

مؤسسه علمی آموزشی
فرهیختگان راه دانش



آمار و روش تحقیق

درسنامه - نکات کلیدی - تست های فصل به فصل



مؤلف: مهدی رضایی
دکترای روانشناسی عمومی

به نام خالق

آمار و روش تحقیق

تألیف و گردآوری:

مهدی رضایی

«دکتری روان شناسی عمومی»

مقدمه:

مؤسسه علمی آموزشی فرهیختگان راه دانش با هدف ارائه کیفی ترین خدمات آموزشی و با تلاش گسترده توانست مجموعه‌ای از خدمات آموزشی را که از نظر علمی، به روز بودن مطالب، پوشش دادن مطالب رفرنس‌ها و بازدهی در زمره بهترین‌ها است ارائه دهد.

مشاوره و پشتیبانی تحصیلی:

مشکل عدیده ای که بیشتر داوطلبان با آن مواجه هستند و هر ساله با وجود صرف هزینه‌های مالی و زمان زیاد نمی توانند در آزمون قبول شوند به این دلیل می‌باشد که داوطلبان آگاهی کافی از منابع مطالعاتی، روشهای مطالعه و مرور مطالب صحیح، روشهای تست زنی و مدیریت زمان را ندارند بنابراین مؤسسه فرهیختگان جهت تحکیم رسالت خود که همواره ارتقاء کیفیت آموزش بوده است جمعی از برترین مشاورین و رتبه‌های تک رقمی را به خدمت گرفته است تا با ارائه منابع مطالعاتی کاربردی، آموزش روش‌های مطالعه و مرور مطالب هر درس، نحوه تست‌زنی صحیح و برنامه مطالعاتی روزانه و هفتگی به داوطلبان، آنها را از سردرگمی درآورده و با ایجاد انگیزه و تمرکز در داوطلبان سبب موفقیت آنها در آزمون گردد.

بسته‌های آموزشی مؤسسه:

بسته‌های آموزشی که به داوطلبان ارائه می‌گردد حاصل ماه‌ها تلاش بی پایان گروه علمی مؤسسه (که ترکیبی از رتبه‌های تک رقمی دکتری و کارشناسی ارشد و اساتید دانشگاه‌های تهران) می‌باشد که با در نظر گرفتن منابع وزارت بهداشت تالیف گردیده است. در این بسته‌ها تلاش شده است که درسنامه به صورت شرح جامعی از دروس ارائه گردد و جهت تفهیم بیشتر مطالب، نکات کلیدی منابع وزارت بهداشت و نکات تستی سوالات کنکور سال‌های اخیر نیز به درسنامه اضافه گردیده است و جهت محک و خودآزمایی داوطلبان، تست‌های هر فصل همراه با پاسخنامه گنجانده شده است. به این ترتیب بسته‌های آموزشی مؤسسه فرهیختگان را از نظر پوشش دادن سرفصل‌های آزمون به مجموعه‌ای کم نظیر تبدیل نموده به نحوی که داوطلب با مطالعه و جمع‌بندی بسته‌های آموزشی مؤسسه همراه با مطالعه منابع وزارت بهداشت براحتی پاسخ‌گوی بیشتر سوالات کنکور خواهد بود. بسته‌های آموزشی مؤسسه هر سال ویرایش و به روز گردیده و نکات، مطالب و تست‌های جدید نیز به آن اضافه می‌گردد.

آزمونهای آزمایشی :

داوطلبان رشته‌های مختلف بایستی جهت محک و خودآزمایی خود و جمع‌بندی مطالب بایستی برنامه ریزی مطالعاتی صحیح داشته باشند. مؤسسه با در نظر گرفتن شرایط داوطلبان مختلف اقدام به برگزاری آزمون‌های آزمایشی ۷ مرحله‌ای و ۳ مرحله‌ای در ۲۸ رشته نموده است.

۲ نکته بارزی که آزمون‌های آزمایشی موسسه فرهیختگان را از دیگر موسسات متمایز می‌نماید این است که در آزمون‌های آزمایشی موسسات دیگر، سوالات زبان به صورت جامع و کلی طرح می‌گردد که این موضوع سبب سردرگمی داوطلبان گردیده و داوطلبان نمی‌دانند مطالعه درس زبان انگلیسی را از کدام منبع مطالعاتی شروع کنند، به همین دلیل اکثریت قریب به اتفاق داوطلبان مطالعه درس زبان را رها نموده و این موضوع لطمه بزرگی به داوطلب وارد می‌کند به نحوی که ممکن است داوطلب در چندین درس یک رشته تسلط کافی داشته باشد و در آزمون اصلی نیز درصدهای خوبی را کسب کرده باشد ولی با توجه به اینکه درس زبان را مطالعه نکرده معمولاً این درس را سفید و یا درصد بسیار ضعیفی کسب نماید که این مقوله سبب عدم قبولی داوطلب با وجود شایستگی‌های علمی وی می‌گردد. موسسه فرهیختگان جهت برطرف نمودن این مشکل و چه بسا معضل، اقدام به ارائه طرح درس و سرفصل زبان انگلیسی در آزمون‌های آزمایشی خود نموده تا داوطلبان بتوانند با برنامه ریزی صحیح مطالعه زبان انگلیسی (که ضریب بالایی دارد) را انجام داده و دچار سردرگمی نشوند، این روش سبب می‌شود که داوطلب با طبقه بندی مبحثی، درس زبان را مطالعه نمایند.

نکته دوم اینست که فواصل زمانی آزمون‌های آزمایشی (۴ مرحله طبقه بندی و ۳ مرحله جامع) با توجه به حجم مطالب تنظیم گردیده است، تا داوطلب بتواند با مطالعه بدون استرس و صحیح و مرور و جمع بندی مطالب به آمادگی کامل دست یابد. **داوطلبان می‌توانند بعد از ثبت نام جزوه روش‌های مطالعه صحیح، روشهای مرور و تست‌زنی را به صورت رایگان از موسسه دریافت نمایند.**

کلاسهای آمادگی :

با توجه به این که بیشتر دانشجویان در دانشگاه به دلیل ساعات کلاسی کم، موفق به یادگیری مطالب دروس تخصصی نمی‌شوند و با مطالعه چند باره جزوات نیز، بسیاری از نکات برای آنها قابل فهم و یادگیری نمی‌باشد. موسسه فرهیختگان با در نظر گرفتن شرایط داوطلبانی که امکان استفاده از کلاس‌های آمادگی حضوری را ندارند اقدام به تهیه و تدوین DVDهای آموزشی (با استفاده از تدریس اساتید برتر دانشگاه‌های تهران) در دروس مختلف نموده است. سبک تدریس در این کلاس‌ها بمانند کلاس‌های حضوری شامل شرح درس، نکته گویی و حل تست می‌باشد.

داوطلبان رشته‌های مختلف می‌توانند جهت بهره‌گیری از خدمات آموزشی موسسه (بسته‌های آموزشی، آزمون‌های آزمایشی، کلاس‌های آمادگی و مشاوره و پشتیبانی تحصیلی) می‌توانند به نمایندگی‌های سراسر کشور مراجعه نموده و یا با دفتر مرکزی موسسه ۲۴ ۹۵ ۹۷ ۶۶ - ۰۲۱ تماس حاصل فرمایند.

امید است که در سایه حق تعالی و بهره‌مندی از تلاش خود و خدمات آموزشی موسسه شما عزیزان به موفقیت‌های بزرگتری دست یابید.

با آرزوی موفقیت

مدیریت موسسه فرهیختگان راه دانش

مؤسسه علمی آموزشی
فرهیختگان راه دانش



بخش اول

آمار توصیفی

فصل اول

آمار اندازه گیری و خلاصه بندی داده ها

روشهای آماری دارای دو وظیفه هستند:

۱-۱: روش های آماری در طبقه بندی، خلاصه کردن، توصیف و تفسیر اطلاعات جمع آوری شده و برقراری ارتباط از طریق آنها، به پژوهشگر کمک می کنند. روش های آماری که برای این وظیفه بکار می روند آمار توصیفی^۱ نامیده می شوند.

۲-۲: به پژوهشگران امکان می دهند که با استفاده از اطلاعات جمع آوری شده از نمونه کوچکی از آزمودنی ها، ویژگیهای جامعه ای را که نمونه از آن انتخاب شده است برآورد یا استنباط کنند. روشهای آماری مربوط به این وظیفه را آمار استنباطی^۲ می نامند. در آمار استنباطی هدف پژوهشگر عبارت از تعمیم اصول و یافته ها.

هدف نهائی آمار استنباطی برآورد ویژگیهای جامعه است. ویژگیهای عددی یک جامعه را « پارامتر » می نامند. مثلاً متوسط هوش کودکان کودکان کدکستانی شهر تهران یک پارامتر به حساب می آید. به همین ترتیب اندازه هائی که از یک نمونه بدست می آیند. « آماره » و یا شاخص آماری نامیده می شوند. آماره عبارت از ویژگی های کمی که یک نمونه را توصیف می کند. بطور خلاصه می توان گفت که نسبت آماره به نمونه مثل نسبت پارامتر به جامعه است و آماره ، پارامتر جامعه را برآورد کرد.

جامعه: به گروهی از افراد یا اشیا گفته می شود که حداقل دارای یک ویژگی مشترک هستند.

انواع مقیاس های اندازه گیری

مهمترین انواع اندازه گیری از ضعیف تر به قوی تر به شرح زیر است :

۱-مقیاس های طبقه ای :

این نوع اندازه گیری همان عمل طبقه بندی است. وقتی یک مجموعه اشیا را در چند طبقه تقسیم می کنیم می توانیم هر یک از طبقه ها را با اعداد ۱ ، ۲ ، ۳ و غیره نشان می دهیم. این اعداد مفهومی جز نامیدن طبقه ندارند و از خصوصیات انتزاعی بهمین جهت می توان بجای اعداد از علامت های قراردادی دیگر مانند A ، B و C نیز استفاده کرد. در این مقیاس، طبقات اولاً م/نعه /جمع هستند یعنی یک نفر را نمی توان در بیش از یک طبقه جای داد ، ثانیاً مقیاس باید کامل باشد یعنی در هر طبقه کلیه افراد همنوع قرار داده شود. در این مقیاس فقط از علائم ریاضی = و $\square$ استفاده می شود. در این مقیاس فراوانی طبقه ها با روشهای آماری قابل تحلیل هستند. برای متغیرهای طبقه ای دو نوع مقیاس وجود دارد، مقیاس اسمی^۳ و مقیاس طبقه ای مرتب شده^۴. مقیاس اسمی دارای طبقه های مشخص و

^۱ - descriptive Statistics

^۲ - Inferential Statistics

^۳ -nominal Scale

^۴ -Ordered eategorical scale

متمایزی است که جنبه کیفی دارد و تنها رابطه موجود تفاوت آنها با یکدیگر است رنگ چشم مثالی برای یک متغیر در مقیاس اسمی است. نوع دیگر مقیاس طبقه ای آن در یک معنای کمی بایکدیگر ارتباط دارد و بعلاوه با طبقه های مرتب بیان می شود. مثلاً در پرسشنامه ای که هر سوال دارای طبقه های خیلی زیاد ، زیاد ، متوسط ، کم، خیلی کم است.

۲-۲- مقیاس ترتیبی^۱:

در این مقیاس طبقه های ثابتی وجود ندارد که مشاهده ها را به توان به آنها اختصاص داد. مشاهده ها با یکدیگر مقایسه می شود و به ترتیب معینی مثلاً از بزرگترین تا کوچکترین به آنها رتبه داده می شود. از این نظر داده های مقیاس ترتیبی با داده های مقیاس طبقه ای مرتب شده شباهت دارد. بطور کلی اعداد در مقیاس ترتیبی، فقط سلسله مراتب اشیاء یا رویدادها را از لحاظ خصیصه ای معین یک گروه نشان می دهد و بی آنکه درباره فاصله بین افراد مطلبی بیان کند، وضع نسبی آنها را از لحاظ یک خصیصه مشخص می سازد. در این مقیاس از علائم و قراردادهای ریاضی فقط $=$ ، $\square$ ، $>$ و $<$ استفاده می شود.

۳-۳- مقیاس فاصله ای^۲:

مقیاس فاصله ای علاوه بر طبقه بندی، نامگذاری و مرتب کردن طبقه ها امکان تشخیص فاصله های موجود بین افراد، اشیاء یا رویدادها را نیز فراهم می کند. در این مقیاس علاوه بر مواردی که در مقیاسهای طبقه ای و ترتیبی بیان شد عملیات ریاضی جمع و تفریق نیز انجام پذیر است. به عنوان مثال می توانیم اختلاف نمره های دانش آموزان رتبه سوم و هشتم و یا هر رتبه دیگری را مشخص کنیم.

یکی از محدودیت های مقیاس فاصله ای نداشتن صفر مطلق است. لذا نمی توان فرض کرد که مثلاً در یک آزمون پیشرفت تحصیلی نمره دانش آموز رتبه سوم معادل دو برابر رتبه ششم است. و یا اگر دانش آموزی از یک آزمون هوش نمره صفر بگیرد، به این معنا نیست که او اصولاً هوش ندارد.

۴-۴- مقیاس نسبی^۳:

مقیاس نسبی بالاترین سطح اندازه گیری است که به دلیل داشتن صفر مطلق، در این مقیاس کلیه عملیات ریاضی شامل، جمع، تفریق، ضرب، تقسیم و ... انجام پذیر است استفاده از این مقیاس در علوم فیزیکی کاربرد معنادارتری دارد.

انواع متغیرهای اندازه گیری

متغیر: ویژگی یا صفتی که می تواند اندازه های متفاوتی را به خود اختصاص بدهد.

۱- کمیت های پیوسته^۴: (متغیر متصل)

^۱ - Ordinal Scale

^۲ - Interval Scal

^۳ - Ratio Scale

^۴ - Continuous

این کمیت حاصل اندازه گیری هستند و امکان اینکه بین دو واحد متوالی (عدد صحیح) کمیت‌های دیگر نیز قرار بگیرند وجود دارد مثلاً بین قد ۱۶۰ و ۱۶۱ سانتیمتر ممکن است افرادی دارای قد ۱۶۰/۲۵ ، ۱۶۰/۵ ، ۱۶۰/۸۰ و یا ... باشند.

۲- کمیت‌های ناپیوسته (گسسته)^۱: (متغیر منفصل)

کمیت‌های ناپیوسته به اعدادی اطلاق می شود که از آمارگیری و یا شمارش بدست می آیند و بین دو واحد متوالی آنها عدد یا کمیت دیگری قرار نمی گیرد. تعداد دانشجویان یک کلاس از نوع گسسته اند. کمیت‌های ناپیوسته دامنه ندارند و اعشار نمی پذیرند.

۳- ویژگی‌های متغیر و داده‌های پارامتریک:

الف) با مقیاس نسبی یا فاصله ای اندازه گیری شده اند (فاصله بین واحدهای سازنده کمیت ها برابرند)
ب) توزیع آنها در یک جامعه آماری بهنجار و یا نزدیک بهنجار است.

۴- ویژگی‌های متغیر و داده‌های ناپارامتریک:

در این داده ها دو ویژگی مربوط به داده های پارامتری وجود ندارد و بهمین جهت کمیت هائی که دارای مقیاس اسمی و یا رتبه ای باشند ، داده های پارامتری تلقی می شوند و توزیع داده ها نرمال نیست.

۵- داده‌های کیفی^۲:

گاهی گروه مورد مطالعه براساس یک متغیر، فقط طبقه بندی می شوند به این متغیرها، متغیرهای اسمی یا مقوله ای نیز می گویند. به عنوان مثال طبقه بندی افراد یک گروه از نظر جنس (دو وجهی)، گروه خونی (چند وجهی)، رنگ مو، ملیت داده های کیفی هستند.

۶- داده‌های کمی^۳:

این اعداد حداقل از نظم رتبه ای برخوردار هستند. داده های دارای مقیاس های نسبی، فاصله ای و رتبه ای را در بر می گیرند به عنوان مثال اندازه قد، میزان هوشبهر ، یک داده کمی به حساب می آیند.
مجموع کل یک مقدار ثابت (مقدار ثابت کمیتی است که در یک سری تغییر نمی کند).

الف) جدول بندی داده های مقیاس اسمی و رتبه ای

۱- جدول داده های یک متغیری :

مراد از داده های اسمی یک متغیری این است که فقط یک صفت یا متغیر در مورد افراد نمونه آماری مورد مشابه قرار گرفته و موارد تکرار یا فراوانی آن یادداشت می شود. در مورد داده های مقیاس رتبه ای نیز رتبه ها را همراه با

^۱ - Discrete

^۲ - Quality

^۳ - Quantity



توصیف آنها (در صورت لزوم) به ترتیب نزولی و یا صعودی می نویسیم و پس از استخراج فراوانی هر رتبه ، نسبت و یا درصد مربوط به آن را محاسبه و ثبت می کنیم .

جدول ۱

فراوانی درصدی P%	فراوانی نسبی P	فراوانی F	رشته تحصیلی
۴۵	۰/۴۵	۴۵	مشاوره
۳۵	۰/۳۵	۳۵	روانشناسی
۲۰	۰/۲۰	۲۰	علوم اجتماعی
$\Sigma P\% = 100$	$\Sigma P = 1$	$N = 100$	$\square$

جدول ۲

فراوانی درصدی P%	فراوانی نسبی P	فراوانی F	میزان تمایل به ازدواج
۲۵	۰/۲۵	۲۰	بسیار زیاد
۳۷/۵	۰/۳۸	۳۰	زیاد
۱۹	۰/۱۹	۱۵	تاحدودی
۱۲/۵	۰/۱۲	۱۰	کم
	۰/۰۶	۵	بسیار کم
$\Sigma P \% = 100$	$\Sigma P = 1$	۸۰	$\square$

ب) جدول توزیع فراوانی در داده های دارای مقیاس نسبی و فاصله ای

پس از جمع آوری داده ها، تشکیل جدول توزیع فراوانی اولین مرحله بررسی آماری نتایج تحقیق است . هنگامی که تعداد اعداد یک توزیع و همچنین فاصله بین آنها خیلی زیاد باشد از توزیع فراوانی طبقه بندی شده استفاده می کنیم. بطور کلی اگر دامنه تغییر نمره های یک مجموعه (R) مساوی و یا بزرگتر از ۲۰ باشد توزیع فراوانی را طبقه بندی می کنیم. در این توزیع به جای استفاده از نمره ها به صورت تکی، چند عدد با هم در طبقاتی ناسازگار^۱ هستند قرار می گیرند. عده نمره هایی را که در یک طبقه گروه بندی می شوند، فاصله طبقه می نامیم و آنرا با حرف (i) نشان می دهیم . فاصله طبقه که در هر جدول توزیع فراوانی معمولاً مقداری ثابت است باید متناسب با دامنه تغییر نمره ها (R) باشد، بطوری که خارج قسمت R به I که تعداد طبقات جدول را تعیین می کند از ۱۰ کمتر و از ۲۰ بیشتر نباشد بمنظور ایجاد سهولت در محاسبه شاخصهای آماری بهتر است فاصله طبقه، عددی فرد باشد. اگر فاصله طبقه $i=1$ باشد جدول توزیع فراوانی ساده (یعنی طبقه بندی) در جدول ۳ با در نظر گرفتن $R=51$ $15+1=37$ و فاصله طبقه $i=3$ ، تعداد طبقه ها $\frac{R}{i} = \frac{37}{3} = 13 - 13 = C$ خواهد شد.

جدول ۳

طبقات	طبقات
-------	-------

^۱ - ناسازگار یعنی اینکه یک عدد معین فقط در یک طبقه قرار داده شود.

۵۱-۵۳	۱	۳۰-۳۲	۰
۴۸-۵۰	۲	۲۷-۲۹	۹
۴۵-۴۷	۳	۲۴-۲۶	۲
۴۲-۴۴	۲	۲۱-۲۳	۲
۳۹-۴۱	۳	۱۸-۲۰	۴
۳۶-۳۸	۵	۱۵-۱۷	۳
۳۳-۳۵	۴		
			$N = \sum F = 40$

حدود واقعی طبقه ها

این حدود معمولاً از نیم واحد اندازه گیری بالاتر و نیم واحد پائین تر از ارزش گزارش شده است. بنابراین هنگامی که نتیجه یک اندازه گیری بر حسب سانتیمتر ۴۲ گزارش می شود این است که اگر اندازه گیری دقیق تری بکار می رفت ارزش به دست آمده در دامنه ۴۱/۵ و ۴۲/۵ سانتی متر قرار می گرفت. به همین ترتیب هرگاه نتیجه اندازه گیری براساس ۰/۱ سانتیمتر برابر ۳۰/۵ سانتی متر گزارش شود، چنین ارزشی در دامنه ۳۰/۴۵ و ۳۰/۵۵ سانتی متر قرار می گیرد.

مراحل ساختن توزیع فراوانی طبقه بندی شده

۱- محاسبه دامنه تغییر : دامنه تغییر برابر است با تفاضل بزرگترین عدد از کوچکترین عدد به اضافه یک (اضافه کردن یک به خاطر در نظر گرفتن حدود واقعی اعداد است)

$$R = (X_{Max} - X_{Min}) + 1$$

۲- تعیین تعداد طبقات : تعداد طبقات را معمولاً بین ۱۰ تا ۲۰ در نظر می گیرند. البته می توان تعداد طبقات را با استفاده از فرمول

۳- تعیین فاصله هر طبقه (i) : فاصله و یا اندازه هر طبقه از فرمول زیر محاسبه می شود.

$$i = \frac{R}{K} = \frac{\text{دامنه تغییرات}}{\text{تعداد طبقات}}$$

۴- تعیین نقطه شروع طبقات : معمولاً نوشتن طبقات را از کوچکترین عددی که مضربی از فاصله طبقه است آغاز می کنند.

۵- نوشتن طبقات : در این مرحله از کوچکترین عدد که حد پایین اولین (کوچکترین) طبقه نامیده می شود بزرگترین عدد یا حد بالای این طبقه را محاسبه می کنیم و پس از نوشتن اولین طبقه، را به حد پایین و بالای همین طبقه اضافه می کنیم تا طبقه بعدی بدست می آید.

۶- تعیین فراوانی هر طبقه (F_i) : در این مرحله می توان از زدن خط نشان در مقابل هر طبقه نیز استفاده کرد و تعداد نمره هایی را که در هر طبقه قرار می گیرد مشخص نمود .

۷- تعیین عدد میانی هر طبقه (X_i) : عدد وسط هر طبقه را عدد میانی آن می نامند که برابر نصف مجموع حد بالایی و پایینی طبقه است. مثلاً در جدول (۴) فرض بر این است که در پایین ترین طبقه ۲ نمره ۱۹ وجود دارد و یا نمره ۲ نفری که در این طبقه هستند برابر ۱۹ است.

۸- تعیین حدود واقعی هر طبقه ($R-L$) : همانطوریکه قبلاً اشاره شد حدود دقیق هر طبقه از نیم واحد اندازه گیری پایین تر تا نیم واحد اندازه گیری بالاتر از محدوده طبقه بندی شده آن را در بر می گیرد . جدول (۴) حدود واقعی پایین ترین طبقه از ۱۷/۵ تا ۲۰/۵ است.

۹- محاسبه فراوانی نسبی طبقه (P) : فراوانی نسبی هر طبقه برابر نسبتی از مجموعه نمره های یک جدول توزیع فراوانی است که در آن طبقه قرار دارند فراوانی نسبی هر طبقه مساوی نسبت فراوانی مطلق (F_i) آن به تعداد کل نمره ها (N)

$$P = \frac{f_i}{N}$$

دریک جدول توزیع فراوانی نسبی طبقات مختلف برابر یک است.

$$\sum p = 1$$

۱۰- فراوانی درصدی طبقه ($P\%$) : فراوانی درصدی هر طبقه برابر درصدی از مجموعه نمره های یک جدول توزیع فراوانی است که در آن طبقه قرار دارند و برای محاسبه آن کافی است که فراوانی نسبی را در عدد ۱۰۰ ضرب کنیم $P\% = P \times 100$ و یا از فرمول $P\% = F_i \times \frac{100}{N}$ استفاده کنیم.

۱۱- فراوانی تراکمی یا تجمعی (CF) : فراوانی تراکمی هر طبقه برابر تعداد نمره هایی است که زیر کرانه بالای آن طبقه قرار می گیرند و از طریق جمع بستن فراوانی طبقات زیر آن طبقه و خود آن طبقه بدست می آید . به عنوان مثال در جدول (۴) فراوانی تراکمی طبقه (۳۲ - ۳۰) برابر حاصل جمع فراوانی چهار طبقه زیر آن و خود آن طبقه یعنی $15 = 2 + 3 + 4 + 6$ است یعنی نمره ۱۵ نفر زیر ۲۳/۵ قرار دارند.

۱۲- فراوانی تراکمی یا تجمعی نسبی (RCF) : فراوانی تراکمی نسبی هر طبقه برابر نسبت تعداد نمره هایی است که در یک جدول توزیع زیر کرانه بالای آن طبقه قرار می گیرند و مساوی خارج قسمت فراوانی تراکمی هر طبقه به تعداد کل نمرات است.

$$RCF = \frac{CF}{N}$$

۱۳- فراوانی تراکمی یا تجمعی درصدی ($CF\%$) : فراوانی تراکمی درصدی هر طبقه برابر درصد نمره هایی است که در یک جدول توزیع فراوانی زیر کرانه بالایی آن طبقه قرار می گیرند. برای محاسبه فراوانی تراکمی درصدی هر طبقه فراوانی تراکمی نسبی آنرا در ۱۰۰ ضرب می کنیم و یا از فرمول $CF\% = CF \times \frac{100}{N}$ استفاده می کنیم . مثلاً برای طبقه (۳۰ - ۳۲) در جدول (۴) فراوانی تراکمی درصدی برابر است با ۳۰

بنابراین ۳۰ درصد نمره های جدول (۴) زیر ۲۳/۵ قرار دارند. جدول (۴)

CL	Fi	X_i	R-L	P	P%	C_F	R_{CF}	C_F
		۴					۱	۱
		۹					۰/۹	۰
		۴					۸	۹
		۶					۰/۹	۸
		۴					۲	۹
	۱	۳	۵/۴۷ - ۵/۵۰	۰/۰۲	۲	۵۰	۰/۸	۲
	۳	۴	۴۴/۵ - ۴۷/۵	۰/۰۶	۶	۴۹	۰	۸
۴۸ - ۵۰	۶	۰	۴۱/۵ - ۴۴/۵	۰/۱۲	۱۲	۴۶	۰/۶	۰
۴۵ - ۴۷	۷	۳	۳۸/۵ - ۴۱/۵	۰/۱۴	۱۴	۴۰	۶	۶
۴۲ - ۴۴	۸	۷	۳۵/۵ - ۳۸/۵	۰/۱۶	۱۶	۳۳	۰/۵	۶
۳۹ - ۴۱	۱۰	۳	۳۲/۵ - ۳۵/۵	۰/۲۰	۲۰	۲۵	۰	۵
۳۶ - ۳۸ ۳۳ - ۳۵	۶	۴	۲۹/۵ - ۳۲/۵	۰/۱۲	۱۲	۱۵	۰/۳	۰
۳۰ - ۳۲ ۲۷ - ۲۹ ۲۴ -	۴	۳	۲۶/۵ - ۲۹/۵	۰/۰۸	۸	۹	۰	۳
۲۶ ۲۱ - ۲۳ ۱۸ - ۲۰	۳	۱	۲۳/۵ - ۲۶/۵	۰/۰۶	۶	۵	۰/۱	۰
	۰	۲	۲۰/۵ - ۲۳/۵	۰	۰	۲	۸	۱
	۲	۸	۱۷/۵ - ۲۰/۵	۰/۰۴	۴	۲	۰/۱	۸
		۲					۰	۱
		۵					۰/۰	۰
		۲					۴	۴
		۲					۰/۰	۴
		۱					۴	
		۹						
	N = ۵۰			P = $\Sigma 1$	P % $\Sigma = 100$			

۱- در جدول توزیع فراوانی زیر که مربوط به نمره امتحانی ۴۰ نفر دانشجو است تعیین کنید :

طبقات	Fi
٣٦ - ٣٨	١
٣٣ - ٣٥	٣
٣٠ - ٣٢	٥
٢٩ - ٢٧	٧
٢٤ - ٢٦	١٠
٢٣ - ٢١	٦
١٨ - ٢٠	٥
١٥ - ١٧	٢
١٢ - ١٤	١
	N = ٤٠

- الف (چه نسبتی از افراد نمره ۲۵ گرفته اند. ۲) چند در صد افراد نمره ۳۱ گرفته اند.
- ۳) چند نفر نمره کمتر از ۳۰ گرفته اند. ۴) چند در صد افراد زیر نمره ۳۲/۵ گرفته اند.
- ه (چند در صد افراد بالاتر از ۲۰ گرفته اند. و (چه نسبتی از افراد نمره بین ۱۷ و ۳۵ گرفته اند.
- جواب : (۱) ۲۵/۰ (۲) ۱۲/۵ (ج) ۳۱ (د) ۹۰ (ه) ۸۰ (و) ۹۰/۰

نمودارهای فراوانی^۱

پس از سازمان‌بندی کردن داده‌ها را بهتر و آسان‌تر توصیف کند و آنها را با توزیع‌های دیگر مقایسه کند.

رسم نمودارها در داده های مقیاس نسبی و فاصله ای :

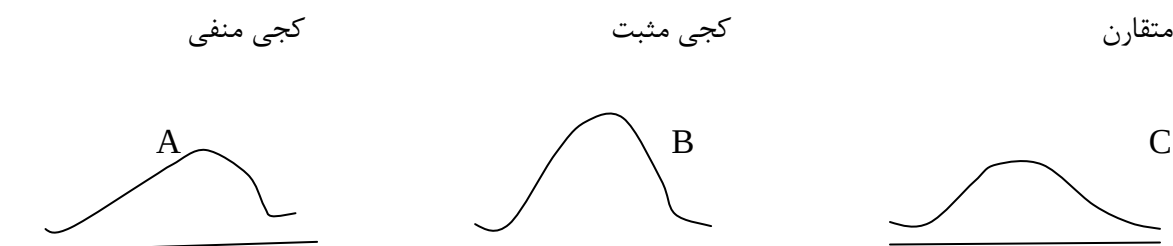
۱- نمودار چند ضلعی: (چند کنج فراوانی)

زمانی که داده ها دارای مقیاس فاصله ای و کمیت ها پیوسته هستند از این نمودار استفاده می شود. ابتدا محورهای مختصات را رسم می کنیم و سپس محور افقی (XX') را بر حسب حد میانی طبقات (Xi) و محور عددی (YY') را بر اساس بیشترین فراوانی مطلق (Fi) طوری تقسیم بندی می کنیم که طول محورهای X و Y به نسبت ۳ به ۲ یا ۴ به ۳ ترسیم شود. اگر جدول توزیع فراوانی مربوط به دو گروه باشد و بنابراین طبقات دارای دو ستون فراوانی مطلق (F_1 و F_2) باشند و بخواهیم که نمره های دو گروه را در روی یک دستگاه محور مختصات رسم و سپس باهم مقایسه کنیم بهترین راه، تبدیل فراوانی های مطلق هر گروه به درصد یا نسبت و ترسیم آنها است (برای تبدیل هر فراوانی مطلق به فراوانی درصدی از رابطه $\% P = \frac{Fi}{N} \times 100$ استفاده می کنیم)

'-Frequency diagram

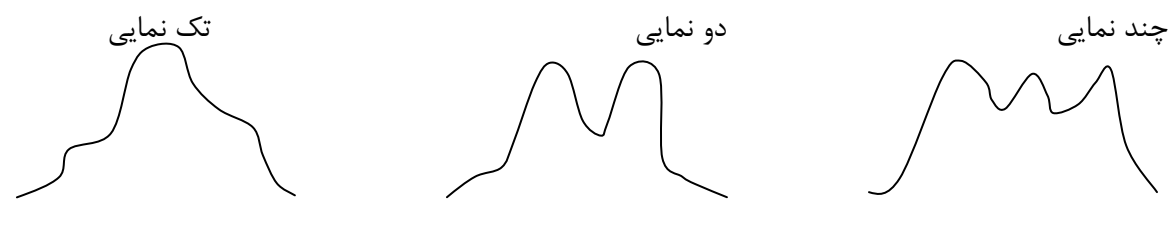
اشکال نمودار چند ضلعی:

نمودار توزیع فراوانی ممکن است دارای شکل های مختلفی باشد که از آن جمله ممکن است متقارن^۱ و یا نامتقارن^۲ باشد. نمودار چند ضلعی وقتی متقارن است که نیمی از آن عیناً همانند نیمی دیگر باشد و در غیر این صورت آنرا نامتقارن نامند. اگر در یک منحنی نامتقارن نقطه برآمدگی ۰ نقطه دارای بیشترین فراوانی (در مرکز منحنی نباشد و دنباله های منحنی نامساوی باشند منحنی را دارای کجی^۳ می نامند. در منحنی دارای کجی اگر دنباله راست منحنی طولانی تر از دنباله چپ آن باشد منحنی را دارای کجی به راست (کجی مثبت) می نامند که در این حالت نمره های زیادتر در انتهای دنباله سمت راست و اگر دنباله چپ منحنی طولانی تر از دنباله راست آن باشد منحنی کجی به چپ (کجی منفی) خواهد بود که در این حالت دنباله طولانی تر منحنی بطرف نمره های خیلی پایین ختم می شود.



۱- نماها در نمودار چند ضلعی :

هرگاه توزیع فقط دارای یک برآمدگی یا نقطه یا حداکثر ارتفاع باشد آن را تک نمایی^۴ گویند در صورتی که توزیع دارای دو برآمدگی یا دو نقطه با حداکثر ارتفاع باشد آنرا دونمایی^۵ گویند و به همین ترتیب چنانچه منحنی توزیع دارای بیش از دو برآمدگی یا دو نقطه با حداکثر ارتفاع باشد آن را چند نمایی^۶ می نامند.



^۱ -Symmetrical

^۲ -Asymmetrical

^۳ -Skewed

^۴ -unimodal

^۵ - bimodal

^۶ - Multimodal



۲- نمودار هیستوگرام (ستونی)

این نمودار برای نمایش داده های ناشی از اجرای متغیرهای پیوسته وسیله مناسب است. در این نمودار هر ستون نشان دهنده یک طبقه از اعداد است . عرض هر ستون که روی محور افقی قرار می گیرد برابر فاصله طبقه و ارتفاع آن مساوی مطلق همان طبقه است.

۳- منحنی درصد فراوانی متراکم (اجیو)

روی محور عمودی بجای مقدر مربوط به فراوانی تراکمی درصدی (CF%) هر طبقه را که از فرمول $CF\% = \frac{CF}{N} \times 100$ قابل محاسبه است مشخص می کنیم بنابراین محور عمودی از صفر تا ۱۰۰ تقسیم می شود.

کاربرد منحنی درصد فراوانی متراکم (اجیو)

(۱) تعیین وضعیت توزیع فراوانی نمره ها:

هر چه شکل منحنی اجیو به () شباهت بیشتری داشته باشد توزیع فراوانی به توزع طبیعی نزدیکتر است.

(۲) تعیین رتبه درصدی^۱ هر نمره (PR)

رتبه درصدی هر نمره خاص نشان می دهد که نمره چند درصد از افراد کمتر از آن است .

(۳) تعیین نقطه درصدی^۲ (P_p)

نقطه درصدی بیان می کند که درصد خاصی از افراد زیر چه نمره ای قرار می گیرند.

باید بدانیم که برخی از نقاط درصدی در بررسی آماری اهمیت زیادی دارند که اهم آنها عبارتند از:

نقطه ۷۵ درصدی (P_{۷۵}) یعنی چارک^۳ سوم (Q_۳)

نقطه ۲۵ درصدی (P_{۲۵}) یعنی چارک اول (Q_۱)

نقطه ۵۰ درصدی (P_{۵۰}) یعنی چارک دوم یا میانه (Q_۲)

نقطه ۹۰ درصدی (P_{۹۰}) یعنی دهک^۴ نهم (D_۹)

نقطه ۱۰ درصدی (P_{۱۰}) یعنی دهک اول (D_۱)

رسم نمودارها در داده های مقیاس اسمی و رتبه ای

۱-نمودار میله ای^۵:

^۱ - Percentile Rank

^۲ -Percentile Point

^۳ - Quarter

^۴ - Decile

^۵ - Bardigram

برای رسم نمودار میله ای روی محور افقی متغیر مستقل (طبقه ها) و روی محور عمودی فراوانی مطلق هر طبقه ، نسبت و یا درصد مربوط به آن نشان داده می شود .

فصل دوم

گرایش های مرکزی

همانطوری که قبلاً اشاره شد اصطلاح « آماره » یا « شاخص آماری » برای آن دسته از داده ها به کار می رود که در مورد یک گروه نمونه به دست آمده است. این شاخص ها در معنای متوسط تحت عنوان اندازه های گرایش مرکزی مطرح می شوند و عبارتند از :

(۱) میانگین
(۲) میانه
(۳) نمایا مد

این شاخص که برای بیان و توصیف داده های کمی بیش از سایر اندازه ها به کار می رود به سه صورت مطرح می شود :

۱- میانگین حسابی^۲ ۲- میانگین هندسی^۳ ۳- میانگین همساز^۴

۱- میانگین حسابی: (معروف ترین شاخص مرکزی)

میانگین حسابی را برای سهولت تنها میانگین می نامند و با X یا M نشان می دهند میانگین از طریق محاسبه خارج خارج قسمت مجموع نمره ها بر تعداد آنها به دست می آید.

ویژگی های میانگین:

(۱) مجموع جبری گشتاور^۵ یا تفاوت نمره ها از میانگین آنها همواره برابر صفر است. به همین جهت میانگین را مرکز ثقل نمره ها می نامند. اگر تفاوت (انحراف) هر نمره از میانگین را به صورت

$\sum x = \sum (X - \bar{X}) = 0$: نشان خواهیم داشت :

(۲) یکتایی: برای مجموعه ای از اعداد تنها یک میانگین وجود دارد.

(۳) مجموع مجذور تفاوت نمره ها از میانگین کمتر از مجموع مجذور تفاوت آنها از هر نقطه دیگری غیر از میانگین است. به عنوان مثال اگر بجای انحراف نمره ها از میانگین تفاوت آنها را از نقطه دیگری (مثلاً A) در نظر

بگیریم خواهیم داشت $\sum (X - \bar{X})^2 < \sum (X - A)^2$

(۴) از همه اعداد تأثیر می پذیرد ؛ به همین دلیل برای داده های پراکندگی زیاد مناسب نیست. مثل مجموعه

اعداد: ۱-۲-۸-۷-۲۰-۱۰۰

' - Mean

 $\bar{y}$ - Arithmetic Mean γ - Geometric Mean^ξ - Harmonic Mean $\circ$ - moment



محاسبه میانگین داده ها :

در محاسبه میانگین داده ها دو حالت کلی وجود دارد.

حالت اول :

وقتی که تعداد نمره ها کم است. اگر نمره ها دارای ضریب اختصاصی نباشند برای محاسبه میانگین از فرمول زیر به صورت $(\bar{X} = \frac{\sum x}{N})$ استفاده می شود، ولی اگر نمره ها دارای ضریب اختصاصی باشند در آن صورت از روش « میانگین وزنی » به صورت فرمول زیر استفاده می شود .

$$\bar{X} = \frac{\sum w.x}{\sum w}$$

در این فرمول X ضریب یا وزن نمره است .

مثال : فرض کنید که در جدول (۱) نمره های امتحانی یک دانشجو در چهار درس مختلف و تعداد واحدهای هر یک از آنها مشخص شده است . چون هر یک از آنها دارای وزن خاص (تعداد واحد) می باشند ، برای محاسبه میانگین از فرمول (۱) استفاده می شود .

جدول (۱)

نام درس	تعداد واحدها W	نمره (X)	W.X
A	۲	۱۵	۳۰
B	۳	۱۲	۳۶
C	۲	۱۶	۳۲
D	۱	۱۸	۱۸
	$\square W=8$		$\square WX=116$

$$\bar{X} = \frac{\sum w.x}{\sum w} = \frac{116}{8} = 14.5$$

حالت دوم :

وقتی که تعداد نمره ها زیاد است در این حالت ابتدا نمره ها را طبقه بندی و جدول توزیع فراوانی آنها را تنظیم می کنیم. سپس براساس استفاده از عدد میانی طبقات میانگین به دست می آید.
می دانیم که در اعداد طبقه بندی شده فرض بر این است که نمره افرادی که در هر طبقه قرار می گیرند بین دو حد بالا و پایین آن طبقه بوده و برابر عدد میانی آن طبقه در نظر گرفته می شود ، بنابراین در این حالت نیز عدد میانی هر طبقه از ضریب اختصاصی برابر با فراوانی مطلق آن طبقه برخوردار است ، و به همین جهت می توانیم از روش میانگین وزنی به صورت فرمول زیر استفاده کنیم .

$$\bar{X} = \frac{\sum F.xi}{N}$$

مثال: جدول توزیع فراوانی (۲) مربوط به نمره هوشبهر یک گروه از کودکان ۶ ساله است می خواهیم میانگین هوشبهر آنها را محاسبه کنیم .



جدول (۲)

طبقات	Fi	X _i	FXi
۱۲۰-		۱	
۱۲۲		۲۱	
۱۱۷-		۱	
۱۱۹		۱۸	
۱۱۴-	۲	۱	۲۴۲
۱۱۶	۴	۱۵	۴۷۲
۱۱۱-	۵	۱	۵۷۵
۱۱۳	۸	۱۲	۸۹۶
۱۰۸-	۷	۱	۷۶۳
۱۱۰	۶	۰۹	۶۳۶
۱۰۵-	۵	۱	۵۱۵
۱۰۷	۳	۰۶	۳۰۰
۱۰۲-		۱	
۱۰۴		۰۳	
۹۹-		۱	
۱۰۱		۰۰	
	N= ۴۰		F. Xi = ۴۳۹۹

$$\bar{X} = \frac{4399}{40} = 109.98$$

۲- میانگین هندسی (G_M) :

این شاخص در آمار تربیتی و روانی به ندرت استفاده می شود و در آمار اقتصادی بیشتر مورد استفاده قرار می گیرد که یکی از موارد کاربرد آن در پسیکوفیزیک است میانگین هندسی دو عدد A و B برابر جذر حاصل ضرب آنها است .

میانگین هندسی N عدد برابر N ام حاصل ضرب آنها است.

$$G_M = \sqrt[N]{X_1 \cdot X_2 \cdot X_3 \dots X_N}$$

مثال : مطلوب است محاسبه میانگین هندسی اعداد ۲۷ ، ۱۲۵ و ۸

$$G_M = \sqrt[3]{125 \times 27 \times 8} = 5 \times 3 \times 2 = 30$$

۳- میانگین همساز (H_M) یا هارمونیک

میانگین همساز مقدر X عبارت است از عکس میانگین حسابی معکوس مقادیر مختلف X (منظور از عکس یا معکوس هر عدد یک تقسیم بر آن عدد است). فرمول محاسبه میانگین همساز به صورت زیر است .

$$H_M = \frac{1}{\frac{1}{N} \left(\frac{1}{X_1} + \frac{1}{X_2} + \dots + \frac{1}{X_N} \right)} = \frac{N}{\sum \frac{1}{X_i}}$$

مثال : سه نفر به تنهایی کاری را در مدت ۴ ، ۵ و ۶ ساعت انجام می دهند. بطور متوسط کار را در چند ساعت به انجام می رسانند؟

$$H_M = \frac{1}{\frac{1}{3} \left(\frac{1}{4} + \frac{1}{5} + \frac{1}{6} \right)} = \frac{3}{\frac{15+12+10}{60}} = \frac{180}{37} = 4/86$$

بنابراین مورد استفاده عمده میانگین همساز وقتی است که بخواهیم تعیین کنیم که برای انجام واحد کار معینی بطور متوسط چه مقدار زمان لازم است. در یک مجموعه اعداد خاص بین میانگین حسابی، هندسی و همساز همواره رابطه زیر برقرار است :

$$\bar{X} > G_M > H_M$$

میانگین مرکب: دو حالت وجود داد: ۱- اگر گروه ها به تعداد مساوی باشند مجموع میانگین ها را به تعداد گروه ها (نه تعداد افراد) تقسیم می کنیم: $(\bar{X}_t = \frac{\sum X}{N})$.

۲- اگر تعداد گروه ها نامساوی باشند میانگین هر گروه را به تعداد همان گروه ضرب کرده و به تعداد گروه ها تقسیم می کنیم: $(\bar{X}_t = \frac{\sum X.n}{N})$.

(۲) میانه^۱

میانه عبارت است از ارزش عددی نقطه ای که پنجاه درصد بالایی نمرات را از پنجاه درصد پایین جدا می کند. بنابراین میانه نقطه ۵۰ درصدی است که مقادیر را به گروه مساوی بالا و پایین تقسیم می کند میانه را با Md یا Mn نشان می دهند.

ویژگی های میانه:

(۱) مجموع قدر مطلق انحراف نمره ها از میانه کوچکتر یا مساوی مجموع قدرمطلق انحراف های نمره ها از هر عدد دیگری است یعنی:

$$\sum |X - Md| \leq \sum |X - C|$$

(۲) برعکس میانگین از همه ی اعداد تأثیر نمی پذیرد. و برای اعدادی که پراکندگی زیادی دارند مناسب تر از میانگین است.

(۳) با متغیر های رتبه ای به کار برد می شود

به دلیل عدم تعریف میانه به صورت جبری نمی توان آن را در محاسبات دیگر مورد استفاده قرار داد. به عنوان مثال میانه گروه های مختلف را نمی توان برای محاسبه گروهی که از ترکیب این گروهها به دست می آید به کار برد.

^۱ - Median



مثلاً فرض کنید میانه سن دانشجویان دختر و پسر یک دانشکده به ترتیب ۲۲ و ۲۵ سال باشد. هیچ راهی وجود ندارد که با استفاده از میانه های فوق میانه سن کل دانشجویان این دانشکده را محاسبه کرد. برای بدست آوردن میانه سن کل دانشجویان باید توزیع سن این گروه را داشته باشیم.

محاسبه میانه: در محاسبه میانه اعداد دو حالت کلی وجود دارد:

حالت اول: وقتی که تعداد داده ها کم است و در مقیاس فاصله ای یا نسبی اندازه گیری شده اند. برای محاسبه میانه در این حالت به ترتیب زیر عمل می کنیم.

- مرتب کردن داده ها (به ترتیب صعودی و یا نزولی)

- تعیین محل میانه (تعیین عددی که در مرتبه $\frac{N+1}{2}$ قرار دارد)

- تعیین میانه

نمره امتحانی ۸ نفر دانشجو به ترتیب زیر است میانه نمره ها چقدر است ؟

۲۰-۲۰-۱۹-۱۵-۱۴-۱۷

اعداد را مرتب می کنیم.

۱۶-۱۷-۱۸-۱۹-۲۰-۲۰

$$\frac{N+1}{2} = \frac{6+1}{2} = 3.5$$

محل میانه را تعیین می کنیم.

بنابراین میانه در مرتبه ۵/۴ قرار می گیرد و برابر میانگین ۱۷ و ۱۸ یعنی ۱۷/۵ خواهد بود .

محاسبه میانه در اعداد تکراری:

اگر میانه توزیع نمره ها بین عددهای تکراری قرار گیرد در آن صورت محاسبه میانه به ترتیب زیر محاسبه می

شود. مطلوب است محاسبه میانه اعداد « ۳-۴-۵-۵-۵-۵ » $\frac{N+1}{2} = \frac{6+1}{2} = 3.5$ محل میانه

بنابراین میانه در مرتبه ۳/۵ ام یعنی وسط ۵ اول و دوم یعنی بین عددهای تکرار قرار می گیرد.

نکته: فرمول محاسبه میانه را برای این حالت خاص می توان به صورت زیر در نظر گرفت.

$$M_n = \text{تکرار پائین کرانه} + \frac{\frac{N}{2} \text{ تعداد موارد زیر تکرار}}{\text{تعداد تکرار}}$$

زیر عدد ۴/۵ (کرانه پائین ۵) مورد و در فاصله ۴/۵ تا ۵/۵ (یعنی دامنه یک نمره) ۴ مورد قرار دارد ولی میانه نقطه ای است که ۳ مورد زیر آن بوده و برابر است با :

$$M_n = 4/5 + \frac{3-2}{4} \times 1 = 4/5 + 0.25 = 4.75$$

مطلوب است محاسبه میانه اعداد : « ۷-۱۰-۱۰-۱۰-۱۰-۱۰-۱۲ »

$$\text{محل میانه} = \frac{N+1}{2} = \frac{7+1}{2} = 4$$

بنابراین میانه در مرتبه چهارم و روی ۱۰ سوم قرار می گیرد. یک مورد زیر ۹/۵ و ۵ مورد نیز بین ۹/۵ تا ۱۰/۵ قرار دارند بنابراین میانه برابر :

$$Mn = 9/5 + \frac{3/5-1}{5} \times 1 = 10$$

حالت دوم : محاسبه میانه در داده های طبقه بندی شده :

وقتی که تعداد داده ها زیاد بوده و به صورت توزیع فراوانی طبقه بندی شده باشند. نقاط درصدی و از جمله میانه را می توان از طریق فرمول زیر بطور نسبتاً دقیق برآورد کرد. در این فرمول :

$$Mn = L + \left(\frac{\frac{N}{2} - CF_b}{Fi} \right) \times i$$

کرانه پائین طبقه میانی فراوانی تراکمی طبقه زیر و یا قبل از طبقه ای که میانه در آن قرار دارد. CF_b :

F_i فراوانی مطلق طبقه میانه

i فاصله طبقات

- محاسبه میانه با استفاده از فرمول بالا

$$Mn = 20/5 + \frac{12-5}{8} \times 3 = 23/12$$

جدول (۳)

CL	Fi	C F	
۳۰-	۲	۲۴	
۲۳	۳	۲۲	
۲۷-	۶	۱۹	
۲۹	۸	۱۳	
۲۴-	۴	۵	طبقه
۲۶	۱	۱	میانه
۲۱-			
۲۳			
۱۸-			
۲۰			
۱۵-			
۱۷			



	N=۲		
	۴		

محاسبه نقاط درصدی (صدکها^۱ و دهکها^۲)

میانة نقطه ۵۰ درصدی است بنابراین محاسبه سایر نقاط درصدی نیز با فرمول مشابه آنچه که در مورد میانة گفته شد انجام می گیرد . نقطه درصدی نیز نمره ای است که درصد معینی از افراد زیر آن قرار دارد فرمول کلی محاسبه نقاط درصدی زیر است:

در این فرمول P_p نقطه درصدی مورد نظر و $P.N$ نیز تعداد افرادی است که زیر نقطه درصدی قرار می گیرند.

- محاسبه صدک ششم (P_6) جدول ۳:

$$P.N = 0.60 \times 24 = 14.4$$

بنابراین صدک ششم در طبقه ۲۴-۲۶ قرار دارد زیرا CF این طبقه از $P.N$ بیشتر است.

$$P_6 = 23/5 + \frac{14 - 13}{6} \times 3 = 24/20 \quad \text{نقطه ۶۰ درصدی}$$

- محاسبه دهک نهم (P_9)

$$P.N = 0.90 \times 24 = 21.6$$

بنابراین دهک نهم در طبقه ۲۷-۲۹ قرار دارد زیرا CF آن از $P.N$ بیشتر است.

$$P_9 = 26/5 + \frac{21 - 19}{3} \times 3 = 29/1 \quad \text{نقطه ۹۰ درصدی (دهک نهم)}$$

محاسبه رتبه درصدی^۳ یک نمره مفروض:

رتبه درصدی (PR) یک نمره خاص بیان می کند که نمره چند درصد از افراد زیر آن نمره قرار دارند برای محاسبه رتبه درصدی از فرمول استفاده می کنیم.

$$PR_{X=\frac{100}{N}} \left[CF_b + \frac{X - L}{i} \times Fi \right]$$

به عنوان مثال در جدول (۳) رتبه درصدی نمره ۲۵ به صورت زیر محاسبه می شود.

$$PR_{25} = \frac{100}{24} \left[13 + \frac{25 - 23/5}{3} \times 6 \right] = 66/66\%$$

(۳) نما

در یک توزیع نما عبارت از عددی است که بیش از سایر اعداد تکرار شده باشد. نما یا مد که آن را با Mo ^۱ نشان می دهند. ساده ترین شاخص گرایش مرکزی است . در یک جدول توزیع فراوانی نما برابر ارزش عددی نقطه میانی

^۱ -Percentiles

^۲ -deciles

^۳ -Percentile Rank

طبقه ای است که فراوانی آن از سایر طبقات بیشتر باشد. بنابراین در توزیع فراوانی (۳)، نما برابر عدد میانی طبقه ۲۱-۲۳ یعنی ۲۲ است که بالاترین فراوانی (۸) را داراست. اگر در توزیع فراوانی دو طبقه مجاور هم دارای حداکثر فراوانی باشند در این صورت نما بین آن دو طبقه قرار خواهد گرفت (یعنی حد بالائی طبقه پائین تر یا حد پائینی طبقه بالاتر). بیشتر توزیع ها یک نمائی^۱ هستند ولی ممکن است توزیع دونمائی^۲ و یا چند نمائی باشد. در اعداد طبقه بندی شده مقدار نما که وسیله عددی که فراوانی آن بیشتر از سایر طبقات است تعیین می شود نمای ظاهر ی نامیده می شود و با نمای حقیقی که تخمینی از عددی است که فراوانی آن حداکثر است تفاوت دارد. در توزیع دونمایی متقارن میانگین و میانه برابر هستند و نماها در طرفین آنها واقع هستند.

مقایسه شاخص های گرایش مرکزی :

در توزیع متقارن اندازه های گرایش مرکزی باهم برابرند. در صورتی که توزیع دارای کجی (چولگی) باشد در آن صورت نمره ها بیشتر در یک طرف توزیع متمرکز می شود و سه شاخص مرکزی برابر نخواهند بود و بنابراین در سه نقطه کاملاً مختلف در توزیع قرار خواهند گرفت. بطور کلی می توان گفت که در منحنی دارای کجی مثبت میانگین دارای بالاترین، نما دارای پائین ترین مقدار است و میانه در بین این دو (به فاصله $\frac{1}{3}$ از میانگین $\frac{2}{3}$ از نما) قرار دارد و در توزیعهای کجی منفی هستند نما دارای بالاترین و میانگین دارای پائین ترین مقدار است و میانه نیز بین این دو (به فاصله $\frac{1}{3}$ از میانگین $\frac{2}{3}$ از نما) قرار دارد. با توجه به موقعیت این سه شاخص می توان رابطه های زیر را در مورد آنها استخراج کرد.

$$Mo = 3Mn - 2\bar{X}$$

$$Mn = \frac{MO + 2\bar{X}}{3}$$

$$\bar{X} = \frac{3Mn - Mo}{2}$$

بهرحال میانگین معتبرترین و پایدارترین شاخص مرکزی است چنانچه از جامعه ای نمونه های زیادی انتخاب شود اختلاف بین میانگین نمونه های مختلف کمتر از اختلاف بین میانه های همین نمونه ها است و به همین ترتیب اختلاف بین میانه های نمونه های مختلف کمتر از اختلاف بین نماهای همین نمونه ها است. بنابراین برای برآورد جامعه از نمونه، میانگین معتبرترین برآورد از میانگین جامعه است. در یک توزیع فراوانی اگر نمره ها را با یک عدد ثابت جمع کنیم به میانگین، میانه و نمای نمره ها، همان عدد ثابت افزوده می شود. این قاعده ترکیب با مقدار ثابت C در مورد تفریق و ضرب و تقسیم اعداد نیز صادق است. در جدول ذیل ویژگیها و موارد استفاده هر یک از شاخص های گرایش مرکزی بیان شده است.

^۱ -Mode

^۲ -Unimodal

^۳ -Bimodal



شاخص	ویژگی ها	موارد استفاده
میانگین	۱- یک شاخص فاصله ای است که با مقیاسهای نسبی و فاصله ای به کار برده می شود. ۲- مجموع انحراف از میانگین همیشه صفر است. ۳- در محاسبه آن از همه اعداد یا نمره ها استفاده می شود و نمره های خیلی کوچک یا بزرگ در آن تأثیر دارد. ۴- با ثبات ترین شاخص مرکزی است. ۵- در محاسبه آن از عملیات ریاضی بیشتری استفاده می شود.	۱- وقتی که سوال پژوهشی به صورت کمی باشد و پژوهشگر علاقمند باشد داده ها را به صورت فاصله ای تفسیر کند. ۲- وقتی که محاسبات دیگر آماری مانند اندازه های پراکندگی یا ضریب همبستگی مورد نیاز باشد. ۳- زمانی که توزیع داده ها نرمال و یا نزدیک به نرمال باشد. ۴- وقتی که مایل به دانستن مرکز ثقل نمونه ای باشیم. ۵- وقتی که معتبرترین معرف از اندازه های گرایش به مرکز لازم باشد.
میان	۱- یک شاخص ترتیبی است که با مقیاس های ترتیبی، فاصله ای و نسبی قابل محاسبه است. ۲- تحت تأثیر ارزش عددی کلیه نمره ها قرار ندارد. ۳- در محاسبه آن از عملیات ریاضی کمتری استفاده می شود. ۴- ثبات آن از میانگین کمتر ولی از نما بیشتر است. ۵- مقدار آن به وسیله عددی که در وسط توزیع قرار دارد تعیین می شود	۱- وقتی که توزیع در مقیاس نسبی یا فاصله ای بوده و دارای کجی قابل ملاحظه باشد، یعنی در توزیع نمره ها اعداد خیلی بزرگ یا خیلی کوچک وجود داشته باشد. ۲- زمانی که توزیع نمره ها ناقص باشد. ۳- وقتی که بخواهیم بدانیم نتیجه اندازه گیری در نیمه بالا یا در نیمه پائین قرار می گیرد. ۴- وقتی که مقیاس اندازه گیری ترتیبی باشد.
میزان	۱- یک شاخص اسمی است که در مقیاسهای اسمی، فاصله ای و نسبی قابل محاسبه است. ۲- عددی است که دارای بیشترین فراوانی است و غالباً در وسط توزیع نمره ها قرار دارد. ۳- در بعضی توزیع ها بیش از یکی وجود دارد. ۴- اعمال ریاضی را نمی توان با آن انجام داد. ۵- دارای ثبات یا پایداری خیلی کمی است.	۱- وقتی که مقیاس اندازه گیری اسمی باشد. ۲- وقتی که تخمین فوری از اندازه های گرایش به مرکز لازم باشد. ۳- وقتی که پژوهشگر بخواهد عددی را که بیشتر تکرار شده باشد پیدا کند.



سوالات خودسنجی

- ۱- در توزیع داده های مقیاس ترتیبی شاخص متوسط قابل محاسبه است: میانه و نما
- ۲- میانگین نمره های دروس X و Y در یک گروه از دانشجویان به ترتیب $۱۶/۵$ و ۱۵ است:
- (۱) اگر نمره های X و Y را برای هر دانشجو رویهم جمع کنیم میانگین نمره های جدید چقدر است.
- (۲) اگر نمره های Y هر دانشجو را از X مربوط به آن کم کنیم میانگین نمره های حاصل چقدر خواهد شد.
- جواب : $۳۱/۵$ (۱) $۱/۵$ (۲)

فصل سوم

شاخص پراکندگی^۱

شاخص های پراکندگی به طور کلی به دو دسته تقسیم می شوند:

شاخص هایی که مربوط به داده های دارای مقیاس یا فاصله ای هستند. مهمترین این شاخص ها عبارتند از: دامنه تغییرات، دامنه صدکها و دهک ها، انحراف متوسط از میانگین، واریانس، انحراف استاندارد و بالاخره ضریب تغییرپذیری. شاخص هایی که مربوط به داده های دارای مقیاس اسمی یا رتبه ای هستند که از جمله آنها می توان به نسبت واریانس یا پراکندگی (V.R) اشاره کرد. البته انحراف صدک ها و دهک ها را در مورد داده های رتبه ای نیز می توان تعیین کرد.

۱- دامنه تغییرات (R):

دامنه تغییرات که آنرا به R نشان می دهند، ساده ترین شاخص پراکندگی است و مقدار آن در یک توزیع برابر است با تفاضل بین بزرگترین و کوچکترین عدد. البته گاهی هم آن را به اضافه ۱ می کنند.

$$R = X_H - X_L = X_{max} - X_{min}$$

$$R = ۱۹ - ۵ + ۱ = ۱۵ \quad \text{بنابراین دامنه تغییرنمره های «۷-۱۵-۱۶-۱۸-۱۹» برابر است با}$$

ویژگی های دامنه تغییرات: دامنه تغییرات چون فقط از دو مقدار انتهائی تأثیر می پذیرد سرعت قابل محاسبه است اما قادر به توصیف توزیع نمره ها به صورت حقیقی نیست و نمی تواند به عنوان یک شاخص پایدار پراکندگی مورد استفاده قرار گیرد. ضمناً همانطوریکه قبلاً اشاره شد استفاده از دامنه تغییرات مستلزم داشتن مقیاس فاصله ای است.

دامنه بین صدکها و دامنه بین دهک ها

برای اینکه اندازه پراکندگی تنها از دو مقدار انتهائی ناشی نشود می توان دامنه تغییر بین دهک ها یا دهکهای معینی را محاسبه کرد. مثلاً دامنه تغییر میان نقاط ۷۵ درصدی (چارک سوم = Q_3) و ۲۵ درصدی (چارک اول = Q_1) را به دست می آوریم و آنرا دامنه چارک ها می نامیم. بنابراین دامنه چارک ها، دامنه تغییرات ۵۰ درصدی میانی توزیع است. نصف دامنه چارک ها را چارک متوسط (نیم دامنه چارکی) می نامند.

$$\frac{Q}{2} = \frac{Q_3 - Q_1}{2} \quad \text{چارک متوسط}$$

همچنین دامنه تغییر میان نقاط ۹۰ درصدی (دهک نهم = D_9) و ۱۰ درصدی (دهک اول = D_1) را به دست می آوریم و آن را دامنه بین دهکها می نامیم.

$$D = D_9 - D_1$$

بنابراین دامنه دهکها، دامنه تغییرات ۸۰ درصد میانی توزیع است.

^۱ -Variability

محاسبه چارک ها

(۱) در اعداد طبقه بندی نشده :

برای محاسبه چارک اول و سوم ابتدا نمره ها را از کوچک به بزرگ مرتب و سپس با تعیین میانه توزیع را به دو قسمت تقسیم می کنیم . میانه های نیمه اول و دوم ب ترتیب چارک اول و سوم خواهند بود . در مورد اعداد «۱۲-۷-۱۰-۱۵-۱۸-۱۱-۱۹-۲۰» مطلوب است محاسبه دامنه چارک ها و چارک متوسط را بدس آوریم:

$$۷-۱۰-۱۱-۱۲-۱۵-۱۸-۱۹-۲۰$$

اول اعداد را مرتب می کنیم.

$$Md = \frac{۱۲+۱۵}{۲} = ۱۳/۵$$

میانه توزیع را محاسبه می کنیم.

$$Q = \frac{۷+۱۳/۵}{۲} = ۱۰/۲۵$$

میانه نیمه اول (چارک اول) را محاسبه می کنیم.

$$Q_۲ = \frac{۱۳/۵+۲۰}{۲} = ۱۶/۷۵$$

میانه نیمه دوم (چارک سوم) را محاسبه می کنیم.

$$Q = Q_۲ - Q_۱ = \frac{۱۶}{۵} - \frac{۱۰}{۲۵} = ۶/۵$$

$$\frac{Q}{۲} = \frac{۶/۵}{۲} = ۳/۲۵$$

چارک متوسط

بنابراین دامنه تغییر ۵۰ درصدی میانی توزیع برابر ۶/۵ نمره است.

(۲) در اعداد طبقه بندی شده :

برای محاسبه چارکها از فرمول زیر که روش مشابه میانه است استفاده می کنیم.

$$P_p = L \frac{PN - CF_b}{Fi} \times i$$

در یک آزمون پیشرفت تحصیلی توزیع فراوانی نمره های ۴۰ نفر دانش آموز بشرح جدول زیر است دامنه چارکها و چارک متوسط نمرات را محاسبه کنید.

طبقات	Fi	CF
۸۰-۸۴	۲	۴۰
۷۵-۷۹	۳	۳۸
۷۰-۷۴	۷	۳۵
۶۵-۶۹	۱۲	۲۸
۶۰-۶۴	۷	۱۶
۵۵-۵۹	۶	۹
۵۰-۵۴	۳	۳
	N=۴۰	

$$PN = ۰/۲۵ \times ۴۰ = ۱۰$$

(۱) محاسبه اول

باتوجه به فراوانیهای تراکمی از پائین به بالا مشخص می شود که چارک اول در طبقه (۶۴-۶۰) قرار دارد .

$$P_{۲۵} = Q_۱ = ۵۹/۵ + \frac{۱۰-۹}{۷} \times ۵ = ۶۰/۲۱$$

چارک اول

$$PN = ۰/۷۵ \times ۴۰ = ۳۰$$

۲) محاسبه چارک سوم

بنابراین چارک سوم در طبقه ۷۴-۷۰ قرار دارد.

$$P_{۷۵} = Q_۳ = ۶۹/۵ + \frac{۳۰-۲۸}{۷} \times ۵ = ۷۰/۹۳$$

چارک سوم

$$Q = Q_۳ - Q_۱ = ۷۰/۹۳ - ۶۰/۲۱ = ۱۰/۷۲$$

دامنه چارکها

بنابراین دامنه نمره ۵۰ درصد میانی گروه در توزیع فراوانی مذکور برابر ۱۰/۷۲ است.

$$\frac{Q}{۲} = \frac{Q_۳ - Q_۱}{۲} = \frac{۱۰/۷۲}{۲} = ۵/۳۶$$

چارک متوسط

برای محاسبه دامنه دهک ها مشابه آنچه که در مورد دامنه چارک ها محاسبه شد عمل می کنیم یعنی نخست

نقاط $P_۹$ (یعنی $D_۱$) و $P_۱$ (یعنی $D_۱$) را با استفاده از فرمول مذکور محاسبه می کنیم و سپس دامنه دهک ها رااز فرمول $D = D_۹ - D_۱$ بدست می آوریم.

انحراف متوسط از میانگین (A.D)

به فاصله بین هر عدد از یکی از شاخص های مرکزی انحراف گفته می شود و معدل قدرمطلق انحراف^۱ نمره های

یک توزیع را از میانگین انحراف متوسط می نامند.

$$A.D = \frac{\sum |X - \bar{X}|}{N} = \frac{\sum |X|}{N}$$

مثال : مطلوب است محاسبه انحراف متوسط از میانگین نمره های «۱۷-۱۱-۸۶-۱۰-۱۴» را بدست آوریم

X	X=X- $\bar{X}$
۱۷	۱۷-۱۲=۵
۱۴	۱۴-۱۲=۲
۱۱	۱۱-۱۲=-۱
۱۰	۱۰-۱۲=-۲
۸	۸-۱۲=-۴
X=۶۰	

$$\bar{X} = \frac{\sum X}{N} = \frac{۶۰}{۵} = ۱۲ \quad A.D = \frac{۱۴}{۵} = ۲/۸۰$$

بنابراین نمره های جدول از میانگین بطور متوسط ۲/۸۰ نمره انحراف دارند.

۴- واریانس^۲ و انحراف استاندارد^۳^۱-Average deviation^۲-Variance^۳- Standard deviation

واریانس (پراش) عبارت است از معدل، مجذورات انحراف نمره‌ها از میانگین آنها. واریانس اگر در مورد نمونه باشد آنرا به s^2 و اگر درباره جامعه باشد با σ^2 نشان می‌دهند، جذر واریانس را انحراف استاندارد (انحراف معیار) می‌نامند. بنابراین انحراف معیار عبارت است از جذر معدل مجذورات انحراف نمره‌ها از میانگین آنها

فرمول محاسبه واریانس نمونه

$$s^2 = \frac{\sum X^2}{N} = \frac{\sum (X - \bar{X})^2}{N}$$

واریانس جامعه از فرمول $\sigma^2 = \frac{\sum (X - \mu)^2}{N}$ قابل محاسبه است ولی چون میانگین و واریانس جامعه معمولاً نامعلوم است باید براساس داده‌های نمونه برآورد شود. برای برآورد واریانس جامعه براساس واریانس نمونه از فرمول زیر استفاده می‌کنیم.

$$\sigma^2 = \frac{\sum X^2}{N-1} = \frac{\sum (X - \bar{X})^2}{N-1}$$

روشهای محاسبه واریانس و انحراف استاندارد

محاسبه واریانس و انحراف معیار در اعداد طبقه بندی نشده :

هر گاه تعداد نمرات کم باشد برای محاسبه واریانس و انحراف معیار از دو روش استفاده می‌شود :

روش غیر مستقیم : استفاده از انحراف نمره‌ها از میانگین با فرمول‌های بالا

روش مستقیم : استفاده از خود نمرات (X) به جای انحراف آنها از میانگین (X)

اگر در فرمول‌های بالا به جای $\bar{X}$ مقدار آن یعنی $\frac{\sum X}{N}$ را قرار دهیم و فرمول‌ها را ساده کنیم خواهیم داشت .

فرمول محاسبه واریانس نمونه

$$s^2 = \frac{\sum X^2 - \frac{(\sum X)^2}{N}}{N} = \frac{N \sum X^2 - (\sum X)^2}{N^2}$$

مثال : نمره‌های امتحانی یک گروه نمونه ۸ نفری $\sigma^2 = \frac{\sum X^2 - \frac{(\sum X)^2}{N}}{N-1} = \frac{N \sum X^2 - (\sum X)^2}{N(N-1)}$ فرمول برآورد واریانس جامعه

از دانشجویان دانشکده ای بشرح زیر است واریانس و انحراف استاندارد نمره‌ها را برای این گروه محاسبه کرده و واریانس جامعه را نیز برآورد کنید (از هر دو روش مستقیم و غیر مستقیم استفاده کنید) « ۱۱-۳-۱۰-۱۲-۷-۵-۶-۱۰ »

X		$X = X - \bar{X}$	
۱۲۱	۱۴۴	۴	۱۶
۱۱	۱۲۱	۳	۹
۱۰	۱۰۰	۲	۴
۱۰	۱۰۰	۲	۴
۷	۴۹	-۱	۱

6	36	-2	4
5	25	-3	9
3	9	-5	25
$\sum X=64$			

روش غیر مستقیم :

$$s^2 = \frac{\sum X^2}{N} = \frac{72}{9} = 8$$

$$S = \sqrt{9} = 3$$

انحراف استاندارد نمونه

$$\partial^2 = \frac{\sum X^2}{N-1} = \frac{72}{11} = 6.545$$

$$\partial = \sqrt{1.0/28} = 3/2$$

ببر آورد انحراف استاندارد جامعه

روش مستقیم :

واریانس نمونه

$$S^r = \frac{N \sum X^r - (\sum X)^r}{N^r} = \frac{1 \times 514 \times 64^r}{64} = 9$$

بر آورد واریانس جامعه

$$\partial^r = \frac{N \sum X^r - (\sum X)^r}{N(N-1)} = \frac{1 \times 514 \times 64^r}{56} = 1.721$$

همانطوری که ملاحظه می شود نتیجه محاسبه از هر دو روش (مستقیم و غیر مستقیم) یکسان است.

نکته ۱: در تمامی موارد اگر از عدد واریانس جذر بگیریم انحراف معیار توزیع فراوانی بدست می آید.

نکته ۲: هر گاه N بزرگ باشد تفاوت حاصل از کاربرد N و $N-1$ در فرمولهای واریانس جزئی خواهد بود و

نتیجه حاصل از دو فرمول (نمونه و جامعه) تقریباً یکسان می شوند . در بسیاری از موقعیتها N بزرگ نیست و تفاوت محاسبات از دو فرمول ناچیز نخواهد بود در این حالت ارزشی که از تقسیم مجموع مجزورات بر N بدست می آید برآوردی اریب دار از ∂^2 (واریانس جامعه) است . این ارزش توصیف کننده ∂^2 است که دارای یک خطای منظم است ، بنابراین نسبت $\frac{\sum X^2}{N-1}$ برآوردی بدون اریب از ∂^2 بدست خواهد داد .

تصحیح شیرد^۱

'-Sheppard Correction

انحراف استاندارد که از داده های طبقه بندی شده کمتر از ۱۲ طبقه به دست می آید با آنچه که در محاسبه روی داده های اصلی محاسبه می شود تا حدی متفاوت است ، علت آن بخاطر استفاده از عدد میانی طبقات به جای داده های اصلی است. بهمین جهت اصلاحی بنام «تصحیح شیرد» روی انحراف استاندارد محاسبه شده صورت می گیرد . این اصلاح به صورت فرمول زیر است.

$$S_C = \sqrt{S^2 - \frac{i^2}{12}}$$

در این فرمول S_C انحراف معیار تصحیح ، S^2 واریانس محاسبه شده و i فاصله طبقات است.

مثال: در یک جدول توزیع فراوانی که دارای ۸ طبقه است انحراف معیار برابر ۵ محاسبه شده است اگر فاصله طبقات ۳ باشد انحراف معیار تصحیح شده را محاسبه کنید .

$$S_C = \sqrt{5^2 - \frac{3^2}{12}} = \sqrt{24/5} = 4/92$$

ضریب تغییرپذیری^۱ (CV)

ضریب تغییرپذیری برابر است با نسبت انحراف استاندارد یک توزیع بر میانگین آن (می توان به جای نسبت درصد محاسبه کرد)

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

از این شاخص می توان برای مقایسه پراکندگی نمونه های مختلف استفاده کرد.

مثال: در جدول زیر میانگین و انحراف استاندارد نمره های امتحانی دانش آموزان پایه پنجم دو مدرسه پسرانه و دخترانه در یک آزمایش نشان داده شده است پراکندگی نمره ها در کدام مدرسه بیشتر است ؟

جنس	شاخص	میانگین	انحراف معیار
پسر		۱۵/۵	۲
دختر		۱۷/۵	۲/۵

$$C.V_G = \frac{2/5}{17/5} = 0/14$$

$$C.V_B = \frac{2}{5/15} = 0/12$$

ضریب پراکندگی

هر چه CV کمتر باشد نمره ها متجاس ترند بنابراین پراکندگی نمره ها در مدرسه دختران بیشتر از پسران است و نمرات در مدرسه دختران نا متجاس ترند.

۲) شاخص های پراکندگی در مورد داده های مقیاس اسمی و رتبه ای

^۱ - Coefficient of Variance

در مورد داده های مقیاس اسمی و رتبه ای می توان نسبت همگونی یا ناهمگونی طبقات را که بنام نسبت پراکندگی^۱ (VR) خوانده می شود محاسبه کرد.

$$VR = 1 - \frac{Fm}{N}$$

در این فرمول منظور از Fm فراوانی ماگزیمم در طبقات جدول است.

مثال: در جدول زیر توزیع فراوانی دانشجویان یک کلاس برحسب گروه خونی نشان داده شده است نسبت پراکندگی توزیع را محاسبه کنید.

فراوانی طبقات

$$V.R = 1 - \frac{Fm}{N} = 1 - \frac{15}{45} = 0.67$$

A ۱۲

B ۸

AB ۱۵

O ۱۰

N=۴۵

هرچه مقدار VR به سمت صفر میل کند پراکندگی می شود در حالت خاصی که در توزیع فقط یک طبقه موجود باشد مقدار VR مساوی صفر خواهد شد.

باید بدانیم که در مورد داده های مقیاس ترتیبی علاوه بر VR می توان دامنه بین صدکها و دهکها را نیز محاسبه کرد. به عنوان مثال در جدول توزیع فراوانی زیر که مربوط به نوع نگرش دانشجویان یک کلاس به نحوه ی مدیریت دانشکده است دامنه چارک ها و دهک ها را محاسبه کنیم.

نوع نگرش	فراوانی	CF
جداً موافقم (۵)	۸	۴۰
موافقم (۴)	۱۰	۳۲
بی نظرم (۳)	۱۱	۲۲
مخالفم (۲)	۷	۱۱
جداً مخالفم (۱)	۴	۴
	N=۴۰	

(۱) محاسبه دامنه چارکها

$$PN = 40 \times \frac{1}{4} = 10$$

(۱) چارک اول

$$Q_1 = (2) \text{ طبقه}$$

بنابراین چارک اول در طبقه (۲) یعنی « مخالفم » قرار می گیرد.

$$PN = 40 \times \frac{3}{4} = 30$$

(۲) چارک سوم

^۱ -Variation ratio

بنابراین چارک سوم در طبقه (۴) یعنی «موافقم» قرار می‌گیرد.
طبقه (۴) $Q_3 =$
دامنه چارکها $Q = Q_3 - Q_1 = 4 - 2 = 2$
بنابراین دامنه نگرش ۵۰ درصد میانی گروه ۲ طبقه فاصله دارد.

۲) محاسبه دامنه دهکها

(۱) دهک اول $PN = 40 \times 0.1 = 4$
بنابراین دهک اول در طبقه (۱) یعنی «جداً مخالفم» قرار دارد.
طبقه (۱) $D_1 =$
(۲) دهک نهم $PN = 40 \times 0.9 = 36$
بنابراین دهک نهم در طبقه (۵) یعنی «جداً مخالفم» قرار دارد.
طبقه (۵) $D_9 =$
دامنه دهکها $D = D_9 - D_1 = 5 - 1 = 4$
پس می‌توان گفت دامنه نگرش ۸۰ درصد میانی گروه حدود ۴ طبقه پراکندگی دارد.

۳) ترکیب نمره ها با یک عدد ثابت و تأثیر آن در شاخص های پراکندگی

۱) تأثیر جمع و تفریق :

اگر به همه اعداد یک توزیع ، یک عدد ثابت اضافه شود به اینکه میانگین نیز به همان اندازه افزایش می‌یابد بنابراین انحراف نمره ها از میانگین و همچنین تفاوت نمره ها از هم نیز تغییر نخواهد کرد و بنابراین در هیچیک از شاخصهای پراکندگی تأثیری نخواهد داشت. در مورد تأثیر تفریق یک عدد ثابت از همه نمره نیز نتیجه مانند تأثیر جمع خواهد بود.

۲) تأثیر ضرب و تقسیم :

اگر همه اعداد یک توزیع در عدد ثابتی ضرب شوند میانگین اعداد نیز در همان عدد ضرب می‌شود. در نتیجه انحراف نمره ها از میانگین و تفاوت نمره ها از هم نیز به همان اندازه بزرگ می‌شود بنابراین همه شاخص های پراکندگی در همان عدد ضرب خواهند شد ولی چون واریانس براساس مجذور انحراف نمره ها محاسبه می‌شود در نتیجه واریانس در مجذور آن عدد ثابت ضرب خواهد شد. در مورد تقسیم همه اعداد بر یک عدد ثابت نیز نتیجه همانند تأثیر ضرب ولی در جهت مقابل خواهد بود یعنی بجای ضرب ، تقسیم می‌شوند.
به عنوان مثال رابطه های مربوط به ترکیب نمره ها با یک عدد ثابت و تأثیر آن در واریانس بصورت فرمول های زیر نشان داده شده است.

$$\begin{aligned} S^2(X + C) &= S_X^2 \\ S^2(X - C) &= S_X^2 \\ S^2(X \cdot C) &= C^2 \cdot S_X^2 \\ S^2\left(\frac{X}{C}\right) &= \frac{S_X^2}{C^2} \end{aligned}$$

کجی

یکی از موارد استفاده از انحراف چارکی تعیین کجی نمره ها است. کجی یعنی انحراف یک منحنی از حالت تقارن. در توزیع متقارن چون فاصله بین چارک اول (Q_1) و چارک سوم (Q_3) تا میانه (چارک دوم Q_2) برابر است بنابراین کجی برابر صفر است. رابطه بین چارکها در منحنی های مختلف به صورت فرمول های زیر است.

$$Q_3 - Q_2 = Q_2 - Q_1 \text{ منحنی متقارن یا کجی صفر}$$

$$Q_3 - Q_2 > Q_2 - Q_1 \text{ کجی مثبت}$$

$$Q_3 - Q_2 < Q_2 - Q_1 \text{ کجی منفی}$$

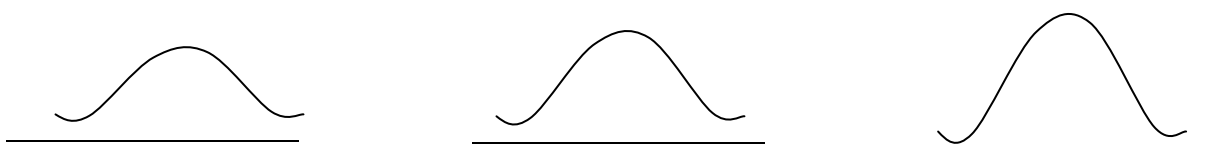
لذا در یک توزیع متقارن $m_3 = 0$ و در نتیجه $SK = 0$ می شود اما در توزیع دارای کجی مثبت مکعب مجموع انحراف نمره ها از میانگین عددی مثبت و در توزیع دارای کجی منفی مکعب مجموع انحراف نمره ها از میانگین عددی منفی خواهد شد.

می توان ضریب کجی را با استفاده از فرمولهای زیر که به فرمول کجی پیرسون معروف است محاسبه کرد.

$$G_1 = SK = \frac{\bar{X} - M_0}{s}$$

کشیدگی

بعضی وقت ها ممکن است میانگین، میانه و نمای گروه مورد مطالعه برابر و شکل منحنی ها متقارن باشد ولی پراکندگی حول میانگین یعنی انحراف معیار در گروه های مختلف تفاوت داشته باشند در این صورت منحنی ها مانند شکل زیر از لحاظ کشیدگی متفاوت خواهند بود مقدار کشیدگی را با K_U یا G_2 نشان می دهند. در منحنی بهنجار $K_U = 0$ است. اگر $K_U > 0$ باشد (در حالت $K_U > 0.5$) منحنی را بر آمده^۱ (در اوج) و اگر $K_U < 0$ (در حالت $K_U < -0.5$) منحنی را پخت^۲ یا خوابیده می نامند.



منحنی با خوابیدگی (پخت)
(تخت قله)

منحنی طبیعی (بهنجار)
(میانه قله)

منحنی برآمدگی (در اوج)
(تیز قله)

بنابراین در حالت $-0.5 < k_u \leq +0.5$ توزیع دارای کشیدگی نزدیک به طبیعی تلقی می شود.

جدول مقایسه خصوصیات شاخص های پراکندگی

شاخص	نقاط قوت	نقاط ضعف
------	----------	----------

^۱-Lepto Kurtic

^۲-Platy Kurtic



۱- تحت تأثیر نمره های خیلی بزرگ یا خیلی کوچک است. ۲- در مورد داده های مقیاس ترتیبی و اسمی قابل محاسبه نیست.	۱- پایدارترین پراکندگی است. ۲- در محاسبه آن از کلیه نمره ها استفاده می شود . ۳- شعاع چرخش گشتاور ماند^۱ دستگاه توزیع فراوانی به دور محور گرانیگاه است.	انحراف استاندارد
۱- در مقایسه با انحراف معیار ثبات کمتری دارد. ۲- در مورد داده های مقیاس اسمی قابل محاسبه نیست.	۱- تحت تأثیر نمره های بزرگ یا کوچکی نیست. ۲- بانمره های حدود ۵۰ درصد میانی مربوط است. ۳- در مورد داده های مقیاس ترتیبی بکار می رود.	دامنه چارک ها
۱- ثبات ندارد ۲- فقط با کمترین و بیشترین نمره ارتباط دارد. ۳- در مورد داده های مقیاس اسمی و رتبه ای قابل محاسبه نیست.	۱- محاسبه آن آسان است. ۲- اطلاعاتی درباره نمره های بالا و پائین به ما می دهد . ۳- در مورد داده های مقیاس فاصله ای و نسبی قابل محاسبه است.	دامنه تغییرات

سوالات خودسنجی

۱- واریانس نمرات یک توزیع برابر ۳۶ است .

(۱) به هریک از نمرات ۵ واحد اضافه می کنیم انحراف استاندارد توزیع جدید را محاسبه کنید ؟

پاسخ : $S=6$

(۲) همه نمره ها را نصف می کنیم انحراف استاندارد توزیع جدید چقدر می شود ؟

پاسخ : ۳

۲- اگر توزیع داده های $X_1, X_2, X_3, \dots, X_N$ انحراف استاندارد برابر ۴ باشد انحراف استاندارد توزیع

داده های $(2X_1 - 3), (2X_2 - 3), \dots, (2X_n - 3)$ چند است ؟

پاسخ : ۸

۳- میزان علاقه گروهی از دانشجویان به رشته تحصیلی خود در جدول زیر نشان داده شده است.

میزان علاقمندی	فراوانی
بسیار زیاد (۵)	۲۰
زیاد (۴)	۳۵
متوسط (۳)	۳۰
کم (۲)	۱۰
بسیار کم (۱)	۵
	$N=100$

^۱ -Moment of Inertia

۱) دامنه دهکها و دامنه چارکها را محاسبه کنید ؟

پاسخ : طبقه $D=$ و طبقه $Q=1$

۲) نسبت واریانس (VR) را بدست آورید ؟

پاسخ : $V.R=0.65$