مدرسان شریف

فصل اول

«آمار»

آمار

آمار، علم توسعه دانش انسان از طریق استفاده از داده‌های تجربی است که بر نظریه‌ی آمار مبتنی است و شاخه‌ای از ریاضیات کاربردی به‌شمار می‌آید. در نظریه‌ی آمار، اتفاقات تصادفی و عدم قطعیت با نظریه احتمال، مدل می‌شود. نخستین جرقه‌های عام آمار را در شمارش‌های آماری باید در آغاز قرن اول میلادی یافت. اما در قرن هجدهم است که این علم، با به کار گرفته شدن در توصیف جنبه‌هایی که شرایط یک وضعیت را مشخص می‌کرد، به عنوان رشته‌ای علمی و مستقل شروع به مطرح شدن کرد. آمار تعاریف متعددی دارد که در زیر بیان شده است:

* **آمار** عبارت است از یک متغیر تصادفی که به نمونه ملاحظه شده بستگی دارد.
* **آمار مجموعه‌ای** از روش‌های جمع‌آوری، تهیه و تنظیم و تجزیه و تحلیل اطلاعات است که برای کسب یک یا چند نتیجه استفاده می‌شود.
* **آمار، علمی** است که به جمع‌آوری، دسته‌بندی، خلاصه‌سازی، تعبیر و تفسیر و نهایتاً برآورد در مورد داده‌ها می‌پردازد.

روش‌های آماری دو وظیفه دارند:

۱ـ به پژوهشگر در طبقه‌بندی، خلاصه‌سازی، توصیف و تفسیر اطلاعات جمع‌آوری شده و برقراری ارتباط از طریق آن‌ها کمک می‌کنند.

۲ـ به پژوهشگران امکان می‌دهند که با استفاده از داده‌های جمع آوری شده از نمونه کوچکی از آزمودنی‌ها، ویژگی‌های جامعه‌ای را که نمونه از آن انتخاب شده است، برآورد یا استنباط کنند.

وظیفه اول آمار را روش‌های آمار توصیفی (Descriptive Statistics) و وظیفه دوم را روش‌های آمار استنباطی (Inferential Statistics) می‌نامند.

آمار تحقیقی ـ آمار کاربردی:

در محافل ورزشی متخصصان دربارهٔ تحقیق آماری و غیرآماری بحث می‌کنند و اغلب این قبیل مباحث صرفاً تخلف این گونه افراد را در به کار بردن آمار می‌رساند. تحقیق آماری به ندرت در مرحله اجرا گذاشته شده است. این‌گونه مطالعات بیشتر موضوع اصلی علم آمار را تشکیل می‌دهند. **در حقیقت، تحقیق آماری به ابداع، توسعه و پیشرفت و به کار بستن شیوه‌ها برای توصیف داده‌ها و استنتاج‌های بعدی مربوط است.** میان تحقیق آماری و کاربرد آمار تفاوت کلی وجود دارد و این آمار کاربردی است که غالباً کارشناسان تربیت بدنی با آن سر و کار دارند. معلمان تربیت بدنی فقط در موارد بسیار نادر تحقیق آماری انجام می‌دهند. آنها باید این آمادگی را داشته باشند تا از این روش‌ها در مطالعات خود بهره گیرند. به واسطهٔ داشتن آگاهی از قوانین احتمالی است که می‌توان نامعلومی استنتاج‌های استقرایی را ارزیابی کنیم. توجه داشته باشید که آمار تحقیقی با تحقیق آماری متفاوت است.

مثال۱: کدام‌یک درباره آمار صحیح می‌باشد؟

۱) روش‌های جمع‌آوری، دسته‌بندی، تهیه گزارش و مستندسازی است.
۲) یک متغیر غیرتصادفی است که بستگی به جامعه دارد.
۳) علم جمع‌آوری، دسته‌بندی، خلاصه‌سازی، تعبیر و تفسیری است.
۴) آمار علمی است‌که به رابطه‌ی علت و معلول پدیده‌ها می‌پردازد.

☑ **پاسخ:** گزینه «۳» آمار علمی است که به جمع‌آوری، دسته‌بندی، خلاصه‌سازی، تعبیر و تفسیر و نهایتاً برآورد در مورد داده‌ها می‌پردازد.

ضرورت استفاده از آمار در تربیت‌بدنی

علی‌رغم این‌که آمار در مسائل روزمره زندگی کاربردهای گوناگونی دارد، هر معلم حرفه‌ای نیز باید به دو دلیل به اطلاعات آماری روی آورد. اول این‌که اگر قرار است معلمی پژوهش‌های تربیتی را از روی بصیرت و منتقدانه، مطالعه و ارزشیابی نماید، به معلومات آماری نیاز دارد و دلیل دوم اهمیت دانستنی‌های آماری برای یک معلم این است که اگر خود معلم قرار باشد، پژوهش‌ها و تحقیقاتی را انجام دهد، درک روش‌های آماری برای این منظور ضروری است. برای

مثال، یک مربی ورزشی برای این‌که بتواند از میان چند نوع روش برای پرورش دانش‌آموزان خود، روشی را انتخاب کند که بهترین برداشت را داشته باشد نیاز دارد که نتایج آن روش‌ها را با هم بررسی کرده و بهترین نتیجه را انتخاب کند و علم آمار به او این امکان را می‌دهد که بتواند به راحتی چند روش را از نظرهای مختلف با هم مقایسه کند و از میان آن‌ها مفیدترین روش را به‌کار گیرد که در کم‌ترین زمان ممکن بهترین بازخورد را به او بدهد.

مثال۲: کمک به طبقه‌بندی، خلاصه‌سازی، توصیف و تفسیر اطلاعات جمع آوری شده از وظایف کدام آمار است؟

۱) آمار تحقیقی ۲) آمار تطبیقی ۳) آمار توصیفی ۴) آمار استنباطی

☑ **پاسخ:** گزینه «۳» آمار توصیفی به پژوهشگر در طبقه‌بندی، خلاصه‌سازی، توصیف و تفسیر اطلاعات جمع آوری شده کمک می‌کند.

اهداف سنجش و اندازه‌گیری

۱) شناسایی نقاط قوت و ضعف

الف) برای کلاس درس: در یک کلاس درس هنگامی که بخواهیم در ابتدای آموزش سطح و شدت برنامه آموزشی یا تجربیات قبلی را تعیین کنیم یا میزان پیشرفت کلاس را در مسیر هدف‌های آموزشی ارزشیابی کنیم.

ب) برای فراگیر: هنگامی که بخواهیم وضعیت شخصی شاگرد را از نظر آموزشی یا تمرین‌های داده شده ارزشیابی کنیم و بر پایه آن برنامه‌های جبرانی، فوق‌العاده و ویژه تجویز نماییم.

۲) طبقه‌بندی افراد در گروه‌های متجانس و همتا

الف) انتخاب و اجرای برنامه آموزشی مناسب برای هر گروه: برای مثال، دبیر تربیت‌بدنی بخواهد از بین چند نوع برنامه آموزشی و چند گروه از دانش‌آموزان که هر کدام توانایی مخصوص به یک رشته خاص را دارند، بهترین برنامه‌ها را انتخاب کرده و به‌طور شایسته روی هر گروه اجرا کند.

ب) انتخاب و اجرای نورم‌گیری روی گروه‌های متجانس و همتا: به‌عنوان مثال، اگر دبیر بخواهد یک نورم‌گیری دقیق روی گروه‌های متجانس انجام دهد، از این روش استفاده می‌کند.

۳) تعیین معافیت افراد از بعضی تجربیات

الف) به منظور همکاری با پزشک برای معافیت‌های پزشکی در تربیت‌بدنی و ورزش.

ب) برخی از شاگردان از انجام قسمتی از برنامه‌های آموزشی معاف می‌شوند که در آن از میانگین سطح کلاس، تبحر و آمادگی بیشتری دارند. برای مثال، اگر فراگیرنده‌ای در دوره‌ی آموزشی که برای وی ترتیب داده‌اند قبلاً شرکت کرده یا مهارت ویژه‌ای را زودتر از دیگر فراگیرنده‌ها کسب نموده و به سطح بالایی از فعالیت برسد، می‌تواند تا زمانی که سطح کل کلاس به او برسد، از انجام دادن آن فعالیت معاف باشد.

ج) به منظور تعیین قبولی در یک قسمت از برنامه و ارتقاء به قسمت‌های بعدی در یک برنامه آموزشی منظم. اگر دبیر برنامه‌ای آموزشی را ترتیب دهد که در آن برای رفتن از یک مرحله به مرحله بعد آزمونی تهیه کند و سطح قبولی برای هر مرحله در نظر بگیرد، فراگیرنده با قبولی در یک مرحله و ارتقاء به مرحله‌ی بعدی از انجام فعالیت‌های مرحله‌ی اول معاف می‌شود.

مثال۳: انتخاب و اجرای برنامه آموزشی برای هر گروه جزو کدام‌یک از اهداف سنجش و اندازه‌گیری می‌باشد؟

۱) شناسایی نقاط قوت و ضعف ۲) طبقه‌بندی افراد در گروه‌های همتا

۳) تعیین معافیت افراد از بعضی تجربیات ۴) راهنمایی آموزشی فراگیرنده

☑ **پاسخ:** گزینه «۲» انتخاب و اجرای برنامه آموزشی جزو طبقه‌بندی افراد در گروه‌های متجانس و همتا می‌باشد.

۴) راهنمای آموزشی فراگیرنده

الف) برای این‌که بتوانند نیرو و زمان خود را نسبت به میزان آگاهی‌ها و اطلاعاتی که دارند روی قسمت‌های مختلف برنامه تنظیم و تقسیم کنند.

ب) برای شناسایی استعدادهای درخشان در بعضی زمینه‌ها و ترغیب آن‌ها به فعالیت‌های جدی‌تر.

ج) برای انتخاب رشته‌های ورزشی، تحصیلی یا شخصی در آینده.

۵) راهنمای دبیر

الف) در برآورد و ارزشیابی روش تدریس خود **ب)** در ارزشیابی و برآورد برنامه آموزشی **ج)** در برآورد و ارزشیابی وسایل کمک آموزشی و شرایط محیطی دیگر **د)** در برآورد و ارزشیابی میزان پیشرفت شاگردان به‌طور عموم

۶) ایجاد انگیزه و رغبت

الف) برای دبیر به منظور ارائه کار بهتر **ب)** برای ایجاد رقابت تحصیلی بین فراگیرنده‌ها و ترغیب آن‌ها برای انجام فعالیت بیشتر **ج)** برای والدین و مسئولین به‌منظور توجه بیشتر به وضعیت موجود و نیازها

۷) تشخیص و تعیین نمرهٔ تحصیلی

الف) برای تعیین مردودی یا قبولی در درس **ب)** برای تدوین کارنامه‌های ورزشی و آموزشی **ج)** برای تهیه گزارش‌های مستند آموزشی و اداری

مثال۴: کدام‌یک جزو اهداف سنجش و اندازه‌گیری نمی‌باشد؟

۱) شناسایی نقاط قوت و ضعف ۲) راهنمای آموزشی دبیر

۳) ایجاد انگیزه و رغبت ۴) تشخیص و تعیین نمره‌ی تحصیلی

☑ **پاسخ:** گزینه «۲» یکی از اهداف سنجش و اندازه‌گیری راهنمای آموزشی فراگیرنده است نه دبیر.

مثال ۵: تشخیص و تعیین نمره تحصیلی در چه موردی استفاده می‌شود؟

۱) تعیین مردودی یا قبولی درس ۲) انتخاب برنامه آموزشی مناسب برای هر گروه

۳) تهیه گزارش‌های مستند آموزشی و اداری ۴) گزینه ۱ و ۳

☑ **پاسخ:** گزینه «۴» هر دو گزینه ۱و۳ مربوط به تشخیص و تعیین نمره تحصیلی می‌باشند؛ ولی گزینه ۲ مربوط به طبقه‌بندی افراد در گروه‌های متجانس و همتاست.

تحلیل داده‌ها

آمار شاخه‌ای از ریاضیات است که اندازه‌گیری‌های کمّی و ارزیابی ما برای رسیدن به تصمیمی در مورد معنی‌دار بودن فرضیه را تحلیل و تفسیر می‌کند. **در تنظیم موضوع تحقیق، محقق نباید فقط درباره روش گردآوری داده‌ها تصمیم بگیرد، بلکه باید روش‌های گوناگون تحلیل این داده‌ها را که به باثبات‌ترین پاسخ‌ها منتهی می‌شود، نیز تعیین کند.** اغلب و پس از اینکه داده‌ها گردآوری شدند، محقق تصمیم می‌گیرد که از چه روش‌هایی برای تحلیل داده‌های خود استفاده کند. این روش‌ها معمولاً روش‌هایی هستند که فرضیه‌های آماری محقق را تأیید می‌کنند یا اینکه حداقل نتایج آماری را در جهت معنی‌داری سوق می‌دهند. در حقیقت، اگر قرار باشد دانش تازه‌ای در زمینهٔ تربیت بدنی یا هر رشتهٔ دیگری شکل گیرد، به کار بردن روش‌های علمی که مبتنی بر تحلیل متناسب داده‌ها باشد از ضروریات است. آگاهی از به‌کار بردن آزمون‌های آماری و شیوه‌های اجرایی برای دسترسی به این اهداف ضروری است.

تفسیر نتایج آزمایش

معمولاً فقط افرادی که دانش لازم در رشته مورد تحقیق را دارند در مقامی هستند که می‌توانند تصمیم نهایی در انتخاب داده‌ها و تفسیر نتایج اتخاذ کنند. هر فرد باید کاملاً مسئولیت تحلیل و تفسیر داده‌های تحقیق خود را بر عهده گیرد. بنابراین محقق تربیت بدنی باید کوشش کند تا روش‌های صحیح آماری و منطقی را که اساس این روش‌ها بر آن استوار است، فراگیرد. این آگاهی بدون تردید برای اجرای تحقیق خود و تفسیر نتایج پژوهش‌های دیگران از ضروریات است.

آیا می‌توان به آمار اعتماد کرد؟

با وجود اینکه امکان دارد حقایق نادرست بسیاری از طریق آمار گفته شود، **در بیشتر موارد عدم آگاهی موجود، سوء استفاده از روش‌های آماری را فراهم می‌کند.** اگر فرض کنیم، محقق فردی درستکار است، آنگاه سوء استفادهٔ غیر عمدی را می‌توان تنها توضیحی برای به کار بردن روش‌های آماری دانست که به نتایج بی‌اعتبار منجر می‌شود، ولی اگر فرض ما درست نباشد، در این صورت، برای محقق بسیار آسان‌تر است که داده‌های خام را دستکاری کند، به طوری که نتیجه نهایی با فرضیه اصلی او توافق حاصل کند.

ضرورت‌های سنجش و اندازه‌گیری

انسان‌ها از هزاران سال پیش به جمع‌آوری اطلاعات و نتیجه‌گیری در مورد این اطلاعات پرداخته‌اند. تاریخچه رشته آمار توصیفی به **تاریخ پیدایش انسان** برمی‌گردد. در اوایل دوران مسیحیت، آمار برای کسب اطلاعات درباره‌ی ماهیت‌ها، جنگ‌ها، محصولات کشاورزی و حتی برای فعالیت‌های ورزشی به کار می‌رفت. در واقع قسمت عمده‌ی علم آمار کنونی محصول تحقیقات مصرانه افراد زیادی در رشته‌های مختلف در چهار صد سال گذشته است.

شاید عطش سیری‌ناپذیر افراد برای قمار بود که پیشرفت نظریه‌ی احتمالات را در مراحل اولیه آن موجب شد. در آن زمان قماربازان به منظور بالابردن میزان برخورد، از ریاضیدان‌ها برای تعیین مناسب‌ترین استراتژی در بازی‌های تابع شانس کمک می‌گرفتند. اکنون ما آمار یا همان سنجش و اندازه‌گیری را به صورت **گزارش نتایج ورزشی، شاخص‌های هزینه زندگی، پیش‌بینی هوا** و... به کار می‌بریم و تصور این‌که بدون در دست داشتن اشکال مختلف اطلاعات آماری بتوانیم به راحتی زندگی کنیم اندکی مشکل به‌نظر می‌رسد.

تاریخچه سنجش و اندازه‌گیری در تربیت‌بدنی

با مطالعه تاریخ آمار در تربیت‌بدنی می‌توان به روش‌های تحقیق گذشتگان در خصوص سنجش در تربیت‌بدنی و هم‌چنین اطلاعاتی درباره‌ی پیشرفت تحقیق در رشته‌ی تربیت‌بدنی در خصوص اندازه‌گیری فاکتورهای آمادگی جسمانی دست یافت. در اواخر دهه‌ی ۱۸۸۰ بود که توجه متخصصان به اندازه‌گیری در تربیت‌بدنی جلب شد و کم‌کم اندازه‌گیری‌های کمی و عینی از وضعیت شاگردان به رشته‌های مختلف راه پیدا کرده و این روش‌ها به سرعت گسترش پیدا کردند. برطبق یافته‌های ون دالن، میشل و بنت روش‌های ارزیابی تربیت‌بدنی از سال ۱۹۲۰ با آزمون‌ها ، روش‌های اندازه‌گیری و ارزشیابی و ابزار اندازه‌گیری، وسیع و پیشرفته‌تر گردید.

به‌طور کلی می‌توان تکامل سنجش و اندازه‌گیری در تربیت‌بدنی را در چندزمینه بررسی کرد که عبارتند از:

۱) اندازه‌گیری شکل ظاهری اجزای بدن:

بین سال‌های ۱۸۸۵ تا ۱۹۹۰ اندازه‌گیری اجزای بدن در تربیت‌بدنی شکوفا گردید. اندازه‌گیری‌های مختلف از قبیل گنجایش شُش‌ها، هم‌چنین قطر و طول اجزای بدن به‌ثبت رسیده است و نیز مقالاتی توسط افرادی به چاپ رسیده است که می‌توان از جزوه‌ای که درباره‌ی اندازه‌های ورزشکاران برجسته و نورم‌های تعیین‌کننده اندازه‌های بدنی زنان و مردان دانشگاهی بود، یاد کرد.

۲) آزمون‌های قدرت بدنی:

در سال ۱۶۹۹ دولاهایر قدرت مردان را در بلند کردن و حمل کردن وزنه‌های سنگین با قدرت اسب‌ها در این مورد مقایسه کرد که از آن پس اندازه‌گیری قدرت و رابطه‌اش با سلامتی عمومی بدن و آمادگی‌های ورزشی، فکر انسان را به خود مشغول کرده است.

اشخاص دیگری نیز در این رابطه فعالیت‌های بسیاری انجام داده‌اند. به طور مثال در سال ۱۹۲۵، مسئله قدرت عضلانی و آزمون‌های آن را به چاپ رساندند یا کراس در سال ۱۹۴۵ تعدادی آزمون برای تعیین حداقل آمادگی‌های عضلانی را همراه با توضیحات اجرایی لازم ارائه داد.

۳) آزمون‌های پیشرفت حرکتی مهارت‌های پایه:

این آزمون‌ها شامل دویدن، پرتاب کردن و پریدن می‌باشد که در مدارس ابتدایی و متوسطه اندازه‌گیری می‌شوند. این آزمون به‌وسیله‌ی گولیک از سال‌های ۱۹۰۴به بعد در آمریکا تنظیم و پیشنهاد گردید.

مواد این آزمون برای پسران عبارت است از: کشش از بارفیکس، پرش ارتفاع، دویدن، دوهای سرعت کوتاه، پرش ارتفاع جفتی، حرکت شنا روی زمین، پرتاب وزنه، بالا رفتن از طناب.

مواد این آزمون برای دختران عبارت است از: پرتاب توپ بسکتبال، پرش، دوهای سرعت کوتاه.

همچنین دکتر سارجنت آزمون جسمانی برای یک مرد را معرفی کرد. در این آزمون قدرت انفجاری پاها مورد سنجش قرار می‌گرفت که بیشتر از همه، توانایی‌های انسان را در رشته‌های دومیدانی نشان می‌دهد که این آزمون به‌نام «پرش سارجنت» معروف است. افرادی در سال ۱۹۲۰ گردهم آمدند و تحقیقاتی را انجام دادند و این نتیجه به‌دست آمد:

متداول‌ترین آزمون‌ها برای پسران عبارتند از: بالارفتن از طناب، پرش طول جفتی، دویدن، پرش ارتفاع، شنا روی زمین، دوهای سرعت کوتاه، پرتاب وزنه، کشش از بارفیکس. **این آزمون‌ها برای دختران عبارتند از:** بالا رفتن از طناب، پرش، پرتاب توپ بسکتبال، دوهای کوتاه سرعتی.

۴) آزمون‌های مهارت ورزشی:

«برایس» در سال ۱۹۲۴ برای ساختن نورم یا استانداردهایی برای ارزشیابی دختران در یک آزمون بسکتبال از مقیاس به‌نام تی (T) استفاده کرد. مدتی بعد سازمان ایفرد کار ساختن آزمون برای ۱۵ رشته ورزشی را در دست گرفت. هدف‌های این آزمون عبارتند از:

تدریس و تعلیم بهتر فراگیران و ایجاد انگیزه برای پیشرفت آنان و ارزیابی بهتر شاگردان در این ۱۵ رشته.

۵) آزمون‌های هماهنگی شخصی و اجتماعی: در سال ۱۹۲۸ لوتر وان باسکیرک آزمونی را ارائه داد که حاوی تعیین اخلاق و رفتار در تربیت‌بدنی می‌باشد کار او را افراد دیگری نیز دنبال کردند.

۶) طبقه‌بندی افراد در گروه‌های متجانس:

در سال ۱۹۱۷ رایلی به عنوان پیشگام و بعد از او، مک کلوی و سارجنت و سپس برایس آزمون‌هایی متناسب با جنس فراگیران، قد، قدرت، وزن و سن ساختند که این آزمون برای طبقه‌بندی افراد در گروه‌های هم قوّه (متجانس) به‌کار می‌رفت.

۷) آزمون‌های آمادگی قلبی و عروقی:

در سال ۱۸۹۲ به‌وسیله‌ی ابداع دستگاه ارگوگراف اندازه‌گیری و سنجش آمادگی‌های قلبی و عروقی پایه‌گذاری شد که وسیله‌ای برای اندازه‌گیری کار انجام شده می‌باشد.

۸) اطلاعات و آگاهی‌های عملی و ورزشی:

در سال ۱۹۲۵، برایس یک آزمون اطلاعات ورزشی در ارتباط با ورزش بسکتبال به چاپ رسانید. بیل نیز برای ورزش تنیس این کار را انجام داد.

۹) آزمون‌های آمادگی حرکتی:

در سال ۱۹۴۱ برای هر یک از شاخه‌های تخصصی ارتش تعدادی از این آزمون‌ها از قبیل کشش از بارفیکس، حرکت شنا روی زمین، دویدن و درازنشست ساخته شد. آزمون‌هایی نیز برای اندازه‌گیری وضعیت بدنی (نیروها) بر پایهٔ روش‌های معتبر آماری تهیه گردید. معروف‌ترین انواع آزمون‌های مرکب به‌نام **آزمون ایفرد** است.

مثال ۶: به طور کلی، تکامل سنجش و اندازه‌گیری را در کدام زمینه می‌توان بررسی کرد؟

۱) آزمون‌های پایه پیشرفت‌های ورزشی ۲) آزمون‌های قدرت بدنی

۳) اندازه‌گیری شکل ظاهری اجزای بدن ۴) تمامی موارد

☑ **پاسخ:** گزینه «۴» هر سه گزینه از موارد بررسی تکامل سنجش و اندازه‌گیری است.

اندازه‌گیری و ارزشیابی

ارزشیابی یعنی دادن مفهوم و محتوا به داده‌هایی که از اندازه‌گیری حاصل شده است و مقصود از اندازه‌گیری گردآوری داده‌ها می‌باشد. هرچه داده‌های گردآوری شده بهتر باشد، موفقیت در ارزشیابی بالاتر است. فرآیند اندازه‌گیری نخستین قدم در ارزشیابی است، یعنی اندازه‌گیری پیشرفته، ما را به ارزشیابی دقیق سوق می‌دهد.

برای لغت اندازه‌گیری تعاریف زیادی قابل بیان می‌باشد که از میان آن‌ها که درباره‌ی ماهیت اندازه‌گیری نوشته شده است، می‌توان تعریف زیر را ارائه کرد.

اندازه‌گیری قاعده‌ای است برای این‌که اعداد را به اشیاء، افراد یا رویدادها اختصاص دهد. این کار را به منظور کمّی‌سازی خصیصه‌های آن‌ها انجام می‌دهیم. قواعدی که به‌منظور کمّی‌سازی صفات اشیاء به‌کار می‌رود، در قلب اندازه‌گیری قرار دارد و برای رسیدن به اهداف اندازه‌گیری باید روش، بدون ابهام باشد. فرآیند اندازه‌گیری با ویژگی‌های خاصی است که با فرد یا شیئی ارتباط دارد. در واقع با صفتی از شیء یا فرد ارتباط دارد و هرگز نمی‌توان خود فرد را اندازه گرفت ولی می توان قد، وزن و غیره را اندازه‌گیری کرد. در این صفات احتمال تغییر وجود دارد.

افراد مختلف تعاریف متعددی برای ارزشیابی ارائه داده‌اند که از بین آنها می‌توان به موارد زیر اشاره کرد. کرونباخ ارزشیابی را گردآوری و کاربرد اطلاعات به منظور تصمیم‌گیری در برنامه‌آموزشی می‌داند و بی‌بای آن را فرآیند نظام‌دار گردآوری و تفسیر شواهدی که به داوری منجر می‌شود و در آخر عمل را در پی دارد، تعریف می‌کند. اما در یک تعریف جامع می‌توان آن را **محتوا دادن به داده‌هایی خامی دانست که به‌وسیله‌ی اندازه‌گیری به‌دست می‌آید** یا این‌که سنجش عملکرد فراگیرندگان و مقایسه نتایج به‌دست آمده با هدف‌های آموزشی از پیش تعیین شده است.

مثال۷: مقصود از اندازه‌گیری ..... می‌باشد.

۱) تفسیر داده‌ها ۲) سنجش و مقایسه ۳) دادن محتوا به اطلاعات ۴) گردآوری داده‌ها

☑ **پاسخ:** گزینه «۴» مقصود از اندازه‌گیری، گردآوری داده‌ها است.

تفاوف بین اندازه‌گیری و ارزشیابی

فرآیند ارزشیابی بهتر و کامل‌تر از فرآیند اندازه‌گیری است؛ به‌طوری که علاوه‌بر جنبه‌های کمّی، جنبه‌های کیفی رفتار را نیز شامل می‌شود. در فرآیند اندازه‌گیری از این جهت که در مورد داده‌های جمع‌آوری شده داوری نمی‌کند، در واقع یک فرآیند غیرارزشی به‌شمار می‌آید. اما در فرآیند ارزشیابی داوری درباره‌ی کارآمدی یک برنامه برای گروه بخصوص از افراد هدایت می‌شود، در حالی که اندازه‌گیری ضمن آن‌که پایه و اساس ارزشیابی درست و دقیق را تشکیل می‌دهد، با واقعیت سر و کار دارد و هرگز شامل داوری ارزشی نیست. در واقع این دو فرآیند در نهایت مکمل یکدیگر بوده‌اند.

مدرسان شریف

فصل دوم

«مفاهیم اندازه‌گیری»

منابع اطلاعاتی مرتبط با رشته تربیت بدنی که با علوم رفتاری و طبیعی هم در ارتباط هستند، به شاخه‌های گوناگون و بسیاری تقسیم‌بندی می‌شوند. برخی از این منابع به بررسی هوش، مهارت‌های عددی و لغوی، مهارت‌های جسمانی، بهداشت، شخصیت، ارزش‌ها، نگرش‌ها و باورها، پیشینه خانوادگی و شرایط اجتماعی، رشد هیجانی و عضوی، پیشرفت تحصیلی و تمایلات در آدمی می‌پردازند.

جامعه، پارامتر، نمونه و آماره

جامعه:

به مجموعهٔ افراد یا اشیایی که می‌خواهیم در مورد یک یا چند خصوصیت مشترک آنها مطالعه و بررسی کنیم، جامعه یا جمعیت آماری گفته می‌شود. جامعه در معنای گسترده آن چیزی فراتر از مجموعه‌ای از افراد است. به عنوان مثال، اگر در هنرستان تربیت بدنی تعداد ۶۰۰ دانش‌آموز وجود داشته باشد که طبق گروه خونی تقسیم‌بندی شده باشند، در آن صورت خواهیم گفت که اندازه‌ی جمعیت ما ۶۰۰نفر است. تجربه تاس‌اندازی، تولید یک جمعیت با اندازه‌ی محدود است. قد شهروندان یک شهر معین، تعداد ماهی‌های یک دریاچه بخصوص، از جمله مثال‌های با اندازه محدود هستند. در هریک از حالات فوق، کل ملاحظات تعداد محدودی است. ملاحظاتی چون به‌دست‌آوردن فشار جو روزانه یک منطقه از گذشته تاکنون و آینده یا تمام اندازه‌گیری‌هایی که بر روی عمق یک دریاچه از یک نقطه فرضی به‌عمل می‌آید، از مثال‌های جمعیت با اندازه نامحدود است. بعضی از جمعیت‌های محدود به قدری وسیع هستند که ما آنها را نامحدود فرض می‌کنیم. مثلاً، اگر جمعیت عمر تمام باطری‌های انبار شده از نوع معینی را که برای توزیع در سرتاسر کشور تولید شده‌اند، در نظر بگیرید، مسئله فوق مصداق پیدا خواهد نمود. به عبارت دیگر، تعریف دیگری که می‌توان از جامعه (جمعیت) کرد، این است:

کل ملاحظاتی که با آن سروکار داریم، اعم از اینکه محدود یا نامحدود باشند، جمعیت نامیده می‌شوند.

متغیر و پارامتر:

خصوصیات مورد مطالعه، از فردی به فرد دیگر یا از شیئی به شیء دیگر در جمعیت آماری تغییر می‌کند که آن را اصطلاحاً متغیر می‌نامیم. این متغیرها را می‌توان به کمیت تبدیل کرد یا برای آنها ارزش تعیین کرد. هنگامی‌که بتوانیم به این متغیرها کمیت دهیم و برای آنها ارزش تعیین کنیم، می‌توانیم آنها را اندازه‌گیری نماییم. حال اندازه‌گیری‌هایی که ویژگی‌های یک جامعه را توصیف می‌کنند، **پارامتر** نامیده می‌شوند. مانند متغیرX که می‌تواند تعداد بازیکنان بسکتبال، تعداد خطاها، تعداد گل‌ها و غیره باشد.

مطالعه ویژگی‌های مورد نظر، موقعی که جمعیت آماری بسیار گسترده‌ای داشته باشیم، مستلزم صرف هزینه و وقت زیادی است و در بسیاری از موارد، این امر اصولاً امکان‌پذیر نیست. بنابراین، در چنین مواردی، برای مطالعه ویژگی مورد نظر، به قسمتی از جمعیت آماری بسنده می‌کنیم.

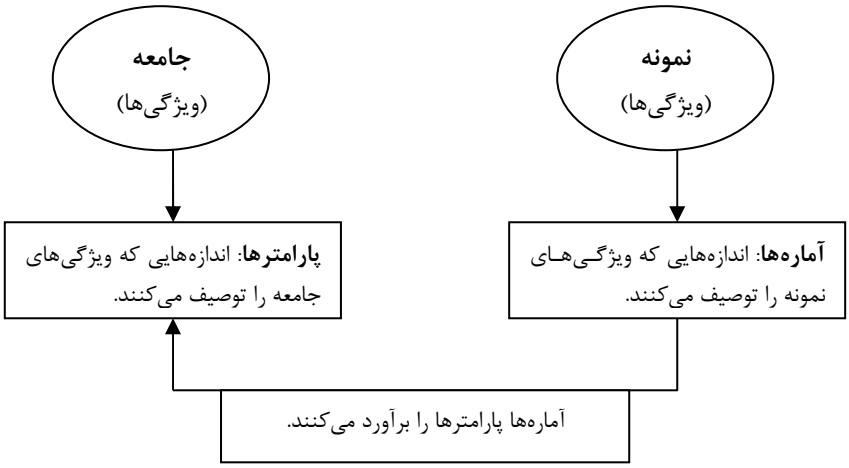
قسمتی از جمعیت را که طبق اصول و قاعده خاصی، برای مطالعه خصوصیتی از جمعیت انتخاب می‌شود، یک نمونه از جمعیت می‌نامیم. این نمونه وقتی مفید و قابل قبول خواهد بود که بتواند نماینده خوبی برای کل جمعیت مورد مطالعه باشد. با توجه به اهمیت این موضوع شاخه‌ای از آمار تحت عنوان نظریه نمونه‌گیری با بررسی نمونه‌ای به این امر مهم می‌پردازد. به عبارت دیگر، نمونه را می‌توان زیرمجموعه‌ای از کل دانست.

نکته۱: تعداد نمونه همیشه کمتر از جمعیت است.

آماره:

اندازه‌هایی که از یک نمونه گرفته می‌شود و از آنها برای **توصیف ویژگی‌های یک جامعه آماری** استفاده می‌شود، آماره نامیده می‌شود. باید توجه داشت که پارامترهای یک جامعه آماری از طریق **آماره** برآورد می‌شوند. به عنوان مثال، نسبت قد به وزن ۷۰ نفر از بازیکنان تیم بسکتبال (آماره) برای برآورد نسبت وزن کل بازیکنان (پارامتر) این تیم مورد استفاده قرار می‌گیرد.

ارتباط بین جامعه، پارامتر، نمونه، آماره در شکل زیر نشان داده شده است.



نکته۲: نسبت آماره به نمونه، مانند نسبت پارامتر به جامعه است.

مثال۱: اندازه‌گیری‌هایی که ویژگی‌های یک جامعه را توصیف می‌کنند، ..... نامیده می‌شوند.

۱) شاخص ۲) پارامتر ۳) مقیاس ۴) آماره

☑ **پاسخ:** گزینه «۲» طبق تعریف پارامتر، اندازه‌گیری‌هایی که ویژگی‌های جامعه را توصیف می‌کند، پارامتر نام دارد.

مثال۲: تعداد ..... همیشه کمتر از ..... می‌باشد.

۱) جامعه ـ نمونه ۲) پارامتر ـ آماره ۳) نمونه ـ جامعه ۴) آماره ـ پارامتر

☑ **پاسخ:** گزینه «۳» از آنجایی که نمونه در داخل جمعیت قرار دارد، همیشه نمونه کمتر از جمعیت است.

آمار توصیفی:

به‌طور کلی، روش‌هایی که به‌وسیله‌ی آنها می‌توان اطلاعات جمع‌آوری‌شده را تنظیم کرد و خلاصه نمود، آمار توصیفی می‌نامیم. **آمار توصیفی عبارت است از روش‌هایی که پردازش داده را فراهم می‌سازد.** به عبارت دیگر، آمار توصیفی به منظور سازمان‌دادن، خلاصه‌کردن و توصیف اندازه‌های نمونه به کار می‌رود و هدف آن پیش‌بینی و درک پارامتر جامعه مورد مطالعه نیست.

آمار استنباطی یا قیاسی:

این نوع از آمار برای پیش‌بینی یا برآورد پارامترهای جامعه آماری از طریق اندازه‌های نمونه به کار برده می‌شود. این عمل با استفاده از **فرآیند استدلال قیاسی و بر اساس نظریه ریاضی احتمالات** صورت می‌گیرد. پارامترهای جامعه آماری را از طریق نمونه نمی‌توان به‌صورت دقیق تعیین نمود و فقط می‌توان با احتمال آن را تعیین کرد. برای آنکه برآورد پیش‌بینی شده به واقعیت نزدیک باشد، باید روش‌های نمونه‌گیری و هم‌چنین روش‌های آماری به کار رفته ایده‌آل باشد. آمار به عنوان یک علم، عبارت است از مطالعه و بررسی برای مفهوم و معنی دادن به داده‌ها.

مثال۳: کدام نوع آمار برای پیش‌بینی پارامترهای جامعه آماری از طریق اندازه‌های نمونه به‌کار برده می‌شود؟

۱) آمار توصیفی ۲) آمار استنباطی ۳) آمار توصیفی و استنباطی ۴) هیچ‌کدام

☑ **پاسخ:** گزینه «۲» زیرا آمار استنباطی است که پیش‌بینی پارامترهای جامعه آماری را انجام می‌دهد.

وجه تمایز احتمال و آمار: فرض کنید کارخانه‌ای توپ تولید می‌کند و تولید روزانه آن حدود یک میلیون عدد است. با توجه به عکس‌العمل بعضی از مشتریان در ارتباط با کیفیت تولید این کارخانه، مسئولان کارخانه علاقمند هستند نسبت توپ‌های معیوب در یک روز معین از تولید کارخانه را تعیین کنند.

این مسئله را می‌توان به دو طریق حل کرد:

۱) یک میلیون توپ تولیدی آزمایش شوند. مسلماً اگر چنین راه‌حلی عملی هم باشد، افزایش قیمت توپ‌ها را دربرخواهد داشت. اصولاً از نظر صرف وقت و هزینه‌های مربوط به آن و غیره، این راه‌حل پیشنهادی مقرون به صرفه نیست.

مدرسان شریف

فصل سوم

« طبقه‌بندی داده‌ها»

در اکثر مسائلی که رودرروی ما قرار دارند، اندازه‌گیری یک متغیر، مستلزم آگاهی و شناخت خاصی است و داده‌های سازمان یافته و طبقه‌بندی شده را به مواردی آشنا و انعطاف‌پذیر تبدیل می‌کند که اعداد نامیده می‌شوند. در سال ۱۹۴۶ استیونز چهار نوع مقیاس برای اندازه‌گیری را معرفی کرده است که به اختصار به آنها اشاره می‌شود.

مقیاس اسمی (Nominal)

این نوع مقیاس اندازه‌گیری عمدتاً برای طبقه‌بندی داده‌ها به کار می‌رود و منظور از آن اطلاق یک عدد طبیعی به داده‌های متفاوت است. این نوع مقیاس ابتدایی‌ترین نوع مقیاس است و ممکن است به طبقه‌بندی افراد، اشیاء یا حوادث بپردازد. شایان ذکر است که این طبقه‌ها باید ناسازگار باشند و معنی این جمله این است که یک نفر را نمی‌توان در چند طبقه جای داد. اگر عددی برای تعیین چنین طبقه‌ای به‌کار ببریم، بدان معنا نیست که این طبقه مفهوم عددی دارد، بلکه تخصیص عدد فقط به معنای نشانه‌گذاری و نام‌گذاری اطلاعات می‌باشد. مانند شماره پشت پیراهن ورزشکاران یا اختصاص اعداد ۱ تا ۴ به گروه‌های خونی O, AB, B, A. توجه داشته باشید که این اعداد را نمی‌توان برای مقایسه به‌کار برد. در مقیاس اسمی هیچ‌کدام از چهار عمل اصلی (جمع، تفریق، ضرب و تقسیم) را نمی‌توان استفاده کرد. هم‌چنین از نظر آماری در آن محدودیت‌های زیادی وجود دارد. در این مقیاس تنها از عملیات آماری تعیین فراوانی و تعیین نما یا مد می‌توان استفاده کرد. به‌طور مثال، نمی‌توان گفت ورزشکاران مرد دو برابر ورزشکاران زن مهارت دارند.

مثال۱: اگر سه گروه ماشین یعنی بنزین سوز، گازسوز و دیزلی داشته باشیم و اعداد ۱ تا ۳ را به آنها اختصاص دهیم، از چه نوع مقیاسی استفاده کرده‌ایم؟

۱) اسمی ۲) ترتیبی ۳) فاصله‌ای ۴) نسبی

☑ **پاسخ:** گزینه «۱» مقیاس اسمی درست می‌باشد؛ زیرا از اعداد ۱ تا ۳ فقط برای طبقه‌بندی ماشین‌ها استفاده کرده‌ایم.

مثال۲: کدام مقیاس به طبقه‌بندی افراد، اشیاء یا حوادث می‌پردازد؟

۱) نسبی ۲) اسمی ۳) ترتیبی ۴) فاصله‌ای

☑ **پاسخ:** گزینه «۲» این مقیاس به طبقه‌بندی افراد، اشیاء و .... می‌پردازد.

مقیاس رتبه‌ای یا ترتیبی (Ordinal)

این نوع مقیاس اندازه‌گیری عموماً برای طبقه‌بندی داده‌ها به منظور نوعی برتری بخشیدن به‌کار می‌رود. این مقیاس علاوه‌بر طبقه‌بندی کردن و نامگذاری به ترتیب طبقه‌ها نیز می‌پردازد. در این مقیاس اعداد تنها برای رتبه‌بندی به‌کار می‌روند. مقیاس رتبه‌ای برای مرتب کردن افراد یا اشیاء به ترتیب از بالاترین به پایین‌ترین بر مبنای ویژگی مورد اندازه‌گیری استفاده می‌شود. به طور مثال، در یک کارخانه ممکن است کارگران را به سه دسته ساده، نیمه ماهر و ماهر تقسیم‌بندی کنیم. اطلاق به ترتیب اعداد ۱ تا ۳ به این سه دسته یک مقیاس ترتیبی است. توجه داشته باشیم که این اعداد تنها برای مقایسه به کار می‌روند، و **نمی‌توان با آنها چهار عمل اصلی را انجام داد.** در تعریف دیگری از مقیاس ترتیبی می‌توان گفت اعدادی که در مقیاس رتبه‌ای به کار می‌روند، مرتبه یا موقعیت افراد را برحسب یک صفت در میان افراد یا اشیاء گروه نشان می‌دهند. به طور مشخص نمی‌توان عنوان کرد که هریک از افراد یا اشیاء به مفهوم واقعی و مطلق چقدر از یک صفت را دارند یا مشخص نمی‌باشد که این افراد یا اشیاء از لحاظ صفت مورد اندازه‌گیری چه مقدار از یکدیگر

فاصله دارند. برای مثال، یک تیم بسکتبال شایان از نظر مهارت در مرتبه اول، سعید در مرتبه دوم و علی در مرتبه سوم قرار دارد. بنابراین تنها می‌توان گفت این افراد از نظر مهارت نسبت به یکدیگر چگونه‌اند، به‌طوری‌که شایان از سعید و سعید از علی برتر است. عبارت دیگر، علی از همه ضعیف‌تر است، اما نمی‌توان گفت علی چقدر از بقیه ضعیف‌تر است. مقیاس رتبه‌ای نسبت به مقیاس اسمی در سطح بالاتری قرار دارد، عملیات مجاز ریاضی و آماری آن متفاوت با مقیاس اسمی است، به‌طوری‌که مقیاس رتبه‌ای ضمن دارا بودن ویژگی مقیاس اسمی، در مورد ترتیب قرار گرفتن ویژگی‌ها در یک پیوستار (یعنی کوچک و بزرگ و مساوی بودن آنها) به ما اطلاعاتی می دهند. عملیات ریاضی که در مقیاس رتبه‌ای قابل کاربرد است فقط علامت $> = <$ است.

مثال۳: در کدام مقیاس فواصل صفت مورد اندازه‌گیری با همدیگر یکسان نیست؟

۱) فاصله‌ای ۲) نسبی ۳) رتبه‌ای ۴) گزینه ۱ و ۳

☑ **پاسخ:** گزینه «۳» در مقیاس رتبه‌ای، مشخص نمی‌باشد که افراد یا اشیاء از لحاظ صفت مورد اندازه‌گیری چه مقدار از یکدیگر فاصله دارند.

مثال۴: در مسابقات تکواندو استان مازندران تیم گاز مقام دوم را به‌دست آورده، برای گزارش نتیجه از چه مقیاسی استفاده شده است؟

۱) اسمی ۲) ترتیبی ۳) فاصله‌ای ۴) نسبی

☑ **پاسخ:** گزینه «۲» در مسابقات تکواندو فاصله بین تیم‌ها دقیقاً مشخص نیست و فقط براساس برتری امتیاز مرتب شده‌اند.

نکته۱: مقیاس رتبه‌ای فاقد نقطهٔ صفر مطلق است.

مقیاس فاصله‌ای (Interval)

این نوع مقیاس اندازه‌گیری عموماً در زمینه‌هایی به‌کار می‌رود که علاوه بر حفظ ترتیب، نحوه‌ی فاصله بین ویژگی‌ها را نیز حفظ می‌کند. به عبارت دیگر، در چنین مقیاسی نسبت تفاضل‌ها ثابت می‌ماند. این نوع مقیاس علاوه‌بر طبقه‌بندی، نامگذاری و مرتب کردن این امکان را برای ما مهیا می‌کند که فاصله بین افراد، اشیاء یا حوادث را به طور دقیق مشخص نماییم در این مقیاس بعضی از عملیات ریاضی را می‌توان اجرا کرد که در مقیاس‌های اسمی و رتبه‌ای قابل اجرا نیست. باید توجه داشت که در این مقیاس **عدد صفر یک مفهوم قراردادی** است؛ یعنی صفر مورد نظر واقعی نمی‌باشد. به عنوان مثال، اندازه‌گیری ضریب هوشی دانش‌آموزان کلاس اول دبستان در شهر اصفهان. در این مقیاس هیچ نوع اطلاعاتی در خصوص مقدار واقعی و مطلق صفت مورد نظر در دست نیست. وجه تمایز این مقیاس از مقیاس ترتیبی این است که در این مقیاس فواصل واحدها معلوم و برابر هستند، درصورتی که در مقیاس ترتیبی فاصله بین واحدها معلوم نیست.

عملیات آماری و ریاضی مقیاس فاصله‌ای بیش از مقیاس رتبه‌ای است و از چهار عمل اصلی می‌توان جمع و تفریق را استفاده کرد، اما کاربرد ضرب و تقسیم در آن مجاز نیست. از لحاظ آماری نیز می‌توان از عملیات محاسبه‌نما، میانه، میانگین، انحراف معیار، ضرایب همبستگی اسپیرمن و پیرسون استفاده کرد.

مثال۵: نمره فرد در درس املاء از چه نوع مقیاسی است؟

۱) اسمی ۲) ترتیبی ۳) فاصله‌ای ۴) نسبی

☑ **پاسخ:** گزینه «۳» نمرات درس املا از صفر تا بیست نمره‌گذاری می‌شود. فاصله نمره‌ی ۱۷ با نمرات ۱۹ و ۱۵ برابر است. اما نمی‌توان گفت کسی که نمره‌ی صفر گرفته است، هیچ اطلاعاتی در خصوص آن ندارد، زیرا آن فرد احتمالاً از بین صد کلمه فقط ۲۰ کلمه را اشتباه نوشته است. زیرا در مقیاس فاصله‌ای صفر واقعی و مطلق وجود ندارد و صفر آن قراردادی و فرضی است.

مثال۶: در آزمون پایان ترم درس روان‌شناسی از کدام مقیاس اندازه‌گیری استفاده می‌شود؟

۱) فاصله‌ای ۲) نسبی ۳) اسمی ۴) رتبه‌ای

☑ **پاسخ:** گزینه «۱» طبق تعریف مقیاس فاصله‌ای نمرات آزمون دروس از نوع مقیاس فاصله‌ای است.

مثال۷: مقیاس اندازه‌گیری موارد زیر، به ترتیب کدام است؟

«مقام تیم ایران در المپیک ریو ـ رکورد دراز و نشست»

۱) فاصله‌ای ـ فاصله‌ای ۲) رتبه‌ای ـ نسبی ۳) اسمی ـ فاصله‌ای ۴) رتبه‌ای ـ فاصله‌ای

☑ **پاسخ:** گزینه «۴» مقیاس مقام تیم ایران در المپیک ریو از نوع رتبه‌ای و مقیاس رکورد درازنشست از نوع فاصله‌ای است.

مقیاس نسبی یا نسبتی (Ratio)

این نوع مقیاس اندازه‌گیری در زمینه‌هایی به‌کار می‌رود که علاوه بر حفظ فاصله، نسبت را نیز حفظ می‌کند، به عبارت دیگر، در این نوع اندازه‌گیری نسبت دو مقدار بستگی به واحد اندازه‌گیری ندارد. این مقیاس بالاترین سطح اندازه‌گیری است و حدود فعالیت آن کلیه عملیاتی است که در سه مقیاس پیشین یعنی اسمی، رتبه‌ای و فاصله‌ای صورت می‌گیرد. در این مقیاس **صفر مطلق** وجود دارد و می‌توان کلیه عملیات ریاضی مثل جمع، تفریق، ضرب و تقسیم را در این مقیاس انجام داد. این مقیاس به‌دلیل داشتن صفر مطلق کاربردهای فراوانی دارد ولی این مقیاس در علوم فیزیکی کاربرد بارزتری دارد. برای مثال، وزن نوزادان تازه به دنیا آمده در یک روز در یک بیمارستان یک شهر. مقیاس نسبی کاملترین مقیاس اندازه‌گیری بوده و نقص‌های مقیاس‌های قبلی را ندارد. این مقیاس مخصوص ویژگی‌های فیزیکی افراد و اشیاء است و در مورد ویژگی‌های روانی افراد و کیفی اشیاء از آن کمتر استفاده می‌شود.

مثال۷: رکورد دوی ۱۲۰۰ متر استقامت با کدام مقیاس اندازه‌گیری می‌شود؟

۱) اسمی ۲) ترتیبی ۳) فاصله‌ای ۴) نسبی

☑ **پاسخ:** گزینه «۴» رکوردهایی که در آن زمان دخیل می‌باشد، از نوع مقیاس نسبی می‌باشند.

مثال۸: مقیاس‌های اندازه‌گیری سانتی‌گراد و کلوین از کدام نوع مقیاس‌ها می‌باشند؟

۱) فاصله‌ای ـ فاصله‌ای ۲) نسبی ـ فاصله‌ای ۳) نسبی ـ نسبی ۴) فاصله‌ای ـ نسبی

☑ **پاسخ:** گزینه «۴» سانتی‌گراد صفر قراردادی دارد، پس فاصله‌ای است. کلوین صفر مطلق دارد که صفر درجه کلوین برابر با ۲۷۳− درجه سانتی‌گراد است، پس نسبی است.

نکته۲: صفر مطلق، صفری است که پایین‌تر از آن عددی وجود نداشته باشد، (مثل اعداد منفی) و از عدم و پوچی حرف بزند.

نمرات آزمون‌ها

می‌خواهیم ببینیم نمرات آزمون‌ها را جزء کدام دسته از مقیاس‌ها می‌توانیم قرار دهیم. کاملاً مشخص است که نمرات آزمون نشانه نیستند؛ زیرا مفهوم نمرات به مراتب بالاتر از نشانه است. از طرف دیگر، نمرات اکثر آزمون‌ها مقیاس نسبی نیز نیستند، زیرا فاقد صفر مطلق هستند. نمرات بعضی از آزمون‌های رغبت به عنوان طبقه به کار می‌رود. نمرات بعضی از آزمون‌ها رتبه‌ها را نشان می‌دهد. نمرات خام و تراز شده در اکثر آزمون‌ها از نوع مقیاس فاصله‌ای می‌باشند. زیرا فقط رتبه نمرات را نشان نمی‌دهد، بلکه تفاوت یا فاصله بین نمرات را نیز نشان می‌دهد. بنابراین نمرات آزمون‌ها ممکن است به صورت طبقه، رتبه یا مقیاس فاصله‌ای، در موارد بسیار نادر به عنوان مقیاس نسبی باشند.

متغیرها به‌وسیله اعداد یا ارزش‌ها مشخص می‌شوند. ماهیت اعداد و ارزش‌ها در گرو این مسئله است که متغیر مورد نظر پیوسته یا گسسته است.

متغیر گسسته:

متغیری است که می‌تواند اعداد یا ارزش‌هایی را به خود اختصاص دهد که مشخص‌کننده یک وجه مشخص و معین از یک مقیاس هستند، در این مقیاس هر چهار عمل اصلی ریاضی کاربرد دارد و فقط از اعداد صحیح استفاده می‌کنیم.

مثال۹: تعداد بازیکنان یک تیم والیبال چه نوع متغیری است؟

۱) گسسته ۲) پیوسته ۳) کیفی ۴) وابسته

☑ **پاسخ:** گزینه «۱» یک تیم والیبال فقط امکان دارد ۱و ۲و ۳و ۴و ۵و ۶ نفر بازیکن داشته باشد و نمی‌توانیم بگوییم ۲/۵ نفر بازیکن دارد.

متغیر پیوسته:

متغیری است که بین دو واحد یا دو نقطه آن، هر نقطه یا ارزشی را می‌توان انتخاب کرد. این متغیر درجه‌های مختلف اندازه‌گیری دارد که به دقت وسیله اندازه‌گیری بستگی دارد. اندازه‌گیری متغیرهای گسسته همواره دقیق و قطعی است، اما اندازه‌های بدست آمده از متغیرهای پیوسته تاحدود زیادی تقریبی هستند. شایان ذکر است که فقط به صورت نظری می‌توان تفاوت بین متغیر گسسته و پیوسته را مشخص کرد و این به‌دلیل کمبود وسایل اندازه‌گیری دقیق و مناسب است. در بسیاری از متغیرهای پیوسته، برای اندازه‌گیری ناچار هستیم اعداد را به‌صورت کلی به کار ببریم. در این مقیاس از هر عددی می‌توانیم استفاده کنیم.

مثال۱۰: نمرات شیرجه نمایشی و رتبه‌های کنکور از چه نوع متغیری می‌باشند؟

۱) پیوسته ـ پیوسته ۲) پیوسته ـ گسسته ۳) گسسته ـ گسسته ۴) گسسته ـ پیوسته

☑ **پاسخ:** گزینه «۲» نمرات شیرجه نمایشی و رتبه‌های کنکور به ترتیب از نوع متغیر پیوسته و گسسته می‌باشند.

مثال ۱۱: طول یا ارتفاع پرش، درصد چاقی و لاغری و درصد تراکم استخوان در تربیت بدنی از کدام نوع متغیر می‌باشند؟

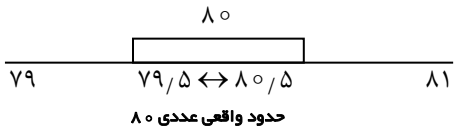
۱) گسسته ۲) پیوسته ۳) کیفی ۴) مستقل

☑ **پاسخ:** گزینه «۲» به‌دلیل اینکه بین دو عدد این متغیرها می‌شود عددی انتخاب نمود از نوع متغیر پیوسته می‌باشند.

محدودیت‌های اعداد:

به دلیل وجود برخی مشکلات در اندازه‌گیری متغیرهای پیوسته، اغلب به متغیرها ارزش عددی گسسته داده می‌شود. برای آنکه این ارزش‌ها را تفسیر کنیم، باید حدود واقعی ریاضی آنها را مورد توجه قرار دهیم. به‌طور کلی حدود واقعی هر عدد، محدوده‌ای است که از **نیم واحد پایین‌تر تا نیم واحد بالاتر** از عدد مورد نظر ادامه دارد. هر عددی که در این محدوده قرار می‌گیرد، آن را به‌صورت گرد شده، یعنی بدون اعشار نشان می‌دهیم. به طور مثال، اگر بعد از امتحان زبان، ببینیم که ۱۰ نفر از افراد نمره ۸۰ گرفته‌اند، احتمال آنکه همه آنها توانایی یکسانی داشته باشند، خیلی کم است. احتمالاً نمره‌ی برخی از این ۱۰ نفر بالاتر و برخی پایین‌تر از ۸۰ بوده است و وسیله‌ی اندازه‌گیری به‌کار برده شده برای اندازه‌گیری به اندازه کافی دقیق نبوده است. حدود واقعی ارزش عددی یک متغیر پیوسته برابر حاصل جمع و تفریق آن عدد با نصف واحد اندازه‌گیری است.

نکته۲: تنها عدد کامل نیست که حدود دارد، بلکه هر عددی حدود واقعی دارد.



مثال۱۲: حدود واقعی عدد ۲۲ کدام گزینه است؟

۱) ۲۱ تا ۲۳ ۲) ۲۱/۵ تا ۲۲/۵ ۳) ۲۰ تا ۲۴ ۴) ۲۲

☑ **پاسخ:** گزینه «۲» حدود واقعی عدد از طریق نیم نمره کم و اضافه کردن به عدد به‌دست می‌آید.

$$\begin{cases} ۲۲+۰/۵=۲۲/۵ \\ ۲۲-۰/۵=۲۱/۵ \end{cases}$$

مثال۱۳: حد پایین و بالای ۱۵/۵ چند می‌باشد؟

۱) ۱۵/۵ـ۱۴/۵ ۲) ۱۶-۱۵ ۳) ۱۵/۵۵ـ۱۵/۴۵ ۴) ۱۶ـ۱۴/۵

☑ **پاسخ:** گزینه «۳»

حد بالا $\Rightarrow ۱۵/۵+۰/۵=۱۵/۵۵$، حد پایین $\Rightarrow ۱۵/۵-۰/۰۵=۱۵/۴۵$

واحد اندازه‌گیری ۰/۱ می‌باشد، پس نیم واحد ۰/۰۵ می‌باشد.

جدول فراوانی‌ها

در علم آمار بعد از جمع‌آوری اطلاعات و دست یافتن به نتایج، محقق موظف است که این اطلاعات را **طبقه‌بندی، خلاصه و برای تجزیه و تحلیل و تفسیر** آماده نماید. این عمل با مهیا ساختن یک جدول یا توزیع فراوانی انجام می‌شود که در این جدول اطلاعات جمع آوری شده به شکل طبقه یا گروه‌های مختلف مرتب می‌شوند. جدول زیر نشان می‌دهد که از بین چه تعداد شرکت کننده، چند نفر قبول و چند نفر در این امتحان رد شده‌اند. در واقع جدول پایین فقط کل فراوانی‌های موجود یا به طور دقیق، کل اطلاعات جمع آوری شده است و از این اطلاعات می‌توانیم به جدول زیر دست پیدا کنیم.

بدین صورت عمل می‌کنیم که ابتدا تعداد کل فراوانی‌ها (تعداد قبولی‌ها + تعدادی ردی‌ها) را به‌دست می‌آوریم: $F=۲۰+۳۰=۵۰ \Rightarrow F=۵۰$

سپس برای مشخص شدن درصد فراوانی هر دسته تناسب می‌گیریم، به این صورت که حساب می‌کنیم ۲۰ نفر قبولی از ۵۰ نفر کل فراوانی‌ها معادل چند درصد می‌شود.

$$\frac{۲۰}{۵۰}=\frac{x}{۱۰۰} \Rightarrow x=\frac{۱۰۰\times ۲۰}{۵۰}=۴۰ \Rightarrow \text{قبولی } x=\%۴۰$$

درصد افراد مردود شده را نیز به صورت افراد قبول شده محاسبه می‌کنیم، در این تناسب درصد ۳۰ نفر مردودی از ۵۰ نفر کل فراوانی‌ها برابر است با:

$$\frac{۳۰}{۵۰}=\frac{y}{۱۰۰} \Rightarrow y=\frac{۱۰۰\times ۳۰}{۵۰}=۶۰ \Rightarrow \text{مردودی } y=\%۶۰$$

طبقات	فراوانی F	درصد فراوانی P
قبول	۲۰	%۴۰
ردی	۳۰	%۶۰

نتایج کل این جدول‌ها نشان می‌دهد:

الف) آزمون بصورت دو گروه یا دو طبقه قبولی و ردی برگزار شده است.

ب) در ستون اول درصد فراوانی افراد و در ستون دوم فراوانی کل افراد قبولی یا ردی حساب شده است.

مثال۱۴ : اگر ۵۰ نفر در آزمون ورودی رشته هندبال شرکت کنند و در این آزمون نتایج را به صورت قبولی و ردی مشخص کنند. یک جدول مانند جدول زیر به دست می‌دهد. چند درصد از افراد قبول و چند درصد مردود شده‌اند؟

ق ر ر ق ق ر ر ق ر ق ر ر ر ق ق ر ر ر ق ر ر ق ر ر ق ر ق ر ق ق ر ر ق ر ق ق ر ر

۱) ۵۰% قبول ـ ۵۰% مردود ۲) ۴۰%مردود ـ ۶۰% قبول ۳) ۴۰% قبول ـ ۶۰% مردود ۴) ۷۰% قبول ـ ۳۰% مردود

☑ **پاسخ:** گزینه «۳»

$$\text{قبولی}=\frac{۲۰}{۵۰}=\frac{x}{۱۰۰} \Rightarrow \frac{۱۰۰\times ۲۰}{۵۰}=\%۴۰$$

$$\text{ردی} =\frac{۳۰}{۵۰}=\frac{x}{۱۰۰} \Rightarrow \frac{۱۰۰\times ۳۰}{۵۰}=\%۶۰$$

بنابراین ۴۰ درصد افراد در آزمون قبول و ۶۰ درصد مردود شده‌اند.

نکته۳: هر گاه تنظیم داده‌ها در جدول آماری براساس فراوانی داده‌ها انجام گیرد، جدول آماری را **جدول فراوانی** می‌نامند.

چگونگی ساخت یک جدول فراوانی

۱) در ابتدا دامنه نمره‌ها را مشخص کرده، پایین‌ترین نمره را از بالاترین نمره کم کرده و حاصل را با ۱ واحد اندازه‌گیری جمع می‌کنیم. اضافه کردن عدد ۱ واحد به خاطر تعیین حدود واقعی اعداد است که این ۱ واحد می‌تواند با توجه به حساسیت ابزار اندازه‌گیری که از روی جدول نمره‌ها مشخص می‌شود. نمره‌ها ۱ و ۰/۱و ۰/۰۱و ... باشد.

$$\{\text{دامنه تغییرات = (بالاترین نمره − پایین‌ترین نمره)+ ۱واحد اندازه‌گیری }\} \Rightarrow \{R=(x_{max}-x_{min})+۱\}$$

۲) تعداد طبقه‌های مورد نیاز را مشخص کنید که معمولاً تعداد طبقه‌ها کمتر از ۱۰و بیشتر از ۲۰طبقه نیستند. دامنه تغییرات را بر تعداد طبقه‌ها تقسیم می‌کنیم، تا بدین وسیله اندازه واقعی هر طبقه را به دست آوریم.

$$\text{اندازه واقعی طبقات}=\frac{\text{دامنه تغییرات}}{\text{تعداد طبقات مورد نظر}}=\frac{R}{\text{تعداد طبقات}}$$

۳) در صورتی که اندازه واقعی طبقه‌ها به صورت اعشاری در آمده، آن را گرد نمایید (عدد باید بدون اعشار باشد) این عمل، عملیات ریاضی را آسان‌تر می‌کند.

مثال۱۵: نمرات ۴۰ دانش‌آموز را که در آزمون آمادگی جسمانی به‌دست آمده، در ۱۱ دسته طبقه‌بندی نمایید.

۸	۹	۹	۱۲	۲۳	۸	۱۰	۱۴
۲۹	۲۰	۲۹	۳۹	۱۰	۲۸	۲۳	۴۰
۲۰	۳۳	۱۲	۲۱	۴۰	۲۶	۲۸	۳۰
۳۲	۳۸	۳۰	۳۵	۱۸	۴۰	۱۳	۳۰
۲۸	۳۵	۲۰	۱۰	۳۴	۲۷	۲۶	۳۲

مدرسان شریف

فصل پنجم

«شاخص‌های پراکندگی»

پس از آشنایی با مفاهیم پایه مطرح‌شده در مباحث آماری، در بخش جاری به توضیح برخی از مفاهیم توصیفی آماری و کاربردهای آن می‌پردازیم. گفته شد که یکی از وظایف آمار، طبقه‌بندی کردن داده‌های مشاهده‌شده می‌باشد. این کار سرعت طبقه‌بندی و درک و فهم داده‌ها را بیشتر می‌کند و سبب صرفه‌جویی در وقت می‌شود. به‌طور کلی اطلاعاتی را که در آمارگیری‌ها جمع‌آوری می‌کنیم، با اعداد نشان می‌دهیم. این اعداد حکم مواد خام در آمار را دارند. برای اینکه نتایج مناسب و مطلوبی از این اطلاعات به‌دست آید، باید:

۱ـ غیرواقع یا غلط نباشند؛ زیرا اعداد، نماینده واقعی مشاهدات می‌باشند.

۲ـ به نحو صحیحی تجزیه و تحلیل گردند.

۳ـ به‌نحو مفیدی تهیه و تنظیم شوند.

۴ـ قابل نتیجه‌گیری صحیح باشند.

مطالب فوق موضوع علم آمار را تشکیل می‌دهند. به‌طور کلی آمار توصیفی عبارت است از: **روش‌هایی که به‌وسیله آنها می‌توان اطلاعات جمع‌آوری شده را تنظیم و خلاصه کرد یا مجموعه روش‌هایی است که سبب پردازش داده‌ها می‌شود.**

معیارهای پراکندگی

باوجود این‌که در بسیاری از موارد، میانگین توصیف نسبتاً کاملی از مجموعه‌ی داده‌ها ارائه می‌دهد ولی گاهی وجود اطلاعات بیشتر در مورد داده‌ها ضروری است. ما برای اینکه چگونگی تغییر افراد و گروه‌ها را بدانیم، به اندازه‌های پراکندگی نیاز داریم. درواقع پراکندگی **به حدودی که افراد یا نمره‌ها با یکدیگر اختلاف دارند، گفته می‌شود.** پراکندگی زیاد نمرات، نشانه‌ی نبودن هماهنگی و یکنواختی بین آنهاست و عکس آن نشان می‌دهد که افراد به هم نزدیک‌ترند و هنگامی که نمره‌های افراد با هم مساوی باشند، پراکندگی وجود نخواهد داشت. دو یا چند توزیع ممکن است که معدل یکسانی داشته باشند اما پراکندگی و سایر خصوصیات بین آنها متفاوت باشد.

مثال ۱: میانگین و پراکندگی دو گروه زیر را با هم مقایسه کنید؟

الف) ۱۲، ۱۳، ۱۵، ۱۷، ۱۸ ب) ۲۱،۲۱،۱۹،۱۵،۱۱،۹،۹

۱) میانگین مساوی، پراکندگی الف بیشتر است. ۲) میانگین نامساوی، پراکندگی الف بیشتر است.

۳) میانگین مساوی، پراکندگی ب بیشتر است. ۴) میانگین نامساوی، پراکندگی ب بیشتر است.

☑ **پاسخ:** گزینه «۳» میانگین هر دو گروه ۱۵ است اما پراکندگی گروه ب بیشتر است.

$$\bar{X}_{\text{الف}} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{۱۸+۱۷+۱۵+۱۳+۱۲}{۵} = \frac{۷۵}{۵} = ۱۵ \Rightarrow \bar{x}_{\text{الف}} = ۱۵ \;،\; R = ۱۸-۱۲+۱ = ۷$$

$$\bar{X}_{\text{ب}} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{۲۱+۲۱+۱۹+۱۵+۱۱+۹+۹}{۷} = \frac{۱۰۵}{۷} = ۱۵ \Rightarrow \bar{x}_{\text{ب}} = ۱۵ \;،\; R = ۲۱-۹+۱ = ۱۳$$

دامنه تغییرات (Range)

دامنه تغییرات، اختلاف بین کوچک‌ترین و بزرگ‌ترین داده‌هاست. **زمانی از دامنه تغییرات استفاده می‌شود که داده‌ها پراکندگی زیادی دارند یا به اطلاعاتی کلی درباره پراکندگی نمره‌ها نیاز داریم.**

این شاخص از طریق کم‌کردن کوچک‌ترین داده از بزرگ‌ترین داده و اضافه‌کردن ۱ واحد اندازه‌گیری به‌دست می‌آید.

R = بزرگ‌ترین داده ـ کوچک‌ترین داده + ۱ واحد اندازه‌گیری

نکته۱: دامنه‌ی تغییرات، ساده‌ترین شاخص پراکندگی است.

نکته۲: دامنه‌ی تغییرات، اطلاعاتی درباره‌ی چگونگی توزیع نمره‌ها ارائه نمی‌دهد؛ بنابراین استفاده از آن محدود است و اعتبار کمتری به‌عنوان شاخص پراکندگی دارد.

نکته۳: زمانی از دامنهٔ تغییرات استفاده می‌شود که نمره‌ها دارای پراکندگی خیلی زیادی هستند و یا هنگامی که به اطلاعات کلی دربارهٔ پراکندگی نمره‌ها نیاز داریم.

مثال ۲: در ۲، ۲، ۹، ۹، ۹، ۷، ۵، ۱۰، ۱۰، ۱۱، ۱۲، ۱۸ دامنه‌ی تغییرات داده‌ها را محاسبه نمایید؟

۱) ۱۰ ۲) ۱۲ ۳) ۱۴ ۴) ۱۷

☑ **پاسخ:** گزینه «۴» دامنه تغییرات این داده‌ها $R = ۱۸ - ۲ + ۱ = ۱۷$ می‌باشد.

◆ ◆ ◆ ◆

مثال ۳: در ۳، ۵، ۸، ۱۰، ۱۲، ۱۵، ۱۶ دامنه‌ی تغییرات کدام است؟

۱) ۱۳ ۲) ۱۴ ۳) ۱۵ ۴) ۱۶

☑ **پاسخ:** گزینه «۲» $R =$ بزرگ‌ترین داده $-$ کوچک‌ترین داده + ۱ واحد اندازه‌گیری $= ۱۶ - ۳ + ۱ = ۱۳ + ۱ = ۱۴$

نکته۴: در محاسبه دامنه‌ی تغییرات اگر بین داده‌ها عددی تا یک رقم اعشار وجود داشت، به فرمول دامنه تغییرات ۰/۱ اضافه می‌کنیم، اگر در داده‌ها عددی تا دو رقم اعشار وجود داشت، به فرمول دامنه‌ی تغییرات ۰/۰۱ اضافه می‌کنیم و ...

◆ ◆ ◆ ◆

مثال ۴: در داده‌های روبرو دامنه‌ی تغییرات را محاسبه نمایید؟ (۶/۳ و ۶/۷۳ و ۷/۳ و ۵/۹۶ و ۴/۴۵ و ۴/۳)

۱) ۳ ۲) ۳/۱ ۳) ۳/۰۱ ۴) ۳/۰۰۱

☑ **پاسخ:** گزینه «۳» چون عددی تا دو رقم اعشار موجود است به فرمول ۰/۰۱ اضافه می‌کنیم.

$$R = (X_H - X_L) + ۰/۰۱ = (۷/۳ - ۴/۳) + ۰/۰۱ = ۳ + ۰/۰۱ = ۳/۰۱ \Rightarrow R = ۳/۰۱$$

انحراف متوسط Average deviation

انحراف به زبان ساده عبارت است از فاصله هر نمره از میانگین توزیع آن. داده‌هایی که در بالای میانگین قرار می‌گیرند، دارای انحراف مثبت و داده‌هایی که پایین‌تر از میانگین هستند، دارای انحراف منفی هستند.

فرمول محاسبه انحراف متوسط $(A.D)$ به شرح مقابل است:

$$A.D = \frac{\Sigma|d|}{n} \quad \text{یا} \quad \frac{\sum|x - \bar{x}|}{n}$$

برای محاسبه انحراف متوسط به طریق زیر عمل می‌کنیم:

$d = x - \bar{x}$

نمره: x میانگین: $\bar{x}$ انحراف متوسط: $A.d$

۱ـ علامت انحراف هر عدد از میانگین (d) را نادیده بگیرید. (نماد $|d|$ به معنای قدر مطلق عدد است؛ یعنی d بدون در نظر گرفتن علامت مثبت یا منفی آن).

۲ـ مجموع داده d را (Σd) حساب کنید.

۳ـ مجموع محاسبه‌شده در مرحله دوم را بر (n) تعداد اعداد توزیع تقسیم کنید.

نکته۵: انحراف از میانگین، شاخصی را برای تعیین مقدار پراکندگی نمره‌ها از میانگین فراهم می‌سازد.

مثال۵: جمع انحراف از میانگین داده‌های مقابل را به‌دست آورید؟ **۳,۵,۷,۸,۱۰,۳**

۱) ۶ ۲) ۲ ۳) صفر ۴) ۴

☑ **پاسخ:** گزینه «۳» جمع انحراف از میانگین داده‌ها همیشه برابر با صفر است.

نکته۶: چون نمره‌ها در اطراف میانگین، به‌صورت مثبت و منفی پراکنده هستند، میزان متوسط واقعی انحرافات برابر با صفر خواهد بود.

مثال ۶: توزیع ۱۰,۹,۹,۸,۷,۶,۴,۳ را در نظر بگیرید. انحراف متوسط این داده‌ها کدام گزینه می‌باشد؟

۱) ۲ ۲) ۳ ۳) ۵ ۴) ۶

☑ **پاسخ:** گزینه «۱» در محاسبه انحراف متوسط، تمام نمره‌ها دخالت دارند. ابتدا میانگین داده‌ها را از فرمول $\overline{x}=\frac{\Sigma x}{n}$ محاسبه می‌کنیم:

$$\overline{x}=\frac{\Sigma x}{n}\Rightarrow \overline{x}=\frac{۵۶}{۸}=۷\Rightarrow \overline{x}=۷$$

سپس اختلاف تک‌تک داده‌ها را از میانگین به‌دست می‌آوریم و در جدول می‌نویسیم. برای این کار از فرمول $d=x-\overline{x}$ استفاده می‌کنیم، سپس $\Sigma|d|$ را به‌دست آورده و در فرمول قرار داده، انحراف متوسط را به‌دست می‌آوریم:

x	$d=x-\overline{x}$
۱۰	۳
۹	۲
۹	۲
۸	۱
۷	۰
۶	-۱
۴	-۳
۳	-۴
	$\Sigma\|d\|=۱۶$

$$A.D=\frac{\Sigma|d|}{n}\Rightarrow A.D=\frac{۱۶}{۸}=۲\Rightarrow A.D=۲$$

نکته۷: طریقه‌ی محاسبه‌ی انحراف متوسط داده‌های طبقه‌بندی شده مانند داده‌های طبقه‌بندی نشده است؛ با این تفاوت که در داده‌های طبقه‌بندی‌شده انحراف نقاط میانی طبقات از میانگین محاسبه می‌شود.

نکته۸: مزیت انحراف متوسط به دامنه تغییرات این است که در محاسبه انحراف متوسط تمام نمره‌ها دخالت دارند.

انحراف معیار یا انحراف استاندارد (Standard deviation)

تفاوت بارز بین انحراف معیار و انحراف متوسط در محاسبه‌ی آنهاست. در محاسبه‌ی انحراف متوسط، علائم اعداد d نادیده گرفته می‌شود و مجموع آنها (حاصل‌جمع قدر مطلق آن‌ها) بر تعداد آن تقسیم می‌شود اما در محاسبه انحراف معیار مشکل علائم حل می‌شود.

الف) محاسبه انحراف معیار در داده‌های طبقه‌بندی نشده:

$S.d$ = انحراف معیار n = تعداد داده‌ها $d^۲$ = مجذور انحراف هر نمره از میانگین

انحراف معیار از فرمول روبرو محاسبه می‌شود: $s.d=\sqrt{\frac{\Sigma d^۲}{n-۱}}$ یا $s.d=\sqrt{\frac{\Sigma(x-\overline{x})^۲}{n-۱}}$

مثال ۷: نمره آزمون دو و میدانی ورزشکاران در جدول زیر نشان داده شده است. باتوجه‌به این نمره‌ها انحراف معیار را محاسبه کنید.

۱) ۲/۴ ۲) ۲/۸ ۳) ۳/۲ ۴) ۳/۶

مدرسان شریف

فصل هفتم

« ضریب همبستگی»

در جهان ما حوادث متعددی رخ می‌دهد که با یکدیگر ارتباط دارند؛ برای مثال می‌توان میزان بارندگی و عمق رودخانه را با هم و یا درجه حرارت زمین و وضعیت تابش خورشید را با هم مرتبط دانست که تغییر در یکی باعث ایجاد تغییر در دیگری می‌شود. افزایش بارندگی با افزایش آب رودخانه همراه خواهد شد و کاهش بارندگی نیز با کاهش آب و عمق رودخانه همراه است، درجه حرارت زمین با وضعیت تابش خورشید نیز چنین ارتباطی دارد؛ پس ما می‌توانیم بگوییم که بارندگی با آب رودخانه و درجه حرارت زمین با وضعیت تابش خورشید همبستگی دارند. همبستگی را می‌توان به صورت‌های زیر تعریف کرد:

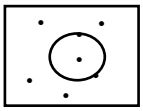
۱- رابطه بین دو یا چند جفت متغیر و یا رابطه میان دو یا چند دسته از داده‌ها.

۲- رابطه‌ی بین دو یا چند متغیر که قابل تبدیل به کمیت باشند.

۳- به‌صورت ساده‌تر رابطه‌ی بین متغیرها را همبستگی می‌گویند.

ضریب همبستگی را با r نشان می‌دهند. ضریب همبستگی بین دو یا چند متغیر بیان‌کننده‌ی این مطلب است که آیا تغییر در یک متغیر، همگام با تغییر در متغیر دیگر است یا خیر. ضریب همبستگی نشان‌دهنده‌ی رابطه علت و معلولی بین متغیرها نیست. **دامنه‌ی تغییرات ضریب همبستگی بین ۱+ و ۱− است** که به این دو ضریب، ضریب همبستگی کامل گفته می‌شود. ۱+ضریب همبستگی کامل مستقیم (مثبت) و ۱- ضریب همبستگی کامل معکوس (منفی) را نشان می‌دهد. (ضریب همبستگی کامل به‌ندرت یافت می‌شود). حال اگر بین متغیرها وابستگی وجود نداشته باشد، ضریب همبستگی صفر خواهد بود.

۱- = ضریب همبستگی کامل معکوس



۰ = ضریب همبستگی صفر

۱+ = ضریب همبستگی کامل مستقیم

اگر بالا بودن اندازه یک صفت با بالا بودن اندازه‌ی صفت دیگر و یا پایین بودن اندازه یک صفت با پایین بودن اندازه صفت دیگر همراه شود، **ضریب همبستگی مثبت** را خواهیم داشت و اگر بالا بودن اندازه‌ی یک صفت با پایین بودن صفت دیگر و یا پایین بودن یک صفت با بالا بودن صفت دیگر همراه شود، **ضریب همبستگی منفی** را خواهیم داشت؛ برای مثال می‌توان به میزان بارندگی و افزایش آب رودخانه اشاره کرد. با افزایش میزان بارندگی، آب رودخانه هم افزایش خواهد یافت و در نتیجه عمق رودخانه بیشتر می‌شود، پس بین این دو متغیر همبستگی مستقیم یا مثبت وجود دارد. از طرف دیگر می‌توان به افزایش سن و کاهش حافظه اشاره کرد. هر چه سن افزایش یابد، حافظه نیز کاهش می‌یابد؛ یعنی با افزایش یک متغیر، متغیر دیگر کاهش می‌یابد؛ پس می‌توان گفت بین این دو متغیر ضریب همبستگی معکوس یا منفی وجود دارد. حال اگر بین متغیرها ارتباطی وجود نداشته باشد، (بین اندازه‌ی کفش و حقوق دریافتی ماهانه) می‌توان گفت ضریب همبستگی بین این دو متغیر از نوع صفر است (هیچ ارتباطی بین این دو متغیر نداریم).

هنگام متغیر ضریب همبستگی لازم است ماهیت جامعه‌ای را که دو متغیر در آن مورد مشاهده یا اندازه‌گیری قرار گرفته‌اند، بررسی کرد. ضریب همبستگی بین دو متغیر از جامعه‌ای به جامعه دیگر فرق می‌کند زیرا:

۱) اساس رابطه، از جامعه‌ای به جامعه دیگر فرق می‌کند.

۲) پراکندگی متغیرها در جوامع مختلف متفاوت است.

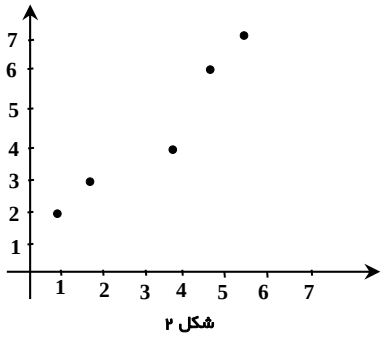
۳) همبستگی بین دو متغیر می‌تواند تحت تأثیر همبستگی آنها با متغیر ثالثی قرار بگیرد.

مثال۱: کدام مورد دارای شدت همبستگی بالاتری می‌باشد؟

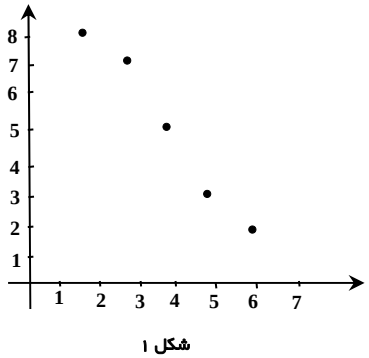
۱) ۰/۷۰+ ۲) ۰/۸۹− ۳) ۰/۸۹+ ۴) ۰/۹۲−

☑ **پاسخ:** گزینه «۴» چون ۰/۹۲ بزرگتر از همه می‌باشد. در شدت به منفی یا مثبت بودن همبستگی کاری نداریم.

مثال۲: در شکل‌های زیر مشخص کنید کدام‌یک دارای ضریب همبستگی مثبت و کدام‌یک منفی هستند؟



شکل ۲



شکل ۱

۱) شکل ۱ منفی ـ شکل ۲ منفی
۲) شکل ۱ مثبت ـ شکل ۲ مثبت
۳) شکل۱ منفی ـ شکل ۲ مثبت
۴) شکل ۱ مثبت ـ شکل ۲ منفی

☑ **پاسخ:** گزینه «۳» در شکل اول مشاهده می‌کنیم که با افزایش یک متغیر، متغیر دیگر کاهش می‌یابد؛ پس ضریب همبستگی معکوس یا همان منفی است ولی در شکل دوم با افزایش یک متغیر، متغیر دیگر هم در حال افزایش است، پس ضریب همبستگی مستقیم و یا مثبت است.

ضریب همبستگی رتبه‌ای اسپیرمن (r_S)

ضریب همبستگی اسپیرمن برای توصیف میزان همبستگی بین دو متغیر که با استفاده از مقیاس ترتیبی اندازه‌گیری شده باشند، به‌کار می‌رود. ضریب همبستگی رتبه‌ای اسپیرمن (spearman Rank – order correlation)، روش مناسبی است که میزان همبستگی و رابطه بین رتبه‌ها را نشان می‌دهد. برای محاسبه‌ی ضریب همبستگی رتبه‌ای اسپیرمن از فرمول زیر استفاده می‌کنیم:

$$r_S = ۱ - \frac{۶\sum D^۲}{N(N-۱)(N+۱)} \quad \text{یا} \quad r_S = ۱ - \frac{۶\sum D^۲}{n(n^۲+۱)}$$

در فرمول بالا، $\sum D^۲$ مجموع مجذورات تفاوت بین جفت رتبه‌ها، D تفاوت بین جفت رتبه‌ها، N تعداد جفت رتبه‌ها است. با ذکر یک مثال به توضیح بیشتر این مبحث خواهیم پرداخت:

مثال۳: دو مربی پس از تست آمادگی جسمانی، ۵ ورزشکار را به‌صورت زیر رده‌بندی کرده‌اند. ضریب همبستگی اسپیرمن بین دو رتبه‌بندی چقدر است؟

ورزشکاران	مربی اول	مربی دوم
رضا	۱	۱
علی	۳	۲
محمد	۴	۵
سیاوش	۲	۳
امیر	۵	۴

۱) ۰/۷۰ ۲) ۰/۷۳ ۳) ۰/۸۰ ۴) ۰/۸۳

☑ **پاسخ:** گزینه «۳» اولین کاری که برای حل این گونه مسائل انجام می‌دهیم، این است که در کنار جدول اصلی یک ردیف دیگر به عنوان D ایجاد کرده و اختلاف بین دو رتبه دانش‌آموزان را محاسبه می‌کنیم، سپس مجموع Dها را در ردیف آخر یعنی $\sum D$ محاسبه می‌کنیم.

◆ ◆ ◆ ◆

در مرحله بعد، یک ردیف دیگر تحت عنوان $D^۲$ ایجاد کرده، مقادیر به‌دست آمده D را به توان ۲ می‌رسانیم و سپس مجموع آن‌ها را نیز در ردیف آخر $\sum D^۲$ می‌نویسیم.

$D^۲$
۰
+۱
+۱
+۱
+۱
$\sum D^۲ = ۴$

D
۰
+۱
−۱
−۱
+۱
$\sum D = ۰$

حال پس از محاسبه این مقادیر، فرمول همبستگی رتبه‌ای اسپیرمن را نوشته و اعداد به‌دست آمده را در فرمول جایگذاری می‌کنیم:

$$r_s = ۱ - \frac{۶\sum D^۲}{N(N-۱)(N+۱)}$$

که در فرمول فوق $\sum D^۲ = ۴$ و $N = ۵$ تعداد یک دسته از داده‌ها

$$r_s = ۱ - \frac{۶ \times ۴}{۵(۵-۱)(۵+۱)} = ۱ - \frac{۲۴}{۱۲۰} \Rightarrow ۱ - ۰/۲ = ۰/۸ \Rightarrow r_s = ۰/۸$$

ضریب همبستگی ۰/۸است و این نشان از توافق زیاد بین نظر دو مربی است.

نکته۱: $\sum D$ همیشه برابر با صفر است.

مثال۴: در یک آزمون تربیت‌بدنی به ۱۰نفر از ورزشکاران توسط ۲ داور، رتبه‌های زیر داده شده است. ضریب همبستگی رتبه‌ای اسپیرمن داده‌های زیر کدام است؟

داور ۲	داور ۱	آزمودنی‌ها
۸	۱۰	۱
۹	۷	۲
۳	۶	۳
۷	۸	۴
۶	۵	۵
۱۰	۹	۶
۲	۲	۷
۵	۳	۸
۱	۱	۹
۴	۴	۱۰

۱) ۰/۵۸ ۲) ۰/۸۵ ۳) ۰/۶۸ ۴) ۰/۷۸

☑ **پاسخ:** گزینه «۲» مطابق مثال قبل، اولین کار، ایجاد ستون D (اختلاف بین رتبه‌ها) و $D^۲$ (به توان ۲ رساندن اختلاف بین رتبه‌ها) در کنار جدول اصلی است. پس مطابق مرحله اول، رتبه‌ای را که دو داور به آزمودنی‌ها داده‌اند از هم کم کرده و در ردیف D می‌نویسیم سپس D (اختلاف بین رتبه‌ها) را به توان ۲ می‌رسانیم و در ردیف $D^۲$ قرار می‌دهیم و سپس $\sum D^۲$ (مجموع توان دوم اختلاف بین رتبه‌ها) را نیز محاسبه می‌کنیم.

پس از به‌دست آوردن موارد فوق، فرمول همبستگی رتبه‌ای اسپیرمن را نوشته و این اطلاعات را در آن قرار می‌دهیم و ضریب همبستگی را حساب می‌کنیم.

$D^۲$
۴
۴
۹
۱
۱
۱
۰
۴
۰
۰
$\sum D^۲ = ۲۴$

D
+۲
−۲
+۳
+۱
−۱
−۱
۰
−۲
۰
۰
$\sum D = ۰$

$$r_s = ۱ - \frac{۶\sum D^۲}{N(N+۱)(N-۱)} \Rightarrow r_s = ۱ - \frac{۶\times ۲۴}{۱۰(۹)(۱۱)} \Rightarrow r_s = ۱ - \frac{۱۴۴}{۹۹۰} \Rightarrow r_s = ۰/۸۵$$

این ضریب همبستگی بالا، از توافق نظر بین دو داور را نشان می‌دهد.

مثال۵: اگر مجموع اختلاف رتبه‌ها بین ۸ نفر برابر با ۳۶ باشد، همبستگی رتبه‌ای کدام می‌باشد؟

۱) ۰/۴۳ ۲) ۰/۵۷ ۳) ۰/۶۷۱ ۴) ۰/۴

پاسخ: گزینه «۲» $r_s = ۰/۵۷$ می‌باشد.

$$r_s = ۱ - \frac{۶\sum D^۲}{N(n-۱)(n+۱)} \Rightarrow ۱ - \frac{۶(۳۶)}{۸(۷)(۹)} \Rightarrow ۱ - \frac{۲۱۶}{۵۰۴} \Rightarrow ۱ - ۰/۴۳ = ۰/۵۷$$

ضریب همبستگی گشتاوری پیرسون

برای به‌دست آوردن ضریب همبستگی گشتاوری پیرسون از دو روش می‌توان استفاده کرد. اگر داده‌ها به‌صورت اعداد خام و مقیاس اندازه‌گیری فاصله‌ای و یا نسبی باشد، این روش سودمند است و از فرمول زیر محاسبه می‌شود:

روش اول:

$$r_{xy} = \frac{n\Sigma xy - (\Sigma x)(\Sigma y)}{\sqrt{[n\Sigma x^۲ - (\Sigma x)^۲][n\Sigma y^۲ - (\Sigma y)^۲]}} \quad \text{(ضریب گشتاوری پیرسون)}$$

که در فرمول فوق n تعداد کل داده‌ها، x داده‌های مربوط به اولین متغیر، y داده‌های مربوط به دومین متغیر و Σ مجموع داده‌ها است.

مثال۶: در جدول زیر نمره‌های ده ورزشکار در دو آزمون انعطاف‌پذیری و چابکی به‌دست آمده است. ضریب همبستگی پیرسون برای این دو دسته نمرات کدام است؟

چابکی	انعطاف‌پذیری (x)
۴	۴۳
۵	۵۵
۶	۶۷
۴	۳۸
۵	۴۹
۷	۷۰
۹	۸۰
۵	۶۲
۶	۷۳
۹	۸۳
$\sum y = ۶۰$	$\sum x = ۶۲۰$

۱) ۰/۸۱

۲) ۰/۹۱

۳) ۰/۱۹

۴) ۰/۱۸

☑ **پاسخ:** گزینه «۲» **روش اول:** برای به‌دست آوردن ضریب همبستگی گشتاوری پیرسون لازم است در جدول فوق $x^۲, y^۲$ و xy را نیز محاسبه کنیم، پس مطابق مثال‌های قبلی در جدولی سه ستون جدید تحت عنوان $x^۲$ (که xها را به توان ۲ می‌رسانیم)، $y^۲$ (که yها را به توان ۲ می‌رسانیم) و xy (که حاصل‌ضرب xها در yها می‌باشد) ایجاد می‌کنیم و در آخر Σ (مجموع) هر یک از موارد فوق را نیز محاسبه کرده، در فرمول ضریب همبستگی گشتاوری پیرسون (مطابق زیر) قرار می‌دهیم:

$x^۲$
$(۴۳)^۲ \rightarrow ۱۸۴۹$
$(۵۵)^۲ \rightarrow ۳۰۲۵$
$(۶۷)^۲ \rightarrow ۴۴۸۹$
$(۳۸)^۲ \rightarrow ۱۴۴۴$
$(۴۹)^۲ \rightarrow ۲۴۰۱$
$(۷۰)^۲ \rightarrow ۴۹۰۰$
$(۸۰)^۲ \rightarrow ۶۴۰۰$
$(۶۲)^۲ \rightarrow ۳۸۴۴$
$(۷۳)^۲ \rightarrow ۵۳۲۹$
$(۸۳)^۲ \rightarrow ۶۸۸۹$
$\sum x^۲ = ۴۰۵۷۰$

$y^۲$
$(۴)^۲ \rightarrow ۱۶$
$(۵)^۲ \rightarrow ۲۵$
$(۶)^۲ \rightarrow ۳۶$
$(۴)^۲ \rightarrow ۱۶$
$(۵)^۲ \rightarrow ۲۵$
$(۷)^۲ \rightarrow ۴۹$
$(۹)^۲ \rightarrow ۸۱$
$(۵)^۲ \rightarrow ۲۵$
$(۶)^۲ \rightarrow ۳۶$
$(۹)^۲ \rightarrow ۸۱$
$\sum y^۲ = ۳۹۰$

xy
$۴۳ \times ۴ \rightarrow ۱۷۲$
$۵۵ \times ۵ \rightarrow ۲۷۵$
$۶۷ \times ۶ \rightarrow ۴۰۲$
$۳۸ \times ۴ \rightarrow ۱۵۲$
$۴۹ \times ۵ \rightarrow ۲۴۵$
$۷۰ \times ۷ \rightarrow ۴۹۰$
$۸۰ \times ۹ \rightarrow ۷۲۰$
$۶۲ \times ۵ \rightarrow ۳۱۰$
$۷۳ \times ۶ \rightarrow ۴۳۸$
$۸۳ \times ۹ \rightarrow ۷۴۷$
$\sum xy = ۳۹۵۱$

$$r_{xy} = \frac{n\sum xy - (\sum x)(\sum y)}{\sqrt{\left[n\sum x^۲ - (\sum x)^۲\right]\left[n\sum y^۲ - (\sum y)^۲\right]}} = \frac{۱۰(۳۹۵۱) - (۶۲۰)(۶۰)}{\sqrt{\left[۱۰(۴۰۵۷۰) - (۶۲۰)^۲\right]\left[۱۰(۳۹۰) - (۶۰)^۲\right]}}$$

$$r_{xy} = \frac{۳۹۵۱۰ - ۳۷۲۰۰}{\sqrt{[۴۰۵۷۰۰ - ۳۸۴۴۰۰][۳۹۰۰ - ۳۶۰۰]}} = \frac{۲۳۱۰}{\sqrt{۶۳۹۰۰۰۰}} = \frac{۲۳۱۰}{۲۵۲۸} \Rightarrow r_{xy} = ۰/۹۱$$

روش دوم: این روش نسبتاً کوتاه‌تر و ساده‌تر از روش اول است. در این روش میانگین را حساب کرده و باتوجه‌به انحراف هر یک از نمره‌های x و y از میانگین، ضریب همبستگی را محاسبه می‌کنیم:

$$r_{xy} = \frac{\sum xy}{\sqrt{(\sum x^۲)(\sum y^۲)}}$$ (ضریب همبستگی پیرسون)

که در فرمول فوق x انحراف هر یک از نمره‌های X از میانگین آنها $(x - \overline{x})$، y انحراف هر یک از نمره‌های Y از میانگین آنها $(y - \overline{y})$ و Σ مجموع داده‌ها می‌باشد.

◆ ◆ ◆ ◆

مثال ۷: ضریب همبستگی پیرسون را بااستفاده‌از روش دوم برای داده‌های جدول زیر محاسبه و گزینه صحیح را انتخاب کنید.

انعطاف‌پذیری X	چابکی Y
۴۳	۴
۵۵	۵
۶۷	۶
۳۸	۴
۴۹	۵
۷۰	۷
۸۰	۹
۶۲	۵
۷۳	۶
۸۳	۹
$\sum x = ۶۲۰$	$\sum y = ۶۰$
$\overline{x} = \frac{۶۲۰}{۱۰} = ۶۲$	$\overline{Y} = \frac{۶۰}{۱۰} = ۶$

۱) ۰/۷۱

۲) ۰/۸۱

۳) ۰/۶۱

۴) ۰/۹۱

☑ **پاسخ:** گزینه «۴» در ابتدا میانگین x و y ها را محاسبه کرده، سپس اختلاف هر یک از نمره‌ها را با میانگین آنها به‌دست می‌آوریم و در ستون جدید x و y قرار می‌دهیم (مطابق بالا). سپس x و y های جدید را به توان ۲ می‌رسانیم و مجموع هر یک از آنها را نیز محاسبه می‌کنیم:

	x	x^2		y	y^2	xy
۴۳ − ۶۲	−۱۹	۳۶۱	۴ − ۶	−۲	۴	۳۸
۵۵ − ۶۲	−۷	۴۹	۵ − ۶	−۱	۱	۷
۶۷ − ۶۲	+۵	۲۵	۶ − ۶	۰	۰	۰
۳۸ − ۶۲	−۲۴	۵۷۶	۴ − ۶	−۲	+۴	۴۸
۴۹ − ۶۲	−۱۳	۱۶۹	۵ − ۶	−۱	۱	۱۳
۷۰ −۶۲	+۸	۶۴	۷ − ۶	۱	۱	۸
۸۰ −۶۲	+۱۸	۳۲۴	۹ − ۶	۳	۹	۵۴
۶۲ − ۶۲	۰	۰	۵ − ۶	−۱	۱	۰
۷۳ − ۶۲	۱۱	۱۲۱	۶ − ۶	۰	۰	۰
۸۳ − ۶۲	+۲۱	۴۴۱	۹ − ۶	+۳	۹	۶۳
	$\sum x = ۰$	$\sum x^2 = ۲۱۳۰$		$\sum y = ۰$	$\sum y^2 = ۳۰$	$\sum xy = ۲۳۱$

$$r_{xy} = \frac{\sum xy}{\sqrt{(\sum x^2)(\sum y^2)}} \Rightarrow \frac{۲۳۱}{\sqrt{(۲۱۳۰)(۳۰)}} = \frac{۲۳۱}{\sqrt{۶۳۹۰۰}} = \frac{۲۳۱}{۲۵۲} = ۰/۹۱ \Rightarrow r_{xy} = ۰/۹۱$$

◆ ◆ ◆ ◆

مثال۸: نمره‌های ۵ نفر از دانشجویان که به‌طور تصادفی انتخاب شده‌اند، در دو متغیر روزهای غیبت (x) و میزان اضطراب (y) در جدول زیر آمده است. ضریب همبستگی بین این دو متغیر کدام گزینه است؟

میزان اضطراب (Y)	روزهای غیبت (X)
۱۴	۲
۱۰	۰
۸	۳
۶	۶
۲	۴
$\sum y = ۴۰$	$\sum x = ۱۵$

۱) ۰/۵۱
۲) −۰/۷۱
۳) −۰/۶
۴) ۰/۱۶

☑ **پاسخ:** گزینه «۳» ما برای به‌دست آوردن ضریب همبستگی می‌توانیم از دو روش استفاده کنیم که هر دو راه، جواب یکسانی را به ما می‌دهند. از روش اول استفاده کرده و در کنار جدول بالا، سه ستون جدید تحت عنوان x^2، y^2 و xy ایجاد می‌کنیم و مجموع هر یک را $(\sum x^2, \sum y^2, \sum xy)$ محاسبه می‌کنیم. بعد از انجام دادن عملیات فوق، اعداد جدید به‌دست آمده را در فرمول زیر قرار می‌دهیم و ضریب همبستگی را محاسبه می‌کنیم:

x^2	y^2	xy
۴	۱۹۶	۲۸
۰	۱۰۰	۰
۹	۶۴	۲۴
۳۶	۳۶	۳۶
۱۶	۴	۸
$\sum x^2 = ۶۵$	$\sum y^2 = ۴۰۰$	$\sum xy = ۹۶$

$$r_{xy}=\frac{N\sum xy-(\sum x)(\sum y)}{\sqrt{\left[N\sum x^{۲}-(\sum x)^{۲}\right]\left[N\sum y^{۲}-(\sum y)^{۲}\right]}}=\frac{۵(۹۶)-(۱۵)(۴۰)}{\sqrt{\left[۵(۶۵)-(۱۵)^{۲}\right]\left[۵(۴۰۰)-(۴۰)^{۲}\right]}}$$

$$r_{xy}\frac{۴۸۰-۶۰۰}{\sqrt{[۳۲۵-۲۲۵][۲۰۰۰-۱۶۰۰]}}=\frac{-۱۲۰}{\sqrt{۴۰۰۰۰}}\Rightarrow\frac{-۱۲۰}{۲۰۰}=-۰/۶\Rightarrow r_{xy}=-۰/۶$$

برای توصیف و تفسیر تقریبی ضریب همبستگی از جدول زیر می‌توان استفاده کرد:

توصیف و تفسیر	r
همبستگی بسیار پایین و ضعیف	۰ – ۰/۱۹
همبستگی پایین	۰/۲۰ – ۰/۳۹
همبستگی متوسط	۰/۴۰ – ۰/۶۹
همبستگی بالا	۰/۷۰ – ۰/۸۹
همبستگی بسیار بالا و قوی	۰/۹۰ – ۱

مثال۹: کدام‌یک از گزینه‌های زیر ضریب همبستگی بالاتری را نسبت به دیگر گزینه‌ها نشان می‌دهد؟

۱) ۱/۹ ۲) ۰/۹ – ۳) ۰/۹ + ۴) گزینه ۲ و ۳

☑ **پاسخ:** گزینه «۴» ضریب همبستگی بین ۱+ و ۱– قرار می‌گیرد که منفی آن نشان‌دهنده رابطه معکوس متغیرها با هم و مثبت آن نشان‌دهنده رابطه مستقیم آنها با هم است.

نکته۲: متداول‌ترین شاخص محاسبه همبستگی، ضریب همبستگی گشتاوری پیرسون است. در ضریب پیرسون، دو متغیر با مقیاس فاصله‌ای یا نسبی اندازه‌گیری می‌شوند.

تفسیر ضریب همبستگی

ضریب همبستگی جهت و شدت رابطه بین دو متغیر را تعیین می‌کند. اگر چه این ضریب به‌صورت اعشاری بیان می‌شود، ولی تفسیر آن نباید برحسب درصد باشد. ضریب همبستگی را نمی‌توان به‌صورت نسبت مقایسه و تفسیر کرد؛ به‌عنوان مثال نمی‌توان گفت ضریب همبستگی ۰/۹ دقیقاً دو برابر ضریب همبستگی ۰/۴۵ است.

مسأله دیگری که باید مورد توجه قرار گیرد این است که همبستگی بین چند متغیر ضرورتاً نشان‌دهنده رابطه علت و معلولی بین متغیرها نیست و همبستگی بین دو متغیر نباید به این صورت تفسیر شود که یک متغیر، تنها علت متغیر دیگر است. همبستگی بین متغیرها این مطلب را بیان می‌کند که تغییر در یک متغیر هماهنگ با تغییر در متغیر دیگر هست یا خیر.

مثال۱۰: کدام عبارت صحیح نمی‌باشد؟

۱) ضریب همبستگی نشان‌دهنده رابطه علت و معلولی نیست.
۲) ضریب همبستگی جهت و شدت رابطه را تعیین می‌کند.
۳) ضریب همبستگی ۰/۸، دو برابر ضریب ۰/۴ است.
۴) گزینه‌های ۱ و ۲

☑ **پاسخ:** گزینه «۳» ضریب همبستگی را نمی‌توان به‌صورت نسبت مقایسه و تفسیر کرد.

ضریب تعیین (تشخیص)

ضریب همبستگی، اندازه و مقدار همبستگی بین دو متغیر را نشان می‌دهد؛ در حالی که درباره ماهیت همبستگی اطلاعات زیادی به ما نمی‌دهد. ضریب تشخیص از فرمول $V=r^{۲}(۱۰۰)$ محاسبه می‌شود که تعیین می‌کند چنددرصد از کل واریانس x، ناشی از واریانس y است و بالعکس. به عبارت دیگر ضریب همبستگی که به‌صورت درصد اعلام می‌شود، **ضریب تشخیص** نامیده می‌شود. در تربیت‌بدنی، از ضریب تشخیص به‌عنوان یک شاخص مهم بین

مدرسان شریف

فصل نهم

«ویژگی‌های آزمون و کلیات نمره‌دهی در تربیت بدنی»

روایی یا اعتبار

هر آزمون یا ابزار اندازه‌گیری باید دارای ویژگی‌هایی باشد تا بتواند توسط محقق استفاده شود. سه ویژگی اصلی یک ابزار یا آزمون عبارت‌انداز: روایی، پایایی و عینیت. آزمون وقتی دارای روایی (Validity) است که رفتار یا خصیصه معینی را که به‌منظور اندازه‌گیری آن ساخته شده است، اندازه بگیرد. روایی آزمون نشان می‌دهد که آیا آزمون، منحصراً آنچه را که موردنظر است اندازه می‌گیرد یا نه؟ و آیا خصوصیت موردنظر را به‌طور کامل اندازه‌گیری می‌کند یا نه؟ انواع روایی عبارت است از:

۱ـ روایی وابسته به ملاک (Criterion-Related validity) که به دو دسته تقسیم می‌شود:

۱ـ۱ـ روایی پیش‌بین: نشان می‌دهد که نتیجه آزمون تا چه اندازه می‌تواند موفقیت افراد را در زمینه‌ای که آزمون برای اندازه‌گیری آن ساخته شده است، پیش‌بینی کند. برای برآورد روایی پیش‌بین آزمون، می‌توان ضریب همبستگی نتایج آن آزمون را در مورد گروه نمونه‌ی آماری با میزان موفقیت آنها در آینده و در همان زمینه محاسبه کرد. معمولاً از این نوع روایی برای آزمون‌هایی که **جهت استخدام و گزینش افراد** تهیه می‌شوند، استفاده می‌کنند. یکی از مشکلات روایی پیش‌بین، ارزیابی ملاک پیش‌بین است؛ مثلاً اندازه‌گیری موفقیت در یک تیم، کار دشواری است. مربی ممکن است در ارزیابی میزان کارایی یک ورزشکار نظرات متفاوتی ابراز کند، به علاوه موفقیت افراد در یک تیم تحت‌تأثیر عوامل بسیاری مانند رضایت شغلی، محیط تیمی، روابط بازیکنان با مربی، دستمزد، تشویق و ... قرار می‌گیرد. در پیش‌بینی میزان موفقیت تحصیلی نیز مشکل فوق وجود دارد. موفقیت تحصیلی افراد بر مبنای نمرات درسی آنها تعیین می‌شود، درحالی‌که روش نمره‌دهی از معلمی به معلم دیگر فرق می‌کند و نیز موفقیت دانش‌آموزان تحت‌تأثیر عوامل زیادی قرار می‌گیرد که معمولاً خارج از کنترل محققان و طراحان آزمون است.

نکته۱: برای تعیین روایی پیش‌بین می‌توان از روش ضریب همبستگی رتبه‌ای اسپیرمن استفاده کرد.

۱ـ۲ـ روایی همزمان: در هر دو نوع روایی همزمان و پیش‌بین، روش ضریب همبستگی نتایج آزمون با نتایج ملاکی در همان زمینه محاسبه می‌گردد. **تفاوت عمده این دو روش در زمان به‌دست آوردن** نتایج ملاکی است. برای تعیین روایی همزمان، نتیجه اجرای آزمون را با نتایج ملاکی که در همان زمان وجود دارد مقایسه می‌کنند؛ برای مثال، برای محاسبه ضریب روایی همزمان یک آزمون جدید می‌توان ضریب همبستگی نمرات آن را در مورد یک گروه نمونه آماری با نمرات همان افراد در آزمون دیگری که همان خصیصه را اندازه می‌گیرد و روایی پیش‌بین آن قبلاً تعیین شده است، محاسبه کرد. در این مورد نتایج آزمون و نیز نتایج ملاکی تقریباً در یک زمان به‌دست می‌آید. از روایی همزمان بیشتر برای **آزمون‌های روانی و آزمون‌هایی که کاربرد بالینی دارند،** استفاده می‌کنند.

نکته۲: با استفاده از روش همبستگی پیرسون، اعتبار هم‌زمان تعیین می‌شود.

مثال۱: از کدام روایی بیشتر برای آزمون‌های روانی و بالینی استفاده می‌کنند؟

۱) روایی پیش‌بین ۲) روایی صوری ۳) روایی محتوا ۴) روایی همزمان

☑ **پاسخ:** گزینه «۴» از روایی هم‌زمان بیشتر برای آزمون‌های روانی و آزمون‌هایی که کاربرد بالینی دارند، استفاده می‌کنند.

◆ ◆ ◆ ◆

۲ـ روایی محتوا(Content Validity): روایی محتوا به‌طورکلی به بررسی منظم محتوای آزمون به‌منظور پاسخ دادن به این سؤال است که آیا سؤالات آزمون، محتوای کامل موضوعی را که آزمون برای سنجش و اندازه‌گیری آن ساخته شده است، دربرمی‌گیرد یا خیر. این روش بیشتر برای آزمون‌های فراگیری(پیشرفت تحصیلی) و آموزشی کاربرد دارد.

آزمون فراگیری موقعی دارای روایی محتواست که سؤالات ما نمونه خوبی از منبع سؤالات باشد. روایی پیش‌بین براساس روش‌های آماری انجام می‌شود و نتیجه به‌صورت یک ضریب عددی تعیین می‌گردد و ظاهراً جنبه‌ی عینی دارد، درحالی‌که تعیین روایی محتوا از طریق تجزیه و تحلیل سؤالات با استفاده از جدول دو بُعدی (هدف‌ها ـ محتوا) صورت می‌گیرد که معمولاً جنبه عقلی و منطقی دارد و تعیین آن براساس قضاوت ذهنی و فردی است.

دو نوع روایی محتوا داریم:

۱ـ۲ـ روایی صوری: اگر فرد یا افراد متخصصی آزمونی را بررسی کنند و نتیجه بگیرند که آزمون، صفت موردنظر را اندازه‌گیری می‌کند، روایی صوری حاصل شده است. این روایی یک بررسی نسبتاً سطحی از ظاهر سؤالات و آزمون می‌باشد.

۲ـ۲ـ روایی منطقی: مشتمل‌بر تعریف دقیق از حیطه رفتار مورد اندازه‌گیری آزمون و طرح منطقی سؤال‌هایی که کلیه بخش‌های عمده حیطه موردنظر را می‌پوشاند.

مثال۲: کدام روایی از طریق تجزیه و تحلیل سؤالات صورت می‌گیرد و جنبه عقلی دارد؟

۱) روایی پیش بین ۲) روایی محتوا ۳) روایی همزمان ۴) گزینه‌های ۱ و ۳

☑ **پاسخ:** گزینه «۲» روایی پیش‌بین براساس روش‌های آماری انجام می‌گیرد و جنبه‌ی عینی دارد؛ درحالی‌که روایی محتوا از طریق تجزیه و تحلیل سؤالات با استفاده از جدول دوبُعدی صورت می‌گیرد.

۳ـ روایی سازه (ساختاری ـ مفهومی) (construct validity): روایی سازه معمولاً برای آزمون‌ها و متغیرهایی که ملاک خارجی ندارند، به‌کار می‌رود؛ مثل هوش، پرخاشگری نهفته، اضطراب، انگیزه و غیره. تعیین روایی سازه همانند روایی محتوا به‌صورت ضریب همبستگی محاسبه نمی‌شود و ضریب روایی برای آن تعیین نمی‌شود، بلکه از طریق فرض، آزمایی و آزمون پیش‌بینی‌ها در مورد یک نظریه، روایی سازه را مشخص می‌سازیم. به عبارت دیگر روایی سازه در رابطه با نظریه است. روایی سازه به دو شکل واگرا (طراحی ۴ آزمون برای سنجش ۴ صفت) و همگرا (طراحی ۴ آزمون برای سنجش یک صفت) وجود دارد. **در روایی واگرا، همبستگی پایین بین آزمون‌ها و در روایی همگرا، همبستگی بالا بین آزمون‌ها شرط روایی آزمون می‌باشد.**

این روایی معمولاً بیانگر آن است که آزمون تا چه اندازه، ویژگی یا صفت خاصی را اندازه می‌گیرد. برای تعیین این روایی، آزمون خاصی روی دو گروه آزمون‌شونده کاملاً متفاوت (مثلاً یک گروه مبتدی و یک گروه پیشرفته) اجرا شده، نتایج دو اجرا از طریق محاسبه تست t مستقل با یکدیگر مقایسه می‌شوند. هراندازه که اختلاف این دو گروه نمره معنی‌دارتر باشد، نشانه بالاتر بودن روایی سازه آزمون است و نشان می‌دهد آزمون یا ابزار، قدرت تشخیص گروه‌های متفاوت را از یکدیگر دارد.

عوامل تأثیرگذار بر روایی یا اعتبار:

۱) فاصله زمانی تکرار آزمون

۲) آمادگی و عدم آمادگی آزمون‌شونده

۳) میزان یادگیری ایجادشده ناشی از اجرای آزمون در مرحله اول

۴) انگیزه و علاقه آزمون‌شونده

۵) شرایط محیطی متفاوت

نکته۳: فرمول محاسبه ضریب اعتبار: $r_{xy} = \sqrt{(r_{xx})(r_{yy})}$

r_{xy} = همبستگی بین آزمون x و y

r_{xx} = ضریب پایایی آزمون x

r_{yy} = ضریب پایایی آزمون y

مثال۳: از کدام نوع روایی برای آزمون‌های استخدام و گزینش افراد استفاده می‌شود؟

۱) روایی هم‌زمان ۲) روایی پیش‌بین ۳) روایی منطقی ۴) روایی صوری

☑ **پاسخ:** گزینه «۲» از روایی پیش‌بین برای آزمون‌هایی که جهت استخدام و گزینش افراد تهیه می‌شوند، استفاده می‌کنند.

پایایی (ثبات) آزمون (Reliability)

منظور از پایایی آزمون، دقت اندازه‌گیری، ثبات و پایداری آن است؛ به عبارت دیگر پایایی آزمون نشان می‌دهد که آزمون، صفت مورد اندازه‌گیری را با چه دقت و صراحتی اندازه می‌گیرد و نتیجه اندازه‌گیری تا چه‌حد دارای ثبات و پایداری است. به عبارت دیگر پایایی آزمون تفاوت نمره مشاهده‌شده آزمودنی را از نمره‌ی واقعی او نشان می‌دهد. از نظر تئوری اندازه‌گیری، نمره‌ی آزمون، که در اصطلاح به آن نمره مشاهده‌شده می‌گویند، از دو اندازه تشکیل شده است:

۱ـ نمره واقعی ۲ـ خطاهای اندازه‌گیری.

منظور از نمره مشاهده‌شده، نمره‌ای است که دانش‌آموز در نتیجه گذراندن آزمونی به‌دست می‌آورد که از هر نوع خطای اندازه‌گیری دور باشد. هرقدر خطای اندازه‌گیری در یک آزمون کمتر باشد. به‌همان‌نسبت دقت اندازه‌گیری و ثبات و پایداری نمره آزمون بیشتر است و بالعکس. دو دسته از عوامل سبب به‌وجود آمدن خطای اندازه‌گیری می‌شوند:

۱ـ عوامل بیرونی: که به شرایط و نحوه اجرای آزمون مرتبط است و عبارت‌اند از: نامساعد بودن جلسه آزمون، بیماری، عدم تعادل روانی آزمودنی،حدس زدن پاسخ‌ها، اشتباه در نمره‌گذاری پاسخ‌ها، خستگی و... .

۲ـ عوامل درونی:که مستقیماً به کیفیت آزمون مربوط است و عبارت‌اند از: محدود بودن سؤالات آزمون، عدم تجانس یا همگنی سؤالات با یکدیگر، طولانی بودن آزمون، دشواری آزمون و ... .

مثال۴: کدام‌یک از موارد زیر ثبات آزمون را افزایش می‌دهد؟

۱) اضافه کردن سؤالات با ضریب تمیز بالا و ضریب دشواری پایین
۲) اضافه کردن سؤالات با ضریب تمیز بالا و ضریب دشواری بالا
۳) اضافه کردن سؤالات با ضریب تمیز بالا و ضریب دشواری متوسط
۴) اضافه کردن سؤالات با ضریب تمیز متوسط و ضریب دشواری بالا

☑ **پاسخ:** گزینه «۳» یکی از مواردی که باعث افزایش ثبات آزمون می‌شود، اضافه کردن سؤالاتی با ضریب تمیز بالا و ضریب دشواری متوسط است.

◆ ◆ ◆ ◆

برای برآورد پایایی آزمون از روش‌های زیر استفاده می‌شود:

۱ـ بازآزمایی (آزمون ـ آزمون مجدد):

ساده‌ترین روش تعیین پایایی یک آزمون، بازآزمایی (**test − retest**) است. در این روش، آزمون را در دو نوبت از آزمودنی‌ها به عمل می‌آورند و نمرات حاصل‌شده را با هم مقایسه می‌کنند. ضریب همبستگی بین نمرات حاصل از دو بار اجرای آزمون، ضریب پایایی آزمون می‌باشد، البته در این روش عواملی چون فراموشی، یادگیری‌های تازه، عامل رشد (اگر فاصله دو آزمون بسیار زیاد باشد) ، شرایط متفاوت اجرای آزمون و ... در برآورد پایایی تأثیر می‌گذارند.

مثال۵: منظور از ..... آزمون، دقت اندازه‌گیری، ثبات و پایداری آن است.

۱) روایی
۲) Validity
۳) ثبات
۴) عینیت

☑ **پاسخ:** گزینه «۳» منظور از پایایی (ثبات) آزمون (Reliability) دقت، ثبات و پایداری اندازه‌گیری است.

◆ ◆ ◆ ◆

۲ـ فرم‌های موازی (همتا):

این روش مستلزم اجرای دو فرم موازی (**Parallel forms**) یک آزمون در مورد گروهی واحد و محاسبه ضریب همبستگی بین نمرات آزمون است. کاربرد دو فرم موازی به جای بازآزمایی این است که از تغییرات ناشی از یادآوری و تمرین قبلی جلوگیری می‌شود. در این روش اگر اجرای دو آزمون بدون فاصله‌ی زمانی انجام شود، ممکن است عامل خستگی در نتیجه آزمون تأثیر بگذارد. در روش فرم‌های موازی معمولاً فرم اولیه آزمون ملاک درنظر گرفته می‌شود و هدف، تعیین پایایی فرم دوم است.

۳ـ تعیین همسانی درونی آزمون: که با سه روش زیر انجام می‌شود:

۱ـ۳ـ روش ضریب انسجام آزمون (کودر- ریچاردسون ۲۰ و ۲۱):

در این روش: آزمون فقط یک‌بار اجرا می‌شود و میزان همبستگی درونی یا تجانس سؤالات مورد بررسی قرار می‌گیرد. برای محاسبه همسانی درونی آزمون در صورتی‌که سؤالات از لحاظ دشواری همگن نباشند، از فرمول کودر ـ ریچاردسون ۲۱ استفاده می‌کنیم:

$$R_{۲۱} = \frac{k}{k-۱}\left[۱ - \frac{\bar{x}(k-\bar{x})}{\bar{x}S^{۲}}\right]$$

در فرمول فوق K تعداد سؤالات، $\bar{x}$ میانگین نمرات و $S^{۲}$ واریانس کل نمرات است.

مثال۶: اگر تعداد سؤالات آزمونی ۱۰۰ و میانگین نمرات ۲۰ و واریانس نمرات ۴۵ باشد، ضریب همبستگی درونی سؤالات کدام گزینه است؟

۱) ۱/۷۷ ۲) −۰/۷۷ ۳) ۱/۰۱ ۴) ۰/۳۳

☑ **پاسخ:** گزینه «۲» $R = ?$: خواسته مسأله، $S^2 = 45$ واریانس کل نمرات، $\overline{X} = 20$ میانگین نمرات، $K = 100$ تعداد سؤالات: داده‌های مسأله. با توجه به اطلاعات داده‌شده از فرمول کودر ـ ریچاردسون ۲۱ استفاده می‌کنیم:

$$R_{21} = \frac{k}{k-1}\left[1-\frac{\overline{x}(k-\overline{x})}{\overline{x}s^2}\right] \Rightarrow R_{21} = \frac{100}{100-1}\left[1-\frac{20(100-20)}{20\times 45}\right] \Rightarrow R_{21} = -0/77$$

نمره‌گذاری در این روش تابع قانون همه یا هیچ است؛ به این صورت که پاسخ درست، ارزش یک و پاسخ غلط و یا جواب‌نداده ارزش صفر دارد. زمانی که سؤالات آزمون از نظر دشواری همگن و یکسان باشد، از فرمول کودر ـ ریچاردسون ۲۰ استفاده می‌کنیم:

$$R_{20} = \frac{k}{k-1}\left[1-\frac{\sum pq}{S^2}\right]$$

در فرمول فوق k تعداد سؤالات، p نسبتی از آزمودنی‌ها که توانسته‌اند به آن سؤال خاص پاسخ درست دهند، q نسبتی از آزمودنی‌ها که نتوانسته‌اند به آن سؤال خاص پاسخ درست دهند و S^2 واریانس کل نمرات است.

نکته۱: هر چقدر دشواری مواد با هم ناهمگن‌تر باشد، واریانس کل و پراکندگی بیشتر است.

۲ـ۳ـ روش دو نیمه کردن آزمون(Split half):

گاهی اوقات از یک آزمون، دو فرم موازی در دست نیست و یا اینکه فرم موازی وجود دارد اما امکان اجرای آن در دو نوبت وجود ندارد. در این‌گونه موارد برای تعیین پایایی آزمون از روش زوج و فرد استفاده می‌کنیم. بدین ترتیب که وقتی آزمون به‌طور طبیعی اجرا و هر سؤال نمره‌گذاری شد، به‌جای محاسبه نمره‌ی کامل، آزمون را به دو قسمت تقسیم می‌کنیم و نمره هر قسمت را جداگانه به‌دست می‌آوریم. سپس بین دو سری نمرات، ضریب همبستگی می‌گیریم. ضریب به‌دست آمده، پایایی نصف آزمون خواهد بود. برای محاسبه‌ی پایایی تمام آزمون از فرمول ضریب همبستگی اسپیرمن استفاده می‌کنیم:

$$r_t = \frac{2r_{xy}}{1+r_{xy}} \quad \text{(ضریب پایایی کل آزمون)}$$

در فرمول فوق r_{xy} ضریب همبستگی بین دو نیمه آزمون است.

در این روش باید دقت شود که:

۱ـ هرگز نباید نتیجه یک آزمون را با نیمه دوم مقایسه کنیم، به‌خصوص اگر زمان اجرای آزمون محدود باشد. ۲ـ در این روش نباید تعداد سؤالات بسیار محدود و اندک باشد.

مثال۷: اگر ضریب همبستگی بین دو نیمه آزمون ۰/۶ باشد ، ضریب اعتبار آزمون کدام گزینه است؟

۱) ۰/۷۰ ۲) ۰/۸۰ ۳) ۰/۷۵ ۴) ۰/۸۵

☑ **پاسخ:** گزینه «۳» در مثال فوق ضریب همبستگی ۰/۶ است و ضریب اعتبار را می‌خواهیم. پس می‌توانیم با استفاده از فرمول دو نیمه کردن آزمون جواب را به‌دست آوریم. پس:

$$r_t = \frac{2r_{xy}}{1+r_{xy}} \Rightarrow r_t = \frac{2\times 0/6}{1+0/6} = \frac{1/2}{1/6} = 0/75 \Rightarrow r_t = 0/75$$

۳ـ۳ـ روش آلفای کرونباخ:

درصورتی‌که واریانس نمره‌های نیمه‌ها، مساوی نباشند و یا نشانه‌هایی درباره موازی نبودن نیمه‌ها وجود داشته باشد، از روش آلفای کرونباخ برای برآورد پایایی کل آزمون استفاده می‌کنیم. غالباً برای سنجیدن همسانی درونی بخش‌های مختلف پرسشنامه از آلفای کرونباخ استفاده می‌کنند. برخلاف روش کودر ـ ریچاردسون در نمره‌دهی، اینجا قانون همه یا هیچ مصداق پیدا نمی‌کند و پاسخ‌ها می‌توانند علاوه‌بر صفر و یک، اعداد دیگر را نیز بپذیرند.

$$r_{tt} = \frac{k}{k-1}\left(1-\frac{\sum S_i^2}{S^2}\right) \quad , \quad \sum S_i^2 = \sum pq$$

در فرمول فوق k تعداد سؤالات، S_i^2 واریانس هر سؤال و S^2 واریانس کل آزمون است.

ضریب آلفا مقادیر ۰ تا ۱ را می‌تواند اختیار کند. ارزش مقادیر مختلف ضریب آلفا در جدول زیر آمده است:

تفسیر	غیرقابل قبول	نامطلوب	حداقل پذیرفته شده	قابل قبول	خیلی خوب	عالی
α	$<0/6$	۰/۶۵ تا ۰/۶	۰/۷۵ تا ۰/۶۵	۰/۸۰ تا ۰/۷۰	۰/۹۰ تا ۰/۸۰	$>0/90$

مثال۸: در کدام‌یک از روش‌های تعیین Reliability، تمرین بیشترین تأثیر را در پراش خطای آزمون دارد؟

۱) split half ۲) parallel forms ۳) test - retest ۴) همه موارد

☑ **پاسخ:** گزینه «۳» در آزمون‌های بازآزمایی یا همان آزمون ـ آزمون مجدد (test - retest) فرد می‌تواند با استفاده از تمرین در آزمون دوم نتیجه بهتری را کسب کند.

رابطه بین روایی و پایایی

روایی یک آزمون نشان می‌دهد که آن آزمون تا چه‌اندازه برای اندازه‌گیری رفتاری که برای ارزیابی آن ساخته شده است، مناسب است و پایایی آزمون، دقت اندازه‌گیری، ثبات و پایداری آن را نشان می‌دهد. **داشتن پایایی شرط روایی آزمون است؛** به عبارت دیگر، آزمون نمی‌تواند دارای روایی باشد مگر آن که دارای پایایی کافی باشد. از طرف دیگر، بالا بودن ضریب پایایی آزمون، روایی آن را تضمین نمی‌کند؛ برای مثال، اگر ضریب پایایی آزمون، بالا (مثلاً ۰/۹۰) باشد، نمی‌توان گفت که این آزمون دارای روایی نیز هست. به‌طور کلی در ارزشیابی آزمون‌های استانداردشده، توجه به دو شرط اساسی آن یعنی **روایی و پایایی** اهمیت خاصی دارد.

مثال۹: برای سنجش همسانی درونی بخش‌های مختلف پرسشنامه از ..... استفاده می‌کند.

۱) فرم‌های موازی ۲) بازآزمایی ۳) آلفای کرونباخ ۴) هیچ‌کدام

☑ **پاسخ:** گزینه «۳» غالباً برای سنجیدن همسانی درونی بخش‌های مختلف پرسشنامه از آلفای کرونباخ استفاده می‌کنند.

عینیت آزمون

عینیت را می‌توان **توافق نزدیک بین دو یا چند داور در نمره دادن به هر آزمودنی** تعریف کرد؛ مانند داوران ژیمناستیک یا وقت‌نگهداران دوی سرعت ۱۰۰ متر. عینیت، ویژگی مهم آزمون یا ابزار اندازه‌گیری است که به دو عامل مرتبط با هم بستگی دارد:

۱ـ مشخص بودن روش نمره‌گذاری و دیگری درجه و شیوه‌ای که داور می‌تواند نمره‌های دقیق بدهد. آزمون‌های معینی مثل دوی ۱۰۰ متر و پرش طول، روش‌های نمره‌گذاری مشخصی (عینی) دارند. در این آزمون‌ها، نمره‌گذار به‌راحتی می‌تواند نمره‌های فرد را تعیین کند. در مقابل، آزمون‌های نوشتاری یا اجرای یک حرکت موزون ژیمناستیک، روش نمره‌گذاری دقیقی ندارند و به قضاوت نمره‌گذار در دادن امتیاز بستگی دارد.

۲ـ اگر داوران، نحوه‌ی نمره دادن را ندانند، مسلماً کار را به‌دقت انجام نخواهند داد. برای مثال وقت نگهداری که با زمان‌سنج‌ها آشنایی ندارد، احتمالاً رکورد دقیقی در یک مسابقه دوی سرعت ثبت نخواهد کرد. زمانی که دو یا چند نفر یک آزمون را اجرا می‌کنند، درجه بالایی از عینیت ضروری است. هرچقدر امتیازاتی که به هر آزمودنی توسط آزمون‌گیرنده‌ها داده می‌شود به‌هم نزدیک‌تر باشد؛ آزمون موردنظر عینی‌تر است؛ به عبارت دیگر نمرات با تغییرات مصحح دستخوش تغییر نمی‌شوند. با تعیین همبستگی بین نمرات به‌دست‌آمده از آزمون‌گیرنده‌ها دررابطه با یک گروه، می‌توان ضریب عینیت آزمون را به‌دست آورد.

مثال۱۰: اگر دو داور در انجام یک مهارت ورزشی به ورزشکار نمره نزدیکی دهند، بیان‌گر این است که:

۱) آزمون پایایی دارد. ۲) آزمون عینیّت دارد.

۳) آزمون هم پایایی دارد و هم عینیّت. ۴) هیچ‌کدام از موارد صحیح نیست.

☑ **پاسخ:** گزینه «۲» همان‌طور که گفته شد عینیّت را می‌توان توافق بین دو یا چند داور در نمره دادن به هر آزمودنی تعریف کرد. زمانی که دو داور در انجام یک مهارت ورزشی به ورزشکار نمره‌ی نزدیکی دهند، نشان‌دهنده این است که آزمون عینیّت دارد.

نکته ۲: عینیت آزمون به مهارت و توانایی آزمون‌کنندگان و عینی‌بودن آزمون در کسب نمره‌های یکسان در یک گروه مشابه بستگی دارد.

ضمیمه

جدول I: سطح زیر منحنی نرمال (Z)

1 z	2 A سطح تا میانگین	3 B سطح بزرگتر	4 C سطح کوچکتر	5 y ارتفاع
0.00	.0000	.5000	.5000	.3989
0.01	.0040	.5040	.4960	.3989
0.02	.0080	.5080	.4920	.3989
0.03	.0120	.5120	.4880	.3988
0.04	.0160	.5160	.4840	.3986
0.05	.0199	.5199	.4801	.3984
0.06	.0239	.5239	.4761	.3982
0.07	.0279	.5279	.4721	.3980
0.08	.0319	.5319	.4681	.3977
0.09	.0359	.5359	.4641	.3973
0.10	.0398	.5398	.4602	.3970
0.11	.0438	.5438	.4562	.3965
0.12	.0478	.5478	.4522	.3961
0.13	.0517	.5517	.4483	.3956
0.14	.0557	.5557	.4443	.3951
0.15	.0596	.5596	.4404	.3945
0.16	.0636	.5636	.4364	.3939
0.17	.0675	.5675	.4325	.3932
0.18	.0714	.5714	.4286	.3925
0.19	.0753	.5753	.4247	.3918
0.20	.0793	.5793	.4207	.3910
0.21	.0832	.5832	.4168	.3902
0.22	.0871	.5871	.4129	.3894
0.23	.0910	.5910	.4090	.3885
0.24	.0945	.5948	.4052	.3876
0.25	.0987	.5987	.4013	.3867
0.26	.1026	.6026	.3974	.3857
0.27	.1064	.6064	.3936	.3847
0.28	.1103	.6103	.3897	.3836
0.29	.1141	.6141	.3859	.3825
0.30	.1179	.6179	.3821	.3814
0.31	.1217	.6217	.3783	.3802
0.32	.1255	.6255	.3745	.3790
0.33	.1293	.6293	.3707	.3778
0.34	.1331	.6331	.3669	.3765
0.35	.1368	.6368	.3632	.3752
0.36	.1406	.6406	.3594	.3739
0.37	.1443	.6443	.3557	.3725
0.38	.1480	.6480	.3520	.3712
0.39	.1517	.6517	.3483	.3697
0.40	.1554	.6554	.3446	.3683
0.41	.1591	.6591	.3409	.3668
0.42	.1628	.6628	.3372	.3653
0.43	.1664	.6664	.3336	.3637
0.44	.1700	.6700	.3300	.3621