفاوت بین آماره نمونه و پارامتر جامعه آماری (که اغلب دانسته نیست) تغییرپذیری آماره‌های نمونه‌های مختلف از یک جامعه آماری را برآورد کند. از آنجا که هر آماره نمونه یک برآورد از پارامتر همان جامعه آماری محسوب می‌شود، پس هر تفاوتی بین آماره‌های نمونه‌ها را باید به خطای نمونه‌گیری نسبت داد.

// طبیعت قانونمند خطاهای نمونه‌گیری

با توجه به اینکه نمونه‌های تصادفی مختلف که از یک جامعه آماری انتخاب شده باشند، با یکدیگر تفاوت خواهند داشت، آیا استفاده از یک نمونه برای نتیجه‌گیری در مورد جامعه آماری، واقعاً بهتر از گمانه زنی است؟ بله، بهتر است؛ چون خطاهای نمونه‌گیری به شیوه‌ای قانونمند و قابل پیش‌بینی عمل می‌کنند. قوانین حاکم بر خطای نمونه‌گیری از طریق منطق قیاسی به دست آمده و از طریق تجربه تأیید شده‌اند.

اگر چه ما نمی‌توانیم ماهیت و مقدار خطا در یک نمونه منفرد را پیش‌بینی کنیم، اما می‌توانیم ماهیت و مقدار خطاهای نمونه‌گیری را در کل پیش‌بینی کنیم. اجازه دهید این موضوع را با مراجعه به خطاهای نمونه‌گیری مربوط به میانگین روشن کنیم.

/ خطاهای نمونه‌گیری میانگین

هنگامی که از میانگین یک نمونه $\bar{X}$ برای برآورد میانگین یک جامعه آماری (μ) استفاده می‌شود، همیشه می‌توان انتظار مقداری خطای نمونه‌گیری داشت. اگر چه، چنین برآوردی در عمل مبتنی بر میانگین یک نمونه منفرد است، اما فرض کنیم که نمونه‌های تصادفی متعددی را از یک جامعه آماری انتخاب و میانگین هر یک را محاسبه کرده‌ایم. ما درخواهیم یافت که این میانگینها با یکدیگر و همچنین با میانگین جامعه آماری (اگر آن را بدانیم) تفاوت دارند. این گوناگونی میانگینها ناشی از خطای نمونه‌گیری است که در میانگین هر نمونه تصادفی به عنوان برآورد میانگین جامعه آماری وجود دارد. خطاهای نمونه‌گیری میانگین به دقت مطالعه شده‌اند و این نتیجه به دست آمده است که آنها از قوانین مشخصی پیروی می‌کنند.

امید ریاضی میانگین خطاهای نمونه‌گیری صفر است اگر تعداد نامحدودی نمونه

تصادفی داشته باشیم که از یک جامعه آماری انتخاب شده باشند، می‌توان انتظار داشت که خطاهای مثبت، به گونه‌ای خطاهای منفی را متعادل کنند که میانگین خطاهای نمونه‌گیری صفر شود. برای مثال، اگر میانگین قد دانشجویان سال اول دانشگاه ۵ فوت و ۹ اینچ باشد و نمونه‌های تصادفی متعددی از این جامعه آماری انتخاب شوند، انتظار خواهیم داشت که برخی از نمونه‌ها، میانگین قد بیشتر از ۵ فوت و ۹ اینچ و بعضی نمونه‌ها، میانگین قد کمتر از ۵ فوت و ۹ اینچ داشته باشند. به هر روی، در بلند مدت، خطاهای نمونه‌گیری مثبت و منفی یکدیگر را تعدیل خواهند کرد. اگر ما تعداد نامحدودی نمونه تصادفی با اندازه یکسان داشتیم و میانگین هر یک از این نمونه‌ها را حساب می‌کردیم، سپس میانگین همه این میانگین‌ها را محاسبه می‌کردیم، میانگین میانگین‌ها با میانگین جامعه آماری برابر می‌شد.

از آنجا که خطاهای مثبت با خطاهای منفی برابر هستند، احتمال این که میانگین یک نمونه منفرد، برآورد بزرگتر یا برآورد کوچکتری از میانگین جامعه آماری بدهد، به یک اندازه است. بنابراین، ما می‌توانیم بگوییم که میانگین یک نمونه برآورد نااریب و معقول از میانگین جامعه آماری است.

خطای نمونه‌گیری تابع معکوس اندازه نمونه است هر قدر اندازه نمونه افزایش یابد، مقادیر میانگین نمونه‌های متعدد، نوسان کمتری خواهند داشت. به سخن دیگر، هر قدر اندازه نمونه افزایش یابد، خطای نمونه‌گیری مورد انتظار کاهش می‌یابد. نمونه‌های کوچکتر بیشتر از نمونه‌های بزرگ مستعد خطای نمونه‌گیری هستند. محقق انتظار خواهد داشت که میانگینهای مبتنی بر نمونه‌های ده نفری نوسان بسیار بیشتری از میانگینهای مبتنی بر نمونه‌های ۱۰۰ نفری داشته باشند. در مثالی که درباره قد دانشجویان زدیم، احتمال بسیار بیشتری وجود دارد که در یک نمونه ۴ نفری، ۳ دانشجو بلندتر از متوسط قد و ۱ دانشجو کوتاه‌تر از متوسط قد باشند، تا این که در یک نمونه ۴۰ نفری، ۳۰ دانشجو بالای متوسط قد و ۱۰ دانشجو پایین متوسط قد قرار بگیرند. هر قدر اندازه نمونه افزایش یابد، احتمال نزدیک شدن میانگین نمونه به میانگین جامعه آماری نیز افزایش می‌یابد. بین اندازه نمونه و خطای نمونه‌گیری یک رابطه ریاضی وجود دارد. ما بعداً نشان خواهیم داد که این رابطه چگونه در قالب فرمولهای استنباطی جای گرفته است.

خطای نمونه‌گیری تابع مستقیم انحراف معیار جامعه آماری است بین اعضای یک جامعه آماری، هر قدر پراکندگی یا تنوع بیشتری داشته باشیم، در میانگینهای نمونه نیز انتظار پراکندگی زیادتری خواهیم داشت. برای مثال، میانگین وزن یک نمونه تصادفی ۲۵ نفری، که از

جامعه آماری سوارکاران حرفه‌ای انتخاب شده باشند، از میانگین وزن نمونه تصادفی ۲۵ نفری آموزگاران مدارس، خطای نمونه‌گیری نسبتاً کمتری را نشان خواهد داد. وزنهای سوارکاران حرفه‌ای در یک دامنه محدود قرار می‌گیرند، اما این امر در مورد وزنهای آموزگاران صدق نمی‌کند. بنابراین، اگر اندازه نمونه‌ها یکسان باشد، خطای نمونه‌گیری مورد انتظار برای وزن آموزگاران بیشتر از خطای نمونه‌گیری مورد انتظار برای وزن سوارکاران حرفه‌ای خواهد بود.

خطاهای نمونه‌گیری به شکل نرمال یا نزدیک به نرمال حول امید ریاضی میانگین صفر توزیع می‌شوند میانگینهای نمونه نزدیک به میانگین جامعه آماری اغلب بیش از میانگینهای نمونه دور از میانگین جامعه آماری به دست می‌آیند. ما هر چه از میانگین جامعه آماری دورتر و دورتر می‌شویم، می‌بینیم که میانگینهای نمونه کمتر و کمتری به دست می‌آیند. هم نظریه و هم تجربه نشان داده‌اند که میانگینهای نمونه‌های تصادفی به شکل نرمال یا نزدیک به نرمال حول میانگین جامعه آماری توزیع می‌شوند.

از آنجا که در چنین حالتی خطای نمونه‌گیری، تفاوت بین میانگین نمونه و میانگین جامعه آماری است، توزیع خطاهای نمونه‌گیری نیز از نظر شکل، نرمال یا نزدیک به نرمال است. بنابه تعریف، هر دو توزیع یکسان هستند، به استثنای اینکه توزیع میانگینهای نمونه‌ها، خود میانگینی برابر با جامعه آماری دارد، در حالی که میانگین خطاهای نمونه‌گیری صفر است.

حتی اگر جامعه‌ای که نمونه‌ها از آن انتخاب شده‌اند، به شکل نرمال توزیع نشده باشد، توزیع میانگینهای نمونه‌ها به منحنی نرمال شباهت خواهد داشت. برای مثال، در یک مدرسه ابتدایی نوعی، تعداد کودکان سنین مختلف تقریباً با هم مساوی است، بنابراین نمودار چند ضلعی سن کودکان اساساً به شکل مستطیل خواهد بود. اگر از مدرسه‌ای که تعداد دانش‌آموزان ۶ تا ۱۱ ساله آن با هم مساوی هستند، نمونه‌های تصادفی ۴۰ نفری بگیریم، درخواهیم یافت که بسیاری از نمونه‌ها دارای میانگین سنی نزدیک به میانگین جامعه آماری، یعنی ۸/۵، هستند. نمونه‌هایی که میانگین ۸ یا ۹ دارند، کمتر خواهند بود و نمونه‌های دارای میانگین ۷ و یا ۱۰ به ندرت یافت خواهند شد.

خطای معیار میانگین

از آنجا که گستره و توزیع خطاهای نمونه‌گیری قابل پیش‌بینی است، ما می‌توانیم برای استنباط میانگین جامعه آماری، با یک اطمینان قابل پیش‌بینی، از میانگین نمونه استفاده کنیم. به

هر روی، هنگامی که از میانگین نمونه به عنوان برآوردی از میانگین جامعه آماری استفاده می‌کنیم، به برآوردی از مقدار خطای نمونه‌گیری که در میانگین نمونه مستتر است، نیاز داریم. برای این منظور یک وسیله مهم در دست داریم و آن **خطای معیار میانگین** است.

قبلاً گفته شد که خطای نمونه‌گیری خود را در تغییر پذیری میانگینهای نمونه‌ها نشان می‌دهد. بنابراین، اگر انحراف معیار مجموعه میانگینهای نمونه‌های تصادفی یک جامعه آماری را محاسبه کنیم، برآوردی از مقدار خطای نمونه‌گیری به دست خواهیم آورد. قبلاً دیدیم که دو چیز بر مقدار خطای نمونه‌گیری تأثیر می‌گذارند: اندازه نمونه و انحراف معیار در جامعه آماری. هنگامی که از این دو آگاه باشیم، می‌توانیم انحراف معیار خطاهای نمونه‌گیری را پیش‌بینی کنیم. این امید ریاضی انحراف معیار خطاهای نمونه‌گیری میانگین، خطای معیار میانگین نامیده می‌شود و با نماد σ_x نشان داده می‌شود. با منطق قیاسی نشان داده شده است که خطای معیار میانگین مساوی است با انحراف معیار جامعه آماری (σ) تقسیم بر جذر تعداد موجود در هر نمونه ($\sqrt{n}$). اگر به صورت فرمول بنویسیم:

$$\sigma_x = \frac{\sigma}{\sqrt{n}} \quad (۱-۶)$$

σ_x = خطای معیار میانگین

σ = انحراف معیار جامعه آماری

n = تعداد در هر نمونه

در فصل ۵ دیدیم که انحراف معیار (σ) شاخصی است که درجه پراکندگی در بین افراد یک جامعه آماری را نشان می‌دهد. به همان شیوه، خطای معیار میانگین (σ_x) نیز شاخص پراکندگی مورد انتظار در بین میانگینهای نمونه‌هایی است که از یک جامعه آماری انتخاب شده باشند. همان گونه که خواهیم دید، تفسیر σ و σ_x با هم کاملاً مشابه هستند.

چون میانگینهای نمونه‌های تصادفی توزیع تقریباً نرمال دارند، می‌توانیم از مدل منحنی نرمال نیز برای استنباط میانگینهای جامعه آماری استفاده کنیم. با فرض این که امید ریاضی میانگین میانگینهای نمونه با میانگین جامعه آماری مساوی است و میانگینهای نمونه‌های تصادفی به صورت نرمال توزیع شده‌اند، می‌توانیم یک نمره z برای میانگین نمونه محاسبه

کنیم و آن را به جدول منحنی نرمال ببریم تا احتمال کسب یک میانگین نمونه را به طور اتفاقی، که تا این اندازه دور یا دورتر از میانگین جامعه آماری باشد، برآورد کنیم. نمره z با کم کردن میانگین جامعه آماری از میانگین نمونه و تقسیم حاصل آن بر خطای معیار میانگین به دست می‌آید.

$$z = \frac{\bar{X} - \mu}{\sigma_x} \quad (۶-۲)$$

برای آنکه مطلب روشن شود، یک مسئول پذیرش دانشکده را در نظر بگیرید که از خود می‌پرسد آیا جامعه آماری داوطلبان ورود به دانشکده ما در سطح متوسط قرار دارند یا زیر متوسط؟ میانگین عمومی (سراسری) آزمون ورودی دانشگاهها^(۱) ۵۰۰ و انحراف معیار آن ۱۰۰ است. او یک نمونه تصادفی ۶۴ نفری از جامعه آماری خود می‌گیرد و در می‌یابد که میانگین نمونه ۴۷۰ است. سپس از خود می‌پرسد که چقدر محتمل است که یک نمونه تصادفی ۶۴ نفری با میانگین ۴۷۰ از یک جامعه آماری با میانگین ۵۰۰ انتخاب شده باشد. مسئول پذیرش، با استفاده از فرمول ۱-۶، خطای معیار میانگین را ۱۲/۵ محاسبه می‌کند.

$$\sigma_x = \frac{\sigma}{\sqrt{n}}$$

$$= \frac{100}{\sqrt{64}}$$

$$= 12/5$$

او از طریق فرمول ۲-۶ نمره z را برای میانگین نمونه محاسبه می‌کند.

$$z = \frac{\bar{X} - \mu}{\sigma_x}$$

$$= \frac{470 - 500}{12/5}$$

$$= -2/4$$

بنابراین، میانگین نمونه به اندازه $2/4$ واحد خطای معیار از میانگین جامعه آماری منحرف می‌شود. احتمال داشتن یک میانگین نمونه که به این اندازه ($2/4 \sigma_{\bar{x}}$) یا بیشتر، از میانگین جامعه آماری منحرف شود، چقدر است؟ تنها لازم است به منحنی نرمال مراجعه کند تا این انحراف (z) را بر حسب احتمالات بیان نماید. او با مراجعه به جدول منحنی نرمال در می‌یابد که احتمال یک $2/4 - z =$ یا کمتر، $0/082$ است. این به معنای آن است که یک نمره z با این مقدار یا کمتر از این مقدار، به طور اتفاقی، تنها ۸ بار در ۱۰۰۰ بار رخ خواهد داد. از آنجا که احتمال کسب یک میانگین نمونه که این قدر از میانگین جامعه آماری دور باشد، اندک است، او نتیجه می‌گیرد که میانگین نمونه احتمالاً از جامعه آماری با میانگین ۵۰۰ نیامده است و بنابراین میانگین جامعه آماری او، یعنی داوطلبان ورود به دانشکده او، احتمالاً از ۵۰۰ کمتر است.

/// استراتژی آمار استنباطی

آمار استنباطی، علم تصمیم‌گیری منطقی با اطلاعات محدود است. ما از آنچه که در نمونه‌ها مشاهده می‌کنیم و از آنچه که درباره خطای نمونه‌گیری می‌دانیم، استفاده می‌کنیم تا به تصمیمهای اشتباه پذیر، اما منطقی، در مورد جامعه آماری دست یابیم. شیوه‌های آماری که قبل از دستیابی به این تصمیمها به اجرا در می‌آیند، آزمونهای معناداری نامیده می‌شوند. یک ابزار اساسی این آزمونهای آماری فرضیه صفر است.

// فرضیه صفر

فرض کنید ۱۰۰ دانش‌آموز کلاس چهارم در یک آزمایش در مورد آموزش مفاهیم ریاضی شرکت کرده‌اند. بعلاوه، فرض کنید که فرضیه تحقیق ما این است که آموزش به روش ب بیش از روش الف، تسلط بر این مفاهیم را حاصل می‌کند. ما به طور تصادفی ۵۰ دانش‌آموز را به آموزش این مفاهیم با روش الف و ۵۰ دانش‌آموز دیگر را به آموزش با روش ب تخصیص می‌دهیم. شرایط محیطی را به گونه‌ای فراهم می‌آوریم که تنها تفاوت بین دو گروه، روش آموزش آنها باشد. در پایان آزمایش، از دانش‌آموزان امتحانی می‌گیریم که تعریف مناسبی از تسلط بر مفاهیم ریاضی مورد نظر باشد. بعد از برگزاری امتحان، در می‌یابیم که میانگین دانش‌آموزانی که با روش ب آموزش دیده‌اند، بالاتر از میانگین دانش‌آموزانی است که این مفاهیم را با روش الف یاد

گرفته‌اند. این تفاوت را چگونه تفسیر می‌کنیم؟

با فرض اینکه مراقب بوده‌ایم که برای دو گروه شرایط یادگیری، به استثنای روش آموزش، مساوی باشد، می‌توانیم علت این تفاوت را ناشی از این دو عامل بدانیم: (۱) روش آموزش این تفاوت را موجب شده است، یا (۲) این تفاوت بر حسب تصادف ایجاد شده است. حتی اگر آزمودنیها به طور تصادفی به تیمارها تخصیص یافته باشند، امکان دارد که بر حسب اتفاق، گروه ب شامل دانش‌آموزانی باشد که باهوش‌تر هستند، انگیزه بیشتری برای یادگیری دارند، یا به دلایل دیگری مفاهیم ریاضی را بهتر از دانش‌آموزان گروه الف یاد گرفته‌اند، فارغ از اینکه با چه روشی آموزش دیده‌اند.

بنابراین تفاوت بین گروهها می‌توانست یا نتیجه یک رابطه بین متغیرها باشد (روش آموزش و تسلط بر مفاهیم) یا تنها نتیجه یک اتفاق باشد (خطای نمونه‌گیری). ما چگونه بدانیم که کدامیک از این دو تبیین درست است؟ در نهایت ما قادر به دانستن این امر نیستیم. بنابراین آنچه که ما انجام می‌دهیم، برآورد این امر است که چقدر احتمال دارد که تفاوت حاصله ناشی از یک تصادف باشد و تعیین می‌کنیم که در نتیجه این برآورد کدام تبیین را می‌پذیریم.

تبیین تصادفی بودن، تفاوت را به عنوان فرضیه صفر می‌شناسند. همان‌طور که از فصل ۴ به خاطر خواهید آورد، فرضیه صفر حکمی است در مورد اینکه هیچ رابطه واقعی بین متغیرها وجود ندارد و هر رابطه مشاهده شده‌ای تنها تابعی از یک تصادف است. در مثال ما، فرضیه صفر می‌گوید که هیچ رابطه‌ای بین روش آموزش و تسلط بر مفاهیم ریاضی وجود ندارد.

راه دیگر بیان فرضیه صفر در مثال ما اظهار این امر است که میانگین دانش‌آموزانی که با روش الف آموزش دیده‌اند، برابر با میانگین دانش‌آموزانی است که با روش ب یاد گرفته‌اند. اگر فرضیه صفر را به شکل فرمول بنویسیم و نماد μ را برای میانگین جامعه آماری بکار ببریم، این حکم به صورت زیر در می‌آید.

$$H_0 : \mu_A = \mu_B$$

H_0 = فرضیه صفر

μ_A = میانگین دانش‌آموزان کلاس چهارم که با روش الف آموزش دیده‌اند

μ_B = میانگین دانش‌آموزان کلاس چهارم که با روش ب آموزش دیده‌اند

پژوهشگران در تفسیر تفاوت مشاهده شده بین گروه‌ها، باید بین تبیین تصادفی بودن این تفاوت (فرضیه صفر) و تبیینی که می‌گویند بین متغیرها رابطه‌ای وجود دارد (فرضیه تحقیق)، یکی را انتخاب کنند. پژوهشگر باید این انتخاب را بدون آگاهی از واقعیت نهایی مربوط به جامعه‌های آماری مورد نظر به عمل آورد. پژوهشگر بر مبنای اطلاعات ناکامل تصمیم‌گیری می‌کند و بنابراین در معرض ارتکاب اشتباه احتمالی است.

// اشتباه نوع اول و اشتباه نوع دوم

محقق می‌تواند فرضیه صفر را هم به عنوان فرضیه صفر بپذیرد و هم آن را رد کند. هر یک از این دو تصمیم ممکن است درست یا غلط باشد. اگر فرضیه صفر درست باشد، محقق در صورت پذیرفتن آن در مسیر صحیحی گام برداشته و در صورت رد کردن آن قدم در راهی اشتباه گذاشته است. رد کردن یک فرضیه صفر درست، اشتباه نوع اول نامیده می‌شود.

اگر فرضیه صفر غلط باشد، محقق در صورت پذیرفتن آن در اشتباه است و اگر آن را رد کند، کار درستی انجام داده است. پذیرفتن یک فرضیه صفر غلط، اشتباه نوع دوم نامیده می‌شود. چهار حالت ممکن در جدول ۲ - ۶ خلاصه شده است.

اکنون بیاپید برخی از نتایج حاصله از دو نوع اشتباه در مثال را مورد توجه قرار دهیم.

/ اشتباه نوع اول

محقق اعلام می‌کند که بین روش آموزش و تسلط بر مفاهیم ریاضی رابطه وجود دارد. او روش ب را به عنوان روش بهتر توصیه می‌کند. مدارس کتابها و سایر موادی را که روش الف را متجلی می‌کنند، به کناری می‌نهند و موادی را که بر مبنای روش ب تهیه شده‌اند، خریداری می‌کنند. برای آشنایی معلمان با شیوه آموزش روش ب آموزشهای حین خدمت برگزار می‌شود. بعد از هزینه این همه پول و زمان، مدارس پیشرفتی در یادگیری مفاهیم ریاضی مشاهده نمی‌کنند. آزمایشهای بعدی نیز نتایجی را که در تحقیق اصلی مشاهده شده است، منعکس نمی‌کنند. اگر چه درستی و نادرستی فرضیه صفر هنوز نامشخص است، اما شواهدی که فرضیه صفر را تأیید می‌کنند، غلبه دارند. محقق اصلی بر آشفته و تحقیر می‌شود.

جدول ۱ - ۶ ارائه شماتیک اشتباه نوع اول و اشتباه نوع دوم

وضعیت واقعی در جامعه آماری			
تصمیم‌گیری محقق پس از انجام آزمون معناداری	فرضیه صفر را رد می‌کند	فرضیه صفر نادرست است	فرضیه صفر درست است
		اشتباه نوع اول	درست
	فرضیه صفر را می‌پذیرد	درست	اشتباه نوع دوم

/ اشتباه نوع دوم

محقق نتیجه می‌گیرد که تفاوت بین دو گروه می‌تواند تنها به طور اتفاقی پیش آمده باشد و فرضیه صفر احتمالاً درست است. او اعلام می‌کند برای نتیجه‌گیری در مورد این که یک روش بهتر از روش دیگر است، شواهد کافی در دست نیست.

محققان بعدی به این نتیجه می‌رسند که روش ب از روش الف بهتر است و مدارسی که روش آموزش خود را از الف به ب تغییر داده‌اند، گزارشهای گيرایی از دستاوردهای خود در زمینه یادگیری دانش‌آموزان ارائه می‌کنند. اگر چه واقعیت نهایی نامشخص باقی می‌ماند، اما کوهی از شواهد، فرضیه تحقیق را تأیید می‌کنند. محقق اصلی بر آشفته می‌شود (اما احتمالاً تحقیر نمی‌شود).

اشتباهات نوع اول عموماً به تغییراتی منجر می‌شوند که تضمینی برای آنها وجود ندارد. اشتباهات نوع دوم، معمولاً هنگامی که تضمینی برای تغییر وجود دارد، به حفظ وضع موجود منتهی می‌شوند. پیامدهای یک اشتباه نوع اول، به طور کلی، جدی‌تر از پیامدهای یک اشتباه نوع دوم پنداشته می‌شوند، اگرچه استثناهایی نیز بر این امر وجود دارد.

/// سطح معنی داری

به خاطر داشته باشید که نتیجه‌گیریهای علمی بیش از آنکه احکامی از واقعیت محض باشند، حکمهایی هستند که از احتمال بالایی برای درست بودن برخوردارند. این احتمال چقدر باید بالا باشد، تا محقق بخواهد اعلام کند که بین متغیرها رابطه‌ای وجود دارد؟ به سخن دیگر، فرضیه

صفر چقدر باید نا محتمل باشد تا کسی آن را رد کند؟ پیامدهای نسبی اشتباه نوع اول و اشتباه نوع دوم را می‌سنجند و قبل از انجام آزمایشهای خود، تصمیم می‌گیرند که شواهد تا چه حد باید قوی باشند تا آنها فرضیه صفر را رد کنند. این سطح از پیش تعیین شده که در آن فرضیه صفر رد خواهد شد، سطح معنی‌داری نامیده می‌شود. احتمال یک اشتباه نوع اول مستقیماً در کنترل محقق است؛ زیرا سطح معنی‌داری را از مطابق با نوع اشتباهی که می‌خواهد خود را در مقابل آن حفاظت کند، تعیین می‌کند.

البته، می‌توان همیشه فرضیه صفر را پذیرفت و از خطای نوع اول اجتناب کرد یا همیشه فرضیه صفر را رد کرد و از خطای نوع دوم اجتناب کرد. هیچیک از این دو روش سازنده نیست. اگر پیامدهای یک اشتباه نوع اول بسیار مهم باشد، اما یک اشتباه نوع دوم پیامد کمی داشته باشد، محقق تنها ممکن است در صورتی که احتمال برآورد شده اتفاقی بودن رابطه مشاهده شده یک در هزار یا کمتر باشد، خطر یک اشتباه نوع اول را بپذیرد. این کار را آزمون فرضیه در سطح معنی‌داری $0/001$ می‌نامند، و این سطح یک سطح کاملاً محافظه کارانه محسوب می‌شود. در چنین مواردی محقق بسیار مراقب است که به هنگامی که هیچ رابطه‌ای وجود ندارد، کشف رابطه‌ای را اعلام نکند. به هر حال، این تصمیم‌گیری به معنای پذیرش احتمال بالایی از یک اشتباه نوع دوم است، یعنی اعلام این که هیچ رابطه‌ای وجود ندارد، در حالی که به واقع رابطه وجود دارد.

اگر به نظر برسد که پیامدهای یک اشتباه نوع اول جدی نخواهد بود، امکان دارد محقق، در صورتی که احتمال اتفاقی بودن رابطه مشاهده شده یک در ده یا کمتر باشد، تصمیم بگیرد که وجود رابطه را اعلام کند. این را آزمون فرضیه در سطح معنی‌داری $0/1$ می‌نامند. در اینجا محقق نوعی دوراندیشی اعتدالی در برابر اشتباه نوع اول اعمال می‌کند، در حالی که در برابر اشتباه نوع دوم خطر بزرگی متقبل نمی‌شود.

سطح معنی‌داری احتمال یک اشتباه نوع اول است که محقق مایل است با رد فرضیه صفر خطر آن را بپذیرد. اگر محقق سطح معنی‌داری را $0/01$ قرار دهد، به این معنا است که تنها اگر احتمال برآورد شده در مورد پدیده اتفاقی بودن رابطه مشاهده شده یک درصد باشد، فرضیه صفر را رد خواهد کرد. اگر او سطح معنی‌داری را $0/0001$ قرار دهد، فرضیه صفر را تنها در صورتی رد خواهد کرد که احتمال برآورده شده تابعی از اتفاقی بودن رابطه مشاهده شده، یک در

هزار یا کمتر باشد. دو سطح معنی‌داری که در حیطه تعلیم و تربیت بیشترین مورد استفاده را دارند، سطوح ۰/۰۵ و ۰/۰۱ می‌باشند.

محققان به طور سنتی، سطح معنی‌داری را بعد از سنجش اهمیت اشتباه نوع اول و اشتباه نوع دوم، اما قبل از شروع آزمایش تعیین می‌کنند. اگر داده‌هایی که بعد از تکمیل آزمایش به دست آمده‌اند، نشان دهند که احتمال درست بودن فرضیه صفر مساوی با یا کمتر از احتمال از پیش تعیین شده پذیرفتنی است، فرضیه صفر رد و اعلام می‌شود که نتایج از نظر آماری معنی‌دار (معتبر) هستند. اگر این احتمال بزرگتر از احتمال از پیش تعیین شده پذیرفتنی باشد، نتایج را نامعتبر توصیف می‌کنند، یعنی فرضیه صفر به قوت خود باقی می‌ماند.

منظور از واژه «معنی‌دار»^(۱) در آمار این است: «کمتر از یک احتمال از پیش تعیین شده، محتمل است که تابعی از اتفاق باشد». نتایج پژوهشها می‌توانند از نظر آماری معنی‌دار (معتبر) باشند، بدون اینکه به طور ذاتی با معنی یا مهم باشند.

// آزمونهای جهت‌دار و فاقد جهت^(۲)

به طور معمول، در آزمون فرضیه صفر، جهت تفاوت برای ما اهمیت زیادی ندارد. در مقابل علاقه‌مندیم که دوری یا نزدیکی احتمالی آماره نمونه را از پارامتر جامعه آماری تشخیص دهیم. در مثال قبل درباره مسئول پذیرش دانشکده، دلمشغولی اصلی ما تعیین این مسئله بود که آیا تفاوت بین میانگین نمونه و میانگین جامعه آماری، از آن مقداری که می‌توانست بر حسب اتفاق روی دهد، افزونتر است یا خیر. آزمونی که در آنجا انجام شد، آزمون فاقد جهت نامیده می‌شود؛ زیرا محقق به وجود تفاوت در هر دو جهت علاقه‌مند است. محقق تنها اعلام می‌کند که تفاوت وجود خواهد داشت.

به شکل ۱-۶ توجه کنید که ناحیه رد به طور مساوی بین دو دنباله توزیع تقسیم شده است. بنابراین، اگر یک میانگین نمونه به دست آوریم که خیلی بزرگتر یا خیلی کوچکتر از مقدار مفروض باشد، فرضیه صفر رد می‌شود. جهت تفاوت مهم نیست. اگر محقق بر مبنای تجربه، تحقیقات قبلی، یا نظریه، قادر به بیان جهت تفاوت‌های احتمالی باشد، آنگاه آزمون جهت‌دار انجام خواهد داد. یک فرضیه جهت‌دار تعیین خواهد کرد که پارامتر بزرگتر از، یا کوچکتر از، مقدار مفروض است. بنابراین، در آزمونهای جهت‌دار ناحیه رد در یکی از دو دنباله توزیع قرار

می‌گیرد. در یک منحنی نرمال، این ناحیه نقطه‌ای مساوی با $z = 1/645$ است. معنای این گفته آن است که ما در آزمونهای جهت‌دار، برخلاف آزمونهای فاقد جهت، ۵ درصد را بین دو طرف منحنی تقسیم نمی‌کنیم، بلکه به جای این کار، تمام ۵ درصد خطای اتفاقی را در یک طرف منحنی قرار می‌دهیم. آشکار است که آزمون جهت‌دار در مقایسه با آزمون فاقد جهت، برای رد فرضیه صفر در یک سطح مفروض به مقدار z کوچکتری نیاز دارد ($z = 1/645$ را با $z = 1/96$ مقایسه کنید). پس یک آزمون جهت‌دار، رد فرضیه صفر را سهل‌تر می‌سازد و بنابراین این احتمال را که فرضیه صفر رد شود، افزایش می‌دهد.

در واژگان آمار اغلب از آزمون فاقد جهت به عنوان آزمون معنی داری دو دامنه و از آزمون جهت‌دار به عنوان آزمون یک دامنه نام برده می‌شود. تصمیم برای استفاده از هر یک از آزمونهای یک دامنه یا دو دامنه باید در ابتدای مطالعه، قبل از انجام هر گونه آزمون آماری، آغاز شود. محقق منتظر نمی‌ماند تا داده‌ها بررسی بشوند و سپس آزمون یک دامنه یا دو دامنه را انتخاب کند. شیوه‌های متعددی برای آزمون فرضیه صفر وجود دارند. رایجترین این شیوه‌ها، آزمون t ، تحلیل واریانس، و آزمون χ^2 دو می‌باشند.

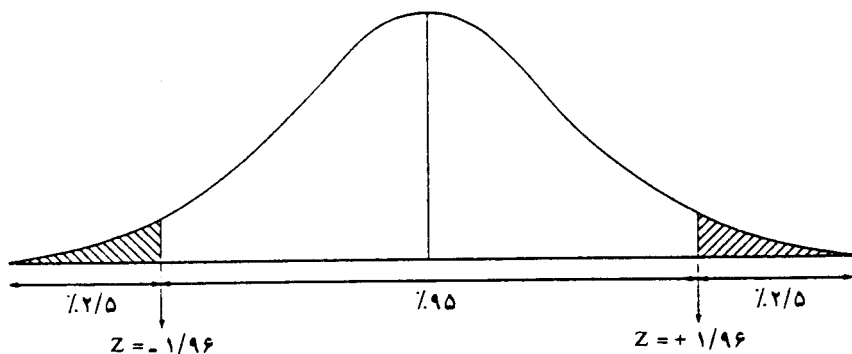
/// آزمون t

ما نشان دادیم که برای مقایسه میانگین نمونه با میانگین جامعه آماری می‌توان منحنی احتمالات نرمال را به کار برد. بدین منظور و برای آنکه دریا بایم آیا میانگین نمونه معرف میانگین جامعه آماری هست یا خیر، از نمره z استفاده می‌کنیم. در مثال قبل، مسئول پذیرش دانشکده خطای معیار میانگین را برای توزیع نمونه‌پیدا کرد، سپس فرمول $X - \mu / \sigma_x$ را به کار برد. در استفاده از این روش تناسب موجود در منحنی احتمالات نرمال نهفته است.

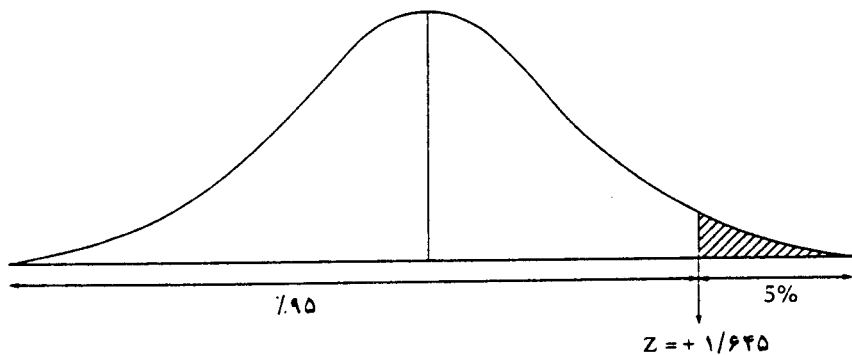
به هر روی، از نظر ریاضی نشان داده شده است که منحنی نرمال تنها هنگامی برای آزمون فرضیه مناسب است که انحراف معیار جامعه آماری مشخص باشد؛ همان طور که در مورد نمره‌های آزمون ورودی دانشگاهها مشخص بود. در بسیاری از موقعیتهای تحقیقاتی انحراف معیار جامعه آماری روشن نیست و باید با فرمول زیر برآورد شود.

$$S = \sqrt{\frac{\sum x^2}{n-1}} \quad (۶-۳)$$

شکل ۱- ۶ منحنی‌های نمایانگر ناحیه‌های رد آزمونهای جهت‌دار و فاقد جهت



الف. منحنی نمایانگر ناحیه رد برای یک آزمون فاقد جهت (آزمون دو دامنه)



ب. منحنی نمایانگر ناحیه رد برای یک آزمون جهت‌دار (آزمون یک دامنه)

برآورد انحراف معیار جامعه آماری S

مجموع مربع نمره‌های انحراف از میانگین یا $\sum (X - \bar{X})^2 = \sum x^2$

تعداد نمونه n

هنگامی که در محاسبه خطای معیار میانگین، این برآورد (S) جانشین انحراف معیار جامعه آماری (σ) می‌شوند، مرسوم است که فرمول ۱ - ۶ به صورت زیر بیان شود.

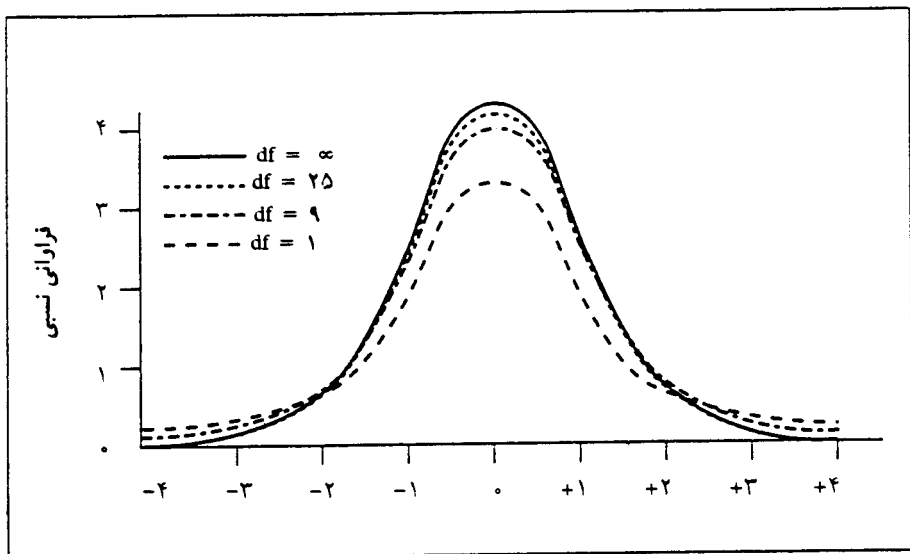
$$\sigma_x = \frac{\sigma}{\sqrt{n}} \longrightarrow s_x = \frac{S}{\sqrt{n}}$$

وقتی از s_x به جای σ_x استفاده شود، هر نمونه محدود، توزیع احتمالات خاص خود را خواهد داشت. این توزیعها به عنوان منحنی t شناخته می‌شوند. هر قدر اندازه نمونه افزایش می‌یابد، این توزیعها بیشتر و بیشتر به منحنی نرمال شباهت پیدا می‌کنند. مجموعه‌ای از توزیعها به نام توزیعهای t برای آزمون فرضیه‌های مربوط به میانگین جامعه آماری، با استفاده از نمونه‌های کوچک تهیه شده‌اند. هر قدر اندازه نمونه کوچکتر شود، توزیع t نیز به طور فزاینده‌ای با توزیع Z متفاوت می‌شود. در تحقیقاتی که ما انجام می‌دهیم، نیازی به دانستن چگونگی محاسبه توزیع t نیست؛ زیرا ضروری‌ترین نتایج این محاسبات در جدول ۲-۱۰ در ضمیمه آمده است. منحنی t کندتر از منحنی نرمال به خط پایه نزدیک می‌شود. در شکل ۲ - ۶، تعدادی از منحنی‌های t ، به همراه منحنی نرمال - که با خط تیره و با عنوان $df = \infty$ مشخص شده - نشان داده شده است.

منحنی‌های t برحسب درجه‌های آزادی آنان، که علامت اختصاری آن df است، نامگذاری می‌شوند. پیش از بحث بیشتر درباره ویژگیهای منحنی‌های t ، اجازه دهید توجه خود را به مفهوم درجه‌های آزادی معطوف کنیم.

درجه آزادی /

عدد درجه آزادی به عددی اطلاق می‌شود که مستقل از هر تغییری حول یک پارامتر ثابت است. برای روشن شدن مفهوم کلی درجه آزادی، تصور کنید که آموزگاری از یک دانش‌آموز می‌خواهد که پنج عدد را که به ذهنش خطور می‌کند، بیان کند. دانش‌آموز آزاد خواهد بود که هر

شکل ۲-۶ منحنی‌های t برای درجه‌های آزادی مختلف

پنج عددی را که خود انتخاب می‌کند، نام ببرد. ما خواهیم گفت که این دانش‌آموز، پنج درجه آزادی دارد. اکنون فرض کنید که آموزگار به دانش‌آموز می‌گوید که پنج عدد را نام ببرد، اما اطمینان حاصل کند که میانگین این پنج عدد برابر ۲۰ است. اکنون این دانش‌آموز آزاد است که برای چهار عدد اول، هر عددی را که انتخاب می‌کند، نام ببرد، اما برای عدد آخر، باید عددی را نام ببرد که مجموع پنج عدد را به ۱۰۰ می‌رساند تا به میانگین ۲۰ دست یابد. اگر دانش‌آموز برای چهار عدد اول ۱۰، ۱۶، ۲۰ و ۳۵ را نام ببرد، پس عدد پنجم او باید ۱۹ باشد. این دانش‌آموز می‌تواند پنج عدد را نام ببرد و تنها یک محدودیت دارد، بنابراین درجه آزادی او، پنج منهای یک، یعنی چهار است ما می‌توانیم این را به صورت فرمول نشان دهیم.

$$df = n - 1$$

$$= 5 - 1$$

$$= 4$$

در نظر بگیرید که آموزگار از دانش‌آموز خواسته بود که از هفت عدد نام ببرد، به گونه‌ای که میانگین سه عدد اول ۱۰ و میانگین هفت عدد ۱۲ باشد. در اینجا او هفت عدد و دو محدودیت دارد، بنابراین درجه آزادی عبارت است از:

$$\begin{aligned}df &= n - 2 \\&= 7 - 2 \\&= 5\end{aligned}$$

هنگامی که انحراف معیار نامعلوم جامعه آماری از انحراف معیار نمونه برآورد می‌شود، یعنی $s_x = \frac{S}{\sqrt{n}}$ ، یک درجه آزادی از دست می‌رود؛ زیرا آماره (های) نمونه از طریق انحراف از میانگین نمونه، که مجموع آن همیشه باید صفر باشد، به دست می‌آیند. بنابراین، همه بجز یکی از انحراف‌های از میانگین، آزادی تغییر دارند؛ به عبارتی $df = N - 1$.

/ آزمون t برای نمونه‌های مستقل

پژوهشگران اغلب دو نمونه را به طور تصادفی از یک جامعه آماری بیرون می‌کشند و تیمار آزمایشی ویژه‌ای را به هر یک تخصیص می‌دهند. این دو گروه، بعد از قرار گرفتن در معرض تیمارها، برای یافتن اثر تیمارها، با توجه به برخی از خصیصه‌ها با هم مقایسه می‌شوند. ممکن است بعد از تیمارها تفاوتی بین دو گروه مشاهده شود، اما امکان دارد این تفاوت از نظر آماری نامعتبر^(۱) باشد؛ یعنی به شانس نسبت داده شود.

در این حالت، شاخصی که برای یافتن معنی‌داری تفاوت بین میانگین‌های دو نمونه مورد استفاده قرار می‌گیرد، آزمون t برای گروه‌های مستقل نامیده می‌شود. این نمونه‌ها را از آن رو مستقل می‌نامند که بدون جور کردن^(۲) یا رابطه دیگری بین دو گروه، به طور مستقل از یک جامعه آماری انتخاب شده‌اند.

اجازه دهید از یک مثال استفاده کنیم. فرض کنید پژوهشگری می‌خواهد در یابد که آیا تنشها بر عملکرد حل مسئله تأثیر دارند یا خیر. اولین مرحله، تخصیص تصادفی دانش‌آموزان یک کلاس به دو گروه است.

از آنجا که افراد دو گروه به طور تصادفی انتخاب شده و به گروه‌ها تخصیص یافته‌اند، میانگین عملکرد دو گروه در تکلیف حل مسئله نباید به صورت معنی‌داری قبل از ارائه تیمار فرق داشته باشد. به هر ترتیب، اگر تنش واقعاً با حل مسئله مرتبط باشد، بعد از ارائه تیمار باید میانگین عملکرد دو گروه به طور معنی‌داری متفاوت باشد. جدول ۲ - ۶، نمره‌های حل مسئله

پس از ارائه تیمار (X)، نمره‌های انحراف از میانگین (x)، و مربع نمره‌های انحراف از میانگین (x^2) افراد دو گروه را که یکی از آنها در شرایط تنش و دیگری در شرایط آرام (بدون تنش) کار کرده‌اند، نشان می‌دهد. میانگین نمره عملکرد آزمودنی‌های گروه تنش ۱۰ و میانگین عملکرد گروه دیگر (گروه بدون تنش) ۱۴ است.

روشن است که بین این دو میانگین تفاوت وجود دارد. ضروری است تعیین کنیم که این تفاوت می‌توانست به آسانی بر حسب اتفاق بوقوع بپیوندد یا خیر. برای این کار ما باید برآورد کنیم که چه مقدار تفاوت بین گروه‌ها تنها بر حسب اتفاق با یک فرضیه صفر درست، مورد انتظار خواهد بود. یک روش مناسب برای انجام این کار خطای معیار تفاوت بین دو میانگین ($s_{x_1 - x_2}$) است. فرمول خطای معیار تفاوت بین دو میانگین در مورد نمونه‌های مستقل عبارت است از:

$$s_{x_1 - x_2} = \sqrt{\frac{\sum x_1^2 + \sum x_2^2}{n_1 + n_2 - 2} \left(\frac{1}{n_1} + \frac{1}{n_2} \right)} \quad (۴-۶)$$

$s_{x_1 - x_2}$ خطای معیار تفاوت بین دو میانگین

$n_1 = ۱$ تعداد موردها در گروه ۱

$n_2 = ۲$ تعداد موردها در گروه ۲

$\sum x_1^2 = ۱$ مجموع مربع نمره انحراف از میانگین در گروه ۱

$\sum x_2^2 = ۲$ مجموع مربع نمره انحراف از میانگین در گروه ۲

گاهی به خطای معیار تفاوت بین دو میانگین، واژه خطا برای آزمون t نیز می‌گویند. در مثال ما این مقدار را چنین محاسبه می‌کنند:

$$\begin{aligned} s_{x_1 - x_2} &= \sqrt{\frac{۱۰۶ + ۴۴}{۱۵ + ۱۵ - ۲} \left[\frac{۱}{۱۵} + \frac{۱}{۱۵} \right]} \\ &= \sqrt{\frac{۱۵۰}{۲۸} \left[\frac{۲}{۱۵} \right]} \\ &= \sqrt{۰/۷۱۴} \\ &= ۰/۸۴ \end{aligned}$$

جدول ۲-۶ محاسبه مقدار t برای دو میانگین نمونه

گروه ۱			گروه ۲		
گروه بدون تنش			(گروه با تنش)		
X_1	x_1	$\dot{x}_1$	X_2	x_2	$\dot{x}_2$
۱۸	+۴	۱۶	۱۳	+۳	۹
۱۷	+۳	۹	۱۲	+۲	۴
۱۶	+۲	۴	۱۲	+۲	۴
۱۶	+۲	۴	۱۱	+۱	۱
۱۶	+۲	۴	۱۱	+۱	۱
۱۵	+۱	۱	۱۱	+۱	۱
۱۵	+۱	۱	۱۰	۰	۰
۱۵	+۱	۱	۱۰	۰	۰
۱۴	۰	۰	۱۰	۰	۰
۱۴	۰	۰	۱۰	۰	۰
۱۳	-۱	۱	۹	-۱	۱
۱۲	-۲	۴	۹	-۱	۱
۱۱	-۳	۹	۸	-۲	۴
۱۰	-۴	۱۶	۷	-۳	۹
۸	-۶	۳۶	۷	-۳	۹
$\Sigma \dot{x}_1 = ۱۰۶$			$\Sigma \dot{x}_2 = ۴۴$		

این محاسبه، تفاوت بین دو میانگین را، که اگر فرضیه صفر درست باشد صرفاً به طور اتفاقی مورد انتظار خواهد بود، به ما می‌دهد. به سخن دیگر، عدد $۰/۸۴$ تفاوتی است که ما بین میانگینهای نمره‌های عملکرد دو گروه انتظار خواهیم داشت؛ اگر آن دو گروه به طور تصادفی از یک جامعه آماری بیرون کشیده شوند و در معرض تیمارهای متفاوت قرار نگیرند. در چنین حالتی، اگر تعداد نامحدودی نمونه داشته باشیم، انتظار خواهیم داشت که در ۶۸ درصد محاسبات تفاوت بین این گروههای تصادفی تفاوت کمتر از $۰/۸۴$ و در ۳۲ درصد دیگر آنها

تفاوت بیشتر از ۰/۸۴ را مشاهده کنیم (دلیل اینکه چرا کار برد فرمول خطای معیار تفاوت بین میانگینها تفاوت برآورد شده مناسبی را به دست می‌دهد که ناشی از اتفاق خواهد بود، خارج از حوصله این متن است).^(۱)

در مثال ما، با یک فرضیه صفر درست، باید انتظار تفاوت ۰/۸۴ را بر حسب اتفاق داشته باشیم. دیدیم که تفاوت بین دو میانگین ۴ است. آیا تفاوت مشاهده شده به اندازه کافی بزرگتر از تفاوت مورد انتظار هست که ما را به رد فرضیه صفر قادر سازد؟
برای پاسخ به این سؤال، ابتدا نسبت دو عدد را به دست می‌آوریم. این نسبت را نسبت t می‌نامند. فرمول آن به صورت زیر است:

$$t = \frac{\bar{X}_1 - \bar{X}_2}{s_{x_1 - x_2}} \quad (5-6)$$

نسبت $t = t$

$\bar{X}_1 - \bar{X}_2$ = تفاوت مشاهده شده بین دو میانگین

$s_{x_1 - x_2}$ = خطای معیار تفاوت بین دو میانگین (تفاوت مورد انتظار بین دو

میانگین هنگامی که فرضیه صفر درست است)

ما می‌توانیم فرمول نسبت t را با گنجاندن فرمول خطای معیار تفاوت بین دو میانگین در آن به صورت کاملتری بنویسیم:

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\left[\frac{\sum x_1^2 + \sum x_2^2}{n_1 + n_2 - 2} \right] \left[\frac{1}{n_1} + \frac{1}{n_2} \right]}} \quad (6-6)$$

در مثال ما مقدار نسبت t عبارت است از:

$$\frac{14 - 10}{0.84} = 4.76$$

۱ - برای اطلاعات بیشتر در مورد منطق این شیوه مراجعه کنید به:

D. Ary and L. C. Jacobs (1976), *Introduction to Statistics*, New York: Holt, Rinehart and Winston.

تفاوت مشاهده شده ما $4/76$ بار بزرگتر از تفاوت مورد انتظار با یک فرضیه صفر درست است. آیا این مقدار به اندازه کافی بزرگ هست که فرضیه صفر در سطح $0/05$ رد شود؟ برای پاسخ به این سؤال ما تنها نیاز داریم که درجه آزادی را محاسبه و به جدول t مراجعه کنیم. درجه آزادی برای آزمون t مستقل عبارت است از تعداد موردها در گروه اول، بعلاوه تعداد موردها در گروه دوم، منهای ۲.

$$df = n_1 + n_2 - 2$$

در مثال ما درجه آزادی $28 = 15 + 15 - 2$ است. اکنون می‌توانیم از جدول t - A ضمیمه، برای تعیین معنی‌داری نتایج استفاده کنیم. در این جدول ستون اول درجه آزادی نام نهاده شده است. محقق با جایگذاری درجه آزادی که در مطالعات خود به دست آورده است، ردیف مناسب را پیدا می‌کند. در مثال ما، به ردیف مربوط به درجه آزادی ۲۸ مراجعه خواهیم کرد. ستونهای بعدی مقادیر t همبسته با احتمالات معین برای آزمونهای جهت‌دار و فاقد جهت را نشان می‌دهند. از آنجاکه در مثال ما، متغیر مستقل می‌تواند بر عملکرد حل مسئله، هم در جهت مثبت و هم در جهت منفی، تأثیر بگذارد، ضروری است که یک آزمون فاقد جهت انجام دهیم. در ردیف مربوط به درجه آزادی ۲۸، در ستونی که عنوان $0/1$ برای آزمون فاقد جهت دارد، عدد $1/701$ را پیدا می‌کنیم. این عدد به ما می‌گوید که با فرضیه صفر درست و درجه آزادی ۲۸، یک نسبت t $1/701$ یا بیشتر یا $-1/701$ یا کمتر، در هر ده بار یک بار، برحسب اتفاق به وقوع خواهد پیوست. عدد $2/048$ در ستونی که عنوان $0/05$ دارد، گویای آن است که طبق فرضیه صفر درست و با درجه آزادی ۲۸، یک نسبت t $2/048 \pm$ یا بیشتر در پنج درصد مواقع بر حسب اتفاق به وقوع خواهد پیوست.

نسبت مشاهده شده ما $4/76$ بزرگتر از $2/048$ است. معنای این امر آن است که تفاوت بین گروههای ما بزرگتر از مقدار لازم برای رد فرضیه صفر در سطح معنی‌داری $0/05$ است. احتمال برآورد شده درست بودن فرضیه صفر کمتر از ۵ درصد است ($P < 0/05$). اگر چه ما به یقین نمی‌دانیم که متغیرها، تنش و حل مسئله، با هم رابطه دارند، اما شواهدی که طبق معیارهای از قبل مقرر شده به دست آمده‌اند، به اندازه کافی معتبر هستند که ما را قادر سازند نتیجه بگیریم که رابطه مشاهده شده تنها یک رویداد اتفاقی نیست. اگر نسبت t مشاهده شده کمتر از $2/048$

می‌بود، نتیجه می‌گرفتیم که شواهد به اندازه کافی معتبر نیستند تا ما را به اعلام وجود رابطه‌ای بین متغیرها راهبر شوند و به سخن دیگر، فرضیه صفر را می‌پذیرفتیم.

توجه داشته باشید که ارقام t ، از چپ به راست، مقادیر لازم t را برای رد فرضیه در سطوح بالاتر معنی‌داری نشان می‌دهند. برای $df = 28$ مقدار $2/736$ یا بیشتر به رد فرضیه صفر در سطح $0/01$ رهنمون خواهد شد. مقدار $3/674$ یا بیشتر به رد فرضیه صفر در سطح $0/001$ منجر خواهد شد. بنابراین مقدار t ما، یعنی $4/76$ ، نه تنها در سطح $0/05$ ($P < 0/05$) معنی‌دار است، بلکه همچنین در سطح $0/01$ ($P < 0/01$) و در سطح $0/001$ ($P < 0/001$) نیز معنی‌دار می‌باشد.

/ منطقی آزمون t

اگر نسبت t را در نظر بگیریم، صورت آن تفاوت واقعی مشاهده شده بین دو گروه است. مخرج کسر $s_p^2 - s_x^2$ برآوردی است از این امر که انتظار می‌رود این دو گروه بر حسب اتفاق چقدر با هم تفاوت داشته باشند، یعنی تفاوتی را نشان می‌دهد که انتظار می‌رود بین دو گروه، که با روش تصادفی از یک جامعه آماری مشترک انتخاب شده‌اند، وجود داشته باشد. این مخرج مبتنی است بر: (۱) تعدادی که در نمونه‌ها هستند، $n_1 + n_2$ (هر چه این تعداد بیشتر باشد تفاوت‌های تصادفی کمتری بین میانگین‌های نمونه انتظار می‌رود)، (۲) تغییر در داخل گروه‌ها، S_p و S_1 (هر چه تغییر در داخل گروه‌ها بیشتر باشد، تفاوت تصادفی بیشتری در بین گروه‌ها انتظار می‌رود). از آنجا که مخرج کسر اندازه‌ای است از این که چه مقدار تفاوت آشکار بر حسب اتفاق می‌توان انتظار داشت، آن را واژه خطای آزمون t می‌نامند.

اگر نسبت تفاوت مشاهده شده (صورت کسر) بر واژه خطا (مخرج کسر) تقسیم شود و مقدار حاصله برابر مقدار معین در جدول مقادیر t ، یا بیشتر از آن باشد، می‌توان فرضیه صفر را در سطح معنی‌داری نشان داده شده رد کرد.

/ آزمون t برای گروه‌های نامستقل

تاکنون بحث ما حول مقایسه میانگینهای به دست آمده از دو گروه نمونه مستقل تمرکز یافته بود. در یک نمونه مستقل، هر نفر به طور تصادفی از جامعه آماری انتخاب می‌شود و ترکیب یک

گروه، اثری بر گروه دیگر ندارد. به هر حال، در برخی از مواقع ممکن است محقق بخواهد که در بعضی از صفات، که برای هدف تحقیق اهمیت دارند، آزمودنیهای دو گروه را با هم جور کند، یا بخواهد میانگینهایی را که از یک گروه تحت دو شرایط آزمایشی متفاوت به دست آمده است، با هم مقایسه کند. در چنین حالتی دیگر گروهها مستقل نیستند؛ زیرا ترکیب یک گروه با ترکیب گروه دیگر مرتبط است. در چنین حالتی ما انتظار خواهیم داشت نمره‌های متغیر وابسته همبستگی داشته باشند. بنابراین باید آزمون t برای میانگین‌های نامستقل یا همبسته به کار برده شود؛ این آزمون به نام آزمون t همبسته یا وابسته نیز شناخته می‌شود. اندازه‌ای که با آزمون t نامستقل تحلیل می‌شود، تفاوت بین نمره‌های جفت شده است.

بیایید مثالی را در نظر بگیریم. فرض کنید ما می‌خواهیم بدانیم که آیا گذراندن یک درس روش تحقیق، بر نگرش دانشجویان نسبت به تحقیق تأثیر می‌گذارد؟ برای بررسی این مسئله ما یک کلاس روش تحقیق را انتخاب می‌کنیم و نگرش دانشجویان را در روز اول و روز آخر کلاس نسبت به امر تحقیق می‌سنجیم. فرض کنید چنین داده‌هایی را جمع‌آوری کرده‌ایم و نتایج آنرا در جدول ۶-۳ آورده‌ایم. ستونهای ۲ و ۳ نمره‌های هر دانشجو را در اولین و دومین آزمون نشان می‌دهند. ستون ۴ تفاوت بین نمره‌های اول و دوم هر دانشجو را ارائه می‌کند. مجموع این تفاوتها به ۳۰+ می‌رسد. میانگین این تفاوتها، ۲+، با تقسیم $(\sum D)$ بر n ، تعداد مشاهده‌های جفت شده یا ۱۵، به دست می‌آید. ستون ۵ مربع تفاوتها را نشان می‌دهد. فرمول آزمون t نامستقل عبارت است از:

$$t = \frac{\bar{D}}{\sqrt{\frac{\sum D^2 - \frac{(\sum D)^2}{N}}{N(N-1)}}} \quad (6-7)$$

t مقدار t برای میانگینهای نامستقل (همبسته)

D تفاوت بین نمره‌های جفت شده

$\bar{D}$ میانگین تفاوتها

$\sum D^2$ مجموع مربع تفاوتها

N تعداد زوجها

جدول ۳- ۶ نمره‌های قبل - و - بعد ۱۵ دانش آموز در کلاس مقدماتی تحقیق

شماره آزمودنی	پیش آزمون	پس آزمون	D	D ^۲
۱	۱۰	۱۲	+۲	+۴
۲	۹	۱۳	+۴	+۱۶
۳	۸	۱۲	+۴	+۱۶
۴	۱۱	۹	-۲	+۴
۵	۱۰	۸	-۲	+۴
۶	۷	۹	+۲	+۴
۷	۱۰	۱۲	+۲	+۴
۸	۹	۱۱	+۲	+۴
۹	۸	۱۰	+۲	+۴
۱۰	۶	۱۰	+۴	+۱۶
۱۱	۱۰	۱۲	+۲	+۴
۱۲	۷	۱۳	+۶	+۳۶
۱۳	۱۰	۶	-۴	+۱۶
۱۴	۹	۱۳	+۴	+۱۶
۱۵	۱۰	۱۴	+۴	+۱۶
			$\Sigma D = +۳۰$	$\Sigma D^2 = +۱۶۴$

با جایگزینی مقادیر جدول ۳- ۶ در فرمول ۷- ۶ به دست می آوریم:

$$t = \frac{\frac{30}{15}}{\sqrt{\frac{164 - \frac{(30)^2}{15}}{15(15-1)}}} = \frac{2}{\sqrt{\frac{164 - 60}{210}}} = \frac{2}{\sqrt{\frac{104}{210}}} = \frac{2}{\sqrt{0.4952}} = \frac{2}{0.704} = 2.84$$

نسبت t می گوید که تفاوت مشاهده شده ۲/۸۴ بار بزرگتر از تفاوتی است که با یک فرضیه صفر درست مورد انتظار خواهد بود. اکنون برای تعیین معنی داری آماری نسبت مشاهده شده

باید به جدول مقادیر t (جدول ۲- A ضمیمه) مراجعه کنیم.

عدد درجه آزادی برای آزمون t نامستقل برابر است با $1 - N$ که N تعداد زوجهای مشاهده است. ما در مثال خود درجه آزادی $14 = 15 - 1$ را داریم. در جدول مقادیر t به هنگام انجام یک آزمون فاقد جهت می‌بینیم که با درجه آزادی ۱۴، برای آن که t در سطح 0.05 معنی‌دار باشد، به عدد 2.145 و برای آن که در سطح 0.01 معنی‌دار باشد، به عدد 2.977 نیاز است. مقدار به دست آمده t ، 2.84 ، افزونتر از مقدار داده شده در جدول برای سطح 0.05 است، اما به مقدار داده شده برای سطح 0.01 نمی‌رسد. معنای این امر آن است که تفاوت بین دو میانگین در سطح 0.01 معنی‌دار نیست. اگر ما سطح معنی‌داری را 0.05 تعیین کرده بودیم، می‌توانستیم نتیجه بگیریم که نگرش دانشجویان در مورد تحقیق تغییر کرده است.

/ آزمون t برای ضریب همبستگی r پیرسون

کاربرد مهم دیگر آزمون t ، آزمون فرضیه‌های مربوط به همبستگی جامعه آماری (p) است. قبل از آن که محقق بتواند از داده‌های نمونه، در مورد همبستگی جامعه آماری نتیجه‌گیری کند، به آزمون معنی‌داری رابطه بین متغیرها نیاز دارد. در چنین مواردی رایجترین فرضیه صفر می‌گوید که همبستگی جامعه آماری صفر و همبستگی مشاهده‌شده در نمونه (r) تابعی از اتفاق است. برای مثال، یک محقق می‌تواند یک نمونه ۲۵ نفری از دانشجویان سال اول بگیرد، آزمونهای وازگان و موقعیت‌یابی فضایی را با آنها اجرا کند، و عدد 0.2 را برای r پیرسون بین آن دو آزمون به دست آورد. گام بعدی، تصمیم‌گیری در مورد این امر است که آیا این ضریب همبستگی می‌توانست یک نتیجه اتفاقی در جامعه آماری باشد که همبستگی واقعی (p) در آن صفر است؟ توزیع t مدل مناسبی برای آزمون فرضیه صفر است که می‌گوید ضریب همبستگی جامعه آماری صفر است. آماره t را می‌توان با استفاده از فرمول زیر محاسبه کرد:

$$t = r \sqrt{\frac{N-2}{1-r^2}} \quad (8-6)$$

t = مقدار

r = ضریب همبستگی بدست آمده

N = تعداد مشاهده‌های زوج شده

با کاربرد فرمول ۸ - ۶ در مورد مثال بالا، در می‌یابیم که مقدار t ، ۱/۰۲ است:

$$t = 0/20 \sqrt{\frac{27 - 2}{1 - (0/20)^2}} = 1/02$$

مقدار t محاسبه شده با مقادیر جدول t در درجه آزادی $N - 2$ مقایسه می‌شوند. اگر مقدار مشاهده شده در سطح معنی‌داری مورد نظر با مقدار جدول مساوی یا از آن بیشتر باشد، یک رابطه معنی‌دار وجود دارد. جدول $A - 2$ ضمیمه نشان می‌دهد که وقتی یک آزمون فاقد جهت انجام می‌دهیم، برای معنی‌داری در سطح ۰/۰۵ با درجه آزادی (df) ۲۵، به عدد t ، ۲/۰۶ نیاز داریم. چون t محاسبه شده، یعنی ۱/۰۲ کمتر از مقدار جدول است، محقق باید نتیجه بگیرد که برای بیان وجود رابطه بین دو متغیر شواهد کافی در دست ندارد.

جداولی تهیه شده‌اند که مقدار بحرانی لازم ضریب همبستگی r را برای رد فرضیه صفر در سطوح معنی‌داری مختلف نشان می‌دهند (به جدول $A - 5$ ضمیمه مراجعه کنید). یک r معنی‌دار آن است که مقدار آن مساوی با، یا بزرگتر از مقدار جدول، با درجه آزادی $N - 2$ ، باشد (N تعداد زوجهای نمره است). جدول $A - 5$ نشان می‌دهد که هنگامی که یک آزمون دو دامنه انجام می‌دهیم، برای رد فرضیه صفر، با $df = 25$ ، به r پیرسون بیشتر از ۰/۳۸۰۹ + یا کمتر از ۰/۳۸۰۱ - نیاز داریم.^(۱)

بنابراین، همبستگی ۰/۲ به‌دست آمده در بررسی مربوط به دانشجویان سال اول معنی‌دار نیست. اگر نمونه نسبتاً بزرگی داشته باشیم، ممکن است یک ضریب همبستگی از نظر مقدار کم، اما از نظر آماری معنی‌دار باشد. از آنجا که مقدار همبستگی میزان رابطه بین متغیرها را نشان می‌دهد، یک همبستگی کم همیشه نمایانگر یک رابطه ضعیف است؛ حتی هنگامی که همبستگی از نظر آماری معنی‌دار باشد. برای مثال جکسون و لاوهاردن با یک نمونه ۱۴۴ نفری از دختران کلاس ششم، همبستگی ۰/۲۵ + را بین پاسخ دانش‌آموزان به پرسشنامه رضایت از مدرسه و پیش‌بینی آموزگاران آنها در مورد نحوه پاسخگویی دانش‌آموزان به این پرسشنامه به دست آوردند.^(۲) از

۱ - از یک آزمون یک دامنه یا دو دامنه می‌توان استفاده کرد. اگر محقق تنها به یک ضریب مثبت یا منفی نظر داشته باشد، از آزمون یک دامنه استفاده می‌کند. به هر روی، اگر هر دو ضریب مثبت یا منفی مورد نظر باشند، آزمون دو دامنه مورد استفاده قرار می‌گیرد.

2- Jackson, P.W., and H.M. Lahaderne (1967). Scholastic success and attitude toward school in a population of sixth graders. *Journal of Educational Psychology*, 58, 15 - 18.

آنجا که اندازه نمونه بزرگ بود، همبستگی اگر چه پایین، اما از نظر آماری در سطح ۰/۰۱ معنی‌دار بود. یافته‌ها نشان می‌دهند که در جامعه آماری‌ای که نمونه معرف آن است، همبستگی احتمالاً صفر نمی‌شود. هر چند، همبستگی کم خود گویای آن است که اگر چه آموزگاران می‌توانند رضایت دانش‌آموزان را در سطحی بالاتر از تصادف و اتفاق پیش‌بینی کنند، اما پیش‌بینی آموزگاران تنها رابطه ضعیفی با رضایت واقعی دانش‌آموزان دارد.

همچنین، آزمون t می‌تواند برای آزمون فرضیه‌هایی درباره همبستگیهای به غیر از صفر جامعه آماری به کار رود. همچنین، از آن می‌توان برای آزمون این فرضیه که همبستگی مشاهده شده در دو نمونه می‌توانسته است از یک جامعه آماری نشئت گرفته باشد، استفاده کرد. از آنجا که این کتاب، یک کتاب مقدماتی است، از گنجانیدن چنین آزمونهایی در آن خودداری کردیم. توصیف سودمندی از این آزمونها را می‌توان در کتاب گلاس و هاپکینز^(۱) و بسیاری دیگر از کتابهای آماری یافت.

/// تحلیل واریانس

در تحلیل واریانس (ANOVA) نظیر آزمون t نسبت تفاوت مشاهده شده به واژه خطا، برای آزمون فرضیه، مورد استفاده قرار می‌گیرد. این نسبت که نسبت F نامیده می‌شود، از واریانس (σ^2) میانگینهای گروهی به عنوان اندازه‌ای از تفاوتهای مشاهده شده در بین گروهها، استفاده می‌کند. معنای این امر آن است که ANOVA روشی متنوعتر از آزمون t است. از آزمون t تنها برای آزمون تفاوت بین دو میانگین می‌توان استفاده کرد. برخی از آمار دانان هرگز از آزمون t استفاده نمی‌کنند؛ زیرا از ANOVA می‌توان در هر موقعیتی که آزمون t به کار می‌آید، استفاده کرد و بیش از آن می‌توان کارهای بسیاری صورت داد که از آزمون t بر نمی‌آید.

منطق اساسی ANOVA آن است که در یک آزمایش، واریانس کل همه آزمودنیها را می‌توان به دو منبع تقسیم کرد. واریانس بین گروهی و واریانس درون گروهی.

واریانس بین گروهی در نسبت F بخشی از صورت کسر است. واریانس درون گروهی قسمتی از واژه خطا یا مخرج کسر است؛ همان گونه که در آزمون t نیز هست. هر چقدر واریانس

بین گروهی افزایش می‌یابد، F نیز زیاد می‌شود. هر قدر واریانس درون‌گروهی افزایش می‌یابد، از نسبت F کاسته می‌شود. تعداد آزمودنیها بر نسبت F تأثیر می‌گذارد. این تعداد هر چه بزرگتر باشد، نسبت F بزرگتر می‌شود. هنگامی که صورت و مخرج کسر مساوی هستند، تفاوتهای بین میانگینهای گروهی بزرگتر از مقداری نیست که تنها بر حسب اتفاق انتظار می‌رود. اگر صورت کسر بزرگتر از مخرج آن باشد، محقق به جدول مقادیر F (در ضمیمه) مراجعه می‌کند تا تعیین کند که آیا نسبت F به اندازه کافی بزرگ هست که او را به رد فرضیه صفر در یک سطح از پیش تعیین شده، قادر سازد.

// محاسبه نسبت F (تحلیل واریانس ساده)

فرض کنید ما سه شرایط آزمایشی تنش زیاد، تنش متوسط و بدون تنش داریم و می‌خواهیم عملکرد سه گروه از افراد را که به طور تصادفی به این شرایط سه گانه تخصیص یافته‌اند، در یک عمل ساده حل مسئله مقایسه کنیم. فرض کنید داده‌هایی که در جدول ۴-۶ ارائه شده، خلاصه مشاهدات ما از عملکرد این سه گروه است و اکنون می‌خواهیم فرضیه صفر را، مبنی بر این که تفاوت معنی‌داری میان این مشاهدات نیست، در سطح ۰/۰۱، آزمون کنیم.

می‌توان دید که میانگینها با یکدیگر و با میانگین کل ۳۰ آزمودنی (X ، میانگین کل) تفاوت دارند. آیا تفاوتهای بین میانگینها به اندازه کافی بزرگ هستند که از نظر آماری معنی‌دار باشند یا احتمال دارد که این تفاوتها به طور اتفاقی به وقوع پیوسته باشند؟ برای پاسخ به این سؤال نسبت F را محاسبه می‌کنیم.

اولین گام یافتن مجموع مربعات انحراف نمره‌های فردی از میانگین کل است. این شاخص مجموع مربعات کل نامیده می‌شود و تمام اثرات تیمار و خطای نمونه‌گیری را منعکس می‌کند.

$$SS_t = \sum x_t^2 = \sum X_t^2 - \frac{(\sum X_t)^2}{N} \quad (۹-۶)$$

در مثال ما این مقدار عبارت است از:

$$SS_t = ۶۸۳۰ - \frac{(۴۴۴)^2}{۳۰} = ۲۵۸/۸$$

جدول ۴-۶ اندازه‌های حاصله از سه نمونه تصادفی مبتنی بر عملکرد آنها در شرایط تنش زیاد، تنش متوسط و بدون تنش

گروه ۱ تنش زیاد		گروه ۲ تنش متوسط		گروه ۳ بدون تنش	
X_1	X_1^2	X_2	X_2^2	X_3	X_3^2
۱۹	۳۶۱	۲۲	۴۸۴	۱۵	۲۲۵
۱۸	۳۲۴	۲۰	۴۰۰	۱۴	۱۹۶
۱۷	۲۸۹	۱۹	۳۶۱	۱۴	۱۹۶
۱۶	۲۵۶	۱۸	۳۲۴	۱۳	۱۶۹
۱۵	۲۲۵	۱۷	۲۸۹	۱۳	۱۶۹
۱۵	۲۲۵	۱۶	۲۵۶	۱۲	۱۴۴
۱۴	۱۹۶	۱۶	۲۵۶	۱۲	۱۴۴
۱۳	۱۶۹	۱۵	۲۲۵	۱۱	۱۲۱
۱۲	۱۴۴	۱۴	۱۹۶	۱۱	۱۲۱
۱۱	۱۲۱	۱۲	۱۴۴	۱۰	۱۰۰
$\Sigma X_1 = 150$	$\Sigma X_1^2 = 1310$	$\Sigma X_2 = 169$	$\Sigma X_2^2 = 2935$	$\Sigma X_3 = 125$	$\Sigma X_3^2 = 1585$
$\bar{X}_1 = 15/0$		$\bar{X}_2 = 16/9$	$\bar{X}_2 = 12/5$	$\bar{X}_3 = 14/8$	
$\Sigma X_1 = 444$		$\Sigma X_2 = 6830$			

سپس مجموع کل مربعاتی انحراف میانگینهای گروهی از میانگین کل را پیدا می‌کنیم. این شاخص، مجموع مربعاتی بین گروهی نامیده می‌شود (اگر بیش از دو گروه داشته باشیم، از نظر دستوری درست‌تر است بگوییم مجموع مربعاتی بین گروهها. به هر حال، استفاده از اصطلاح مجموع مربعاتی بین گروهی یک سنت دیر پا است، و برای این که با سایر کتابهای درسی آمار در تضاد نباشیم، این اصطلاح را بکار می‌بریم). این شاخص با استفاده از فرمول زیر به دست می‌آید.

$$SS_b = \Sigma \bar{x}_b = \frac{(\Sigma X_1)^2}{n_1} + \frac{(\Sigma X_2)^2}{n_2} + \dots - \frac{(\Sigma X)^2}{N} \quad (6-10)$$

در مسئله ما این مقدار عبارت است از:

$$SS_b = \sum x_b^2 = \frac{(150)^2}{10} + \frac{(169)^2}{10} + \frac{(125)^2}{10} - \frac{(444)^2}{30} = 97/4$$

سپس مجموع کل مربعاتی انحراف نمره‌های فردی از میانگین گروه خود را پیدا می‌کنیم. این شاخص، مجموع مربعاتی درون گروهی نامیده می‌شود و با کاربرد فرمول نمره‌های خام برای مجموع مربع انحراف از میانگین هر گروه و بعد جمع همه گروه‌ها به دست می‌آید $(SS_w = SS_1 + SS_2 + SS_3)$.

$$SS_w = \sum x_w^2 = \sum X_1^2 - \frac{(\sum X_1)^2}{n_1} + \sum X_2^2 - \frac{(\sum X_2)^2}{n_2} + \dots \quad (6-11)$$

در مسئله ما این مقدار عبارت است از:

$$SS_w = \sum x_w^2 = 2310 - \frac{(150)^2}{10} + 2935 - \frac{(169)^2}{10} + 1585 - \frac{(125)^2}{10} = 161/4$$

همچنین می‌توانستیم مجموع مربعاتی درون گروهی را با منها کردن مجموع مربعاتی بین گروهی از مجموع مربعاتی کل به دست آوریم، یعنی:

$$SS_w = SS_t - SS_b \quad \text{یا} \quad \sum x_w^2 = \sum x_t^2 - \sum x_b^2 \quad (6-12)$$

در مورد مثال ما

$$SS_w = 258/8 - 97/4 = 161/4$$

آزمون معنی‌داری F

جدول ۵-۶ نتایج محاسباتی را که تاکنون انجام داده‌ایم، همراه با نتایج محاسبات بعدی، در خود جای داده است. ستون اول این جدول سه منبع واریانس را فهرست کرده است: واریانس بین گروهی، واریانس درون گروهی و واریانس کل. ستون دوم مجموع مربع‌ها را شامل است، که ما قبلاً آنها را محاسبه کردیم. ستون سوم این جدول عدد درجه آزادی همبسته با هر منبع واریانس را فهرست کرده است. عدد درجه آزادی برای واریانس بین گروهی برابر است با

(۱ - G) که G تعداد گروه‌ها است. در مثال ما این مقدار عبارت است از $2 = 3 - 1$. درجه‌های آزادی برای واریانس درون گروهی برابر است با $... + 2 - n_2 + 1 - n_1$ (یا $[n - G]$ ، تعداد کل نمره‌ها $[n]$ منهای تعداد گروه‌ها $[G]$). در مثال ما این مقدار عبارت است از $27 = 30 - 3$ یا $27 = 10 - 1 + 10 - 1 + 10 - 1$. عدد درجه آزادی برای واریانس کل برابر است با $1 - N$ ، در مثال ما، $29 = 30 - 1$. این عدد آخر را می‌توانستیم با جمع کردن درجه‌های آزادی بین‌گروهی و درون‌گروهی به دست آوریم.

سپس، گام بعدی محاسبه میانگین مربع بین‌گروهی و درون‌گروهی است. این مقادیر با تقسیم مجموع مربعات بین‌گروهی و درون‌گروهی به درجه‌های آزادی آنها به دست می‌آید. مقادیر حاصله میانگین مربع‌ها هستند. در مثال ما میانگین مربع بین‌گروهی عبارت است از $48/7 = 97/4 \div 2$ و میانگین مربع درون‌گروهی عبارت است از $5/98 = 161/4 \div 27$. میانگین مربع درون‌گروهی واژه خطا برای نسبت F است. سرانجام، با کاربرد فرمول زیر روش محاسبه تحلیل واریانس، نسبت F، به پایان می‌رسد:

جدول ۵-۶ خلاصه تحلیل واریانس سه گروه

(۱)	(۲)	(۳)	(۴)	(۵)	(۶)
منبع	SS	df	MS	F	سطح
واریانس					معنی‌داری
بین‌گروهی	۹۷/۴	۲	۴۸/۷	۸/۱۴	۰/۰۱
درون‌گروهی	۱۶۱/۴	۲۷	۵/۹۸		
کل	۲۵۸/۸	۲۹			

$$F = \frac{MS_b}{MS_w} = \frac{SS_b / df_b}{SS_w / df_w} \quad (۶-۱۳)$$

در مثال ما این مقدار عبارت است از:

$$F = \frac{48/7}{5/98} = 8/14$$

اکنون برای تعیین معنی داری نسبت F حاصله از نظر آماری، به جدول ۳-۱ ضمیمه مراجعه می‌کنیم. ستونی را که در راس آن درجه آزادی بین‌گروهی (صورت کسر) آزمایش ما قرار دارد، پیدا می‌کنیم و در این ستون رو به پایین می‌رویم تا به ردیفی که مدخل آن منطبق با عدد درجه آزادی درون گروهی (مخرج کسر است) است، برسیم. در اینجا دو مقدار پیدا خواهیم کرد: یک عدد با حروف ریز و یک عدد با حروف درشت.

اگر نسبت F ما برابر یا بزرگتر از مقداری باشد که با حروف ریز نوشته شده، در سطح $0/05$ معنی دار است. اگر نسبت F به دست آمده برابر یا بزرگتر از مقداری باشد که با حروف درشت نوشته شده، در سطح $0/01$ نیز معنی دار است. در مثال ما، با درجه‌های آزادی ۲ و ۲۷، برای رد فرضیه صفر در سطح $0/05$ به یک نسبت F ۳/۳۵ و در سطح $0/01$ به یک نسبت F ۵/۴۹ نیاز داریم. از آنجا که نسبت F به دست آمده بزرگتر از این دو مقدار است، در سطح $0/01$ معنی دار است و فرضیه صفر در این سطح رد می‌شود.

روش تحلیل واریانس بر این مفروضه مبتنی است که اگر گروه‌هایی که قرار است با هم مقایسه شوند، از یک جامعه آماری انتخاب شده‌اند، میانگین مربعهای بین گروهی نباید بیش از مقداری که به طور اتفاقی انتظار خواهیم داشت، با میانگین مربعهای درون گروهی تفاوت داشته باشند.

بنابراین، ما با یک فرضیه صفر درست، انتظار خواهیم داشت که نسبت F تقریباً مساوی ۱ باشد. از سوی دیگر، اگر فرضیه صفر نادرست باشد، تفاوت بین میانگین گروهها باید بزرگتر از آن چیزی باشد که برحسب اتفاق انتظار داریم، به گونه‌ای که میانگین مربع بین گروهی افزونتر از میانگین مربع درون گروهی باشد. در چنین مواردی، نسبت F یعنی میانگین مربع بین گروهی تقسیم بر میانگین مربع درون گروهی، بزرگتر از ۱ خواهد بود. سپس به جدول مقادیر F مراجعه می‌کنیم (جدول ۳-۱) تا ببینیم این نسبت برای داده‌های ما به اندازه کافی بزرگتر از ۱ هست تا بتوانیم فرضیه صفر را در سطح مورد نظر رد کنیم.

همچنانکه تفاوت بین این میانگین مربعها افزایش می‌یابد، نسبت F نیز افزایش می‌یابد و احتمال درستی فرضیه صفر کاهش می‌یابد.

هنگامی که فرضیه صفر به عنوان نتیجه این روش تحلیل واریانس رد می‌شود، نمی‌توانیم چیزی بیش از این بگوییم که اندازه‌هایی که از این گروهها به دست آورده‌ایم، با هم تفاوت دارند و

این تفاوتها بزرگتر از مقداری است که تنها به طور اتفاقی انتظار خواهیم داشت.

یک نسبت F معنی‌دار لزوماً به این مفهوم نیست که تمام گروهها به طور معنی‌داری با تمام گروههای دیگر تفاوت دارند. F معنی‌دار ممکن است نتیجه تفاوت موجود بین یک گروه با بقیه گروهها باشد. برای نمونه امکان دارد در مسئله ما گروه ۳ به طور معنی‌داری با گروه ۱ و گروه ۲ تفاوت داشته باشد، اما گروه ۱ و ۲ با هم تفاوت معنی‌داری نداشته باشند. آزمونهای آماری بسیاری وجود دارند که می‌توان آنها را برای یافتن مکان معنی‌داری تفاوتها به کار برد؛ آزمونهایی که به وسیله تاکي^(۱) و شفه^(۲) تهیه شده‌اند، در این زمینه خیلی سودمند هستند.^(۳)

در مثال ما، سه گروه را به طور تصادفی از یک جامعه آماری برگزیدیم و بنابراین می‌توانیم فرض کنیم که قبل از آن که تیمارهای آزمایشی را در مورد آنها اعمال کنیم، با هم تفاوتی فراتر از آنچه که معلول تصادف است، ندارند. معنی‌داری نسبت F نشان می‌دهد که تفاوت یافت شده بین این گروهها، بعد از تیمار، فراتر از آن است که حاصل اتفاق باشد. ما این تفاوت را به تیمار آزمایشی نسبت می‌دهیم و نتیجه می‌گیریم که سطح تنش بر عملکرد افراد در تکالیف ساده حل مسئله تأثیر می‌گذارد.

این تمام آن چیزی است که ما در تفسیر این نسبت F می‌توانیم بگوییم. اگر به تحلیل آماری بیشتری نیاز داشته باشیم، می‌توانیم از آزمونهای تاکي، شفه، یا سایر آزمونها برای تعیین معنی‌داری بین گروههای مشخص یا ترکیبهای گروهها استفاده کنیم. این روشها (تکنیکها) به ما می‌گویند که چگونه شرایط تنش خاصی بر عملکرد تأثیر می‌گذارند و می‌توانند به سؤالاتی نظیر «آیا در شرایط تنش متوسط و زیاد، تنش متوسط و بدون تنش، تنش زیاد و بدون تنش، در نمره‌های عملکرد تفاوتی وجود دارد؟»، پاسخ گویند.

// تحلیل واریانس چند عاملی

امکان دارد بخواهیم اثر ترکیبی سطح تنش و نیاز پیشرفت در عملکرد یک تکلیف مسئله گشایی را بررسی کنیم. برای بررسی این مسئله هم سطح تنش و هم نیاز پیشرفت را تغییر خواهیم داد. طرح یک آزمایش که اثرهای ترکیبی دو یا چند متغیر مستقل را بررسی می‌کند، یک طرح عاملی خوانده می‌شود و نتایج آن به وسیله تحلیل واریانس چند عاملی بررسی می‌شود.

اجازه دهید فرض کنیم که این آزمایش را با استفاده از پنج آزمودنی در هر گروه انجام داده‌ایم و داده‌های جدول ۶-۶ خلاصه‌ای از مشاهدات ما را در مورد عملکرد آزمودنیها منعکس می‌کند. کاربرد تحلیل واریانس چند عاملی ما را قادر می‌سازد که این موارد را پیدا کنیم: (۱) آیا تفاوت معنی‌داری بین عملکرد آزمودنیهایی که در شرایط تنش زیاد قرار دارند و آنها که در شرایط تنش کم هستند، وجود دارد، (۲) آیا تفاوت معنی‌داری بین عملکرد آزمودنی‌ها با نیاز پیشرفت زیاد و آزمودنیها با نیاز پیشرفت کم، وجود دارد و (۳) آیا دو متغیر، تنش و نیاز پیشرفت، بر عملکرد آزمودنیها اثر ترکیبی دارند؟ اثرهایی که با تحلیل‌های اول و دوم بررسی شدند، اثرهای اصلی نامیده می‌شوند، در حالی که سومی اثر تعاملی خوانده می‌شود. حاصل نهایی این تحلیل‌ها سه نسبت F خواهد بود، دو تای آنها معنی‌داری دو اثر اصلی و سومی معنی‌داری اثر تعاملی را نشان می‌دهد.

جدول ۶-۶ اندازه‌های تکلیف مسئله‌گشایی آزمودنیهایی با نیاز پیشرفت کم و زیاد تحت شرایط تنش زیاد و کم

تنش			
	زیاد	کم	
زیاد	۲۰	۲۳	$\sum X_{r1} = 200$ $\bar{X}_{r1} = 20$
	۲۰	۲۲	
	۱۹	۲۱	
	۱۹	۲۰	
	۱۷	۱۹	
	$\sum X_{c1} = 95$	$\sum X_{c2} = 105$	
کم	۲۲	۱۸	$\sum X_{r2} = 175$ $\bar{X}_{r2} = 17.5$
	۲۱	۱۶	
	۲۰	۱۵	
	۱۹	۱۴	
	۱۸	۱۲	
	$\sum X_{c1} = 100$	$\sum X_{c2} = 75$	
$\sum X_{c1} = 195$		$\sum X_{c2} = 180$	$\sum X = 375$ کل
$\bar{X}_{c1} = 19.5$		$\bar{X}_{c2} = 18$	$\bar{X} = 18/75$ (میانگین کل)

محاسبه این نسبت‌های F شامل گامهای زیر است:

- ۱- مجموع مربعات کل، مجموع مربعات بین گروهی و مجموع مربعات درون‌گروهی را با استفاده از همان روشها و فرمولهایی که در تحلیل واریانس ساده به کار رفت، پیدا کنید. این مقادیر، که از داده‌های جدول ۶-۶ استخراج شده‌اند، عبارتند از:

$$SS_t = 7181 - \frac{(375)^2}{20} = 149/75$$

$$SS_b = \frac{(95)^2}{5} + \frac{(105)^2}{5} + \frac{(100)^2}{5} + \frac{(75)^2}{5} - \frac{(375)^2}{20} = 103/75$$

$$SS_w = 149/75 - 103/75 = 46$$

- ۲- مجموع مربعات بین گروهی را به سه مجموع مربعات جداگانه بشکنید (تبدیل کنید):
(الف) مجموع مربعات بین ستونها، (ب) مجموع مربعات بین ردیفها و (ج) مجموع مربعات برای تعامل بین ستونها و ردیفها.
- الف - مجموع مربعات بین ستونها معرف مجموع توان دوم انحرافهای ناشی از تفاوت بین میانگینهای ستونی و میانگین کل است. این مجموع را با استفاده از فرمول زیر پیدا می‌کنند:

$$SS_{bc} = \frac{(\sum X_{c1})^2}{n_{c1}} + \frac{(\sum X_{c2})^2}{n_{c2}} + \dots - \frac{(\sum X)^2}{N} \quad (6-14)$$

با استفاده از این فرمول، مجموع مربعات بین ستونها برای داده‌های جدول ۶-۶ عبارت است از:

$$SS_{bc} = \frac{(195)^2}{10} + \frac{(180)^2}{10} - \frac{(375)^2}{20} = 11/25$$

- ب - مجموع مربعات بین ردیفها عبارت است از مجموع توان دوم انحرافهای ناشی از تفاوت بین میانگینهای ردیفی و میانگین کل. این مجموع را با کاربرد فرمول زیر به دست می‌آورند:

$$SS_{br} = \frac{(\sum X_{r1})^2}{n_{r1}} + \frac{(\sum X_{r2})^2}{n_{r2}} + \dots - \frac{(\sum X)^2}{N} \quad (6-15)$$

برای داده‌های جدول ۶-۶، این مقدار عبارت است از:

$$SS_{br} = \frac{(200)^2}{10} + \frac{(175)^2}{10} - \frac{(375)^2}{20} = 31/25$$

ج - تعامل مجموع مربعات بخشی از انحراف بین میانگینهای گروهی و میانگینهای کلی است که نه از تفاوت‌های ردیفی و نه از تفاوت‌های ستونی ناشی می‌شود. به سخن دیگر، این مجموع، تفاوت بین کل مجموع مربعات بین گروهها و مجموع مربع‌های بین ردیفها است، یعنی:

$$SS_{int} = SS_b - (SS_{bc} + SS_{br}) \quad (۶-۱۶)$$

اگر به زبان کلمات بیان کنیم، مجموع مربعات تعاملی برابر است با مجموع مربع‌های بین گروهی منهای مجموع مربع‌های بین ستونی و مجموع مربعات بین ردیفی برای داده‌های جدول ۶-۶، مقدار این تعامل عبارت است از:

$$SS_{int} = 103/75 - (11/25 + 31/25) = 61/25$$

۳- عدد درجه آزادی همبسته با هر منبع تغییر را تعیین کنید. درجه‌های آزادی را به صورت زیر پیدا می‌کنیم:

df برای مجموع مربعات بین ستونها = ۱ - C

df برای مجموع مربعات بین ردیفها = ۱ - R

df برای تعامل = (۱ - R) (۱ - C)

df برای مجموع مربعات بین گروهها = ۱ - G

df برای مجموع مربعات درون گروهها = N - G

df برای مجموع مربعات کل = ۱ - N

C = تعداد ستونها

R = تعداد ردیفها

G = تعداد گروها

N = تعداد آزمودنی‌ها در همه گروها

۴- مقادیر میانگین مربعها را با تقسیم هر مجموع مربع بر عدد درجه آزادی همبسته آن پیدا کنید.

۵- نسبتهای F را برای اثرهای اصلی و تعاملی با تقسیم میانگین مربعهای بین‌گروهی بر میانگین مربع درون‌گروهی برای هر یک از سه جزء محاسبه کنید.

۶- نتایج محاسباتی که بر مبنای داده‌های جدول ۶-۶ انجام گرفته، به طور خلاصه در جدول ۶-۷ آمده است. سه نسبت F در این جدول فهرست شده است. برای یافتن معنی‌داری هر یک از این مقادیر، مانند قبل، به جدول مقادیر F مراجعه می‌کنیم. در مراجعه به این جدول از عدد درجه آزادی همبسته با هر نسبت F (df برای صورت کسر) و عدد درجه آزادی همبسته با میانگین مربع درون‌گروهی (df برای مخرج کسر) استفاده می‌کنیم. برای مثال، نسبت F بین ستونهای ما ۳/۹۱۳ است. با مراجعه به جدول F، در می‌یابیم که با درجه‌های آزادی ۱ و ۱۶، یک نسبت F ۴/۴۹ یا بیشتر برای معنی‌داری در سطح ۰/۰۵، مورد نیاز است. چون نسبت F ما کمتر از مقداری است که در جدول نشان داده شده است، پس معنی‌دار نیست.

جدول ۶-۷ خلاصه یک تحلیل واریانس چند عاملی ۲ × ۲

منبع واریانس	SS	df	MS	F	سطح معنی‌داری
بین ستونها (تنش)	۱۱/۲۵	۱	۱۱/۲۵	۳/۹۱۳	-
بین ردیفها (نیاز پیشرفت)	۳۱/۲۵	۱	۳۱/۲۵	۱۰/۸۶۹	۰/۰۱
ستونها در ردیفها (تعامل)	۶۱/۲۵	۱	۶۱/۲۵	۲۱/۳۰۴	۰/۰۱
بین گروها	۱۰۳/۷۵	۳	۳۴/۵۸۳		
درون گروها	۴۶	۱۶	۲/۸۷۵		
کل	۱۴۹/۷۵	۱۹			

نسبت F برای بین ردیفها، با درجه‌های آزادی ۱ و ۱۶، برای معنی‌دار بودن باید به $۴/۴۹$ (در سطح $۰/۰۵$) یا $۸/۵۳$ (در سطح $۰/۰۱$) برسد. چون مقدار F به دست آمده‌ما، $۱۰/۸۶۹$ ، از هر دوی این مقادیر بیشتر است، در سطح $۰/۰۱$ معنی‌دار است. برای تعامل بین ستونها و ردیفها، با درجه‌های آزادی ۱ و ۱۶ یک نسبت F $۴/۴۹$ (در سطح $۰/۰۵$) یا $۸/۵۳$ (در سطح $۰/۰۱$) مورد نیاز است. مقدار F به دست آمده‌ما، $۲۱/۳۰۴$ ، از هر دوی این مقادیر بیشتر است، بنابراین در سطح $۰/۰۱$ معنی‌دار است.

/ تفسیر نسبت‌های F

اولین نسبت F در جدول ۷ - ۶ (بین ستونها) معنی‌دار نیست و نشان می‌دهد که شرایط تنش‌زا در تأثیرشان بر عملکرد آزمودنیها در آزمایش، به طور معنی‌دار با هم تفاوت ندارند. این تجزیه و تحلیل مقایسهٔ عملکرد ترکیبی گروههای ۱ و ۲ با عملکرد ترکیبی گروههای ۳ و ۴ است. ما می‌توانستیم با استفاده از روش آزمون t به همین نتیجه برسیم.

نسبت F دوم (بین ردیفها)، که در سطح $۰/۰۱$ معنی‌دار است، بر مبنای مقایسه عملکرد آزمودنیهای گروههای ۱ و ۳ با آزمودنیهای گروههای ۲ و ۴ بنا نهاده شده است. ما از معنی‌داری این نسبت F می‌توانیم استنباط کنیم که تفاوت بین عملکرد آزمودنیهای با نیاز پیشرفت زیاد و آزمودنیهای با نیاز پیشرفت کم فراتر از یک تفاوت اتفاقی است. با بررسی داده‌های جدول ۷ - ۶ در می‌یابیم که گروههای با نیاز پیشرفت زیاد میانگین ترکیبی ۲۰ را در مقایسه با میانگین $۱۷/۵$ گروههای با نیاز پیشرفت کم کسب کرده‌اند. از آنجا که برای این تفاوت یک نسبت F معنادار داریم، نتیجه می‌گیریم که در شرایطی مشابه شرایط آزمایش ما، از آزمودنیهایی که نیاز پیشرفت زیاد داشته باشند، سطح بالاتری از عملکرد را می‌توان انتظار داشت.

نسبت F سوم اثر تعاملی بین دو متغیر، سطح تنش و نیاز پیشرفت را نشان می‌دهد. در این مورد معنی‌داری نسبت F ، یعنی اثر سطح تنش بر عملکرد یک کار مسئله‌گشایی به درجه نیاز پیشرفت بستگی دارد.

مقایسه نتایج مشاهده شده با نتایج مورد انتظار در صورت وجود تعامل بین دو متغیر مستقل، این پدیده را با وضوح بیشتری نشان می‌دهد.

اکنون اجازه دهید میانگین چهار گروه را در صورتی که هیچ تعاملی وجود نمی‌داشت، محاسبه کنیم. میانگین تمام آزمودنیها ۱۸/۷۵ است. میانگین برای ده آزمودنی که در شرایط تنش زیاد قرار دارند، ۱۹/۵ است که ۰/۷۵ بیشتر از میانگین کل است، در حالی که میانگین ده آزمودنی که در شرایط تنش کم قرار دارند، ۰/۷۵ کمتر است. میانگین ده آزمودنی با نیاز پیشرفت زیاد، ۲۰ است که ۱/۲۵ بیشتر از میانگین تمام آزمودنیها است، در حالی که میانگین ده آزمودنی که نیاز پیشرفت کم دارند، ۱/۲۵ کمتر است.

ما می‌توانیم برای هر گروه میانگینی را محاسبه کنیم که اگر هیچ تعاملی وجود نمی‌داشت، انتظار می‌رفت آن گروه چنین میانگینی را دارا باشد. ما این کار را با افزودن تفاوت مربوط به ستونی که گروه در آن قرار دارد و تفاوت مربوط به ردیفی که گروه در آن است، به میانگین کل انجام می‌دهیم.

اگر تعاملی وجود نمی‌داشت، انتظار می‌داشتیم که میانگین گروه ۱ چقدر باشد؟ با میانگین کل، ۱۸/۷۵، شروع می‌کنیم، ۰/۷۵ را به آن اضافه می‌کنیم؛ زیرا آزمودنیها در شرایط تنش زیاد قرار داشتند و سپس ۱/۲۵ را به آن می‌افزاییم، زیرا آنها نیاز پیشرفت زیاد داشتند. بعد از این افزایشها یک میانگین ۲۰/۷۵ را به دست می‌آوریم.

با پیروی از همین روش برای هر چهار گروه، مقادیر مورد انتظار زیر را به دست خواهیم آورد:

مقدار مورد انتظار	=	تفاوت مربوط به نیاز پیشرفت	+	تفاوت مربوط به تنش	+	میانگین رویهم
۲۰/۷۵		+ ۱/۲۵		+ ۰/۷۵		۱۸/۷۵
۱۸/۲۵		- ۱/۲۵		+ ۰/۷۵		۱۸/۷۵
۱۹/۲۵		+ ۱/۲۵		- ۰/۷۵		۱۸/۷۵
۱۶/۷۵		- ۱/۲۵		- ۰/۷۵		۱۸/۷۵
گروه ۱						گروه ۱
گروه ۲						گروه ۲
گروه ۳						گروه ۳
گروه ۴						گروه ۴

اکنون میانگینهای گروهی واقعی را با این میانگینهای گروهی مورد انتظار مقایسه می‌کنیم:

واقعی		مورد انتظار	
		زیاد	کم
زیاد	کم	گروه ۱ $X=۱۹$	گروه ۳ $X=۲۱$
	زیاد	$X=۲۰$	$X=۲۰$
کم	کم	گروه ۲ $X=۲۰$	گروه ۴ $X=۱۵$
	زیاد	$X=۱۷/۵$	$X=۱۷/۵$
$\bar{X}=۱۹/۵$		$\bar{X}=۱۸$	$\bar{X}=۱۸/۷۵$

		زیاد	کم
آسان	کم	گروه ۱ $X=۲۰/۷۵$	گروه ۳ $X=۱۹/۲۵$
	زیاد	$X=۲۰$	$X=۲۰$
دشوار	کم	گروه ۲ $X=۱۸/۲۵$	گروه ۴ $X=۱۶/۷۵$
	زیاد	$X=۱۷/۵$	$X=۱۷/۵$
$\bar{X}=۱۹/۵$		$\bar{X}=۱۸$	$\bar{X}=۱۸/۷۵$

(توجه کنید که ما می‌توانستیم از تفاوت‌های بین مقادیر مورد انتظار و مقادیر مشاهده‌شده برای محاسبه مجموع مربعات برای تعامل، مستقیماً استفاده کنیم. میانگین واقعی هر گروه با میانگین مورد انتظار آن $۱/۷۵$ تفاوت دارد. این مقدار را به توان دو می‌رسانیم و در تعداد موردها ضرب می‌کنیم تا $۶۱/۲۵ = ۲۰ \times ۱/۷۵^2$ را به دست آوریم)

می‌بینیم که گروه ۱ بدتر از آن که ما انتظار داشتیم، عمل کردند و می‌دانیم که آنها در شرایط تنش زیاد هستند و نیاز پیشرفت زیاد دارند. گروه ۲، که نیاز پیشرفت کم داشتند و در شرایط تنش زیاد بودند، بهتر از انتظار ما بودند. اگر گروه‌های تحت تنش کم را در نظر بگیریم، در می‌یابیم که گروه ۳ با نیاز پیشرفت زیاد بهتر از انتظار عمل کرده، در حالی که گروه ۴ با نیاز پیشرفت کم، کمتر از انتظار عمل کرده است.

از آنجا که آزمون F نشان داد که تعامل معنی‌دار است، نتیجه می‌گیریم که تنش زیاد هنگامی که با نیاز پیشرفت کم ترکیب شود، نمره‌های بهتری ایجاد می‌کند تا اینکه با نیاز پیشرفت زیاد ترکیب شود، در حالی که تنش کم زمانی که با نیاز پیشرفت زیاد همراه می‌شود، نسبت به ترکیب آن با نیاز پیشرفت کم نمره‌های بهتری را موجب می‌شود.

در تحقیقات آموزشی استفاده از تحلیل چند عاملی از ارزش زیادی برخوردار است؛ زیرا بسیاری از سؤالاتی که دست اندرکاران آموزش و پرورش، تحقیق در مورد آنها را ضروری می‌بینند، به طور ذاتی دارای طبیعت پیچیده‌ای هستند. این روشها ما را قادر می‌سازد که اثرات ترکیبی دو یا چند متغیر مستقل را در رابطه با یک متغیر وابسته تحلیل کنیم. برای مثال، یک مقایسه ساده میانگینهای متغیر وابسته دو گروه از دانش‌آموزان که با روشهای متفاوتی آموزش دیده‌اند، ممکن است نتایج نامعتبر حاصل کند. اما اگر هوش را به عنوان یک متغیر مستقل

اندازه‌گیری شده در آزمایش دخالت دهیم، ممکن است دریابیم که یک روش با دانش‌آموزان کم‌هوش‌تر، نتیجه بهتری می‌دهد، در حالی که روش دیگر برای دانش‌آموزان باهوش‌تر مؤثرتر است. تحلیل واریانس چند عاملی، مانند مثال ما، به دو متغیر مستقل محدود نیست. امکان داخل کردن هر تعداد از متغیرهای مستقل در این روش وجود دارد. کتابهای آمار متعددی، از جمله کتاب ادواردز، محاسبه و تفسیر این روشها را توصیف کرده‌اند.^(۱)

/// آزمون معنی‌داری خی دو^(۲)

گاهی از اوقات نیازمند آن هستیم که معنی‌داری تفاوت میان نسبتهای آزمودنیها، اشیاء، رویدادها و نظایر آنها را که در طبقه بندیهای مختلفی جای می‌گیرند، پیدا کنیم. آزمون آماری که در چنین مواردی به کار گرفته می‌شود، آزمون خی دو (χ^2) نامیده می‌شود.

در آزمون خی دو دو مجموعه از فراوانیها با هم مقایسه می‌شوند: فراوانی مشاهده شده و فراوانی مورد انتظار. فراوانی مشاهده شده، همان گونه که از نام آن آشکار است، فراوانی واقعی است که در نتیجه مشاهده به دست آمده است. فراوانی مورد انتظار یک فراوانی نظری است که برای مقایسه به کار می‌رود.

این فرضیه را در نظر بگیرید که نسبت دانش‌آموزان پسر به دختر در درس آمار، با نسبت کل دانش‌آموزان پسر و دختر در مدرسه متفاوت است. اگر بدانیم که ۴۰ درصد از کل نام نویسی در مدرسه، مربوط به پسران است و ۳۰۰ دانش‌آموز در درس آمار نام نوشته‌اند، فراوانیهای مورد انتظار عبارت خواهند بود از:

$$\left. \begin{array}{l} ۱۲۰ \text{ دانش‌آموزان پسر} \\ ۱۸۰ \text{ دانش‌آموزان دختر} \end{array} \right\} ۳۰۰$$

اکنون فرض کنید که فراوانیهای مشاهده شده ما چنین باشد.

$$\left. \begin{array}{l} ۱۴۰ \text{ دانش‌آموزان پسر} \\ ۱۶۰ \text{ دانش‌آموزان دختر} \end{array} \right\} ۳۰۰$$

1- Edwards, A. L. (1984). *Experimental Designs in Psychological Research* (5 th ed.,; ch. 11 and 12). New York: Harper & Row.

می‌خواهیم تعیین کنیم آیا تفاوت بین فراوانیهای مورد انتظار و مشاهده شده، از نظر آماری معنی‌دار است؟ برای تعیین این مسئله فرمول خ‌ی دو را به کار می‌بریم، که به صورت زیر است:

$$\chi^2 = \sum \frac{(f_o - f_e)^2}{f_e} \quad (۱۷ - ۶)$$

χ^2 = مقدار خ‌ی دو

f_o = فراوانی مشاهده شده

f_e = فراوانی مورد انتظار

اگر داده‌ها را در این فرمول بگذاریم، خواهیم داشت:

$$\chi^2 = \frac{(۱۴۰ - ۱۲۰)^2}{۱۲۰} + \frac{(۱۶۰ - ۱۸۰)^2}{۱۸۰} = ۵/۵۵$$

برای تعیین اینکه این مقدار خ‌ی دو معنی‌دار است، به جدول مقادیر χ^2 در ضمیمه مراجعه می‌کنیم (جدول ۴ - A). اولین ستون این جدول عدد درجه آزادی مربوط به مسئله خ‌ی دو را نشان می‌دهد. ستونهای دیگر، مقادیر مورد نیاز برای سطوح مختلف معنی‌داری را ارائه می‌کنند. عدد درجه آزادی، همان گونه که قبلاً بحث کردیم، مبتنی است بر تعداد مشاهده‌های مستقل از تغییر، به هنگامی که محدودیتهای خاصی در مورد داده‌ها اعمال شده است. هنگامی که تعداد ثابتی مشاهده داشته باشیم که تنها به دو طبقه تقسیم شده‌اند، به محض اینکه تعداد موجود در یک طبقه مشخص شد، دیگری نیز ثابت می‌شود. بنابراین، زمانی که دانش‌آموزان پسر را ۱۴۰ نفر تعیین کردیم، تعداد دانش‌آموزان دختر در کل ۳۰۰ نفر، باید ۱۶۰ نفر باشد. در این مثال تنها یک درجه آزادی وجود دارد. در مسائلی نظیر این، عدد درجه آزادی برابر است با $K - ۱$ وقتی K ، تعداد طبقه‌های مورد استفاده برای طبقه‌بندی باشد. با مراجعه به جدول χ^2 در می‌یابیم که مقدار مشاهده شده ۵/۵۵، در سطح ۰/۰۵ (و ۰/۰۲) از نظر آماری معنی‌دار است.

با تفسیر این نتیجه اکنون می‌توانیم بگوییم که در مدرسه ما، با سطح اطمینان ۰/۰۵، نسبت پسرانی که در درس آمار شرکت می‌کنند، از دختران بیشتر است. سطح معنی‌داری ۰/۰۵ به این معنا است که در هر ۱۰۰ بار مشاهده، چنین تفاوتی بین نسبت‌های دانش‌آموزان پسر و دختر، کمتر از ۵ بار آن می‌تواند نتیجه اتفاق باشد. بنابراین داده‌ها، فرضیه تحقیق ما را مبنی بر اینکه

نسبت دانش‌آموزان پسری که در درس آمار نام‌نویسی می‌کنند، بیشتر از دانش‌آموزان دختر است، تأیید می‌کنند.

استفاده از آزمون خی دو تنها به موقعیتهایی که دارای دو طبقه‌بندی باشند، محدود نمی‌شود، بلکه همچنین می‌توان آن را در آزمون فرضیه صفر، مبنی بر عدم تفاوت معنی‌دار بین آزمودنیهایی که در طبقه‌بندیهای مختلفی قرار دارند، به کار برد. برای مثال، فرض کنید از یک نمونه ۱۲۰ نفری دانشجویان دوره لیسانس سؤال کرده‌ایم آیا ترجیح می‌دهند در خوابگاه زندگی کنند، یا در خانه شخصی، و یا برای آنها فرقی نمی‌کند. نتایج این پرسش در جدول ۸-۶ آمده است. اگر هیچ تفاوتی بین سه طبقه از پاسخها وجود نداشته باشد، در هر طبقه ۴۰ پاسخ خواهیم داشت. این مقادیر، فراوانیهای مورد انتظار ما خواهند بود، همان‌گونه که در جدول ۹-۶ نشان داده شده است.

مقایسه‌ای بین دو مجموعه فراوانی ارائه شده در جداول ۸-۶ و ۹-۶ نشان می‌دهد که بین داده‌های مورد انتظار و مشاهده شده تفاوتی وجود دارد. برای تعیین اینکه آیا این تفاوتها معنی‌دار هستند، آزمون خی دو را به کار می‌بریم. با استفاده از فرمول ۱۷-۶ مقدار χ^2 برای این داده‌ها عبارت خواهد بود از:

$$\chi^2 = \frac{(40 - 40)^2}{40} + \frac{(50 - 40)^2}{40} + \frac{(30 - 40)^2}{40} = 5$$

درجه آزادی برابر است با تعداد طبقه‌ها منهای یک ($K - 1$) یادراین مورد، $2 = 3 - 1$. با مراجعه به جدول χ^2 می‌بینیم که با درجه آزادی ۲، برای معنی‌داری در سطح ۰/۰۵ به خی دو ۵/۹۹۱ یا بیشتر نیاز است. به هر ترتیب مقدار χ^2 کمتر از آن مقدار است و بنابراین از نظر آماری معنی‌دار نیست. معنای این امر آن است که تفاوت مشاهده شده بین طبقه‌های ما به آسانی می‌توانست در اثر اتفاق به وقوع بپیوندد. در نتیجه فرضیه صفر مبنی بر عدم تفاوت معنی‌دار بین فراوانیهای سه طبقه‌بندی ما نمی‌تواند رد شود. به سخن دیگر، اگر نسبتهای رجحان برای این سه طبقه‌بندی در کل جامعه آماری دانشجویان دوره لیسانس مساوی باشند، انتظار خواهیم داشت که تفاوت بین نمونه‌ها، به همان بزرگی که در مورد نمونه ما وجود داشت، بیش از ۵ بار در هر ۱۰۰ بار نمونه‌گیری، در نتیجه اتفاق روی بدهد.

جدول ۸ - ۶ فراوانیهای مشاهده شده در پاسخهای ۱۲۰ دانشجوی دربارهٔ رجحان آنها در مورد مکان زندگی

کل	فرقی نمی‌کند	خانه شخصی	خوابگاه	آزمودنی
۱۲۰	۳۰	۵۰	۴۰	دانشجویان دوره لیسانس

جدول ۹ - ۶ فراوانیهای مورد انتظار پاسخهای ۱۲۰ دانشجوی دربارهٔ رجحان آنها در مورد مکان زندگی

کل	فرقی نمی‌کند	خانه شخصی	خوابگاه	آزمودنی
۱۲۰	۴۰	۴۰	۴۰	دانشجویان دوره لیسانس

// آزمون ناوابستگی خی دو

تا اینجا، تنها مثالهایی را در نظر گرفتیم که در آنها مشاهده‌ها در جهت یک بعد منفرد طبقه‌بندی شده‌اند. به هر حال، گاهی می‌خواهیم از بیش از یک بعد برای طبقه‌بندی استفاده کنیم. برای مثال، فرض کنید بعد دیگری به مسئله قبلی اضافه می‌کنیم و هم از دانشجویان دوره لیسانس و هم از دانشجویان دوره فوق لیسانس می‌خواهیم که رجحانهای خود را در مورد مکان زندگی بیان کنند. فرض کنید فراوانیهای جدول ۱۰ - ۶ نتیجه این پرسش است.

در این مورد فرضیه صفر ما این است که رجحان برای مکان زندگی در مورد دانشجویان فوق لیسانس همان است که برای دانشجویان لیسانس بود، یعنی متغیرهای پایگاه اجتماعی دانشجوی با رجحان در مورد مکان زندگی نامرتب یا مستقل هستند. در این آزمون ناوابستگی خی دو، فرضیه صفر همیشه می‌گوید که متغیرها در جامعه آماری مستقل یا ناوابسته هستند.

مشاهدات ما نشان می‌دهد که ۳۰ درصد از کل دانشجویان، خوابگاه و ۴۵ درصد خانه شخصی را ترجیح می‌دهند و ۲۵ درصد نیز بین این دو فرقی نمی‌گذارند. اگر فرضیه صفر درست باشد، ما انتظار خواهیم داشت که بین دانشجویان لیسانس و فوق لیسانس نسبتهای مساوی پیدا

کنیم، همان گونه که در جدول ۱۱ - ۶ نشان داده شده است. ما می‌توانیم فراوانی طبقات مورد انتظار را با ضرب کردن فراوانی ردیف همبسته با طبقه در فراوانی ستون همبسته با همان طبقه و سپس تقسیم حاصلضرب بر کل ($E = f_r \cdot f_c / N$) محاسبه کنیم.

برای مثال، فراوانی مورد انتظار پاسخهای دانشجویان دوره لیسانس که می‌خواهند در خوابگاه زندگی کنند، عبارت است از $36 = 120 \times 60 / 200$ ، برای آنها که ترجیح می‌دهند در خانه شخصی زندگی کنند، عبارت است از $54 = 120 \times 90 / 200$ و برای دانشجویان دوره فوق لیسانس که می‌خواهند در خوابگاه زندگی کنند، عبارت است از $24 = 80 \times 60 / 200$. با استفاده از این روش، فراوانی مورد انتظار برای هر طبقه را پیدا می‌کنیم.

جدول ۱۰ - ۶ فراوانیهای مشاهده شده در پاسخهای ۲۰۰ دانشجوی دوره‌های

لیسانس و فوق لیسانس دربارهٔ رجحان آنها در مورد مکان زندگی

کل	فرقی نمی‌کند	خانه شخصی	خوابگاه	آزمودنی
۱۲۰	۳۰	۵۰	۴۰	دانشجویان دوره لیسانس
۸۰	۲۰	۴۰	۲۰	دانشجویان دوره فوق لیسانس
۲۰۰	۵۰	۹۰	۶۰	کل

جدول ۱۱ - ۶ فراوانیهای مورد انتظار پاسخهای ۲۰۰ دانشجوی دوره‌های

لیسانس و فوق لیسانس دربارهٔ رجحان آنها در مورد مکان زندگی

کل	فرقی نمی‌کند	خانه شخصی	خوابگاه	آزمودنی
۱۲۰	۳۰	۵۴	۳۶	دانشجویان دوره لیسانس
۸۰	۲۰	۳۶	۲۴	دانشجویان دوره فوق لیسانس
۲۰۰	۵۰	۹۰	۶۰	کل

توجه کنید که در جدول ۱۱-۶، اعداد کل ردیف و ستون، دقیقاً همان اعدادی هستند که در جدول ۱۰-۶ نیز آمده‌اند. اکنون ما در پی یافتن این هستیم که آیا فراوانیهای مشاهده شده به حد کفایت، با فراوانیهای مورد انتظار تفاوت دارند که ما را به رد احتمال بروز اتفاقی چنین تفاوتی قادر سازند. با کاربرد فرمولی که به دست آورده‌ایم، این امر را آزمون می‌کنیم.

$$\chi^2 = \frac{(40 - 36)^2}{36} + \frac{(50 - 54)^2}{54} + \frac{(30 - 30)^2}{30} + \frac{(20 - 24)^2}{24} + \frac{(40 - 36)^2}{36} + \frac{(20 - 20)^2}{20}$$

$$\chi^2 = 1/8518$$

برای یافتن عدد درجه آزادی در مورد یک جدول دو طرفه از فرمول زیر استفاده می‌شود:

$$df = (C - 1)(R - 1) \quad (18 - 6)$$

عدد درجه آزادی

C = تعداد ستونها

R = تعداد ردیف‌ها

با کاربرد این فرمول عدد درجه آزادی برای مسئله مورد نظر ما، عبارت خواهد بود از:

$$df = (3 - 1)(2 - 1) = 2$$

با مراجعه به جدول ۴ - A می‌بینیم که مقدار χ^2 لازم برای معنی داری در سطح ۰/۰۵ با دو درجه آزادی ۵/۹۹۱ است، اما مقدار χ^2 به دست آمده ما، یعنی ۱/۸۵۱۸، کمتر از مقدار جدول است و بنابراین معنی دار نیست. این امر به معنای آن است که تفاوت بین فراوانیهای مورد انتظار و مشاهده شده فراتر از آن مقداری نیست که به طور اتفاقی انتظار می‌رود. به سخن دیگر، ما در جامعه آماری که از آن نمونه‌گیری کردیم، شواهد معتبری در مورد رابطه بین متغیر «پایگاه اجتماعی دانشجوی» و متغیر «رجحان مکان زندگی» در دست نداریم. فرضیه صفر را نمی‌توان رد کرد.

/ مفروضه‌های خی دو

استفاده از خی دو آن قدر آسان است که شاید فراموش کنیم برای تفسیر معتبر آن، باید

مفروضه‌هایی را مدنظر قرار دهیم:

- ۱- مشاهده‌ها باید مستقل باشند، یعنی در هر نمونه، آزمودنیها باید به طور تصادفی و مستقلاً انتخاب شوند.
- ۲- طبقه‌بندیها باید نامتداخل باشند، هر مشاهده‌ای می‌تواند در یکی و تنها در یکی از طبقه‌بندیهای جدول ما قرار بگیرد.
- ۳- مشاهده‌ها برحسب فراوانی اندازه‌گیری می‌شوند.

/// خلاصه

پژوهشگران آرزو دارند که با مطالعه گروههایی از افراد، که از یک جامعه آماری انتخاب شده‌اند، به تعمیم‌هایی در مورد آن جامعه آماری دست پیدا کنند. این تعمیم‌ها تنها هنگامی بی‌نقص خواهند بود که گروههای انتخابی مورد استفاده در این مطالعات (نمونه‌ها) معرف گروههای بزرگتری که این گروههای انتخابی از میان آنها برگزیده شده‌اند (جامعه آماری)، باشند. مابین دو نوع اصلی نمونه‌گیری تمایز قابل شده‌اند: نمونه‌گیری احتمالاتی و نمونه‌گیری غیر احتمالاتی. مشخصه نمونه‌گیری احتمالاتی انتخاب تصادفی اعضای جامعه آماری است. در نمونه‌گیری غیراحتمالاتی، قضاوت محقق جای انتخاب تصادفی را می‌گیرد. نمونه‌گیری تصادفی ساده، نمونه‌گیری طبقه‌ای، نمونه‌گیری خوشه‌ای و نمونه‌گیری سیستماتیک (نظام‌مند) اشکال نمونه‌گیری احتمالاتی هستند. در نمونه‌گیری تصادفی ساده، همه اعضای جامعه آماری برای قرار گرفتن در نمونه، بخت مساوی دارند. در نمونه‌گیری طبقه‌ای، نمونه‌های مستقل، از زیرگروه‌ها یا طبقات جامعه آماری انتخاب می‌شوند. در نمونه‌گیری خوشه‌ای، گروههایی که به طور طبیعی در جامعه آماری تشکیل شده‌اند، انتخاب می‌شوند. سپس تمام افرادی که در خوشه‌های انتخابی قرار دارند، به عنوان نمونه مورد استفاده قرار می‌گیرند. سرانجام، در نمونه‌گیری سیستماتیک، موردهای K ام از فهرست جامعه آماری، به عنوان اعضای نمونه انتخاب می‌شوند. اشکال نمونه‌گیری غیر احتمالاتی عبارتند از نمونه‌گیری اتفاقی، نمونه‌گیری هدفمند و نمونه‌گیری سهمیه‌ای. در نمونه‌گیری اتفاقی، موردهای قابل دسترسی به عنوان نمونه مورد استفاده قرار می‌گیرند. در نمونه‌گیری هدفمند، موردهای نوعی جامعه آماری مورد نظر، بنابر قضاوت محقق، نمونه را تشکیل می‌دهند. سرانجام، در نمونه‌گیری سهمیه‌ای، به طبقات

مختلف جامعه آماری سهمیه تخصیص می‌یابد، سپس در درون هر طبقه موردهایی که معرف آن طبقه تشخیص داده شوند، به عنوان نمونه انتخاب می‌شوند.

آمار استنباطی ابزاری به دست پژوهشگران می‌دهد تا به وسیله آن بتوانند میزان اطمینان را در پاسخگویی به این پرسش که «اگر پدیده‌ای در نمونه مشاهده شد، آیا در جامعه آماری مربوط به آن نمونه نیز مشاهده خواهد شد»، برآورد کنند. به سخن دیگر، آمار استنباطی ما را قادر می‌سازد تا مقدار اعتبار مشاهدات خود را برآورد کنیم.

یک راهبرد اصلی در آمار استنباطی، محاسبه مقدار تفاوت مابین مشاهداتی است که احتمالاً تنها به طور اتفاقی پدید خواهند آمد. نتیجه این محاسبه را معمولاً واژه خطا می‌نامند. تفاوت موجود بین مشاهده‌ها با واژه خطا مقایسه می‌شود. اگر تفاوت‌های موجود مشابه تفاوت‌هایی باشند که می‌توانست به طور اتفاقی پدید آید، محقق نخواهد توانست این احتمال را که تفاوت‌های موجود صرفاً تابعی از اتفاق هستند، رد کند. اگر تفاوت‌های موجود بزرگتر از واژه خطا باشند؛ محقق برای آن که تعیین کند که نسبت مشاهده به خطا، آن قدر بزرگ هست که وقوع اتفاقی پدیده را در سطح معنی‌داری معین رد کند، به جداول آماری مراجعه می‌کند.

شاخصهایی که بیشترین موارد استفاده را در آمار استنباطی دارند، عبارتند از: آزمون t ، تحلیل واریانس، و آزمون χ^2 . برای تعیین این که آیا تفاوت بین دو میانگین نمونه از نظر آماری معنی‌دار است یا نه، استفاده می‌شود. دو نوع t وجود دارد: (۱) آزمون t برای گروه‌های مستقل، هنگامی مورد استفاده قرار می‌گیرد که هر دو نمونه به طور مستقل از یک جامعه آماری انتخاب شده باشند، و (۲) آزمون t برای گروه‌های نامستقل، هنگامی به کار می‌رود که آزمودنی‌های دو نمونه با هم زوج شده باشند یا آزمودنی‌های یک نمونه دو بار اندازه‌گیری شده باشند. تحلیل واریانس برای مقایسه میانگین‌های دو یا چند نمونه و برای آزمون فرضیه صفر که می‌گوید تفاوت معنی‌داری بین میانگین‌های حاصل از این نمونه‌ها موجود نیست، مورد استفاده قرار می‌گیرد. تحلیل واریانس چند عاملی به ما این توانایی را می‌دهد که اثر بیش از یک متغیر مستقل و اثر تعاملی چنین متغیرهایی را آزمون کنیم.

آماره χ^2 دو شاخصی است که برای یافتن معنی‌داری تفاوت بین نسبت‌هایی از آزمودنی‌ها، اشیاء، رویدادها و غیره که در طبقه‌بندی‌های مختلفی می‌گنجند، از طریق مقایسه فراوانی‌های مشاهده شده و مورد انتظار به کار می‌رود.

// مفاهیم اصلی

آزمون t برای نمونه‌های مستقل	آزمون t برای نمونه‌های مستقل
آزمون دو دامنه	آزمون جهت‌دار (یک دامنه)
آزمون معنی‌داری	آزمون فاقد جهت
آزمون یک دامنه	آزمون F
آماره	آزمون z
پارامتر	اثر اصلی
تحلیل واریانس چند عاملی	تحلیل واریانس (ANOVA)
جامعه آماری	تعامل
جامعه آماری هدف	جامعه آماری قابل دسترسی
خطای معیار میانگین	جدول اعداد تصادفی
خطای نوع I و II	خطای نمونه‌گیری
درجه آزادی	خی دو
طرح عاملی	سطح معنی‌داری
فراوانی مورد انتظار	فراوانی مشاهده شده
نمونه اریب	فرضیه صفر
نمونه‌گیری احتمالاتی	نمونه‌گیری اتفاقی
نمونه‌گیری سهمیه‌ای	نمونه‌گیری خوشه‌ای
نمونه‌گیری طبقه‌ای	نمونه‌گیری سیستماتیک
نمونه‌گیری هدفمند	نمونه‌گیری غیراحتمالاتی

// تمرینها

- ۱- آیا دقت یک نمونه در زمینه نمایاندن ویژگیهای جامعه آماری مربوط به خود، همیشه با اندازه نمونه افزایش می‌یابد؟ شرح دهید.
- ۲- از شما خواسته شده است که تعیین کنید آیا آموزگاران ناحیه آموزشی مرکزی، از مفهوم «تحصیل در تمام طول سال» پشتیبانی می‌کنند یا خیر. از آنجا که این بخش نسبتاً بزرگ است، از

شما درخواست می‌شود که تنها با ۵۰۰ آموزگار تماس بگیرید. تعدادی را که شما از هر یک از سطوح زیر برای گرفتن یک نمونه تصادفی طبقه‌ای متناسب انتخاب خواهید کرد، تعیین کنید.

تعداد کل	سطح
۳۵۰۰	دوره دبستان
۲۱۰۰	دوره راهنمایی
۱۴۰۰	دوره دبیرستان
۷۰۰۰	کل

۳- از شما خواسته می‌شود یک نظرسنجی در دانشگاهی با جامعه آماری ۱۵۰۰۰ دانشجو انجام دهید. شما برای گرفتن یک نمونه معرف از این دانشجویان برای انجام پیمایش خود چگونه عمل خواهیم کرد؟

۴- هیأت سردبیری یک مجله که در سراسر کشور منتشر می‌شود و یک میلیون نفر مشترک دارد، می‌خواهد بداند کدام یک از جنبه‌های مجله مورد علاقه خوانندگان است و کدام یک نیست؟ هیأت سردبیری به این نتیجه می‌رسد که مصاحبه فردی بهترین روش برای کسب اطلاعات است. به دلایل اجرایی و اقتصادی تنها با ۵۰۰ نفر در پنج شهر مصاحبه خواهد شد. در چنین وضعیتی، مشخص کنید:

الف - جامعه آماری هدف

ب - جامعه آماری قابل دسترسی

ج - نمونه

۵- کدام یک از نمونه‌گیرهای زیر، نمونه‌گیری احتمالاتی است؟ کدام یک نمونه‌گیری غیراحتمالاتی است؟

الف - نمونه تصادفی

ب - نمونه اتفاقی

ج - نمونه خوشه‌ای

د - نمونه طبقه‌ای

ه - نمونه هدفمند

و - نمونه سهمیه‌ای

ز - نمونه سیستماتیک

۶- پژوهشگران می‌خواهند درباره این مسئله که «آیا مو طلایی‌ها از زندگی خود لذت بیشتری می‌برند؟» مطالعه کنند.

الف - در این مسئله فرضیه صفر چیست؟

ب - در این مورد خطای نوع I چه خواهد بود؟

ج - در این مورد خطای نوع II چه خواهد بود؟

د - اگر در بررسی این مسئله یک محقق از سطح معنی‌داری 0.05 و محقق دیگر از سطح معنی‌داری 0.01 استفاده کند، کدام یک بیشتر محتمل است مرتکب خطای نوع اول شود؟

ه - اگر در بررسی این مسئله یک محقق از سطح معنی‌داری 0.05 و محقق دیگر از سطح معنی‌داری 0.01 استفاده کند، کدام یک بیشتر محتمل است مرتکب خطای نوع دوم شود؟

۷- آمار استنباطی محقق را قادر می‌سازد:

الف - به نتایج بدون اشتباه برسد.

ب - با اطلاعات ناکامل به نتایج منطقی برسد.

ج - به آنچه که صرفاً حدس و گمان است، ظاهر واقعیت بدهد.

۸- برای آن که یک خطای نوع I واقع شود، کدام دو شرط ضرورت دارد؟

۹- کدام یک از عبارات زیر نقش فرضیه صفر را در تحقیق توصیف می‌کند؟

الف - فرضیه صفر ما را قادر می‌سازد هنگامی که هیچ رابطه واقعی بین متغیرها موجود نیست، احتمال وقوع یک رویداد را به طور اتفاقی، تعیین کنیم.

ب - فرضیه صفر ما را قادر می‌سازد وجود یک رابطه واقعی بین متغیرها را ثابت کنیم.

ج - فرضیه صفر ما را قادر می‌سازد عدم یک رابطه واقعی بین متغیرها را ثابت کنیم.

۱۰- یک خطای نوع II هنگامی اتفاق می‌افتد که محقق:

الف - یک فرضیه صفر نادرست را رد می‌کند.

ب - یک فرضیه صفر درست را رد می‌کند.

ج - قبلاً یک خطای نوع I را مرتکب شده است.

د - یک فرضیه صفر نادرست را حفظ می‌کند.

ه - یک فرضیه صفر درست را حفظ می‌کند.

۱۱ - اصطلاح سطح معنی‌داری به کدام یک از عبارات زیر اطلاق می‌شود:

الف - احتمال وقوع یک رویداد تنها بر حسب اتفاق، که بعد از تحلیل داده‌های یک آزمایش محاسبه می‌شود.

ب - احتمال یک خطای نوع I که محقق می‌خواهد آن را بپذیرد.

ج - احتمال واقعی یک خطای نوع II

د - احتمال یک خطای نوع II که محقق می‌خواهد آن را بپذیرد.

۱۲ - یک محقق سطح معنی‌داری را برای استفاده در یک آزمایش چگونه تعیین می‌کند؟

۱۳ - یک صاحب کارخانه دخانیات، محقق را استخدام کرده است تا نرخ بروز سرطان ریه را در بین سیگاریها و غیر سیگاریها مقایسه کند. با در نظر گرفتن نتایج تحقیقات قبلی درباره این مسئله، صاحب کارخانه محقق را ترغیب خواهد کرد تا از کدام یک از خطاهای زیر اجتناب کند:

الف - خطای نوع I

ب - خطای نوع II

۱۴ - وقتی که نتایج یک مطالعه از نظر آماری معنی‌دار نباشد، این امر بر چه چیزی دلالت دارد؟

۱۵ - نمونه‌گیری طبقه‌ای را با نمونه‌گیری سهمیه‌ای مقایسه کنید.

۱۶ - مقدار χ^2 دو را برای مجموعه داده‌های زیر پیدا کنید. اگر مقدار χ^2 دو در سطح 0.05 معنی‌دار است، آن را نشان دهید.

۲	۸	۱۰
۶	۴	۱۰
۸	۱۲	

۱۷ - ما پاسخهای دو گروه از دانشجویان (رشته تعلیم و تربیت و رشته‌های دیگر) را به یک پرسش مقیاس نگرش نوع لیکرت داریم. مقدار χ^2 دو را برای این مسئله پیدا کنید و نشان دهید آیا مقدار χ^2 دو حاصله از نظر آماری معنی‌دار است یا نه.

کاملاً موافق	موافق	نامشخص	مخالف	کاملاً مخالف	
۷	۸	۱۰	۹	۶	دانشجویان رشته تعلیم و تربیت
۸	۱۰	۷	۸	۷	دانشجویان رشته های دیگر

۱۸- کدام یک از شیوه های آماری زیر برای آزمون فرضیه های الف تا ز مناسب هستند؟

آزمون t برای میانگین های مستقل

آزمون t برای میانگین های نامستقل

تحلیل واریانس عاملی

خی دو

الف - نسبت دانشجویان دوره دکترا در کلاس، که نمره الف می گیرند، از نسبت دانشجویان دوره فوق لیسانس که نمره الف می گیرند، بزرگتر خواهد بود.

ب - نمره میانگین آن گروه از دانشجویان که در ساعت ۱ بعد از ظهر امتحان نهایی می دهند، از نمره میانگین دانشجویانی که در ساعت ۷ بعد از ظهر امتحان می دهند، به طور معنی داری بیشتر خواهد بود.

ج - برای دانشجویان دختر، شرکت در کلاس عصر، مزیتی در جهت کسب نمره بهتر است، در حالی که برای دانشجویان پسر مانعی برای کسب نمره بهتر محسوب می شود.

د - اگر همان امتحان آماری را که در اولین روز کلاس از دانشجویان گرفته اند، در روز آخر کلاس نیز تکرار کنند، دانشجویان پیشرفت تحصیلی معنی داری را نشان خواهند داد.

ه - بین پذیرفته شدن یا رد شدن در کلاس و اظهار رضایت یا عدم رضایت از دوره درسی رابطه وجود دارد.

و - هر تفاوتی بین میانگین نمره های کلاس ترم پاییز و کلاس ترم تابستانی یک استاد را می توان بسادگی برحسب اتفاق دانست.

ز - بین تجربه تدریس و تجربه تحقیق و موفقیت در یک دوره درسی تحقیق، اثر تعاملی معنی داری وجود دارد.

۱۹- تفاوت یک آزمون جهت دار با یک آزمون فاقد جهت در این است که در آزمون جهت دار

محقق به تغییراتی علاقه مند است که:

الف - تنها در جهت مثبت صورت می‌گیرد.

ب - تنها در جهت منفی صورت می‌گیرد.

ج - در هر دو جهت مثبت و منفی صورت می‌گیرد.

د - یا در جهت مثبت یا در جهت منفی صورت می‌گیرد.

۲۰- دو گروه که به طور تصادفی انتخاب شده‌اند، در یک آزمایش شرکت کردند که در آن آزمایش، گروه I تیمار را دریافت کرده و گروه II تیمار را دریافت نکرده است. فرضیه محقق این است که عملکرد متوسط گروه I بالاتر از عملکرد متوسط گروه II خواهد بود. آزمون t را در مورد اطلاعات زیر به کار ببرید و بگویید که آیا فرضیه محقق می‌تواند تایید شود یا خیر.

$\sum x^2$	N	X	
۳۸۲/۰۲	۳۰	۴۵/۳۲	گروه I
۲۶۴/۳۲	۳۰	۴۱/۷۸	گروه II

۲۱- محقق می‌خواهد این فرضیه را که همبستگی بین متغیر A و متغیر B به طور معنی‌داری بزرگتر از صفر است، آزمون کند. او با استفاده از ۲۲ آزمودنی $r = ۰/۲۱$ را به دست آورده است. برای آن که دریابید آیا فرضیه مزبور را می‌توان در سطح $۰/۰۵$ رد کرد (یک دامنه)، از جدول ۵-A استفاده کنید.

۲۲- جدول زیر داده‌های یک طرح آزمایشی ۳×۲ است. برای آزمون معنی‌داری اثرهای اصلی و اثر تعاملی، تحلیل واریانس چند عاملی را به کار ببرید.

ستونها

	A	B	C
ردیف‌ها	۲۵ و ۲۳ و ۲۰ و ۱۷ و ۱۵	۲۲ و ۲۰ و ۱۸ و ۱۶ و ۱۴	۲۰ و ۱۸ و ۱۶ و ۱۴ و ۱۲
	۱۶ و ۱۴ و ۱۲ و ۱۰ و ۸	۱۸ و ۱۶ و ۱۴ و ۱۲ و ۱۰	۱۹ و ۱۸ و ۱۶ و ۱۴ و ۱۳

جدولی تهیه کنید که در آن مجموع مربعات (SS)، درجه آزادی (df)، میانگین مربعات (MS) و مقادیر F نشان داده شده است. سپس به پرسشهای زیر پاسخ دهید:

الف - کدام مقادیر F و در چه سطحی معنی دار هستند؟

ب - در این مسئله چند فرضیه صفر آزمون شدند؟

ج - ما چند تا از این فرضیه‌ها را می‌توانیم رد کنیم؟

۲۳- فهرستی از دانش‌آموزان یک دبیرستان دارید که اعداد ۱ تا ۱۰۰۰ به آنها تخصیص داده شده است. برای انتخاب یک نمونه ۵۰ نفری از این فهرست فرضی از جدول اعداد تصادفی ضمیمه کتاب استفاده کنید. اعدادی را که برای نمونه انتخاب کرده‌اید، ذکر کنید.

/// پاسخها

۱- یک نمونه تصادفی بزرگتر، بیشتر از یک نمونه تصادفی کوچکتر محتمل است که معرف جامعه آماری باشد. یک نمونه بزرگ که روش انتخاب آن اجازه آریبی نظام‌مند را بدهد، بیشتر از یک نمونه آریب کوچک، معرف جامعه آماری نخواهد بود.

۲- برای به دست آوردن یک نمونه طبقه‌ای متناسب، ۵۰۰ آموزگار را به نسبت بازنمایی آنها در جامعه آماری تقسیم کنید، به صورت زیر عمل کنید:

$$\begin{array}{rcl} 3500 / 7000 \times 500 = 250 & \text{دوره دبستان} \\ 2100 / 7000 \times 500 = 150 & \text{دوره راهنمایی} \\ 1400 / 7000 \times 500 = 100 & \text{دوره دبیرستان} \\ \hline & \text{کل نمونه} \\ & 500 \end{array}$$

۳- فهرستی از تمام دانشجویان تهیه کنید و به اسامی آنها شماره بدهید. سپس با استفاده از جدول اعداد تصادفی، یک نمونه تصادفی را به اندازه مورد نظر انتخاب کنید. یک نقطه جدول را به طور تصادفی انتخاب کنید و سپس روی ستون اعداد به بالا یا پایین بروید و دانشجویانی را که شماره آنها در ستون اعداد تصادفی وجود دارد، انتخاب کنید.

۴- الف - تمام مشترکین مجله

ب - مشترکین در پنج شهر

ج - ۵۰۰ نفر که مصاحبه می‌شوند.

۵- نمونه‌های احتمالاتی: الف، ج، د، ز

نمونه‌های غیر احتمالاتی: ب، ه، و

۶- الف - هیچ رابطه‌ای بین رنگ مو و لذت بردن از زندگی وجود ندارد.

ب - محققان دچار خطای نوع I می‌شوند، اگر اعلام کنند موطلایی‌ها بیشتر از غیر موطلایی‌ها از زندگی لذت می‌برند یا موطلایی‌ها کمتر از غیر موطلایی‌ها از زندگی لذت می‌برند، در حالی که در واقع هر دو گروه به یک اندازه از زندگی لذت می‌برند.

ج - محققان دچار خطای نوع II می‌شوند، اگر نتوانند نتیجه بگیرند که موطلایی‌ها لذت بیشتر یا کمتر می‌برند، در حالی که واقعیت چنین باشد.

د - محققى که از سطح معنی‌داری ۰/۰۵ استفاده می‌کند.

ه - محققى که از سطح معنی‌داری ۰/۰۰۱ استفاده می‌کند.

۷- ب

۸- فرضیه صفر باید درست باشد و محقق باید آن را رد کند.

۹- الف

۱۰- د

۱۱- ب

۱۲- با سبک و سنگین کردن پیامدهای خطای نوع I و خطای نوع II

۱۳- الف

۱۴- نتایج بسادگی می‌توانند تابعی از اتفاق باشند، شواهد برای توجیه یک نتیجه‌گیری ناکافی است.

۱۵- در نمونه‌گیری طبقه‌ای، معرف بودن در هر طبقه، با استفاده از انتخاب تصادفی به دست می‌آید، در حالی که در نمونه‌گیری سهمیه‌ای معرف بودن در طبقات مختلف از راه قضاوت حاصل می‌شود.

۱۶- ۳/۳۳ معنی دار نیست.

۱۷- ۰/۹۵۴ معنی دار نیست.

۱۸- الف - خى دو

ب - آزمون t برای میانگینهای مستقل

ج - تحلیل واریانس عاملی

د - آزمون t برای میانگینهای نامستقل

ه - χ^2 دو

و - آزمون t برای میانگینهای مستقل

ز - تحلیل واریانس عاملی

۱۹- د

۲۰- $df = 58$ و $t = 4/11$ ، معنی‌دار در سطح $0/001$

۲۱- با درجه آزادی ۲۲، به یک t مساوی با $0/4227$ یا بزرگتر از آن نیاز داریم تا فرضیه صفر را رد کنیم. بنابراین، فرضیه صفر رد نمی‌شود.

۲۲-

منبع واریانس	SS	df	MS	F
بین ستونها	۰	۲	۰	۰
بین ردیفها	۱۲۰	۱	۱۲۰	۱۱/۳۴
تعامل	۸۰	۲	۴۰	۳/۷۸
بین گروهها	۲۰۰	۵	۴۰	۳/۷۸
داخل گروهها	۲۵۴	۲۴	۱۰/۵۸	-

الف - بین ردیفها در سطح $0/01$ و تعامل در سطح $0/05$

ب - ۳

ج - ۲

بخش چهارم

مبانی اندازه گیری

فصل هفتم

ابزار تحقیق

هدفهای آموزشی

دانشجو بعد از مطالعه این فصل، قادر خواهد بود:

- ۱- نقش اندازه گیری را در تحقیق توضیح دهد.
- ۲- از سالنامه اندازه گیریهای روانی بوروس برای دستیابی به داده های ضروری در ارزشیابی آزمونهای استاندارد شده و دیگر وسائل اندازه گیری استفاده کند.
- ۳- تفاوت بین آزمون و مقیاس را شرح دهد.
- ۴- تفاوت بین آزمونهای هنجاری^(۱) و آزمونهای ملاکی^(۲) را تشخیص دهد.
- ۵- تفاوت بین اندازه های استعداد و پیشرفت را تشخیص دهد.
- ۶- مراحل آماده سازی مقیاس لیکرت^(۳) را برای اندازه گیری نگرشها توصیف کند.
- ۷- مراحل آماده سازی مقیاس تورستون^(۴) را توصیف کند.
- ۸- ویژگیهای مقیاس لیکرت و تورستون را مقایسه کند.

۹- مقیاس گاتمن^(۱) و مزیت‌های آن را توصیف کند.

۱۰- ویژگی‌های مقیاس افتراق معنایی را توصیف کند.

۱۱- انواع خطاهای رایج در مقیاس‌های درجه‌بندی را بیان کند.

۱۲- مزایا و نقایص اندازه‌های خودسنجی شخصیت را بیان کند.

۱۳- حداقل پنج رهنمود را که یک محقق باید به هنگام استفاده از مشاهده مستقیم، به عنوان

یک روش گردآوری داده‌ها، از آنها پیروی کند، فهرست کند.

۱۴- آزمون موقعیتی را تعریف کند و بگوید در تحقیق چه موقع می‌توان از آن استفاده کرد.

۱۵- ویژگی‌های اصلی یک روش فرافکن را بیان کند و حداقل دو روش فرافکن شناخته شده را نام ببرد.

یکی از هدفهای تحقیق در زمینه تعلیم و تربیت، کسب آگاهی بیشتر از روابط بین متغیرها، در جامعه‌های آماری مختلف است. برای مثال، ممکن است کسی بپرسد: «در بچه‌های شش ساله، چه رابطه‌ای بین هوش و خلاقیت وجود دارد؟» ما به طور مستقیم نه می‌توانیم هوش و خلاقیت را مشاهده کنیم و نه می‌توانیم همه کودکان شش ساله را مشاهده کنیم. اما این به آن معنا نیست که باید درباره این مسئله و مسائل شبیه آن در جهل باقی بمانیم. ما نشانگرهایی در دست داریم که ساختار هوش و خلاقیت را به طور تقریب برآورد می‌کنند، یعنی رفتارهای قابل مشاهده‌ای وجود دارند که به عنوان شاخصهای معتبر این ساختارها پذیرفته شده‌اند. یکی از مهم‌ترین وظایف محقق انتخاب یا تهیه مقیاسها و وسایلی است که می‌توانند ویژگیهایی مانند هوش، پیشرفت تحصیلی، شخصیت، انگیزش، نگرشها، استعدادها، علایق، خود پنداره^(۱) و غیره را اندازه بگیرند. برای کمی کردن کیفیت، انواع مختلفی از وسایل مورد استفاده قرار می‌گیرند. برای به دست آوردن این اندازه‌ها دو روش اصلی وجود دارد: (۱) محقق از وسیله‌ای که قبلاً ساخته شده است، استفاده می‌کند و (۲) محقق، خود، وسیله اندازه‌گیری لازم را می‌سازد.

محقق برای انتخاب یک وسیله اندازه‌گیری، باید تحقیقاتی را که در زمینه مورد نظر او منتشر شده‌اند، مطالعه کند. او باید ببیند که محققان دیگر برای اندازه‌گیری سازه مورد نظر او از چه استفاده کرده‌اند. این محققان به طور کلی خاطر نشان خواهند کرد که وسیله اندازه‌گیری آنان خوب عمل کرده است، یا بهتر است از شیوه‌های دیگری استفاده شود. سالنامه اندازه‌گیریهای روانی بوروس و مجموعه آزمونهای منتشره شماره سه، که در فصل سوم توصیف شدند، منابع سودمند دیگری برای شناسایی وسایل اندازه‌گیری منتشر شده می‌باشند. منبع اطلاعاتی مفید دیگر در مورد آزمونها، مجموعه آزمونهای موسسه آزمونهای آموزشی (ETS)^(۲) است. مجموعه آزمونهای ETS، مجموعه‌ای است مشتمل بر بیش از ۱۶۰۰۰ آزمون تجارتي و دیگر وسایل اندازه‌گیری که به منظور فراهم آوردن اطلاعات روز در زمینه آزمون سازی برای محققان آموزشی طراحی شده است. ETS بر خلاف بوروس، هم آزمونهای منتشر شده و هم آزمونهای منتشر نشده را فهرست بندی می‌کند. بخش جدیدی که اخیراً به این مجموعه اضافه شده است،

فهرست تجمعی آزمون‌ها به صورت میکروفیش، ۸۷-۱۹۷۵، نام دارد، که آزمون‌های محققان را برحسب مؤلف، عنوان و موضوع، طبقه‌بندی می‌کند. در فهرست موضوعی بیش از ۷۵۰ عنوان وجود دارد که هر چیزی را، از استدلال انتزاعی تا نگرش‌های مربوط به شغل، شامل می‌شود. اگر محقق نتواند وسیله اندازه‌گیری را که قبلاً ساخته شده است، پیدا کند، آن‌گاه باید خود آن را بسازد. ساختن وسیله اندازه‌گیری مستلزم شناسایی و استفاده از رفتاری است که می‌تواند نشانگر حضور سازه مورد نظر محسوب شود. محقق برای یافتن این نشانگرها، ابتدا باید به نظریه‌ای که پشتوانه تحقیق است، روی آورد. یک نظریه خوب، به طور کلی گویای آن خواهد بود که سازه مورد نظر، خود را چگونه آشکار خواهد کرد و چه تغییراتی را می‌توان در آن مشاهده کرد؛ به عبارت دیگر راه‌هایی را برای اندازه‌گیری سازه (ها) پیشنهاد خواهد کرد. محقق همچنین می‌تواند از تجربه و مهارت خود برای تصمیم‌گیری در مورد نشانگرهای مناسب سازه استفاده کند. ما در بخش‌های بعدی، برخی از انواع وسایل اندازه‌گیری را که در تحقیقات آموزشی مورد استفاده قرار می‌گیرند، مورد بحث قرار خواهیم داد.

/// وسایل اندازه‌گیری // آزمون‌ها

آزمون‌ها وسایل اندازه‌گیری ارزشمندی برای تحقیق در زمینه تعلیم و تربیت هستند. یک آزمون مجموعه‌ای از محرک‌ها است که برای دریافت پاسخ به فرد ارائه می‌شود تا بتوان بر اساس آن پاسخ‌ها یک نمره عددی تعیین کرد. این نمره مبتنی بر یک نمونه معرّف از رفتار فرد و نشانگری است از این که آن فرد تا چه حد خصلت اندازه‌گیری شده را داراست. سودمندی این نمره‌ها، به عنوان نشانگرهای سازه مورد نظر، تا حد زیادی تابع عینیت، اعتبار و پایایی آزمون‌هاست. عینیت به میزان توافق در میان نمره‌ها اطلاق می‌شود. برخی از آزمون‌ها، نظیر آزمون‌های چند گزینه‌ای و درست - نادرست، آزمون‌های عینی به شمار می‌آیند؛ زیرا نمره‌گذاری آنها از طریق مقایسه پاسخ‌های فرد با کلید نمره‌گذاری صورت می‌گیرد و نمره‌گذار خود در این مورد تصمیم نمی‌گیرد. آزمون‌های انشایی از عینیت کمتری برخوردار هستند؛ زیرا نمره‌گذاری آنها تحت تأثیر قضاوت و نظرات نمره‌گذاران قرار می‌گیرد. اعتبار و پایایی در فصل ۸ مورد بحث قرار خواهند گرفت.

/ آزمونهای پیشرفت

متغیر وابسته، در اکثر تحقیقاتی که با مسئله تأثیر روشهای آموزشی سروکار دارند، پیشرفت است. بنابراین آزمونهای پیشرفت به فراوانی در تحقیقات مربوط به تعلیم و تربیت و همچنین در نظامهای مدرسه‌ای به کار می‌رود. این آزمونها تسلط و تبحر افراد را در حیطه‌های مختلف دانش می‌سنجند.

به طور کلی، آزمونهای پیشرفت به دو طبقه استاندارد شده (میزان شده) و ساخته آموزگار یا ساخته پژوهشگر تقسیم می‌شوند. آزمونهای استاندارد شده آزمونهای منتشرشده‌ای هستند که از یک فعالیت دقیق و ماهرانه حاصل شده‌اند و هدفهای آموزشی گسترده‌ای را پوشش می‌دهند که در بسیاری از نظامهای مدرسه‌ای مشترک هستند. برای این آزمونها، هنجارهای تطبیقی استنتاج، اعتبار و پایایی تعیین و رهنمودهای اجرا و نمره‌گذاری تبیین شده است. رهنمودهای اجرایی آزمون در دستورالعمل‌هایی که ناشران آزمون تهیه کرده‌اند، درج شده است.

سازندگان آزمونها، برای آن که هنجارهایی برای آن آزمونها تعیین کنند، آنها را با یک نمونه معرف و مرتبط اجرا می‌کنند. گروه هنجاری ممکن است به گونه‌ای انتخاب شود که معرف تمام کشور یا ایالت، شهر، ناحیه یا مدرسه باشد.

میانگین نمونه برای یک سطح تحصیلی خاص، هنجار آن سطح تحصیلی برای تمام جامعه آماری می‌شود. مهارت‌های اندازه‌گیری شده لزوماً آن چیزهایی نیستند که باید در هر سطح تحصیلی آموخته شوند، بلکه استفاده از هنجارها به متخصصان تعلیم و تربیت مبنایی برای مقایسه گروه خود با برآورد میانگین همه کودکان آن سطح تحصیلی می‌دهد.

آزمونهای استاندارد شده برای موضوعهای مجزای مدرسه‌ای (دروس مجزای کلاسی) نظیر ریاضیات و شیمی و همچنین به شکل مجموعه آزمونهای جامع، که حیطه‌های مختلف پیشرفت را می‌سنجند، در دسترس هستند. برای مثال، آزمون پیشرفت کالیفرنیا شامل آزمونهایی در حیطه‌های خواندن، زبان و حساب است. آزمونهای زنجیره‌ای پیشرفت آموزشی^(۱) شامل آزمونهایی برای هفت حیطه است.

پژوهشگر در انتخاب یک آزمون پیشرفت باید دقت به خرج دهد تا آزمونی را برگزیند که هم پایا و هم برای اندازه‌گیری آن جنبه از پیشرفت که منظور اوست، مناسب باشد. به هر حال، گاهی

از مواقع او قادر به انتخاب آزمون نخواهد بود، بلکه اجبار خواهد داشت از آزمونی که تشکیلات آموزشی قبلاً انتخاب کرده است، استفاده کند. سالنامه اندازه‌گیریهای روانی، فهرست جامعی را ارائه می‌کند که مروری بر آزمونهای پیشرفت موجود را به همراه دارد.

اگر آزمونی که در دسترس است، رفتار مورد نظر را اندازه‌گیری کند و اگر پایایی و هنجارهای آن برای منظور ما کفایت کند، استفاده از این ابزار استاندارد شده مزایایی دربرخواهد داشت. علاوه بر زمان و تلاشی که صرفه‌جویی می‌شود، پژوهشگر مزیت پیوستگی روش آزمون کردن را در خواهد یافت؛ یعنی می‌توان نتایج بررسی او را با نتایجی که دیگران با استفاده از همان ابزار به دست آورده‌اند، مقایسه و تفسیر کرد.

هنگامی که استفاده از یک آزمون پیشرفت استاندارد شده برای هدفهای خاص یک بررسی تحقیقاتی مناسب به نظر نرسد، امکان دارد پژوهشگر خود آزمون لازم را بسازد. ساختن آزمون مورد نظر بسیار بهتر است تا استفاده از آزمون استاندارد شده نامناسب، تنها به این دلیل که به آن دسترسی داریم. در این حالت، پژوهشگر باید دقت زیادی در آماده‌سازی آزمون به کار بندد؛ بویژه در تعیین اعتبار و پایایی آن قبل از این که آزمون را به کار برد. برای اطلاعات بیشتر در زمینه آزمون سازی می‌توان به متون تخصصی در مورد سنجش، نظیر کتابهای گرون لوند، ساکس، یا تورندایک و هاگن^(۱) مراجعه کرد.

آزمونهای استاندارد شده و معلم ساخته را، بر مبنای نوع تفسیری که از آن به عمل می‌آید، می‌توان همچنین به آزمونهای هنجاری و آزمونهای ملاکی طبقه‌بندی کرد.^(۲) آزمونهای هنجاری امکان مقایسه عملکرد یک فرد را با عملکرد افراد دیگر در آزمون فراهم می‌آورند. عملکرد فرد برحسب جایگاه نسبی او در یک گروه مشخص، که گروه هنجاری نامیده می‌شود، تفسیر و برحسب صدکها، نمره‌های استاندارد و غیره گزارش می‌شود.

از سوی دیگر، آزمونهای ملاکی این امکان را فراهم می‌آورند تا بدون آن که به عملکرد دیگران رجوع کنیم، به طور دقیق توصیف کنیم که یک فرد چه می‌تواند انجام دهد. عملکرد بر

۱- Gronlund, N.E. and R.L. Line (1990). *Measurement and Evaluation in Teaching* (6th ed.). New York: Macmillan; Sax, G. (1980). *Principles of Educational and Psychological Measurement and Evaluation in Psychology and Education* (4th ed.). New York: Wiley.

۲- برای مطالعه بیشتر درباره اندازه‌های ملاکی مراجعه کنید به:

R.A. Berk (1984), *A Guide to Criterion Referenced Test Construction*, Baltimore: Johns Hopkins University Press; and W.J. Popham (1981), *Modern Educational Measurement*, Englewood Cliffs, NJ: Prentice - Hall.

حسب سطح تسلط بر یک حیطه محتوایی / مهارتی مشخص گزارش می‌شود. به طور نوعی، سطح تسلط با درصد پاسخهای درست نشان داده می‌شود. در تفسیر عملکرد فرد از نمره‌های حداقل، که از پیش تعیین شده‌اند، استفاده می‌شود.

محقق، قبل از طراحی وسیله اندازه‌گیری، باید از نوع تفسیری که قرار است به عمل آید، آگاه باشد. در آزمونهای هنجاری، پرسشهایی انتخاب می‌شوند که دامنه وسیعی از نمره‌ها را حاصل می‌کنند. دشواری پرسشها و توان ایجاد افتراق پرسشها در بین افراد، باید مورد توجه محقق باشد. در آزمونهای ملاکی، پرسشها صرفاً بر مبنای نحوه اندازه‌گیری مجموعه مشخصی از هدفهای آموزشی انتخاب می‌شوند. آزمون، بسته به این که چه چیزی را اندازه‌گیری می‌کند، می‌تواند آسان یا دشوار باشد. دلمشغولی اصلی، داشتن یک نمونه معرف از پرسشهایی است که هدفهای بیان شده را اندازه‌گیری کنند؛ به گونه‌ای که امکان توصیف عملکرد را برحسب آگاهی و مهارت مشخصی که افراد می‌توانند کسب کنند، به ما بدهد.

گاهی معدل عددی دروس مختلف به عنوان معیار موفقیت و پیشرفت تحصیلی به کار می‌رود. هنگام استفاده از معدل عددی دانش‌آموزان نظامهای مختلف درسی، احتیاط، شرط لازم است. نمرات الفبایی در مدارس مختلف لزوماً معنای همسان ندارند و نمی‌توان اطمینان داشت که تسلط و تبحر یکسانی را نشان می‌دهند.

/ آزمونهای هوش

آزمونهای هوش از این نظر با آزمونهای پیشرفت تفاوت دارند که در آزمونهای هوش، تلاش در جهت اندازه‌گیری عملکرد کلی است؛ در حالی که در آزمونهای پیشرفت، کوشش برای اندازه‌گیری عملکرد در حیطه‌های خاص صورت می‌گیرد. آزمونهای هوش سعی می‌کنند توانایی آزمودنی را در درک رابطه‌ها، حل مسئله‌ها و بکارگیری آگاهیها در زمینه‌های متنوع بسنجند.

آزمونهای هوش را نباید اندازه‌های هوش فطری یا «خالص» تصور کرد. عملکرد یک آزمودنی در چنین آزمونهایی، تا حدودی به زمینه فردی و سابقه تحصیلی او بستگی دارد. در سالهای اخیر استفاده از واژه هوش برای توصیف این آزمونها، به دلیل اختلاف نظر درباره مفهوم هوش و نیز به این دلیل که مردم تمایل دارند هوش را با تواناییهای ارثی مرتبط سازند، رو به کاستی گذارده است. اکنون، اصطلاح استعداد تحصیلی جایگزین واژه هوش شده است - یک

اصطلاح بیشتر توصیفی که به طور مشخص به کارکرد اصلی این آزمونها، که پیش‌بینی عملکرد تحصیلی است، اشاره دارد.

متخصصان تعلیم و تربیت این آزمونها را برای پیش‌بینی موفقیت تحصیلی مفید و به طور کلی معتبر تشخیص داده‌اند. محققان نیز به طور گسترده از این آزمونها استفاده می‌کنند. هوش یک متغیر مستقل است که معمولاً باید در آزمایشهای آموزشی کنترل شود. محقق برای کنترل این متغیر، از نمره‌های چند آزمون استعداد تحصیلی استفاده می‌کند. تعدادی از آزمونهای موجود برای استفاده انفرادی و بقیه، برای استفاده گروهی طراحی شده‌اند.

آزمونهای انفرادی هوش آزمون استانفورد - بینه و سه آزمون وکسلر رایج‌ترین آزمونهای انفرادی هوش هستند. آزمون استانفورد - بینه، که امروز معمول است، حاصل تجدید نظرهای متعدد در آزمونی است که برای اولین بار آلفرد بینه در فرانسه برای شناسایی آن دسته از کودکانی که احتمالاً نمی‌توانند از آموزشهای کلاسی عادی استفاده کنند، تهیه کرد. این آزمون، که اول بار درست در آستانه ورود به قرن حاضر ساخته شد، در اصل برای اندازه‌گیری سن عقلی فرد به کار می‌رفت. بعدها مفهوم **هوشبهر**^(۱) مطرح شد. این بهر، از تقسیم سن عقلی (MA)^(۱) بر سن تقویمی (CA)^(۳) و ضرب آن در ۱۰۰ حاصل می‌شود. نسخه فعلی استانفورد - بینه سن عقلی را به دست می‌دهد، اما نسبت MA/CA را برای تعیین IQ به کار نمی‌برد. به جای آن، IQ را از طریق مقایسه عملکرد یک فرد (نمره او) با هنجارهای حاصله از همان گروه سنی با استفاده از نمره‌های استاندارد تعیین می‌کند (به فصل ۵ مراجعه کنید). یک ویژگی اساسی آزمونهای استانفورد - بینه آن است که یک اندازه عمومی از هوش به دست می‌دهد. این آزمون برای اندازه‌گیری تواناییهای مجزا، آن گونه که برخی دیگر از آزمونها انجام می‌دهند، تلاشی نمی‌کند. آزمونهایی که دیوید وکسلر برای اندازه‌گیری هوش تهیه کرد، اکنون در اشکال مختلفی در دسترس قرار دارد: مقیاس هوش وکسلر برای کودکان (WISC)^(۴)، که در سال ۱۹۷۴ در آن تجدید نظر شد و اکنون WISC - R نامیده می‌شود، مقیاس هوش وکسلر برای بزرگسالان (WAIS)^(۵)، و مقیاس هوش وکسلر برای کودکان پیش دبستانی و دبستانی (WPPSI)^(۶) که برای کودکان ۴ تا ۶/۵ سال عرضه شد.

1- Intelligence Quotient (IQ)

2- Mental age

3- Chronological age

4- Wechsler Intelligence Scale for Children

5- Wechsler Adult Intelligence Scale

6- Wechsler Preschool and Primary Scale of Intelligence

یک خصلت مهم آزمونهای وکسلر تقسیم آنها به دو خرده آزمون^(۱) است که امکان دستیابی به دو نمره برای هر آزمودنی را فراهم می‌آورد، یک نمره برای هوشبهر کلامی و دیگری برای هوشبهر غیر کلامی. این خرده آزمونها سپس به خرده مقیاسهایی^(۲) تقسیم می‌شوند که عملکرد امتحان شونده را در تکالیف ویژه نشان می‌دهند.

آزمونهای گروهی هوش آزمون استانفورد - بینه یا وکسلر را باید یک متخصص روان‌سنجی، که برای این کار تربیت شده است، با یک آزمودنی منفرد اجرا کند و این روشی است که به زمان طولانی و هزینه زیاد نیازمند است. هنگامی که سنجش هوش برای گروه بزرگی از افراد مورد نظر باشد، این آزمونها غیر عملی هستند. در چنین موقعیتی، آزمونهای گروهی هوش مورد استفاده قرار می‌گیرند. اولین آزمون گروهی توانایی ذهنی^(۳) در دوران جنگ اول جهانی برای اندازه‌گیری هوش مردانی که به خدمت نظام احضار می‌شدند، ساخته شد. یک نمونه از این آزمون (آزمون آلفای نظامی)^(۴) بعد از جنگ برای استفاده غیر نظامی منتشر و سرمشقی برای تعدادی آزمون گروهی شد. امروز، آزمونهای گروهی هوش بسیاری در دسترس است. برای اطلاعات بیشتر درباره آزمونهای هوش به کتاب آناستازی^(۵) مراجعه کنید.

// اندازه‌های شخصیت

دستیابی به اندازه‌های شخصیت، حیطه مورد علاقه دیگری برای محققان مسائل تعلیم و تربیت است. برای اندازه‌های شخصیت انواع مختلفی وجود دارد که هر یک دیدگاه نظریه‌ای متفاوتی را منعکس می‌کنند. برخی از آنها انعکاسی از نظریه و ویژگی و نظریه سنخ هستند، در حالی که برخی دیگر ریشه در نظریه‌های روانکاوی و انگیزشی دارند. پژوهشگر باید دقیقاً بداند که چه چیزی را می‌خواهد اندازه بگیرد و سپس ابزار آن را انتخاب کند، در عین حال به شواهد مربوط به اعتبار آن ابزار توجه ویژه‌ای مبذول دارد. دو نوع اندازه شخصیت که گسترده‌ترین استفاده را در تحقیق دارند، عبارتند از: پرسشنامه‌ها و روشهای فراقکن.

1- Subtest

2- SubScale

3- Mental capacity

4- The Army Alpha

5- Anastasi, A. (1988). Psychological Testing (6th ed.). New York: Macmillan.

پرسشنامه‌ها

در یک پرسشنامه، مجموعه جامعی از عباراتی که الگوهای رفتاری را توصیف می‌کنند، به آزمودنی ارائه و از او خواسته می‌شود که نشان دهد آیا هر عبارت جزو ویژگیهای اوست یا خیر. او می‌تواند «بله»، «خیر»، یا «نامشخص» را علامت بزند.

نمره آزمودنی با شمارش پاسخهای موافق در مورد خصوصیتی که آزماینده در تلاش اندازه‌گیری آن است، محاسبه می‌شود. برای مثال، انتظار خواهد رفت که یک شخص پارانویایی به عبارت «مردم همیشه پشت سر من حرف می‌زنند»، پاسخ «بله» و به عبارت «ماموران پلیس عادل و منطقی باشند» پاسخ «خیر» بدهد. البته، چنین پاسخهایی به تنها دو پرسش، تمایلات پارانویایی رانشان نخواهد داد. به هر ترتیب، پاسخهایی این چنین، به تعداد زیادی از این نوع پرسش‌ها، می‌توانست نشانگر پارانویا تصور شود. برخی از پرسشنامه‌های خودسنجی^(۱) تنها یک ویژگی را اندازه‌می‌گیرند، نظیر مقیاس F کالیفرنیا^(۲)، که قدرت طلبی را می‌سنجد. پرسشنامه‌های دیگر، نظیر پرسشنامه شانزده عامل شخصیتی کتل^(۳) تعدادی از ویژگیها را اندازه می‌گیرند. دیگر پرسشنامه‌های مشهور در امر تحقیق عبارتند از: شخصیت سنج چند وجهی مینه سوتا،^(۴) خوی سنج گیلفورد - زیمرمن،^(۵) مشکل سنج مونی^(۶)، مقیاس ادواردز برای رجحانهای شخصی^(۷)، سنخ نمای مایرز - بریگز^(۸) و پرسشنامه رغبت استرانگ.^(۹)

در تحقیقات آموزشی از پرسشنامه‌ها برای دستیابی به توصیف ویژگیهای گروههای معین، نظیر کم‌آموزها، ترک تحصیل کنندگان، اعضای گروههای اقلیت و غیره استفاده شده است. همچنین این پرسشنامه‌ها برای تحقیق در مسائلی مثل روابط متقابل بین صفات شخصیتی و متغیرهایی مانند هوش، پیشرفت و نگرشها مورد استفاده قرار می‌گیرند. پرسشنامه‌ها مزایای اقتصادی بودن، سادگی و عینیت را دارا هستند. بیشترین عیوب آنها به مسئله اعتبار مربوط می‌شود. اعتبار پرسشنامه‌ها، تا حدی به توانایی پاسخگو در خواندن و فهمیدن پرسشها، درک او از خودش و بویژه تمایل او به ارائه پاسخهای صریح و صادقانه بستگی دارد. به عنوان نتیجه، اطلاعاتی که از پرسشنامه‌ها به دست می‌آید، ممکن است سطحی یا اریب باشند. هنگامی که از نتایج حاصله از این ابزارها استفاده می‌کنیم، باید چنین امکانی را در نظر بگیریم.

1- Self - report inventories

3- Cattell's Sixteen Personality Factor Questionnaire

5- Guilford - Zimmerman Temperament Survey

7- Edwards Personal Preference Schedule

9- Strong Interest Inventory

2- California F- Scale

4- Minnesota Multiphasic Personality Inventory

6- Mooney Problem Check list

8- Myers- Briggs Type Indicator

/ روشهای فرافکن

روشهای فرافکن اندازه‌هایی هستند که در آن از یک فرد خواسته می‌شود به یک محرک مبهم یا بی‌ساخت پاسخ دهد. این روشها از آن رو «فرافکن» خوانده می‌شوند که انتظار می‌رود فرد، نیازها، خواسته‌ها، ترسها، اضطرابها و... خود را بر محرک منعکس کند. آزماینده تلاش می‌کند بر مبنای تفسیر و پاسخهای آزمودنی تصویر جامعی از ساخت شخصیتی فرد بنا کند.

روشهای فرافکن را بیشتر روان‌شناسان بالینی برای مطالعه و تشخیص افرادی که مشکلات هیجانی دارند، به کار می‌برند. این روشها اغلب در تحقیقات آموزشی مورد استفاده قرار نمی‌گیرد؛ زیرا برای اجرا و نمره‌گذاری به آموزش تخصصی نیاز دارد و اجرای فردی آن نیز هزینه زیادی به بار می‌آورد. بعلاوه، بسیاری از پژوهشگران احساس می‌کنند که اعتبار آنها به گونه رضایت بخشی تبیین نشده است. دو روش مشهور فرافکن، آزمون رورشاخ^(۱) و آزمون اندریافت موضوع^(۲) است. در آزمون رورشاخ از لکه‌های جوهر به عنوان محرک استفاده می‌شود و در آزمون اندریافت موضوع به پاسخگو تصاویری به عنوان محرک نشان داده و از او خواسته می‌شود درباره هر یک از آنها داستانی تعریف کند. هر گونه بحث بیشتر درباره این آزمونها و روشهای آنها را می‌توان در کتاب آناستازی^(۳) یافت.

مطلب زیر مثالی است از یک بررسی تحقیقاتی که برای اندازه‌گیری متغیرها از پرسشنامه استفاده کرده است.

رابطه بین مقیاس سنخ‌نمای مایرز - بریگز و ارجحیت شیوه مشاوره دانشجویان سال اول^(۴)

پیش زمینه

تحقیقات قبلی نشان داده است که مشاوره تحصیلی در تحصیلات عالی، یک راهبرد مهم برای تداوم تحصیلی دانشجویان است و بهبود کیفیت مشاوره می‌تواند یادداری دانشجویان را افزایش دهد. همچنین

1- Rorschach

2- Thematic Apperception test (TAT)

۳ - Anastasi، همان، فصل ۱۹.

4- Crackett, J.B., and Crawford, R.L (1989), The relationship between Myers-Briggs Type indicator (MBTI) scale scores and advising style preferences of college Freshmen. Journal of College Student Development, 30:154-161

تحقیقات نشان داده است که نیازها و ارجحیت‌های مشاوره‌ای دانشجویان با توجه به نیمرخهای شخصیتی خاص هر دانشجو تغییر می‌کند. برای تحقیق بیشتر در مورد رابطه بین ویژگیهای دانشجویان و ارجحیت‌های شیوه مشاوره، باید وسایلی برای اندازه‌گیری هر دوی این متغیرها پیدا کرد. مرور پیشینه تحقیق نشان داد که سنخ‌نمای مایرز - بریگز (MBTI) یک وسیله اندازه‌گیری است که بخوبی تبیین شده است و به طور گسترده مورد استفاده قرار می‌گیرد.^(۱) یک وسیله اندازه‌گیری دیگر به نام پرسشنامه مشاوره تحصیلی (AAI) برای سنجش ارجحیت‌های شیوه مشاوره تحصیلی انتخاب شد.^(۲)

سؤال تحقیق

(۱) آیا رابطه‌ای بین نمره‌های مقیاس MBTI و ارجحیت‌های شیوه مشاوره وجود دارد؟ (۲) آیا رابطه‌ای بین سن، نژاد، جنس یا سطح زبان انگلیسی دانشجویان به هنگام ورود به دانشگاه، و ارجحیت‌های شیوه مشاوره وجود دارد؟

نمونه

۲۰۱ نفر از دانشجویان سال اول یک دانشگاه بزرگ در جنوب کشور، که هر دو پرسشنامه را تکمیل کردند، گروه نمونه بررسی را تشکیل می‌دادند. پرسشنامه‌ها در کلاسهای زبان انگلیسی سال اول اجرا شد.

وسایل اندازه‌گیری

MBTI پرسشنامه‌ای است که مبنای آن نظریه ارجحیت ها و

1-Briggs,K.C., and I.B. Myers (1983). Myers-Briggs Type Indicator, From G. Palo Alto: Consulting Psychologists Press.

2-Winston,R,B., and J,A,Sandor (1984) Academic Advising Inventory. Athens, GA: Stundent Development Associates.

سنخ‌های کارل یونگ^(۱) می‌باشد. از افراد خواسته شد ارجحیت خود را بین مبنای احساسات و ارزشهای فردی، و قضاوت-ادراک (JP)^(۲) نشان دهند، که ارجحیت را در مورد نتیجه‌گیری و سازماندهی زندگی (قضاوت) یا آزادی ذهن نسبت به اطلاعات دریافتی و تغییرات اندازه‌گیری می‌کند.

پرسشنامه مشاوره تحصیلی یک چشم انداز مشاوره تحصیلی مبتنی بر نظریه دارد که بر این فرض استوار است که ارجحیت‌های مشاوره بر روی پیوستاری قرار دارد که دامنه آن از کاملاً تجویزی تا کاملاً توسعه‌ای است. مشاوره تجویزی نوع سنتی مشاوره است که در آن مشاور، مسئولیت اصلی تصمیمها را بر عهده می‌گیرد. مشاوره توسعه‌ای گستره بزرگتری دارد و تجربه زندگی دانشجویی را در بر می‌گیرد. در این بررسی ۱۴ عبارت زوج شده از پرسشنامه انتخاب و از دانشجویان خواسته شد ارجحیت خود را در مورد هر یک از اجزای زوجها نشان دهند. نمره‌ها در سه مقیاس به دست آمد: (۱) خصوصی کردن تعلیم و تربیت، که گستره فعالیت‌های مشاوره‌ای را اندازه می‌گیرد، (۲) تصمیم‌گیری تحصیلی، که نحوه تقسیم مسئولیت تصمیم‌گیری را بین مشاور و دانشجو اندازه‌گیری می‌کند و (۳) انتخاب دروس، که چگونگی انتخاب دروس را می‌سنجد.

روش

با استفاده از نمره‌های MBTI و AAI و داده‌های مربوط به متغیرهای جمعیت شناختی (نژاد، سن، جنس و سطح زبان انگلیسی دانشجویان به هنگام ورود به دانشگاه) تحلیل‌های همبستگی و رگرسیون چند متغیری انجام شد.

نتایج

مشاوره توسعه‌ای به طور کامل مورد تأیید همه دانشجویان شرکت کننده در بررسی قرار گرفت. نمره‌های مقیاس های SN و TF مایرز - بریگز بطور مثبت و معنادار با ارجحیت‌های شیوه مشاوره دانشجویان سال اول مرتبط بود. دانشجویانی که بینش شهودی و درک احساسی بیشتری داشتند، ارجحیت بالاتری برای مشاوره توسعه‌ای قایل بودند. سطح زبان انگلیسی دانشجویان به هنگام ورود به دانشگاه تنها متغیر جمعیت‌شناختی معنادار مرتبط با ارجحیت شیوه مشاوره بود.

// مقیاسهای نگرش

یک مقیاس، مجموعه‌ای از ارزشهای عددی است که به منظور کمی کردن و اندازه‌گیری کیفیتها، به موضوعها، اشیاء یا رفتارها نسبت داده می‌شود. مقیاسها برای اندازه‌گیری نگرشها، ارزشها و علایق مورد استفاده قرار می‌گیرند. مقیاسها از این نظر با آزمونها تفاوت دارند که نتایج آنها، بر خلاف نتایج آزمونها، موفقیت و شکست، یا قوت و ضعف را نشان نمی‌دهند. مقیاسها این مسئله را می‌سنجند که یک فرد تا چه حد ویژگی مورد نظر را داراست. برای مثال، ما می‌توانیم از یک مقیاس برای سنجش نگرش دانشجویان نسبت به مذهب یا هر موضوع دیگر استفاده کنیم.

پژوهشگران اغلب خود مقیاسهای خویش را برای اندازه‌گیری نگرشها می‌سازند. ما در این مطلب تلاش خواهیم کرد برخی از شیوه‌های مورد استفاده در مقیاس‌بندی نگرشها را معرفی کنیم.

نگرش را می‌توان چنین تعریف کرد: یک عاطفه مثبت یا منفی نسبت به یک گروه، نهاد، مفهوم یا شیء اجتماعی خاص. اندازه‌گیری نگرشها بر این فرض مبتنی است که می‌توان مکان افراد را بر روی یک پیوستار مطلوبیت - عدم مطلوبیت نسبت به شیئی نگرشی تعیین کرد.

چهار نوع اصلی از مقیاسهای نگرش وجود دارند: (۱) مقیاسهای مجموع درجه‌ها (مقیاسهای لیکرت)، (۲) مقیاسهای فاصله‌های یکسان نما (مقیاسهای ترستون)، (۳) مقیاسهای تراکمی (مقیاسهای گاتمن)، و (۴) مقیاسهای افتراق معنایی.

/ مقیاسهای لیکرت: روش مجموع درجه‌ها

مقیاس لیکرت به عنوان یکی از شیوه‌های سنجش نگرش، بسیار مورد استفاده قرار گرفته و نتایج موفقیت آمیزی نیز به بار آورده است.^(۱) یک مقیاس لیکرت نگرش نسبت به یک موضوع را از این طریق می‌سنجد که از پاسخگویان می‌خواهد که نشان دهند در مورد هر عبارت از یک مجموعه عبارات، که درباره آن موضوع باشد، کاملاً موافق، موافق، مردد، مخالف، یا کاملاً مخالف هستند. یک مقیاس لیکرت با جمع‌آوری تعدادی عبارت درباره یک موضوع، که تقریباً نیمی از آنان به وضوح نگرش مطلوب و نیم دیگر به وضوح نگرش نامطلوب را بیان می‌کنند، ساخته می‌شود. آنچه که در اینجا اهمیت دارد، آن است که این عبارات یک نمونه معرف از تمام نظرات یا نگرشهای ممکن درباره موضوع نگرش را به وجود آورند. روشی که می‌تواند در این راه به محقق کمک کند، این است که ابتدا درباره همه زیر موضوعهای مرتبط با موضوع نگرش تفکر کند و سپس در مورد هر زیر موضوع، عبارتهای لازم را بنویسد. اگر محقق از اشخاصی که درباره یک موضوع خاص آگاهی لازم و نگرشهای مشخصی دارند، درخواست کند که تعدادی عبارت مثبت و منفی درباره آن موضوع بنویسند، شاید تهیه آن مجموعه متنوع عبارات سهل تر شود. همچنین، سر مقاله‌هایی که نشریات مختلف درباره موضوع مورد نظر می‌نویسند، می‌تواند منبع خوبی از عبارات بالقوه، برای یک مقیاس نگرش باشد.

عبارات، همراه با مقولات پاسخ (به طور نوعی پنج مقوله)، بر روی یک پیوستار توافق - عدم توافق، به آزمودنی ارائه می‌شوند. ترتیب عبارات باید به طور تصادفی تنظیم شود تا مانع از آن شود که پاسخهای آزمودنی، به دلیل ترتیب عبارت، همه به یک جهت متمایل شوند. به آزمودنیها گفته می‌شود که پاسخی را انتخاب کنند که به بهترین نحو نمایانگر واکنش آنها نسبت به هر عبارت باشد: کاملاً موافق (SA)^(۲)، موافق (A)^(۳)، مردود (U)^(۴)، مخالف (D)^(۵) و کاملاً مخالف (SD)^(۶).

برای نمره دادن به مقیاس باید مقوله‌های پاسخ ارزش گذاری شوند. به مقوله‌های پاسخ عبارات مطلوب یا مثبت، به ترتیب از سمت مثبت آن، ارزشهای عددی ۵، ۴، ۳، ۲، ۱ داده

1- Likert, R. (1932). A Technique for the Measurement of Attitudes. Archives of Psychology, no. 140. This is Likert's Original Monograph.

2- Strongly agree

3- Agree

4- Undecided

5- Disagree

6- Strongly disagree

می‌شود. برای مثال، پاسخ «کاملاً موافقم» یک عبارت مطلوب، ارزش ۵، پاسخ «موافقم» ارزش ۴ و پاسخ «کاملاً مخالفم» ارزش ۱ خواهد گرفت. برای عبارات نامطلوب یا منفی، ارزش گذاری بر عکس می‌شود؛ زیرا از نظر روان شناختی، مخالفت با یک عبارت نامطلوب، معادل موافقت با یک عبارت مطلوب است. بنابراین «کاملاً موافقم» عبارت نامطلوب، ارزش ۱ و «کاملاً مخالفم»، ارزش ۵ خواهد گرفت (برروی مقیاسی که به پاسخگویان ارائه می‌شود، ارزش‌های وزنی درج نمی‌شوند). تصور کنید مثالهای زیر از مقیاسی که نگرش را نسبت به ریاضیات اندازه می‌گیرد، انتخاب شده‌اند:

SA	A	U	D	SD	
۵	۴	۳	۲	۱	ریاضیات درس دلخواه من در مدرسه است
۱	۲	۳	۴	۵	من ریاضیات را به هیچ شکلی دوست ندارم

شخصی که نگرش مطلوبی نسبت به ریاضیات دارد، با عبارت مثبت، موافقت و با عبارت منفی، مخالفت خواهد کرد. شخصی که ریاضیات را دوست ندارد، با عبارت منفی موافقت و با عبارت مثبت، مخالفت خواهد کرد. اگر آزمودنی برای عبارت اول A و برای عبارت دوم D را علامت بزند، مجموع نمره او برای این دو عبارت $4 + 4$ ، یعنی ۸، خواهد بود. در یک مقیاس کامل، مجموع ارزشهای همه عباراتی که آزمودنی علامت زده است، نمره کل او را نشان خواهد داد.

این سیستم ارزش گذاری به معنای آن است که کسب نمره بالا در یک مقیاس نگرش (پاسخ «کاملاً موافقم» به عبارات مطلوب، و پاسخ «کاملاً مخالفم» به عبارات نامطلوب)، نگرش مثبت نسبت به موضوع را نشان می‌دهد. بالاترین نمره ممکن برای مقیاس نگرش، $5 \times N$ ، و پایین ترین نمره ممکن $1 \times N$ است (N تعداد عبارات است).

تحلیل عبارت بعد از آن که مقیاس نگرش با یک گروه مقدماتی از پاسخگویان اجرا شد، برای تشخیص بهترین عبارات، یک تحلیل عبارت انجام می‌شود. تحلیل عبارت، به طور نوعی، سه آماره برای هر عبارت به دست می‌دهد: (۱) شاخص افتراق، (۲) تعداد و یا درصد پاسخگوییانی که گزینه‌های مختلف هر عبارت را انتخاب کرده‌اند، و (۳) میانگین و انحراف معیار

عبارت. شاخص افتراق عبارت، همانند نمره کل، میزان تمایزی را که هر عبارت در بین پاسخگویان ایجاد می‌کند، نشان می‌دهد. اگر دارندگان نمره بالا در یک عبارت، در کل مقیاس نیز نمره بالایی به دست آورده باشند و اگر دارندگان نمره پایین در یک عبارت، در کل مقیاس هم نمره پایینی به دست آورده باشند، پس به همان نحو که نمره کل در بین آزمودنیها تمایز ایجاد کرده است، آن عبارت نیز آزمودنیها را از هم متمایز خواهد کرد. شاخص افتراق عبارات، از طریق محاسبه همبستگی نمره هر عبارت با نمره کل مقیاس به دست می‌آید و این کار معمولاً به وسیله کامپیوتر صورت می‌گیرد. هر عبارت باید حداقل همبستگی 0.25 با نمره کل داشته باشد. عباراتی که با نمره کل، همبستگی پایین یا منفی دارند، باید حذف شوند؛ زیرا آنها همان چیزی را که نمره کل اندازه می‌گیرد، اندازه نمی‌گیرند، از اینرو در اندازه‌گیری نگرش دخالت ندارند. محقق باید آن عباراتی را که متمایز کننده نیستند، بررسی کند. این عبارات ممکن است مبهم یا دو پهلوی (حاوی دو عقیده یا نظر در یک جمله) باشند یا ممکن است جملات واقعی باشند که احساسی را در مورد یک موضوع بیان نمی‌کنند. تجدید نظر در این عبارات، شاید آنها را قابل استفاده سازد.

آماره‌های دیگری که از تحلیل عبارات به دست می‌آیند (۲ و ۳ یاد شده)، میزان استفاده پاسخگویان را از گزینه‌های مختلف نشان می‌دهند. عباراتی که در آنها پراکندگی پاسخگویان بر روی مقوله‌های پاسخ زیاد است، بر عباراتی که در آنها پاسخها بر روی یک یا دو مقوله تجمع یافته‌اند، ارجحیت دارند. محقق پس از آن که بر مبنای نتایج تحلیل، عبارات خوب را انتخاب کرد، باید مقیاس تجدید نظر شده را با گروه دیگری از آزمودنیها اجرا و باز هم عبارات را از نظر ایجاد افتراق بررسی کند.

اعتبار یافتن معیارهایی برای تعیین اعتبار مقیاسهای نگرش، کاری دشوارتر است. برخی از محققان، از مشاهده رفتار واقعی به عنوان معیار نگرشی که اندازه‌گیری شده است، استفاده کرده‌اند، اما این شیوه همیشه قابل استفاده نیست؛ زیرا تعیین این که چه رفتاری بهترین معیار نگرش خواهد بود، دشوار است، و بعلاوه، حصول اطمینان از این که رفتار بدرستی اندازه‌گیری شده است نیز کار آسانی نیست.

یکی از آسانترین راههای گردآوری شواهدی در زمینه اعتبار مقیاسهای نگرش، تعیین میزان توانایی آنها در ایجاد تمایز بین دو گروه از افرادی است که مشخص شده است نگرشهای متفاوتی

دارند. اگر مقیاسی که نگرش را نسبت به مذهب اندازه‌گیری می‌کند، بین اعضای فعال کلیسا و افرادی که به کلیسا نمی‌روند و هیچ‌گونه وابستگی مذهبی ندارند، تمایز قایل شود، می‌توان گفت که آن مقیاس دارای اعتبار است. یک مقیاس معتبر که نگرش را نسبت به سقط جنین اندازه‌گیری می‌کند، باید اعضای گروه‌های ضد آزادی سقط جنین را از گروه‌های طرفدار آزادی سقط جنین متمایز کند، یعنی انتظار خواهد رفت که میانگین نمرات اعضای این دو گروه در مقیاس نگرش مذکور، از نظر آماری با هم تفاوت معنادار داشته باشند. روش دیگر تعیین اعتبار، محاسبه همبستگی بین نمره‌های دو مقیاس نگرش است که هر دو یک سازه را اندازه می‌گیرند و مشخص شده است که یکی از آنها دارای اعتبار است.

/ مقیاسهای ترستون: روش فاصله‌های یکسان نما

ترستون روشی را برای تخصیص ارزشهای مقیاسی ویژه به عبارات نگرشی ابداع کرد.^(۱) در حالی که در شیوه لیکرت از پاسخگویان خواسته می‌شود تا موافقت یا مخالفت خود را با هر یک از عبارات ابراز کنند، مقیاس ترستون مجموعه‌ای از عبارات را که درباره یک موضوع باشند و دامنه آنها از خیلی مطلوب تا خنثی و از خنثی تا خیلی نامطلوب سیر کند، به پاسخگویان ارائه می‌کند و از آنها می‌خواهد که عباراتی را که بیشتر از بقیه با نگرش خود آنها مطابقت دارند، انتخاب کنند. تهیه یک مقیاس ترستون، مراحل زیر را شامل می‌شود:

۱- تعداد زیادی عبارت (۵۰ تا ۱۰۰ عبارت) در مورد موضوع نگرش جمع‌آوری کنید.

۲- عبارتها را در اختیار تعدادی داور بگذارید. معمولاً از پنجاه تا صد داور برای ارزیابی عبارتها استفاده می‌شود. داورها مستقلاً کار می‌کنند و از آنها خواسته می‌شود همه عبارات را بر مبنای درجه «مطلوبیت» هر عبارت در یازده طبقه دسته‌بندی کنند. هر داور عباراتی را که به نظرش مطلوبترین محتوا را نسبت به موضوع نگرش دارند، در مجموعه اول قرار می‌دهد. در مجموعه دوم عبارات کمتر «مطلوب» را جای می‌دهد و الی آخر. مجموعه ششم حاوی عبارات خنثی است و مجموعه یازدهم «نامطلوبترین» عبارات را در خود دارد. نکته مهم در این است که این طبقه‌بندی ربطی به نگرش داورها نسبت به موضوع ندارد، بلکه تنها معرف قضاوت آنها در مورد «مطلوبیت» عبارت‌ها است.

۳- ارزش مقیاسی هر عبارت را از طریق تعیین میانه وزنها یا مکانهای مقیاسی که داوران به هر عبارت تخصیص داده‌اند، پیدا کنید. عبارت‌هایی که تغییر پذیری بسیار زیاد نشان دهند، حذف می‌شوند.

۴- تعداد ۲۰ یا ۳۰ عبارت را که به طور یکنواخت از یک انتها تا انتهای دیگر مقیاس توزیع شده‌اند، انتخاب کنید. این عبارت‌ها مقیاس نگرش را تشکیل می‌دهند. عبارات زیر با ارزش مقیاسی آنها، از مقیاس نگرش ترستون نسبت به کلیسا گرفته شده‌اند.

ارزش مقیاس	عبارت
۰/۲	من بر این باورم که کلیسا در آمریکای امروز بزرگترین نهاد است.
۱/۵	من بر این باورم که عضویت در کلیسا برای گذران زندگی به بهترین نحو، امری اساسی است.
۲/۳	من خدمات کلیسا را هم آرام بخش و هم الهام بخش یافته‌ام.
۳/۳	من از کلیسای خود لذت می‌برم، زیرا فضایی دوستانه بر آن حاکم است.
۴/۵	من به تعلیمات کلیسا معتقدم، اما به شرط عقل.
۵/۶	بعضی اوقات احساس می‌کنم کلیسا و مذهب ضروری هستند و گاهی در مورد آن شک می‌کنم.
۶/۷	من به صداقت و نیکی بدون تشریفات کلیسایی اعتقاد دارم.
۷/۴	من بر این باورم که هر چه تعلیم و تربیت پیشرفت می‌کند، مذهب زمینه خود را از دست می‌دهد.
۸/۳	من فکر می‌کنم که تعلیمات کلیسا بر رویهم بسیار سطحی‌تر از آن است که اهمیت اجتماعی داشته باشد.
۹/۶	من فکر می‌کنم کلیسا مانعی بر سر راه مذهب است، برای این که هنوز به جادو، خرافات و افسانه تکیه می‌کند.
۱۱/۰	من فکر می‌کنم کلیسا انگلی برای جامعه است.

برای اجرای یک مقیاس نوع ترستون، از آزمودنی خواسته می‌شود عباراتی را که معرف موضع او هستند، انتخاب کند، یا سه عبارت را که بیشتر از سایر عبارات به موضع او نزدیک‌اند،

برگزینند. البته ارزشهای مقیاسی در مقیاس نگرش قید نمی‌شوند و عبارات نه بر اساس ارزش مقیاسی، بلکه به طور تصادفی در مقیاس نگاشته می‌شوند. نمره هرازمودنی میانگین ارزشهای عباراتی که او انتخاب کرده است، خواهد بود.

/ مقیاسهای گاتمن: روش تراکمی

منتقدان مقیاسهای نگرش لیکرت و ترستون خاطر نشان کرده‌اند که این مقیاسها حاوی عبارات ناهمگن درباره ابعاد مختلف یک موضوع نگرش است. برای مثال، در مقیاس ترستون که نگرش نسبت به جنگ را می‌سنجد، برای جدا سازی عبارات اخلاقی از عباراتی که درباره نتایج اقتصادی جنگ است یا عباراتی که دیگر جنبه‌های ممکن نگرش نسبت به جنگ را منعکس می‌کنند، هیچ تلاشی صورت نگرفته است. پیامد ترکیب ابعاد مختلف در یک مقیاس، آن است که ارائه تفسیری روشن از نمره‌های حاصله را دشوار می‌سازد. گاتمن روشی را در جهت تلاش برای غلبه بر این مشکل ابداع کرد. روش او، که به عنوان مقیاس تک بعدی مشخص شده است، می‌خواهد تعیین کند که نگرش مورد مطالعه واقعاً دارای یک بعد است. تنها نگرشی تک بعدی محسوب می‌شود که یک مقیاس تراکمی به بار آورد. مقیاسی که عبارات آن به گونه‌ای با هم مرتبط اند که اگر آزمودنی با عبارت ۲ موافق است، همچنین با عبارت ۱ نیز موافقت می‌کند. کسی که با عبارات ۳ موافق است، همچنین با عبارات ۱ و ۲ نیز موافقت می‌کند و الی آخر. بنابراین افرادی که یک عبارت خاص را در این نوع مقیاس تایید می‌کنند، از افرادی که آن عبارت را رد می‌کنند، در کل مقیاس، نمره بالاتری به دست خواهند آورد. برای مثال عبارات زیر را که از پاسخگویان خواسته می‌شود با آنها موافقت یا مخالفت کنند، در نظر بگیرید:

۱- PTA ارزش آن را دارد که برای آن وقت صرف شود.

۲- PTA در امر بهبود وضعیت مدارس، دارای نفوذ بسیار است.

۳- PTA در ایالات متحده بهترین سازمان در زمینه بهبود وضعیت مدارس است.

اگر این یک مقیاس تراکمی باشد، باید امکان مرتب کردن همه پاسخهای پاسخگویان به گونه الگویی که در جدول ۷-۱ نشان داده شده است، وجود داشته باشد. بنابراین اگر نمره یک فرد را بدانیم، این امکان را نیز داریم که دقیقاً بگوییم او کدام عبارتها را تایید کرده است. برای مثال، همه پاسخگویانی که نمره ۲ به دست آورده‌اند، بر این باورند که PTA ارزش صرف وقت را دارد و

دارای نفوذ زیادی در امر بهبود وضعیت مدارس است، اما معتقد نیستند که PTA در ایالات متحده مهمترین سازمان در زمینه بهبود وضعیت مدارس است. آزمودنیها را می توان بر مبنای پاسخهایشان به عبارات مقیاس رتبه بندی کرد.

به هنگام ساختن یک مقیاس تراکمی، ابتدا تعیین این که عبارات آن یک مقیاس تک بعدی را تشکیل می دهند یا خیر، امری اساسی است.^(۱) برای انجام این کار، تجدیدپذیری پاسخها را تحلیل می کنند - یعنی نسبت پاسخهایی را که واقعاً در الگویی شبیه جدول ۱-۷ جای می گیرند، تعیین می کنند. بر مبنای نمره کل، الگوی پاسخها به عبارات خاص، پیش بینی می شود. سپس الگوی واقعی پاسخها مطالعه می شود و بر اساس میزانی که پاسخها از روی نمره کل تجدیدپذیرند، اندازه ای ساخته می شود. یک روش آن است که تعداد کل خطاها بر تعداد کل پاسخها تقسیم و از یک کم شود. گاتمن ۰/۹ را به عنوان حداقل ضریب تجدیدپذیری، که برای شناسایی مجموعه ای از عبارات به عنوان اجزای متشکله یک مقیاس تک بعدی لازم است، پیشنهاد می کند.

جدول ۱ - ۷ مثال روش تراکمی*

نمره	موافقت با عبارت			مخالفت با عبارت		
	۳	۲	۱	۳	۲	۱
۳	×	×	×	۰	۰	۰
۲	۰	×	×	×	۰	۰
۱	۰	۰	×	×	×	۰
۰				×	×	×

* پاسخگو برای هر پاسخ موافق یک نمره می گیرد.

بسیاری بر این باور هستند که مقیاس گاتمن بیش از آن که اهمیت کاربردی داشته باشد، اهمیت نظریه ای دارد؛ زیرا جمع آوری عبارتهایی که معیار تجدیدپذیری را برآورده سازند، دشوار

۱- برای بحث مفصل تر مراجعه کنید به :

A. L. Edwards (1952), *Techniques of Attitude Scale Construction*, New York: Irvington.

است. همچنین این انتقاد بر شیوه گاتمن وارد است که راه یا راههایی را برای تهیه یا انتخاب عبارتها پیشنهاد نمی‌کند. تنها پس از این که عبارتها انتخاب شدند، می‌توان درباره تجدیدپذیری آنها قضاوت کرد.

/ مقیاسهای افتراق معنایی

رویکرد دیگر برای اندازه‌گیری نگرش نسبت به اشیاء، موضوعها و رویدادها روش افتراق معنایی است. این روش را آزگود، سوسی، و تانن‌باوم^(۱) ابداع و از آن استفاده کردند. روش افتراق معنایی بر این مفروضه بنا شده است که اشیاء برای افراد دو معنای مختلف دارند، معنای صریح و معنای ضمنی، که می‌توان آنها را به طور مستقل ارزیابی کرد. معنای صریح یک کلمه به معنایی که در فرهنگ لغات برای آن کلمه آمده است، اطلاق می‌شود، در حالی که معنای ضمنی تداعی‌ها و اشاراتی است که یک کلمه به ذهن متبادر می‌کند. افراد پسادگی می‌توانند معنای صریح یک شیء را بیان کنند، اما در بیان معنای ضمنی آن با مشکل مواجه می‌شوند. اندازه‌گیری معنای ضمنی اشیاء به طور غیر مستقیم، با استفاده از تعدادی صفات دو قطبی و درخواست از افراد برای رتبه بندی اشیاء در قبال این صفات، هم امکان پذیر و هم مفید است. بنابراین، معنای یک شیء برای فرد، الگوی رتبه بندی آن شیء بر روی مقیاسهای صفات دو قطبی خواهد بود.

مقیاسهای نگرش با انتخاب زوجهای صفات دو قطبی که معرف بعد ارزشی باشند، ساخته می‌شوند. زوجهای صفات به همراه یک پیوستار پاسخدهی هفت تایی ارائه می‌شوند و از پاسخگو خواسته می‌شود که با گذاردن × بر روی یکی از فضاها، هفت تایی، نشان دهد که صفت مورد نظر تا چه حد شیء نگرشی را توصیف می‌کند. برای مثال، فرض کنید فردی می‌خواهد نگرشهای دانش‌آموزان دبیرستانی را نسبت به مدرسه اندازه بگیرد.

آزگود، سوسی، و تانن‌باوم از طریق مطالعات تحلیل عاملی، به سه دسته از صفات دست پیدا کرده‌اند: ارزشی^(۲) (شامل صفاتی نظیر خوب و بد، ارزشمند و بی ارزش)، توانایی^(۳) (شامل صفاتی نظیر قوی و ضعیف یا سنگین و سبک)، فعالیت^(۴) (شامل صفاتی نظیر فعال و وارفته یا سریع و کند)، در میان این سه دسته، به نظر می‌رسد که بعد ارزشی برای اندازه‌گیری نگرش، بیشترین اهمیت را دارا باشد.

1- Osgood, G.E., G.J. Suci, and P.H. Tannenbaum (1967). The Measurement of Meaning. Urbana: University of Illinois Press.

2- Evaluative

3- Potency

4- Activity

مدرسه

خوب	----	----	----	----	----	----	بد
وارفته	----	----	----	----	----	----	فعال
خنک	----	----	----	----	----	----	تیزهوش
ناخوشایند	----	----	----	----	----	----	خوشایند
ارزشمند	----	----	----	----	----	----	بی ارزش
نرم	----	----	----	----	----	----	سخت
سبک	----	----	----	----	----	----	سنگین
قوی	----	----	----	----	----	----	ضعیف
کند	----	----	----	----	----	----	سریع

به مقیاس بالا توجه کنید، که در آن، برای این که گرایش پاسخگو به یک سمت مقیاس به حداقل برسد، زوجهای صفات در هر دو جهت مثبت و منفی گنجانده شده‌اند. ممکن است فردی انتهای سمت راست مقیاس را انتخاب کند و برای همه موارد، آنجا را علامت بزند. به هر ترتیب، اگر جهت مقیاس به طور تصادفی تغییر داده شود، به گونه‌ای که سمت راست مقیاس، همیشه پاسخ مثبت را در بر نداشته باشد، پاسخگو وادار می‌شود که هر مورد را ابتدا بخواند و پاسخ خود را برحسب محتوای آن ارائه کند. به هر روی، در نمره دهی مقیاس افتراق معنایی، به پاسخها از ۱ تا ۷ نمره داده می‌شود؛ به طوری که ۷، مثبت‌ترین پاسخ را بنمایاند. بنابراین، مورد اول مقیاس بالا، یعنی صفت بد، اولین گزینه از سمت راست، نمره ۱، دومین گزینه از سمت راست، نمره ۲، و اولین گزینه از سمت چپ، نمره ۷ خواهد گرفت. برای مورد دوم، یعنی صفت فعال، ترتیب نمره دهی بر عکس می‌شود، به طوری که گزینه اول سمت راست، نمره ۷، و گزینه اول سمت چپ، نمره ۱ می‌گیرد. سپس، تمام نمره‌ها با هم جمع می‌شوند و یک نمره میانگین محاسبه می‌شود.

// مقیاسهای درجه‌بندی

یکی از رایجترین ابزارهای اندازه‌گیری، مقیاس درجه‌بندی است. مقیاسهای درجه‌بندی عبارتند از سنجش رفتار یا عملکرد یک فرد به وسیله فردی دیگر. به طور نوعی، از فرد درجه‌بندی کننده خواسته می‌شود که جایگاه فرد درجه‌بندی شونده را بر روی یک پیوستار یا در

یک طبقه‌بندی که ویژگیهای رفتاری او را توصیف می‌کند، مشخص کند. به این درجه‌بندیها ارزش عددی تخصیص می‌یابد. چنین فرض می‌شود که درجه‌بندی‌کنندگان با رفتار نوعی افراد آشنا هستند. درجه‌بندیها در تحقیق بر روی رشد کودکان و سایر جنبه‌های رفتاری مورد استفاده قرار گرفته‌اند.

شکل ۲- ۷ مثالی از یک مقیاس نگراره‌ای		
پایین	متوسط	بالا
		وضع ظاهری
		پذیرش اجتماعی
		مهارتهای بیانی
		همیشه ایده‌های خلاق دارد
		ایده‌های خلاق بسیاری دارد
		گاهی ایده‌های خلاق دارد
		به ندرت ایده‌های خلاق دارد

انواع مختلفی از مقیاسهای درجه‌بندی وجود دارند. یکی از مقیاسهایی که بیشترین مورد استفاده را دارد، مقیاس نگراره‌ای است که در آن درجه‌بندی‌کننده فقط نقطه مناسب را بر روی خطی که دو حد نهایی رفتار مورد نظر را نشان می‌دهد، علامت می‌زند. شکل ۲- ۷ یک مثال نوعی از مقیاس نگراره‌ای است. فردی که درجه‌بندی می‌کند، می‌تواند هر نقطه‌ای را روی خط پیوستار علامت بزند. در برخی از مقیاسهای نگراره‌ای، سازنده آزمون، به نقاط توصیفی، ارزشهای عددی تخصیص می‌دهد. به چنین مقیاسهایی، مقیاسهای درجه‌بندی عددی گفته می‌شود. مورد مهارتهای بیانی در شکل ۲- ۷، در یک مقیاس عددی به صورت زیر در می‌آید.

۱	۲	۳	۴	۵	۶	۷
یکی از			یک سخنران			یکی از
ضعیف‌ترین			متوسط			بهترین
سخنرانان						سخنرانان

نوع دوم مقیاس درجه‌بندی، مقیاس طبقه‌ای است که شامل طبقاتی است که بر مبنای نظم خاصی مرتب شده‌اند. غالباً پنج الی هفت طبقه مورد استفاده قرار می‌گیرند. فرد درجه‌بندی کننده، یکی از طبقات را که به بهترین نحو رفتار فرد درجه‌بندی شونده را توصیف می‌کند، انتخاب می‌کند. فرض کنید تواناییهای یک دانش‌آموز درجه‌بندی شده است و یکی از ویژگیهای درجه‌بندی شده، خلاقیت است. مورد خلاقیت به صورت زیر در می‌آید.

این فرد تا چه حد خلاق است؟ (یکی را علامت بزنید)

فوق العاده خلاق است

خیلی خلاق است

خلاق است

خلاق نیست

ابداً خلاق نیست

درجه‌بندی کنندگان، به هنگام استفاده از مقیاسهای نگاره‌ای و طبقه‌ای، بدون آن که فرد درجه‌بندی شونده را مستقیماً با افراد یا گروههای دیگر مقایسه کنند، قضاوت خود را انجام می‌دهند. از سوی دیگر، در مقیاسهای درجه‌بندی مقایسه‌ای، به درجه‌بندی‌کنندگان آموزش داده می‌شود که با عطف به جایگاه کسانی که قابل مقایسه هستند، قضاوت خود را انجام دهند. در مقیاس درجه‌بندی، جایگاه برحسب یک جمعیت معین با مشخصه‌های شناخته شده، تعریف می‌شود. در شکل ۳-۷ یک مقیاس درجه‌بندی مقایسه‌ای نشان داده شده است. چنین مقایسه‌ای را می‌توان در انتخاب داوطلبان ورود به دانشکده مورد استفاده قرار داد. از درجه‌بندی کنندگان خواسته می‌شود در مورد توانایی داوطلبان برای تحصیل در دانشکده، در مقایسه با همه دانشجویانی که از قبل می‌شناسند، قضاوت کنند.

همه روشهای درجه‌بندی در معرض خطاهای قابل توجهی هستند. از رایجترین خطاهای سیستماتیک در درجه‌بندی افراد «اثر هاله‌ای»^(۱) است. این خطا هنگامی اتفاق می‌افتد که فرد درجه‌بندی کننده اجازه دهد تا برداشت تعمیم یافته او از آزمودنی بر درجه‌بندی جنبه‌های بسیار خاص رفتار آزمودنی تأثیر بگذارد. این برداشت کلی به تمام مقیاس انتقال می‌یابد. برای مثال، ممکن است یک آموزگار، دانش‌آموز درسخوانی را از نظر هوش، محبوبیت، صداقت، امانتداری و همه جنبه‌های دیگر شخصیتی در حد عالی درجه‌بندی کند.

شکل ۳-۷ مثالی از یک مقیاس درجه بندی مقایسه ای

حیطه شایستگی (باید درجه بندی شود)					
شیر قابل قضاوت	دیگران	رفتار تراز	دانش آموزان بهتر از اکثر	میان دانش آموزان حد متوسط در	دانش آموزان ضعیف تر از اکثر
					به طور غیر عادی پایین
۱- آیا این فرد شواهد روشن و شایسته ای از هدفهای حرفه ای نشان می دهد؟					
۲- آیا این فرد با روش سازنده ای به جنگ مشکلات می رود؟					
۳- آیا او انتقادهای مثبت را می پذیرد و از آن به طور سازنده ای استفاده می کند؟					

یک نوع خطای دیگر، «خطای سخاوتمندی»^(۱) است. این خطا، گرایش به در نظر گرفتن سوی مثبت هر شک و تردیدی، در مورد آزمودنی است، یعنی هنگامی که درجه بندی کننده مطمئن نیست، احتمالاً در مورد فردی که درجه بندی می کند، با دید موافق صحبت خواهد کرد. در مقابل آن، «خطای سخت گیری»^(۲) است که گرایش به درجه بندی کردن بسیار پایین همه افراد، در همه ویژگیها دارد. منبع خطای دیگر، «خطای گرایش مرکزی»^(۳) است که به اجتناب از دو حد نهایی مقیاس و درجه بندی کردن همه افراد در حد میانی مقیاس اطلاق می شود.

یک راه برای کاهش چنین خطاهایی آن است که افراد درجه بندی کننده را قبل از انجام کار به طور کامل آموزش دهیم. آنها باید از امکان ارتکاب این نوع خطاها آگاه باشند. این یک امر اساسی است که افراد درجه بندی کننده، قبل از انجام درجه بندی، وقت کافی برای مشاهده رفتار فردی که قرار است درجه بندی شود، داشته باشند. یک راه دیگر برای به حداقل رساندن خطا آن است که مطمئن شویم رفتاری که قرار است درجه بندی شود و نقاطی که روی مقیاس هستند، به طور روشن، تعریف شده اند. نقاط روی مقیاس باید برحسب رفتارهای آشکاری که قابل مشاهده هستند، توصیف شوند، نه برحسب رفتارهایی که نیازمند استنباط از سوی فرد درجه بندی کننده

1- Generosity Error

2- Error of Severity

3- Error of Central tendency

می‌باشند. به خواننده توصیه می‌کنیم برای یک مبحث عالی در مورد شیوه‌های رویارویی با خطاهای فرد درجه‌بندی کننده به کتاب گیلفورد^(۱) مراجعه کند.

معمولاً اگر افراد متعددی به طور مستقل یک فرد را درجه‌بندی کنند، پایایی درجه‌بندی افزایش می‌یابد. این درجه‌بندیهای مستقل یک کاسه می‌شوند، یا میانگین آنها گرفته می‌شود تا یک نمره مستقل به دست آید.

// روشهای گروه سنجی

روشهای گروه سنجی برای مطالعه سازمان گروههای اجتماعی مورد استفاده قرار می‌گیرند. روش اصلی، اگر چه ممکن است به گونه‌های مختلف تغییر یابد، چنین است که از اعضای یک گروه خاص خواسته می‌شود که انتخاب اول، دوم و انتخابهای بعدی خود را از میان سایر افراد گروه، برای همراهی در یک فعالیت خاص و بر مبنای یک معیار ویژه نشان دهند. برای مثال، از افراد یک گروه پرسیده می‌شود: «با چه کسی می‌خواهید درس بخوانید، یا در کنارش بنشینید، یا با او ناهار بخورید، یا بعد از مدرسه با او بازی کنید؟» اساساً این روش، مطالعه انتخابهایی است که هر فرد در یک گروه انجام می‌دهد. انتخابهایی که به دست می‌آید، به صورت یک نگاره گروهی ترسیم می‌شود که الگوی روابط میان فردی را در یک گروه نشان می‌دهد. شکل ۴-۷ یک نگاره گروهی را نشان می‌دهد، که در آن ج بیش از سایر افراد گروه انتخاب شده است. می‌توان از او به عنوان «ستاره» نام برد. توجه کنید که و، ز و ه یکدیگر را انتخاب کرده‌اند. این نوع انتخابها نشانه‌ای از وجود یک دسته است، یعنی سه فرد یا بیشتر که دو بدو یکدیگر را انتخاب می‌کنند. ب اصلاً انتخاب نشده است، اگر چه او خود در گروه انتخابهایی انجام داده است، او یک «منزوی» است. انتخابهایی که بر روی یک نگاره گروهی آشکار می‌شود، می‌تواند به صورت کمی درآید و برای مقاصد تحقیقاتی مورد استفاده قرار گیرد.

روشهای گروه سنجی برای تحقیق در روان‌شناسی اجتماعی و همچنین در تحقیقات آموزشی، که در آن امکان مطالعه رابطه وضع فرد در گروه با سایر متغیرها نظیر توانایی ذهنی، پیشرفت، رجحانهای آموزگاران در مورد کودکان وجود دارد، به طور وسیع مورد استفاده قرار گرفته است.^(۲)

1- Guilford, J.P. (1954). *Psychometric Methods* (ch. 11). New York: McGraw - Hill.

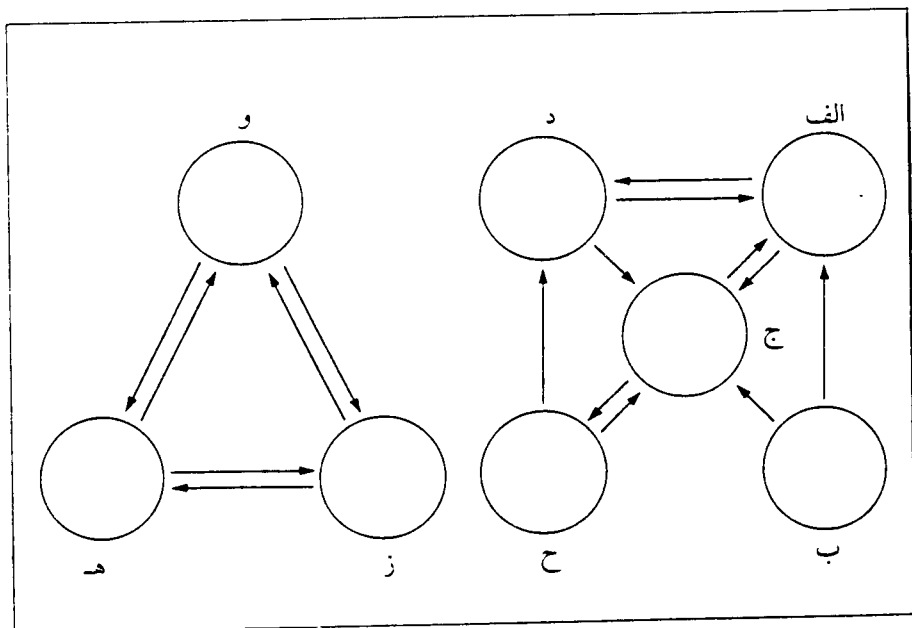
2- G. Lindzey (1968), *Handbook of Social Psychology* (Vol. 1), Reading, MA: Addison - Wesley.

// مشاهده مستقیم

در بسیاری از حالات، مشاهده مستقیم و نظام‌مند رفتار، مطلوب‌ترین روش اندازه‌گیری است. یک پژوهشگر رفتار مورد نظر را معین می‌کند و سپس روش نظام‌مندی برای تشخیص، طبقه‌بندی و ثبت آن رفتار، هم در محیط طبیعی و هم در محیط مصنوعی، ابداع می‌کند. مشاهده نظام‌مند به طور گسترده برای تحقیق بر روی نوزادان و کودکان پیش دبستانی مورد استفاده قرار گرفته است.

تحقیق اوربان^(۱) با عنوان «تغییرات رفتاری حاصله از مطالعه بیماری‌های واگیر» نمونه ممتازی از کاربرد مشاهده مستقیم در محیط طبیعی است. مشاهده گران تعداد موارد رفتارهای نامطلوب، نظیر گذاشتن انگشت یا اشیای دیگر را در دهان و رفتارهای مطلوب، نظیر استفاده از دستمال به

شکل ۴-۷ یک نگاره گروهی



10. Urban, J. (1943). Behavior Changes Resulting From a Study of Communicable Diseases. *Teachers College Contributions to Education*, no. 896. New York: Teachers College Press, Columbia University.

هنگام سرفه یا عطسه را ثبت کردند. در پی آن، یک گروه آزمایشی انتخاب شد و یک دوره شش هفته‌ای را درباره بیماریهای واگیر، که هم برای تغییر رفتار آسکار افراد گروه و هم برای آگاه سازی آنها از این بیماریها طراحی شده بود، گذراند. در پایان دوره، رفتارهای نامطلوب دوباره ثبت شدند. نتایج حاصله مشخص کرد که در گروه آزمایشی رفتارهای نامطلوب به طور چشمگیری کاهش و رفتارهای مطلوب به طرز چشمگیری افزایش یافته‌اند، در حالی که در گروه گواه، که آن دوره را نگذرانده بودند، تغییر کمی صورت گرفته است. مشاهدیهایی که در طول ۱۲ هفته انجام شد، نشان داد که تفاوت‌های بین دو گروه گواه و آزمایشی ماندگار است.

قبل از استفاده از مشاهده مستقیم پنج مرحله مقدماتی مهم باید طی شود.

۱- آن جنبه از رفتار که قرار است مشاهده شود، باید انتخاب شود. از آنجا که یک مشاهده‌گر نمی‌تواند هر چیزی را که اتفاق می‌افتد، مورد توجه قرار دهد، محقق از قبل باید تصمیم بگیرد که کدام رفتارها ثبت و کدام رفتارها حذف شوند.

۲- رفتارهایی که در یک طبقه‌بندی قرار می‌گیرند، باید بروشنی تعریف شوند. مشاهده‌گران باید از پیش بدانند که برای مثال، چه اعمالی به عنوان رفتار جمعی یا رفتار خودخواهانه طبقه‌بندی خواهد شد.

۳- کسانی که کار مشاهده را انجام می‌دهند، باید آموزش داده شوند. مشاهده‌گران باید آموزش ببینند و تمرین کنند تا یکدستی تفسیرها و کاربست همسان طبقه‌بندیهای مشاهده حاصل شود.

۴- برای کمی کردن مشاهده‌ها باید سیستمی تهیه شود. محقق باید یک روش استاندارد برای شمارش رفتارهای مشاهده شده مشخص کند. برای مثال، محقق باید از قبل تعیین کند که آیا یک کنش و واکنش را باید یک رفتار منفرد شمارش کرد یا دو بخش از یک رفتار. یک شیوه پیشنهادی آن است که مدت زمان مشاهده به قطعات زمانی کوتاه تقسیم شود و برای هر دوره زمانی - به عنوان مثال، ۳۰ ثانیه - بروز یا عدم بروز آن رفتار ثبت شود.

۵- برای ثبت رفتار باید شیوه‌های مبسوطی تهیه شود. معمولاً حافظه اکثر مشاهده‌گران برای یک تحقیق هدفمند، به اندازه کافی قابل اعتماد نیست. بهترین راه عبارت است از تهیه یک طرح رمزگردانی که مشاهده‌گران را قادر سازد تا مشاهدات خود را، نه به صورت نوشتاری - که هم وقت مشاهده‌گر را می‌گیرد و هم توجه او را منحرف می‌سازد - بلکه با یک حرف مفرد یا عدد، ثبت کنند.

کاربرد دیگر مشاهده مستقیم، مطالعه رفتار کلاسی است. فلاندرز تعامل کلامی را در کلاس درس مطالعه و یک برنامه رمزگردانی برای ثبت مشاهدات تهیه کرد که آن را می‌توان در شکل ۵-۷ دید. سیستم فلاندرز طبقه‌بندیهای جامع و نامتناخلی ارائه می‌کند، که هر یک را می‌توان با یک عدد مفرد ثبت کرد. مشاهده‌گران آموزش دیده قادرند در هر ۳ ثانیه یک عدد را ثبت کنند. زنجیره اعداد حاصله را می‌توان به آسانی تحلیل کرد تا نه تنها پیشینه نسبتهای رفتار کلامی هر طبقه‌بندی، بلکه همچنین تصویری از تقدم و تأخر آن رفتارها نیز به دست آید.

فلاندرز گفته‌های معلم را به دو بخش گفتار پاسخی، که آن را نفوذ غیر مستقیم نامید، و گفتار آغازی، که آن را نفوذ مستقیم نامید، تقسیم کرد. فلاندرز در یک بررسی، با استفاده از طبقه‌بندی مشاهده خود، تأثیر گفته‌های نوع نفوذ غیرمستقیم معلم را در مقایسه با گفته‌های از نوع نفوذ مستقیم او، مورد مطالعه قرار داد.^(۱) کلاسهای علوم اجتماعی و ریاضیات مورد مشاهده قرار گرفتند و کلاسهایی که معلمان آنها از روش غیرمستقیم استفاده می‌کردند، با کلاسهایی که معلمان آنها از روش مستقیم استفاده می‌کردند، هم از نظر پیشرفت تحصیلی و هم از نظر نگرش نسبت به معلم با هم مقایسه شدند. فلاندرز دریافت که کلاسهای معلمانی که از روش غیرمستقیم استفاده می‌کنند، هم در زمینه نگرش و هم در زمینه پیشرفت تحصیلی، عملکرد بهتری دارند. نتیجه این بررسی آن بود که اگر به معلمان آموزش داده شود که از مقوله غیرمستقیم رفتار کلامی بیشتر استفاده کنند، کار آنها بهبود می‌یابد.

در یک مطالعه کلاسی دیگر، رالینز و دیگران مشاهده نظام‌مند را قبل و بعد از یک دوره آموزشی که طی آن معلمان شیوه‌های تقویت مثبت را فرا گرفتند و به استفاده از این شیوه‌ها تشویق شدند، به کار بردند.^(۲) مشاهده برای این انجام شد که میزان استفاده معلمان از شیوه تقویت مثبت و اثر تقویت مثبت بر رفتار دانش‌آموزان، تعیین شود. مشاهده در دوره‌های ۵ دقیقه‌ای انجام شد. در اولین دوره ۵ دقیقه‌ای رفتارهای تقویتی مثبت و منفی معلمان شمارش شد، در دومین دوره ۵ دقیقه‌ای، رفتارهای اخلاص برانگیز دانش‌آموزان شمارش شد، در سومین دوره ۵ دقیقه‌ای، رفتارهای توجه آمیز دانش‌آموزان مورد مشاهده قرار گرفت. رالینز دریافت

1- Rollins, N.A. (1965). *Teacher Influence, Pupil Attitudes, and Achievement*. U.S. Office of Education Cooperative Research Monograph. Washington: U.S. Government Printing Office.

2- Rolling, H.A., B.R. McCandless, and M. Thompson (1974). Project Success Environment: An Extended Application of Contingency Management in Inner-city Schools. *Journal of Educational Psychology*, 66, pp. 167 - 178.

معلمانی که آموزش دیده بودند، تقریباً چهار برابر معلمان گروه کنترل، نسبت به هر دانش آموز، از رفتار تقویتی استفاده کردند و خیلی کمتر مجبور به استفاده از تنبیه شدند. مشاهداتی که بر روی دانش آموزان صورت گرفت، کاهش چشمگیری در رفتارهای اخلاص برانگیز و افزایش در رفتار توجه آمیز را نشان داد. بعلاوه، دانش آموزانی که در گروه آزمایشی قرار داشتند، در یک دوره ۸ ماهه نمره متوسط بالاتری در آزمون خواندن به دست آوردند. در این مطالعه، از مشاهده نظام مند به طور خیلی مؤثر برای ثبت تغییرات رفتار کلاسی، که با تغییرات چشمگیر در پیشرفت تحصیلی دانش آموزان همبسته بود، بهره گرفته شد.

/ ارزشیابی مشاهده مستقیم

در مشاهده مستقیم ما قادر هستیم که پیشینه‌ای از رفتار واقعی فرد را در موقعیتهای طبیعی به دست آوریم. استفاده از این روش در مورد کودکان کم سن و سال که نمی‌توانند ارتباط زبانی برقرار کنند، بی‌نهایت ارزشمند است.

به هر روی، مشاهده مستقیم به لحاظ مدت زمانی که برای مشاهده نیاز دارد، یک شیوه پرهزینه است. علاوه بر این، همیشه این امکان وجود دارد که حضور مشاهده‌گر رفتار فرد را تغییر دهد. برای غلبه بر این مشکل، در برخی از موقعیتهای، می‌توان از شیشه‌های یک سو به استفاده کرد. در بسیاری از موارد، آزمودنیهایی که مورد مشاهده قرار می‌گیرند، بعد از یک واکنش اصلی، توجه چندانی به مشاهده‌گر نمی‌کنند، بویژه آن که مشاهده‌گر بدون جلب نظر آزمودنی به کار خود مشغول باشد. برای آن که مشاهده‌گران به افراد کارآمدی بدل شوند و بدانند که چه چیزی را مشاهده کنند و نتیجه مشاهده خود را چگونه گزارش دهند، به آموزشهای گسترده‌ای نیاز دارند.

/ مشاهده‌های ابداعی

در مشاهده‌های ابداعی، محقق تدارک کار را به نحوی فراهم می‌آورد که آزمودنیها را در شبیه‌سازی موقعیتهای زندگی واقعی مورد مشاهده قرار دهد. شرایط به گونه‌ای فراهم می‌شود که رفتار دلخواه بروز کند.

یک شکل از مشاهده ابداعی، آزمون موقعیتی است. یک مثال کلاسیک از آزمون موقعیتی - اگر چه در زمان خود به این نام خوانده نشد - در مجموعه مطالعات هارتشورن و می درباره

شکل ۵-۷ طبقه‌بندی تحلیل تعاملی فلاندرز (FIAC)*

گفتار معلم:

۱- احساسات را می‌پذیرد: نگرش یا حال و هوای احساس یک دانش‌آموز را به شیوه‌ای ملاحظه‌آمیز می‌پذیرد و توضیح می‌دهد. احساسات ممکن است مثبت یا منفی باشند. پیش‌بینی و به یادآوری احساسات را نیز شامل می‌شود.

۲- تحسین یا تشویق می‌کند: کنش یا رفتار دانش‌آموز را تحسین یا تشویق می‌کند. لطیفه‌هایی می‌گوید تا تنش را کم کند، اما لطیفه‌های او دانش‌آموز دیگری را نمی‌آزارد؛ سرنگان دادن یا گفتن «وام هوم» یا «ادامه بده» را نیز شامل می‌شود.

پاسخ

۳- ایده‌های دانش‌آموزان را می‌پذیرد یا مورد استفاده قرار می‌دهد: ایده‌های پیشنهادی دانش‌آموز را توضیح می‌دهد، می‌پرواند یا بهبود می‌بخشد. بسط ایده‌های دانش‌آموز توسط معلم را نیز شامل می‌شود، اما اگر ایده‌ها، بیشتر متعلق به خود معلم باشد، به طبقه‌بندی شماره ۵ منتقل می‌شود.

۴- سؤال مطرح می‌کند: درباره شیوه یا محتوا سؤالی مطرح می‌کند که بر ایده‌های خود معلم مبتنی باشد، با این نیت که دانش‌آموز پاسخ دهد.

۵- سخنرانها: واقعیت‌ها یا نظرات را درباره محتوا و شیوه ارائه می‌دهد، ایده‌های خود را بیان می‌کند، توضیحات خود را ارائه می‌دهد، یا منبعی به غیر از یک دانش‌آموز را مثال می‌زند.

۶- رهنمود می‌دهد: رهنمود، فرمان، یا دستوری می‌دهد که انتظار می‌رود دانش‌آموزان از آن اطاعت کنند.

۷- اقتدار را نقد یا توجیه می‌کند: به قصد تغییر رفتار دانش‌آموز، از یک الگوی ناپذیرفتنی به الگوی پذیرفتنی، عباراتی بیان می‌کند، کسی را دعوا می‌کند، دلیل انجام کار خود را شرح می‌دهد، منتها درجه خود مدار است.

۸- پاسخ دانش‌آموز: گفته‌های دانش‌آموزان در پاسخ به معلم. معلم ارتباط را برقرار می‌کند، اظهارات دانش‌آموزان را خواستار می‌شود، یا به موقعیت شکل می‌دهد، آزادی بیان ایده‌های شخصی محدود است.

گفتار دانش‌آموز

۹- آغاز به وسیله خود دانش‌آموز: گفته‌های دانش‌آموزان که خود آغازگر آن باشند. ایده‌های خود را شرح می‌دهند، موضوع جدیدی را آغاز می‌کنند، این آزادی را دارند که نظرات یا یک خط فکری را بسط دهند، نظیر طرح سؤالات فکروانه. از ساختار موجود فراتر می‌روند.

آغاز

۱۰- سکوت یا سردرگمی: دوره‌های کوتاه سکوت و دوره‌های سردرگمی که مشاهده‌گر نمی‌تواند ارتباط را درک کند.

سکوت

بررسی تعلیم منش (CEI) مورد استفاده قرار گرفت.^(۱) این آزمون‌ها برای استفاده در مطالعه بسط ویژگیهای رفتاری، نظیر درستکاری، خویشتنداری، راستگویی و همکاری طراحی شدند. هارتشورن و می مشاهداتی بر روی فعالیتهای روزمره مدرسه‌ای کودکان انجام دادند، اما همچنین موقعیتهایی را نیز برای تمرکز بر رفتار خاص آنان فراهم آوردند. برای مثال، آنها آزمونهای واژگان و خواندن به کودکان دادند، آزمون‌ها را جمع‌آوری کردند و بدون آن که کودکان

* - این مقیاس اسمی است. هر عدد یک طبقه‌بندی را نشان می‌دهد، یعنی نمایانگر یک نوع رویداد ارتباطی است. نوشتن اعداد با این ترتیب، به معنای ارزشگذاری هر رویداد ارتباطی نیست، بلکه این کار صرفاً به منظور طبقه‌بندی کردن این رویدادها صورت می‌گیرد.

Source: From N.A. Flanders (1970). *Analyzing Teaching Behavior* (p. 34). Boston: Addison - Wesley.

1- Hartshorne, H., M. A. May, and F.K. Shuttlesworth. *Studies in the Organization of Character*. New York: Macmillan.

آگاه باشند از پاسخهای آنان رونوشت برداشتند. سپس کلید پاسخها را به کودکان دادند و از آنها خواستند که برگه‌های اصلی آزمون خود را تصحیح کنند و نمره بدهند. تفاوت بین نمره‌هایی که کودکان گزارش دادند و نمره‌های واقعی که از تصحیح رونوشت پاسخها به دست آمد، میزانی برای تقلب به دست داد.

در آزمون دیگر، از کودکان خواسته می‌شد در حالی که چشمان خود را بسته‌اند، در داخل هر یک از ده دایره کوچکی که به طور نامنظم بر برگه آزمون ترسیم شده‌اند، علامتی بگذارند. آزمونه‌های کنترل شده قبلی در شرایطی که از زیرچشمی نگاه کردن کودکان جلوگیری می‌شد، نشان داد که در سه بار آزمون، کسب نمره بیش از ۱۳، یعنی ۱۳ علامتی که به درستی در داخل دایره‌ها جای داده شده باشند، بسیار نامحتمل است. بنابراین، نمره بالاتر از ۱۳ به عنوان نشانه‌ای از زیر چشمی نگاه کردن کودکان ثبت شد.

هارتشورن و می، به طور عملی هیچ همبستگی بین تقلب در موقعیت‌های متفاوت، نظیر تقلب در آزمون و در مسابقات ورزشی، نیافتند. آنها نتیجه گرفتند که پاسخهای کودکان از نظر موقعیتی، خاص هستند. به این معنا که، اعم از این که کودکان بسته به فعالیت خاص، حضور معلم و دیگر موقعیتها تقلب کردند، اما در صفات شخصیتی عمومی، نشانی از تقلب نداشتند.

/// خلاصه

یک وظیفه مهم محققان علوم رفتاری، انتخاب ابزار اندازه‌گیری قابل اعتماد به منظور کمی کردن اطلاعات تحقیق است. در تحقیق آموزشی، آزمونها رایجترین ابزار اندازه‌گیری هستند. یک آزمون، عبارت است از مجموعه‌ای از محرکها که به یک فرد ارائه می‌شود تا پاسخی را از او بگیرد که بر مبنای آن بتوان یک نمره عددی تعیین کرد. آزمونه‌های پیشرفت تحصیلی، نمونه اصلی این نوع ابزار اندازه‌گیری هستند. آزمونه‌های پیشرفت تحصیلی متعددی وجود دارند که هنجارهایی را فراهم می‌آورند که می‌توان از آن هنجارها به عنوان مبنای مقایسه استفاده کرد. آزمونه‌های هوش ابزار سنجش قابلیت‌های کلامی و غیر کلامی یک فرد هستند. پرسشنامه‌های شخصیت برای اندازه‌گیری ویژگیهای فردی و عملکرد نوعی آزمودنی طراحی شده‌اند. مقیاسهای نگرش، ابزاری برای اندازه‌گیری باورها، احساسها و واکنشهای فرد نسبت به موضوعات معین هستند. انواع اصلی مقیاسهای نگرش عبارتند از: مقیاسهای نوع لیکرت، مقیاسهای ترستون،

مقیاسهای گاتمن و افتراق معنایی. روشهای گروه سنجی، ابزار سنجش موقعیت فرد در میان همسالان خود می‌باشند. از طریق این روشها، یافتن افراد محبوب گروهها (ستاره‌ها)، منزویها و گروهکها امکان‌پذیر می‌شود.

تعدادی از روشهایی که رفتار آزمودنیها را به صورت نظام‌مند مورد مشاهده قرار می‌دهند، به عنوان روشهای مشاهده مستقیم و مشاهده ابداعی شکل گرفته‌اند.

// مفاهیم اصلی

آزمون استاندارد شده	آزمون
آزمون پیشرفت تحصیلی	آزمون استعداد تحصیلی
اثر هاله‌ای	آزمون هوش
خطای سخاوتمندی	پرسشنامه
خطای گرایش مرکزی	خطای سختگیری
مشاهده ابداعی	روش فرافکن
مقیاس افتراق معنایی	مشاهده مستقیم
مقیاس ترستون	مقیاس تراکمی
مقیاس گاتمن	مقیاس
مقیاس مجموع درجه‌ها	مقیاس لیکرت
نگاره گروهی	مقیاس نگرش

/// تمرینها

- ۱- معنای اصطلاح استاندارد شده، هنگامی که به ابزارهای اندازه‌گیری اطلاق می‌شود، چیست؟
- ۲- تفاوت بین مقیاسهای درجه‌بندی مقایسه‌ای و مقیاسهای نگاره‌ای و طبقه‌ای چیست؟
- ۳- برخی از منابع رایج ارببی در مقیاسهای درجه‌بندی را فهرست کنید.
- ۴- چرا مقیاس گاتمن، در مقایسه با مقیاسهای ترستون و لیکرت، تک‌بعدی محسوب می‌شود؟

۵- استفاده اصلی مقیاس افتراق معنایی در اندازه‌گیری چیست؟

۶- برخی از شیوه‌های افزایش دقت در روشهای مشاهده مستقیم چه هستند؟

۷- یک مقیاس لیکرت، با پنج عبارت، برای اندازه‌گیری نگرش معلمان در مورد آموزش زبان خارجی در مدارس ابتدایی، بسازید.

- ۸- مزیت‌های اصلی مقیاس لیکرت نسبت مقیاس ترستون چیست؟
- ۹- یک مقیاس درجه‌بندی نگاره‌ای با پنج عبارت بسازید که برای ارزشیابی یک برنامه تحقیقاتی مفید باشد.
- ۱۰- داده‌های زیر، درجه‌بندی‌هایی است که از سوی ۲۰۰ داور به یک عبارت مقیاس ترستون تخصیص یافته است. چه ارزش مقیاسی به این عبارت تعلق خواهد گرفت؟

تعداد داورها	طبقه‌بندی
۰	۱۱
۰	۱۰
۰	۹
۳۰	۸
۵۰	۷
۶۰	۶
۳۴	۵
۱۶	۴
۱۰	۳
۰	۲
۰	۱
۲۰۰	

- ۱۱- اعداد زیر، ارزشهای مقیاسی پنج عبارت یک مقیاس ترستون هستند که دو دانشجو آنها را علامت زده‌اند. نمره‌های نگرش آنها را محاسبه کنید و نمره‌های آنها را برحسب این که نگرشهای آن دو را تا چه حد مطلوب و روشن تعریف می‌کنند، تفسیر کنید.

جان	جین
۱۰/۵	۵/۵
۸/۲	۴/۶
۷/۱	۴/۱
۳/۹	۳/۸
۲/۸	۳/۱

/// پاسخها

۱- استاندارد شده به پرسشنامه‌ای گفته می‌شود که برای آن هنجارهای تطبیقی به دست آمده باشد، پایایی و اعتبار آنها تعیین شده باشد و دستورالعملهایی برای اجرا و نمره دهی آنها مقرر شده باشد.

۲- در مقیاسهای نگاره‌ای و طبقه‌ای، درجه بندی کننده به هنگام داوری بر روی یک فرد، او را با افراد دیگر مقایسه مستقیم نمی‌کند. در مقیاس مقایسه‌ای، به هنگام داوری بر روی یک فرد، درجه بندی کننده باید گروهی را که آن فرد با آن مقایسه می‌شود، بشناسد.

۳- هنگامی که گرایشهایی نظیر اثر هاله‌ای، خطای سخاوتمندی، خطای سختگیری یا خطای گرایش مرکزی تأثیر خود را بر فرد درجه بندی کننده گذارده باشند، داوری او از عینیت کمتری برخوردار خواهد بود.

۴- مقیاسهای ترستون و لیکرت ممکن است حاوی عباراتی در مورد ابعاد مختلف موضوع نگرش باشند. مقیاس گاتمن تک بعدی محسوب می‌شود؛ زیرا سؤالاتی درباره درجات واکنش فرد نسبت به یک موضوع منفرد را مطرح می‌کند. بنابراین، مقیاس گاتمن تراکمی است، یعنی می‌توان پیش‌بینی کرد که اگر فردی با یک عبارت مقیاس موافقت کند، با تمام عبارات بعدی نیز موافقت خواهد کرد.

۵- مقیاس افتراق معنایی برای اندازه‌گیری معنای ضمنی که فرد برای یک شیء نگرشی قایل می‌شود، به کار می‌رود.

۶- رفتارهایی که قرار است مشاهده شوند، باید مشخص گردند، رفتارهایی که در یک طبقه قرار می‌گیرند، باید تعریف شوند، برای کمّی کردن نتایج مشاهده باید سیستمی تهیه شود و مشاهده‌گران باید آموزش ببینند تا کار مشاهده را مطابق با شیوه‌های وضع شده به انجام رسانند.

۹- پاسخها متفاوت خواهند بود.

۱۰- ۶/۱۷

۱۱- جین با میانگین $4/2$ و میانه $4/1$ ، یک نگرش نسبتاً مثبت را نشان می‌دهد. انتخابهای او از نظر ارزش مقیاسی به هم نزدیک هستند، که این نکته، همسانی را نشان می‌دهد. میانگین جان $6/5$ و میانه او $7/1$ است، که نشان‌دهنده نگرش کمی نامطلوب اوست. پراکندگی زیاد ارزشهای مقیاسی، نگرشی را که به طور روشن تعریف نشده است، نمایان می‌کند.

