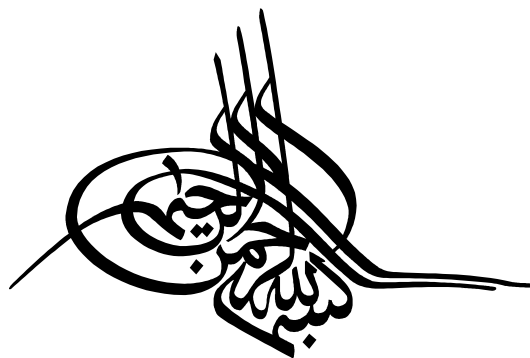




آمار و روش تحقیق

سری کتاب‌های کمک آموزشی کارشناسی ارشد
و دکتری

مجموعه روان‌شناسی

مؤلف: میلاد شهابی

آمار و روش تحقیق / میلاد شهابی / ۱۳۷۴

مشاوران صعود ماهان،

۴۵۶ ص: جدول، نمودار (آمادگی آزمون کارشناسی ارشد و دکتری)

ISBN: ۹۷۸-۶۰۰-۷۴۶۶-۲۴-۷

فهرست نویسی بر اساس اطلاعات فیپا.

فارسی - چاپ اول

۱- آمار و روش تحقیق

۳- آزمون دوره‌های تحصیلات تکمیلی

ج - عنوان

رده‌بندی دیویی:

شماره کتابشناسی ملی:

۲- آزمون‌ها و تمرین‌ها (عالی)

۴- دانشگاه‌ها و مدارس عالی - ایران - آزمون‌ها

مؤسسه آموزش عالی آزاد



نام کتاب: آمار و روش تحقیق

مدیر مسئول: مجید سیاری

مؤلف: میلاد شهابی

مدیر تولید: سمیه بیگی

ناشر: مشاوران صعود ماهان

نوبت و تاریخ چاپ: اول / ۱۴۰۴

تیراژ: ۱۰۰۰ جلد

قیمت: ۵/۵۰۰/۰۰۰ ریال

شابک: ISBN: ۹۷۸-۶۰۰-۷۴۶۶-۲۴-۷

انتشارات مشاوران صعود ماهان: خیابان ولیعصر، بالاتر از تقاطع مطهری،

روبروی قنادی هتل بزرگ تهران، جنب بانک ملی، پلاک ۲۵۰

تلفن: ۸۸۱۰۰۱۱۳-۴

سخن ناشر

آیا آنانکه می‌دانند با آنانکه نمی‌دانند برابرند؟ (قرآن کریم)

پس از حمد و سپاس و ستایش به درگاه بی‌همتای احدیت و درود بر محمد مصطفی، عالی نمونه بشریت که در تاریخ دور تاریخ، بنا به فرمان نافذ صمدیت از میان مردمی برخاست که خود بودند در پست‌ترین حد توحش و ضلال و بربریت و آنگاه با قوانین شامل خویش هم ایشان را راهبری نمود و رهانید از بدویت و استعانت جوییم از قرآن کریم، کتابی که هست جاودانه و بی‌نقص تا ابدیت.

کتابی که در دست دارید آخرین ویرایش از مجموعه کتب خودآموز مؤسسه آموزش عالی آزاد ماهان است که بر مبنای خلاصه درس و تأکید بر نکات مهم و کلیدی و تنوع پرسش‌های چهار گزینه‌ای جمع‌آوری شده است. در این ویرایش ضمن توجه کامل به آخرین تغییرات در سرفصل‌های تعیین شده جهت آزمون‌های ارشد تلاش گردیده است که مطالب از منابع مختلف معتبر و مورد تأکید طراحان ارشد با ذکر مثال‌های متعدد بصورت پرسش‌های چهار گزینه‌ای با کلید و در صورت لزوم تشریح کامل ارائه گردد تا دانشجویان گرامی را از مراجعه به سایر منابع مشابه بی‌نیاز نماید.

لازم به ذکر است شرکت در آزمون‌های آزمایشی ماهان که در جامعه آماری گسترده و در سطح کشور برگزار می‌گردد می‌تواند محک جدی برای عزیزان دانشجو باشد تا نقاط ضعف احتمالی خود را بیابند و با مرور مجدد مطالب این کتاب، آنها را برطرف سازند که تجربه سال‌های مختلف موکد این مسیر به‌عنوان مطمئن‌ترین راه برای موفقیت می‌باشد.

لازم به ذکر است از پورتال ماهان به آدرس www.mahanportal.ir می‌توانید خدمات پشتیبانی را دریافت دارید.

و نیز بر خود می‌بالیم که همه ساله میزان تطبیق مطالب این کتاب با سؤالات آزمون‌های ارشد و دکتری که از شاخصه‌های مهم ارزیابی کیفی این کتاب‌ها می‌باشد- ما را در محضر شما سر بلند می‌نماید.

در خاتمه بر خود واجب می‌دانیم که از همه اساتید بزرگوار و دانشجویان ارجمند از سراسر کشور و حتی خارج از کشور و همه همکاران گرامی که با ارائه نقطه نظرات سازنده خود ما را در پربارتر کردن ویرایش جدید این کتاب یاری نمودند سپاسگزاری نموده و به پاس تلاش‌های بی‌چشمداشت، این کتاب را به محضرشان تقدیم نماییم.

بخش آمار

فصل اول: آمار چیست؟	۶
سوالات طبقه‌بندی شده به همراه پاسخ تشریحی فصل اول	۱۴
فصل دوم: شاخص‌های مرکزی	۱۶
سوالات طبقه‌بندی شده به همراه پاسخ تشریحی فصل دوم	۲۶
فصل سوم: شاخص‌های پراکندگی	۳۴
سوالات طبقه‌بندی شده به همراه پاسخ تشریحی فصل سوم	۴۱
فصل چهارم: نمرات استاندارد و منحنی طبیعی	۴۹
سوالات طبقه‌بندی شده به همراه پاسخ تشریحی فصل چهارم	۵۶
فصل پنجم: همبستگی	۶۲
سوالات طبقه‌بندی شده به همراه پاسخ تشریحی فصل پنجم	۷۰
فصل ششم: رگرسیون	۷۹
سوالات طبقه‌بندی شده به همراه پاسخ تشریحی فصل ششم	۸۴
فصل هفتم: احتمالات	۸۸
سوالات طبقه‌بندی شده به همراه پاسخ تشریحی فصل هفتم	۹۶
فصل هشتم: آمار استنباطی (برآورد)	۱۰۲
سوالات طبقه‌بندی شده به همراه پاسخ تشریحی فصل هشتم	۱۰۸
فصل نهم: آمار استنباطی (آزمون فرضیه)	۱۱۴
سوالات طبقه‌بندی شده به همراه پاسخ تشریحی فصل نهم	۱۲۰
فصل دهم: آزمون‌های ناپارامتریک	۱۲۷
سوالات طبقه‌بندی شده به همراه پاسخ تشریحی فصل دهم	۱۳۶
فصل یازدهم: آزمون پارامتریک	۱۴۳
سوالات طبقه‌بندی شده به همراه پاسخ تشریحی فصل یازدهم	۱۵۱
فصل دوازدهم: آزمون‌های مقایسه چندگروهه	۱۵۶
سوالات طبقه‌بندی شده به همراه پاسخ تشریحی فصل دوازدهم	۱۶۵
فصل سیزدهم: رگرسیون چندمتغیره	۱۷۴
فصل چهاردهم: پیوست	۱۸۰

بخش روش تحقیق

فصل اول: متغیر چیست؟	۲۰۵
سوالات طبقه‌بندی شده به همراه پاسخ تشریحی فصل اول	۲۱۵
فصل دوم: کلیات تحقیق	۲۲۵
سوالات طبقه‌بندی شده به همراه پاسخ تشریحی فصل دوم	۲۳۲
فصل سوم: موضوع تحقیق (مسأله و فرضیه)	۲۳۷
سوالات طبقه‌بندی شده به همراه پاسخ تشریحی فصل سوم	۲۴۴
فصل چهارم: ابزارهای سنجش و اندازه‌گیری	۲۴۹
سوالات طبقه‌بندی شده به همراه پاسخ تشریحی فصل چهارم	۲۵۶
فصل پنجم: روایی و پایایی	۲۶۰

۲۶۵	سوالات طبقه‌بندی شده به همراه پاسخ تشریحی فصل پنجم.....
۲۶۹	فصل ششم: جامعه و نمونه.....
۲۷۶	سوالات طبقه‌بندی شده به همراه پاسخ تشریحی فصل ششم.....
۲۸۵	فصل هفتم: روش تحقیق کمی و کیفی.....
۲۹۴	سوالات طبقه‌بندی شده به همراه پاسخ تشریحی فصل هفتم.....
۲۹۹	فصل هشتم: تحقیقات کیفی.....
۳۰۴	سوالات طبقه‌بندی شده به همراه پاسخ تشریحی فصل هشتم.....
۳۰۷	فصل نهم: تحقیقات کمی (غیرآزمایشی).....
۳۱۱	سوالات طبقه‌بندی شده به همراه پاسخ تشریحی فصل نهم.....
۳۱۵	فصل دهم: تحقیقات کمی (آزمایشی).....
۳۲۴	سوالات طبقه‌بندی شده به همراه پاسخ تشریحی فصل دهم.....
۳۳۳	سوالات و پاسخنامه تشریحی ارشد و دکتری از سال ۹۰ الی ۱۴۰۴.....

فصل اول

آمار چیست؟

آمار چیست؟

آمار شاخه‌ای از روش‌شناسی علمی است که با جمع‌آوری، طبقه‌بندی، توصیف و تفسیر داده‌هایی که از اجرای زمینه‌یابی‌ها و آزمایش‌ها بدست می‌آید، سر و کار دارد.

هدف آمار.....

روش‌های آماری دارای دو هدف (وظیفه) هستند.

الف) به پژوهشگران در طبقه‌بندی، تلخیص (خلاصه کردن: مهم‌ترین کارکرد)، توصیف و تفسیر اطلاعات جمع‌آوری شده و برقراری ارتباط از طریق آن‌ها کمک می‌کند — آمار توصیفی

ب) به پژوهشگران این امکان می‌دهد، که با استفاده از اطلاعات جمع‌آوری شده از نمونه کوچکی از آزمودنی‌ها، ویژگی‌های جامعه‌ای را که نمونه از آن انتخاب شده است، برآورد یا استنباط کند (تبیین و پیش‌بینی) — آمار استنباطی

آمار توصیفی

هنگامی که اطلاعات کمی جمع‌آوری می‌شوند، برای تفسیر آن‌ها ابتدا باید به روش قابل فهم و منظم، سازمان‌بندی و خلاصه شوند، که این کار توسط آمار توصیفی انجام می‌شود.

همانطور که بالاتر به آن اشاره شد، هدف آمار توصیفی: طبقه‌بندی، تلخیص و توصیف اطلاعات است، که از این بین مهم‌ترین هدف آمار توصیفی تلخیص (خلاصه‌سازی) داده‌هاست.

آمار توصیفی شامل مباحث زیر می‌شود:

جدول توزیع فراوانی، نمودارها، شاخص‌های مرکزی، شاخص‌های پراکندگی، نمرات استاندارد و منحنی طبیعی، همبستگی و رگرسیون.

آمار استنباطی

آمار استنباطی به روش‌هایی گفته می‌شود که در آن ویژگی‌های جامعه به کمک ویژگی‌های نمونه برآورد می‌شود (تعمیم).

به طور کلی هدف آمار استنباطی تعمیم اصول و یافته‌ها برای تبیین و پیش‌بینی پدیده‌هاست.

آمار استنباطی به دو دسته آزمون فرضیه و تعمیم تقسیم می‌شود:
تعمیم: به معنای برآورد ویژگی‌های جامعه (پارامتر) از روی ویژگی‌های نمونه (آماره) است. برای توضیح بیشتر در مورد فرآیند تعمیم، ابتدا لازم است با دو مفهوم جامعه و نمونه آشنا شویم.
آزمون فرضیه: آزمون‌های آماری، وظیفه سنجش معناداری فرضیه را به عهده دارند. در فصول آینده انواع آزمون‌های فرضیه معرفی خواهند شد.

جامعه.....

جامعه: عبارت است از گروهی از افراد، اشیاء یا حوادث که حداقل دارای یک صفت یا ویژگی مشترک هستند (عمل تعمیم‌پذیری در مورد آن‌ها صورت بگیرد). ← هدف نهایی آمار استنباطی: برآورد ویژگی‌های جامعه است.

نکته ۱. اصطلاح پارامتر در مورد جامعه به کار می‌رود و به معنای ویژگی عددی یک جامعه است (معمولاً با حرف یونانی نشان داده می‌شوند).

نمونه.....

چون حجم یا اندازه جوامع معمولاً بسیار بزرگ است، اندازه‌گیری تک تک اعضای جامعه امری غیرممکن است و برای تحلیل نیازمند نمونه‌گیری هستیم. نمونه عبارت است از زیرجامعه‌ای که از کل جامعه انتخاب می‌شود و معرف آن است (روش‌های نمونه‌گیری در روش تحقیق به شکل کامل توضیح داده خواهد شد).

نکته ۲. اصطلاح آماره در مورد نمونه به کار می‌رود و عبارتست از ویژگی یا ویژگی‌های کمی که یک نمونه را توصیف می‌کند (معمولاً با حروف انگلیسی نشان داده می‌شوند).
به جدول زیر دقت کنید.

ویژگی شاخص	میانگین	واریانس	انحراف معیار	نسبت	همبستگی	تعداد مشاهدات
آماره	$\bar{X}$	S^2	S	P	r	n
پارامتر	μ (مو)	σ^2	σ (سیگما)	π	ρ (رو)	N

مثال ۱. مهمترین کاربرد آمار توصیفی کدام است؟ (ارشد ۹۲)

- الف) تلخیص داده‌ها
ب) گردآوری اطلاعات
ج) آزمون فرضیه‌ها
د) تعیین شاخصهای گرایش به مرکز
- پاسخ: گزینه الف. مهم‌ترین هدف آمار توصیفی تلخیص (خلاصه‌سازی) داده‌هاست
- مثال ۲.** کدام مورد هدف از محاسبه آماره‌ها (statistics) براساس داده‌های حاصل از نمونه است؟ (ارشد ۹۵)

- الف) برآورد ویژگی‌های جامعه از طریق تعمیم به نمونه مورد نظر
ب) توصیف ویژگی‌های نمونه و جامعه بر اساس آماره‌ها
ج) کاهش حجم داده‌ها و افزایش دقت شاخص‌های آماری
د) توصیف نمونه و فرض آزمایی پیرامون ویژگی‌های جامعه
- پاسخ: گزینه د. هدف محاسبه آماره‌ها، توصیف نمونه (آمار توصیفی) و فرض‌آزمایی (آمار استنباطی) است.

توزیع فراوانی و نمایش نموداری

داده‌های بدست آمده از آزمایش‌ها، غالباً به صورت مجموعه‌ای از اعداد و ارقام هستند. در این بخش نحوه طبقه‌بندی داده‌ها در جدول فراوانی برای سهولت مطالعه و بررسی اعداد آموخته می‌شود. بهتر است قبل از شروع بحث با مفهومی تحت عنوان حدود واقعی اعداد آشنا بشویم.

.....حدود واقعی اعداد.....

طبق قرارداد آمار برای سنجش دقیق تر، هر عدد صحیح را دامنه‌ای از ۰,۵ واحد پایین تر تا ۰,۵ واحد بالاتر از همان عدد، در نظر می گیریم. به این بازه حدود واقعی اعداد گفته می شود. برای مثال حدود واقعی عدد ۱۵، ۱۴,۵ تا ۱۵,۵ است و یا حدود واقعی عدد ۱۵,۵، به علت اینکه واحد اندازه گیری آن ۰,۵ است، از ۱۵,۴۵ تا ۱۵,۵۵ است.

حدود واقعی اعداد زیر به این طریق نوشته می شود.

$$12 = 11.5 \text{ --- } 12.5$$

$$11.50 = 11.45 \text{ --- } 11.55$$

$$11.450 = 11.445 \text{ --- } 11.455$$

نکته ۳. برای راحت بدست آوردن حدود واقعی می توان یک عدد صفر به انتهای عدد (یا اعشار) اضافه کرد و آن را از عدد ۵ کم و اضافه کرد. برای مثال عدد ۱۲ را ۱۲۰ نوشت و آن را از ۵ کم کرد، ۱۱۵ و به آن اضافه کرد ۱۲۵. که نشان دهنده عدد ۱۱,۵ و ۱۲,۵ است و یا عدد ۱۱,۴۵ نیز با اضافه شدن صفر، حدود واقعی اش ۱۱,۴۴۵ و ۱۱,۴۵۵ خواهد شد

مثال ۳. حد پایین عدد ۲,۸ مساویست با؟

الف) ۲,۰۰

ب) ۲,۵

ج) ۲,۷۵

د) ۲,۷۹

پاسخ: گزینه ج. به انتهای عدد صفر اضافه می شود و سپس از عدد ۵ کم می شود، که نتیجه آن به عدد ۲,۷۵ خواهیم رسید.

جدول توزیع فراوانی:

.....جدول توزیع فراوانی طبقه بندی نشده.....

توزیع فراوانی عبارت است از سازمان دادن به اندازه ها یا مشاهده ها به وسیله قرار دادن آن ها در قالب طبقه ها همراه با ذکر فراوانی هر طبقه.

در جدول غیر طبقاتی، به ازای هر داده یک طبقه وجود دارد، که تعداد آن داده را در ستونی تحت عنوان فراوانی ذکر می کنیم. عیب استفاده از این نوع جدول، وقت گیر و زمان بر بودن است.

به عنوان مثال سکه ای را ده بار پرتاب کرده ایم. از این ۱۰ بار ۷ بار شیر و ۳ بار خط آمده است. جدول آن به شکل زیر رسم می شود (گاهی از خط نشان (/)، نیز برای نشان دادن فراوانی استفاده می شود).

x	خط نشان	F (فراوانی)
شیر	////////	۷
خط	///	۳
کل ($\sum f = N$)	////////	۱۰

نکته ۴. دقت کنید برای نشان دادن داده به شکل خام، معمولاً از حرف x استفاده می شود و برای نشان دادن فراوانی از علامت f استفاده می شود.

نکته ۵. علامت $\sum$ به معنای مجموع است. در جدول بالا $\sum f$ به معنای مجموع فراوانی ها، یا به تعریف دیگر به معنای تعداد کل داده ها (N) است.

جدول توزیع فراوانی طبقه‌بندی شده.....

هنگامی که تعداد اعداد یک توزیع و همچنین فاصله بین آن‌ها (پراکندگی) خیلی زیاد باشد، از توزیع فراوانی طبقه‌بندی شده استفاده می‌شود. برای مثال: نمره‌های آمار یک کلاس ۳۵ نفره روانشناسی به شکل زیر است:

۴، ۶، ۷، ۸، ۸، ۱۱، ۱۱، ۱۱، ۱۱، ۱۱، ۱۱، ۱۱، ۱۱، ۱۲، ۱۲، ۱۳، ۱۳، ۱۳، ۱۳، ۱۳، ۱۴، ۱۴، ۱۴، ۱۴، ۱۸، ۱۸، ۱۸، ۱۸، ۱۸، ۱۸، ۱۹، ۱۹، ۱۹، ۱۹، ۱۹

که زمانی که در جدول کشیده می‌شود به صورت زیر است.

F (فراوانی)	x
۱	۵-۱
۴	۱۰-۶
۲۰	۱۵-۱۱
۱۰	۲۰-۱۶
۳۵	کل ($\Sigma f = N$)

عیب جدول طبقه بندی شده، خطای طبقه‌بندی است. هنگام طبقه‌بندی اطلاعات به شکل طبقه‌بندی، خطای طبقه‌بندی رخ خواهد داد که به معنی این است که دقیقاً نمی‌توان مشخص کرد که برای مثال در طبقه دوم (۶-۱۰)، این ۴ عدد، دقیقاً چه اعدادی هستند.

نکته ۶. اگر طول طبقات زیاد شود، خطای طبقه‌بندی افزایش (رابطه مستقیم) و اگر تعداد طبقات زیاد شود، خطای طبقه‌بندی کاهش (رابطه معکوس) پیدا می‌کند.

ساخت توزیع فراوانی طبقه‌بندی شده.....

برای ساخت یک توزیع فراوانی باید مراحل زیر را طی کنیم.

مرحله اول: محاسبه دامنه تغییرات:

تفاضل بزرگترین عدد از کوچکترین عدد به اضافه ۱ (اضافه کردن یک بخاطر در نظر گرفتن حدود واقعی اعداد است).

$$R = \max - \min + 1$$

min = بزرگترین عدد

min = کوچکترین عدد

R = دامنه تغییرات

مرحله دوم: تعیین تعداد طبقات:

معمولاً تعداد طبقات بین ۱۰-۲۰ است. برای مشخص کردن دقیق تعداد طبقات می‌توان از قانونی که به قانون استرژ معروف است، استفاده کرد.

$$k = 1 + 3.3 \log n$$

n = تعداد کل داده‌ها

k = تعداد طبقات

مرحله سوم: تعیین اندازه یا طول هر طبقه است که با حرف i نمایش داده می‌شود.

$$i = \frac{R}{K}$$

R = دامنه تغییرات

k = تعداد طبقات

i = طول طبقات

طول هر طبقه از طریق تفاضل بزرگترین عدد هر طبقه (کران بالای طبقه) از کوچکترین عدد هر طبقه (کران پایین هر طبقه) به اضافه ۱ نیز بدست می آید.

$$i = x_H - x_L + 1$$

x_H = کران بالای طبقه

x_L = کران پایین طبقه

i = طول طبقات

نمونه یک جدول توزیع فراوانی طبقاتی.....

X (نمره خام)	X_c (مرکز دسته)	f (فراوانی مطلق)	Cf (فراوانی تجمعی)	$\frac{f}{N}$ (فراوانی نسبی)	$\frac{f}{N} \times 100$ (درصد فراوانی نسبی)	$\frac{Cf}{N}$ (فراوانی تراکمی نسبی)	$\frac{Cf}{N} \times 100$ (درصد فراوانی تراکمی)
۵-۱	۳	۱	1	0.1	10	0.1	10
۱۰-۶	۸	۴	5	0.4	40	0.5	50
۱۵-۱۱	۱۳	2	7	0.2	20	0.7	70
۲۰-۱۶	۱۸	3	10	0.3	30	1	100
مجموع (Σ)		$N = 10$		1	100		

- ستون مرکز دسته (X_c): در این ستون عددی که دقیقا وسط دو کران بالا و پایین دسته قرار می گیرد، نوشته می شود. برای مثال در ردیف اول، عددی که بین ۱ و ۵ قرار می گیرد، عدد ۳ است.

نکته ۷. فاصله بین مرکز دسته ها، مساوی با طول طبقات است. در جدول بالا طول طبقات ۵ است و فاصله مرکز دسته ها از هم مساوی با ۵ است (مثلا مرکز دسته طبقه اول ۳ و مرکز دسته طبقه دوم ۸ است).

- فراوانی مطلق (f): مساوی با تعداد داده در هر طبقه است.

نکته ۸. همانطور که در جدول بالا معلوم است، تعداد کل داده ها برابر است با مجموع فراوانی مطلق داده ها:

$$\sum f(x) = N$$

- فراوانی تجمعی (Cf): فراوانی تجمعی (تراکمی) مساویست با مجموع فراوانی طبقه موردنظر با فراوانی طبقه های قبلی.

نکته ۹. تعداد کل داده ها برابر است با فراوانی تجمعی دسته آخر. همچنین همواره عدد فراوانی مطلق ردیف اول با فراوانی تجمعی ردیف اول برابر است، چون هیچ طبقه ای قبل از آن وجود نداشته است.

- فراوانی نسبی (f/N): فراوانی نسبی (نسبتی) مساویست با فراوانی طبقه موردنظر تقسیم بر تعداد کل داده ها.

- درصد فراوانی نسبی ($\times 100 f/N$): فراوانی نسبی (نسبتی) مساویست با فراوانی طبقه موردنظر تقسیم بر تعداد کل داده ها ضربدر عدد ۱۰۰.

- فراوانی تراکمی نسبی (Cf/N): فراوانی تراکمی نسبی (نسبتی) مساویست با فراوانی تراکمی طبقه موردنظر تقسیم بر تعداد کل داده ها.

- درصد فراوانی نسبی ($\times 100 cf/N$): فراوانی تراکمی نسبی (نسبتی) مساویست با فراوانی تراکمی طبقه موردنظر تقسیم بر تعداد کل داده ها ضربدر عدد ۱۰۰.

نکته ۱۰. همواره مجموع فراوانی نسبی و فراوانی تراکمی نسبی آخرین دسته در هر توزیعی برابر با عدد ۱ و همواره مجموع درصد فراوانی و درصد فراوانی تراکمی نسبی آخرین دسته نسبی مساوی با ۱۰۰ است.

مثال ۴. در توزیع فراوانی روبه رو، فراوانی نسبی متناظر با عدد ۹ چقدر است؟

X	۳	۵	۷	۹	۱۰
F	۹	۹	۱۲	۱۰	۰

(د) ۰,۲۵

(ج) ۰,۲۴

(ب) ۰,۲۰

(الف) ۰,۱۵

پاسخ: گزینه د. فراوانی روبه روی عدد ۹، ۱۰ است و از آنجایی که جمع تمام فراوانی‌ها برابر با تعداد داده‌ها است، پس داریم:

$$\sum f = N = 9 + 9 + 12 + 10 = 40$$

$$f/N = \frac{10}{40} = \frac{1}{4} = 0,25$$

مثال ۵. فاصله طبقاتی کدام طبقه برابر با ۷ است؟

(د) ۸-۳

(ج) ۹-۳

(ب) ۸-۱

(الف) ۱۰-۳

پاسخ: گزینه ج. برای طول طبقات داریم:

$$i = x_H - x_L + 1 = 9 - 3 + 1 = 7$$

مثال ۶. خطای طبقه بندی هنگامی بیشتر می شود که فاصله طبقات و تعداد طبقات شود؟

(د) کم- زیاد

(ج) زیاد- زیاد

(ب) کم- کم

(الف) زیاد- کم

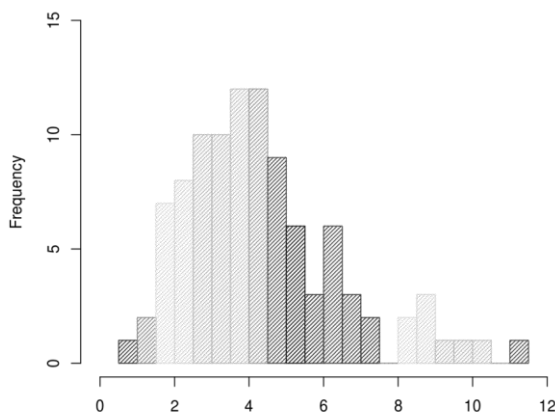
پاسخ: گزینه الف. خطای طبقه بندی با فاصله طبقات رابطه مستقیم و با تعداد طبقات رابطه عکس دارد.

نمودارهای فراوانی:

نمودار ابزاری تصویری است که برای توصیف و نمایش داده‌های جمع آوری شده به کار برده می شود.

انواع نمودارها به صورت زیر است:

نمودار هیستوگرام.....



این نمودار برای نمایش داده‌های ناشی از اجرای متغیرهای

پیوسته (فاصله‌ای و نسبی) بکار می‌رود و کاربرد اصلی آن

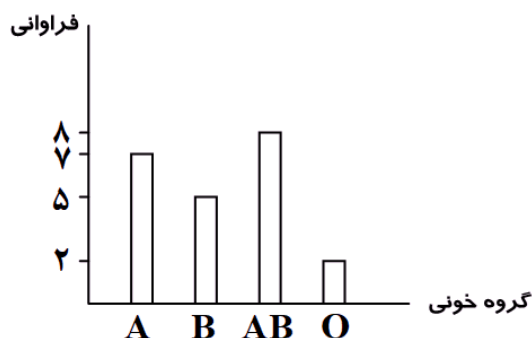
بررسی تقارن است. در هیستوگرام هر ستون، نشان دهنده یک

طبقه از اعداد است. عرض هر ستون برابر با فاصله طبقات و

ارتفاع آن مساوی با فراوانی همان طبقه است.

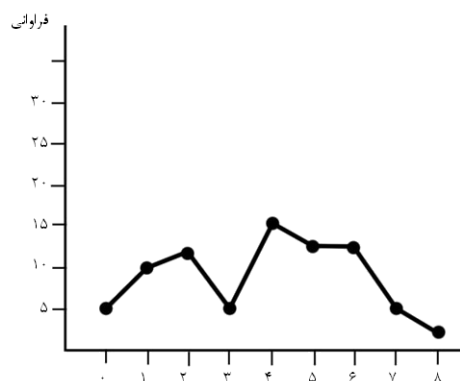
نمودار ستونی (میله‌ای).....

این نمودار برای نمایش داده‌های ناشی از اجرای متغیرهای گسسته (بیشتر متغیر اسمی) بکار می‌رود و کاربرد اصلی آن بررسی تقارن است. عرض هر ستون نشان دهنده طبقه و ارتفاع آن مساوی با فراوانی است.



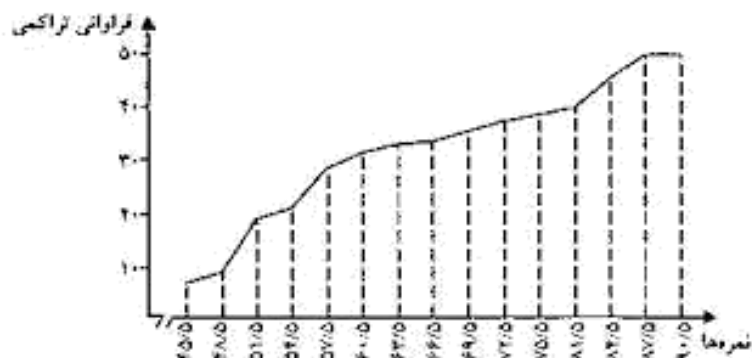
نمودار چند ضلعی.....

این نمودار برای نمایش داده‌های ناشی از اجرای متغیرهای کمی و پیوسته بکار می‌رود و به علت سهولت پرکاربردترین نمودار است. یکی دیگر از کاربردهای اصلی آن مقایسه دو یا چند توزیع است. عرض هر ستون نشان دهنده مرکز دسته و ارتفاع مساوی با فراوانی است.



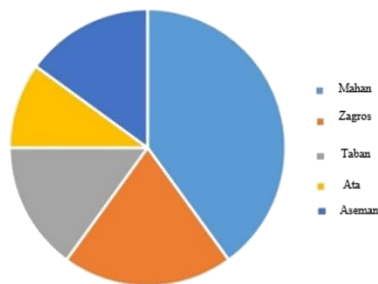
نمودار چند ضلعی تراکمی (اجایو).....

این نمودار برای نمایش داده‌ها، هنگامی که پژوهشگر علاقه‌مند است وضعیت یک نمره یا یک فرد را نسبت به بقیه‌ی نمره‌ها یا افراد مشخص کند. عرض هر ستون نشان دهنده مرکز دسته و ارتفاع مساوی با فراوانی تراکمی است.



نمودار دایره‌ای.....

این نمودار برای نمایش داده‌ها، هنگامی که دارای مقیاس اسمی (گاهی ترتیبی) است، استفاده می‌شود. هدف اصلی آن نمایش نسبت جزء به کل است.



نکته ۱۱. یکی از سوالاتی که ممکن است از نمودار دایره ای شود، اندازه قطاع (برش) است. برای بدست آوردن قطاع، میتوان از ضرب فراوانی نسبی طبقه در ۳۶۰، استفاده کرد. مانند فرمول زیر:

$$\frac{f}{N} \times 360 = \text{اندازه قطاع (برش)}$$

نمودار ساقه و برگ (شاخه و برگ).....

این نمودار برای نمایش داده‌ها، هنگامی که دارای مقیاس فاصله‌ای و نسبی است، استفاده می‌شود. مانند هیستوگرام، اما شاخه و برگ در هنگام اطلاعات کم مفیدتر واقع می‌شود.

ساقه	برگ					
۳	۲	۳	۴	۴	۶	۹
۴	۰	۱	۳	۵	۵	۷
۵	۱	۲	۴	۷	۸	

نکته ۱۲. دقت داشته باشید، می‌توان به وسیله نمودار شاخه و برگ مد (نما) را مشخص کرد. برای مثال در نمودار بالا، دو عدد ۳۴ و ۴۵ نما (مد) نمودار بالا محسوب می‌شود.

مثال ۷: کدام یک از نمودار های زیر برای مقایسه یک نمره با نمرات دیگر مورد استفاده قرار می گیرد؟

الف) هیستوگرام ب) چندضلعی ج) دایره ای د) اجایو

پاسخ: گزینه د. برای مقایسه یک نمره با نمرات دیگر از نمودار اجایو استفاده می شود.

مثال ۸: در جدول زیر اندازه قطاع درس روانشناسی تجربی را مشخص کنید؟

متغیر	فراوانی
روانشناسی عمومی	۲۰
روانشناسی بالینی	۶۰
روانشناسی تجربی	۴۰

الف) ۱۰۰ ب) ۱۲۰

ج) ۱۸۰ د) ۶۰

پاسخ: گزینه ب. با توجه به فرمول قطاع، جواب درست گزینه ب می شود.

$$\frac{f}{N} \times 360 = \frac{40}{120} \times 360 = 120$$

سوالات طبقه‌بندی شده فصل اول آمار

۱- برای ارائه‌ی اطلاعات مربوط به میانگین و انحراف معیار کدام یک از روش‌های زیر مناسب‌تر است؟ (روانشناسی بالینی وزارت بهداشت ۸۵)

(۱) نمودار (۲) جدول (۳) توضیح در متن (۴) نمودار و توضیح در متن

۲- نمودار مناسب برای صفات کمی پیوسته کدام است؟ (روانشناسی بالینی وزارت بهداشت ۸۷)

(۱) دایره‌ای (۲) ستونی (۳) بافت‌نگار (هیستوگرام) (۴) میله‌ای

۳- چنانچه متغیر مورد اندازه‌گیری در یک پژوهش از نوع (ناپیوسته) باشد، از چه نموداری استفاده می‌شود؟

(۱) نمودار هیستوگرام (۲) نمودار ستونی (۳) نمودار چند ضلعی (۴) اوجایو

۴- جایگزین مناسب برای نمودار هیستوگرام کدام نمودار است؟ (روانشناسی بالینی وزارت بهداشت دکتری ۸۸)

(۱) ستونی (۲) ساقه و برگ (۳) دایره‌ای (۴) نمودار پراکنش

۵- گروه خونی ۴۵۰ نفر به صورت نمودار دایره‌ای رسم شده است. زاویه‌ای که برای نمایش نسبت فراوانی گروه خونی AB به دست آمد، برابر ۷۲ درجه بود. تعداد افراد دارای گروه خونی AB چند نفر است؟ (روانشناسی بالینی وزارت بهداشت ۸۹)

(۱) ۹۰ (۲) ۷۲ (۳) ۱۲۰ (۴) ۳۶

۶- تعداد ۱۶ بیمار در یک بخش بستری هستند توزیع تعداد ویزیت بیماران توسط پزشک به صورت جدول زیر است فراوانی نسبی سه دفعه ویزیت کدام است؟ (روانشناسی بالینی وزارت بهداشت ۸۹)

تعداد دفعات ویزیت	۱	۲	۳	۴	۵	(۱) ۳۰
تعداد بیماران	۲	۳	۴	۶	۱	(۲) ۰/۲۵
						(۳) ۰/۳۷
						(۴) ۰/۷۵

۷- مهم‌ترین کاربرد آمار توصیفی کدام است؟ (سراسری ۹۲)

(۱) گردآوری اطلاعات (۲) آزمون فرضیه‌ها

(۳) تلخیص داده‌ها (۴) تعیین شاخص‌های گرایش به مرکز

۸- به داده‌هایی که به‌طور فاحش با داده‌ها توزیع تفاوت دارند چه می‌گویند؟ (دکتری ۹۳)

(۱) پرت (۲) افتاده (۳) انحرافی (۴) گم‌شده

کلید سوالات طبقه بندی شده فصل اول آمار

سوال	پاسخ
۱	۲
۲	۳
۳	۲
۴	۱
۵	۱
۶	۲
۷	۳
۸	۱

پاسخ تشریحی سوالات طبقه بندی شده فصل اول آمار

پاسخ سوال ۱: نکته ۱۳: میانگین و انحراف معیار به ترتیب از شاخص های مرکزی و پراکندگی هستند. و هر دو به آمار توصیفی مربوط می شوند. بهترین ابزار برای ارائه ی اطلاعات در آمار توصیفی جدول است. بنابراین گزینه ی ۲ صحیح است.

پاسخ سوال ۲: نمودار هیستوگرام برای نمایش داده های ناشی از اجرای متغیرهای پیوسته (فاصله ای و نسبی) بکار می رود و کاربرد اصلی آن بررسی تقارن است. در هیستوگرام هر ستون، نشان دهنده یک طبقه از اعداد است. عرض هر ستون برابر با فاصله طبقات و ارتفاع آن مساوی با فراوانی همان طبقه است. بنابراین گزینه ی ۳ صحیح است.

پاسخ سوال ۳: نمودارهای ستونی و دایره ای برای متغیرهای گسسته (ناپیوسته) و کیفی استفاده می شود. بنابراین گزینه ی ۲ صحیح است.

پاسخ سوال ۴: وقتی داده های مورد استفاده کیفی باشند رسم هیستوگرام ممکن نیست. در این مواقع به عنوان جایگزین از نمودار ستونی استفاده می شود. نکته ۱۴: نام دیگر نمودار هیستوگرام، بافت نگار است. بنابراین گزینه ی ۱ صحیح است.

پاسخ سوال ۵: درجه ی هر داده به روی نمودار دایره ایی به شکل زیر محاسبه می شود:

$$\frac{F_i}{n} \times 360 = \frac{F_i}{450} \times 360 = 72 \rightarrow F_i = 90$$

بنابراین گزینه ی ۱ صحیح است.

پاسخ ۶: از تقسیم فراوانی مطلق هر طبقه بر کل داده ها، فراوانی نسبی آن بدست می آید:

$$F_3 = \frac{F_3}{n} \rightarrow \frac{4}{16} = 0.25$$

بنابراین گزینه ی ۲ صحیح است.

پاسخ سوال ۷: گفته شد که مهم ترین هدف آمار توصیفی تلخیص (خلاصه سازی) داده هاست. بنابراین گزینه ی ۳ صحیح است.

پاسخ سوال ۸: نکته ۱۵: داده های پرت یا دور افتاده به داده های گفته می شود که به طور فاحش با سایر داده های توزیع تفاوت دارند. بنابراین گزینه ی ۱ صحیح است.

فصل دوم

شاخص‌های مرکزی

شاخص‌های مرکزی

شاخص‌های مرکزی، شاخص‌هایی هستند که نشان دهنده مرجعیت و مرکزیت داده‌هاست. این شاخص‌ها شامل مد، میانگین و میانگین است.

نما (مد یا M_o)

نما عبارت است از عدد یا نمره‌ای که بیشترین فراوانی را در توزیع فراوانی داراست (بیشترین تکرار). مثلاً در توزیع زیر نما (مد) عدد ۴ است.

۱، ۱، ۲، ۳، ۴، ۴، ۴، ۴، ۴، ۵، ۶

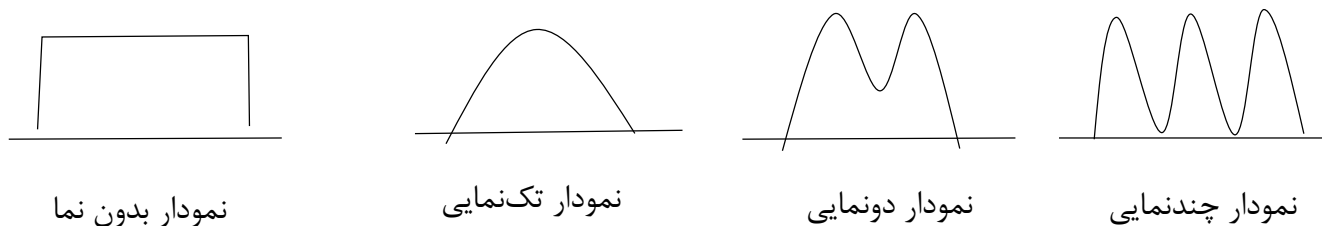
ویژگی‌های نما.....

- نما یک مقیاس اسمی است.
- بی ثبات‌ترین شاخص مرکزی است.
- در نمودار فراوانی، قسمتی از نمودار که بیشترین حجم داده در آنجا قرار دارد، نما نیز در آن منطقه قرار دارد.
- کاربرد نما بیشتر زمانی است که متغیر مورد استفاده اسمی است و یا قصد محاسبه آسان و سریع یک شاخص مرکزی را داشته باشیم.

نکته ۱. در موقعیتی که دو عدد مجاور، بیشترین فراوانی را داشته باشند، نما را می‌توان به طور قراردادی به عنوان میانگین آن دو عدد در نظر گرفت، به عنوان مثال در دنباله زیر نما را می‌توان به طور قراردادی عدد ۴،۵ در نظر گرفت.

۱، ۱، ۲، ۳، ۴، ۴، ۴، ۴، ۵، ۵، ۵، ۵، ۶

نکته ۲. در هنگامی که دو عددی که باهم بیشترین فراوانی را دارند، نزدیک هم قرار نداشته باشند، آنگاه هر دو این اعداد را می‌توان به عنوان نما در نظر گرفت و نمودار مربوطه یک نمودار دونمایی خواهد بود.



مثال ۱. مد یا نما را در توزیع مقابل بدست آورید:

۹، ۵، ۵، ۵، ۳، ۳، ۳، ۳، ۲، ۲، ۲، ۱

الف) ۲ ب) ۳ ج) ۵ د) ۶

پاسخ: گزینه ب. از عدد ۳ بیشترین تعداد وجود دارد، که نشان‌دهنده مد توزیع است.

مثال ۲. مد را در جدول ساقه و برگ زیر را بدست بیاورید؟

ساقه	برگ					
۳	۲	۳	۴	۴	۶	۹
۴	۰	۱	۳	۵	۵	۷
۵	۱	۲	۴	۷	۸	

الف) ۴۵ ب) ۳۴ ج) ۳۶ د) الف و ب

پاسخ: گزینه د. طبق جدول از عدد ۴۵، دو عدد و از عدد ۳۴ نیز دو عدد وجود دارد. پس هر دو مد محسوب می‌شوند.

میانۀ (M_d)

میانۀ نقطه‌ای وسط در توزیع نمره‌هاست. به معنای دیگر نیمی از داده‌ها قبل و نیمی از داده‌ها بعد از آن قرار می‌گیرند.

مثلاً در توزیع زیر، میانۀ مساوی با عدد ۵ است.

۱، ۳، ۵، ۷، ۹

ویژگی‌های میانۀ.....

- میانۀ یک مقیاس ترتیبی (رتبه‌ای) است.
- به علت اینکه این شاخص، همواره وسط داده‌ها را نشان می‌دهد، بنابراین بهترین شاخص مرکزی است، هنگامی که توزیع داده‌ها دارای چولگی (داده پرت) باشند.
- به تغییرات بالاترین عدد یا کم‌ترین عدد واکنشی نشان نمی‌دهد.
- میانۀ در مورد داده‌هایی که ترتیب دارند، به کار می‌رود یعنی ترتیبی، فاصله‌ای و نسبی.
- همواره و در هر توزیعی مجموع قدر مطلق انحراف داده‌ها از میانۀ حداقل است.

$$\sum |x - m_d| \leq \sum |x - A|$$

محاسبه میانۀ.....

- هنگامی که داده‌ها به شکل خام وجود دارند و تعداد داده‌ها فرد است:
- ابتدا داده‌ها را مرتب می‌کنیم و سپس عددی وسطی میانۀ است. مثلاً در دنباله زیر عدد ۳ میانۀ است.
- ۱، ۲، ۳، ۴، ۵
- هنگامی که داده‌ها به شکل خام وجود دارند و تعداد داده‌ها زوج است:
- ابتدا داده‌ها را مرتب می‌کنیم و سپس میانگین دو عدد وسطی میانۀ است. مثلاً در دنباله زیر عدد ۳،۵ میانۀ است.
- ۱، ۲، ۳، ۴، ۵، ۶
- هنگامی که تعداد داده‌ها زوج است و داده تکراری داشته باشیم:

ابتدا داده‌ها را مرتب می‌کنیم و سپس مکان میانه از طریق فرمول $\frac{n}{2}$ پیدا می‌کنیم.
در قدم بعد از طریق فرمول درونیابی خطی، میانه را محاسبه می‌کنیم.

$$M_d = \frac{\text{تعداد تکرارهای قبل از میانه}}{\text{تعداد کل تکرار}} + \text{حد پایین عدد تکراری}$$

برای مثال در دنباله زیر میانه جایی در بین دو عدد ۴ قرار می‌گیرد. در این حالت میانه طبق مراحل زیر عدد ۳٫۸ است.

۱، ۲، ۳، ۴، ۴، ۴، ۵، ۶

$$\frac{n}{2} = \frac{8}{2} = 4$$

$$M_d = 3.5 + \frac{1}{3} = 3.8$$

هنگامی که تعداد داده‌ها فرد است و داده تکراری داشته باشیم، مانند حالت اعداد زوج و تکراری، مراحل را طی می‌کنیم، با این تفاوت که چون میانه روی یک عدد می‌افتد، انگار عدد را به دو نیم تقسیم می‌کند. برای مثال در دنباله زیر میانه مساوی با ۴ است، زیرا یک عدد و نیم عدد قبل از عدد میانه وجود دارد.

۱، ۲، ۳، ۴، ۴، ۴، ۵، ۶

$$\frac{n}{2} = \frac{9}{2} = 4.5$$

$$M_d = 3.5 + \frac{1.5}{4} = 3.87$$

- در جدول فراوانی طبقاتی از طریق فرمول زیر می‌توان میانه را محاسبه کرد.

$$M_d = L + \frac{\frac{n}{2} - cf}{f} \times i$$

L = حد پایین طبقه میانه

cf = فراوانی تراکمی طبقه قبل

f = فراوانی مطلق طبقه میانه

i = فاصله طبقات

n = تعداد کل داده‌ها

مثال ۳. میانه دنباله اعداد زیر را بدست بیاورید؟

۹، ۵، ۵، ۵، ۳، ۳، ۳، ۳، ۲، ۲، ۲، ۱



(د) ۶

(ج) ۵

(ب) ۳

(الف) ۲

پاسخ گزینه ب. طبق فرمول زیر میانه ۳ بدست می‌آید.

$$\frac{n}{2} = \frac{12}{2} = 6$$

$$M_d = \frac{\text{تعداد تکرارهای قبل از میانه}}{\text{تعداد کل تکرار}} + \text{حد پایین عدد تکراری} = 2.5 + \frac{2}{4} = 3$$

مثال ۴. چنانچه قصد داشته باشیم برای توزیع نمرات زیر، مناسب ترین شاخص مرکزی را انتخاب کنیم، کدام یک از شاخص های زیر را انتخاب خواهیم کرد؟
۸-۱۲۹-۹-۱۲۳-۱۳۰-۱۲۵

الف) میانگین (ب) نما (ج) میانه (د) میانگین هندسی
پاسخ: گزینه ج. مناسب ترین شاخص مرکزی هنگامی که داده ها، دارای داده پرت هستند، میانه است. (اعداد ۸ و ۹ داده پرت محسوب می شوند).

مثال ۵. کدام عبارت در مورد میانه صادق است؟

الف) $\sum (x - M_d)^2 = \sum (x - A)^2$ (ب) $\sum |x - M_d| > \sum |x - A|$
ج) $\sum (x - M_d)^2 < \sum (x - M_d)^2$ (د) $\sum |x - M_d| \leq \sum |x - A|$
پاسخ: گزینه د. همانطور که قبل تر داشتیم، مجموع قدرمطلق انحراف داده‌ها از میانه، همواره حداقل است.

میانگین

میانگین یک شاخص مرکزی است. سه نوع میانگین وجود دارد، که مهم‌ترین آن میانگین حسابی است. به ترتیب در ادامه میانگین حسابی، وزنی و هندسی آموزش داده می‌شود.

میانگین حسابی M یا $\bar{x}$

میانگین، معدل حسابی گروهی از نمره‌هاست. میانگین متداول ترین شاخص مرکزی است.

ویژگی‌های میانگین.....

- میانگین یک مقیاس فاصله‌ای است.
- بهترین شاخص مرکزی است، هنگامی که توزیع نرمال باشد و داده پرت یا چولگی وجود نداشته باشند.
- میانگین در مقیاس‌های فاصله‌ای و نسبی مورد استفاده قرار می‌گیرد.
- تحت تاثیر داده‌های پرت یا ارزش عددی یک متغیر قرار می‌گیرد (مرکز ثقل داده‌هاست).
- همواره و در هر توزیعی مجموع انحراف داده‌ها از میانگین برابر صفر است.
 $\sum (x - \bar{x}) = 0$
- همواره و در هر توزیعی مجموع مجذور انحراف داده‌ها از میانگین حداقل (کمینه) است.

$$\sum (x - \bar{x})^2 \leq \sum (x - A)^2$$

نکته ۳. در فصل بعد (شاخص‌های پراکندگی)، با دو شاخص آشنا خواهیم شد، تحت عنوان انحراف چارکی و انحراف معیار. دقت داشته باشید هنگامی که کجی وجود داشته باشد (وجود داده پرت)، بهترین شاخص مرکزی میانه و بهترین شاخص پراکندگی انحراف چارکیست و هنگامی که کجی وجود نداشته باشد (عدم وجود داده پرت)، بهترین شاخص مرکزی میانگین و بهترین شاخص پراکندگی انحراف معیار است.

محاسبه میانگین.....

- هنگامی که داده‌ها به شکل خام وجود دارند.
- میانگین در این حالت از طریق جمع کردن نمره‌ها و تقسیم حاصل جمع بر تعداد کل نمره‌ها به دست می‌آید.

$$\bar{x} = \frac{\sum x}{n}$$

برای مثال در دنباله زیر میانگین مساوی با ۳,۵ است.

۱,۲,۳,۴,۵,۶

$$\bar{x} = \frac{1 + 2 + 3 + 4 + 5 + 6}{6} = \frac{21}{6} = 3.5$$

- در جدول فراوانی طبقاتی از فرمول زیر استفاده می‌شود:

$$\bar{x} = \frac{\sum f x_c}{n}$$

f = فراوانی هر طبقه

x_c = مرکز دسته هر طبقه

n = تعداد کل داده‌ها

میانگین وزنی (میانگین کل، میانگین میانگین‌ها، میانگین مرکب) $\bar{x}_T$

هنگامی که علاقه‌مند به محاسبه میانگین میانگین‌های گروه‌های مختلف هستیم، از میانگین مرکب استفاده می‌کنیم.

محاسبه میانگین وزنی.....

$$\bar{x}_T = \frac{\sum \bar{x}_i n_i}{N_T}$$

مجموع حاصل ضرب میانگین هر گروه در تعداد آن‌ها $\sum \bar{x}_i \cdot n_i$

N_T = مجموع کل نمونه‌ها یا گروه‌ها

میانگین هندسی (جئومتريک) G

میانگین هندسی زمانی یک شاخص مرکزی بحساب می‌آید، که اندازه‌گیری‌ها از نوع نسبت باشند و متغیر از لحاظ سطح اندازه‌گیری نسبی باشد.

محاسبه میانگین هندسی.....

$$G = \sqrt[n]{(x_1)(x_2)(x_3) \cdots (x_n)}$$

به عنوان مثال، میانگین هندسی دو عدد ۲ و ۸ برابر ۴ است.

$$G = \sqrt[2]{(2)(8)} = \sqrt[2]{16} = 4$$

میانگین همساز (هارمونیک) HM

محاسبه میانگین همساز.....

$$HM = \frac{1}{\frac{1}{n} \left(\frac{1}{x_1} + \frac{1}{x_2} + \frac{1}{x_3} + \cdots + \frac{1}{x_n} \right)}$$

نکته ۴. در صورتی که $\bar{x}$ و HM و G به ترتیب برای نمایش میانگین حسابی، هندسی و همساز به کار برده شوند، همیشه رابطه زیر برقرار است:

$$\bar{x} > G > HM$$

نکته ۵. فرمول محاسبه میانگین همساز از طریق میانگین حسابی و میانگین هندسی:

$$HM = \frac{G^2}{\bar{x}}$$

نکته ۶. میانگین دیگری نیز تحت عنوان میانگین پیراسته وجود دارد، که در ابتدا داده‌های پرت کنار گذاشته می‌شوند و سپس میانگین باقی داده‌ها محاسبه می‌شود.

نکته ۷. فرمولی تجربی وجود دارد، تحت عنوان فرمول تجربی پیروسون که رابطه بین m_d و میانگین را نشان می‌دهد:

$$m_o = 3m_d - 2\bar{x}$$

مثال ۶. کلاس ۴۰ نفره ای با میانگین ۱۰ و کلاس ۱۰ نفره ای با میانگین ۱۵ وجود دارد، میانگین کل یا وزنی چقدر است؟

الف) ۱۱ ب) ۱۲ ج) ۱۳ د) ۱۴

پاسخ: گزینه الف. از طریق فرمول زیر می‌توان میانگین وزنی (کل) را محاسبه کرد:

$$\bar{x}_T = \frac{\sum \bar{x}_i n_i}{N_T} = \frac{(40 \times 10) + (15 \times 10)}{40 + 10} = \frac{400 + 150}{50} = \frac{550}{50} = 11$$

مثال ۷. در جدول توزیع فراوانی زیر میانگین کدام است؟

x	1-3	4-6	7-9
f	1	3	1

الف) ۸,۲۵ ب) ۸

ج) ۶,۵ د) ۵

پاسخ: گزینه د. در ابتدا باید مرکز دسته هر خانه را محاسبه کرد و سپس از طریق فرمول، میانگین را محاسبه کرد. پس داریم:

$$\text{مرکز دسته} = \frac{x_L + x_H}{2}$$

مرکز دسته اول = ۲. مرکز دسته دوم = ۵ و مرکز دسته سوم = ۸

$$\bar{x} = \frac{\sum f x_c}{n} = \frac{(2 \times 1) + (5 \times 3) + (8 \times 1)}{1 + 3 + 1} = \frac{2 + 15 + 8}{5} = \frac{25}{5} = 5$$

مثال ۸. اگر میانگین توزیعی ۱۰ باشد و مد آن ۴ آنگاه، میانه آن چقدر است؟

الف) ۴ ب) ۵ ج) ۷ د) ۸

پاسخ: گزینه د. برای محاسبه میانه، می‌توان از فرمول تجربی پیرسون استفاده کرد:

$$m_o = 3m_d - 2\bar{x} = 4 = 3m_d - (2 \times 10) \rightarrow 3m_d = 24 \rightarrow m_d = 8$$

مثال ۹. کدام عبارت زیر ویژگی میانگین است؟

الف) $\sum (x - \bar{x})^2 = \sum (x - A)^2$ ب) $\sum (x - \bar{x})^2 > \sum (x - A)^2$

ج) $\sum (x - \bar{x})^2 < \sum (x - Md)^2$ د) $\sum (x - \bar{x})^2 \leq \sum (x - A)^2$

پاسخ: گزینه د. همانطور که پیش از این گفته شد، همواره و در هر توزیعی مجموع مجذور انحراف داده‌ها از میانگین حداقل است.

مثال ۱۰. با توجه به اطلاعات زیر کدام یک از گزینه‌های زیر میانگین محسوب می‌شود؟

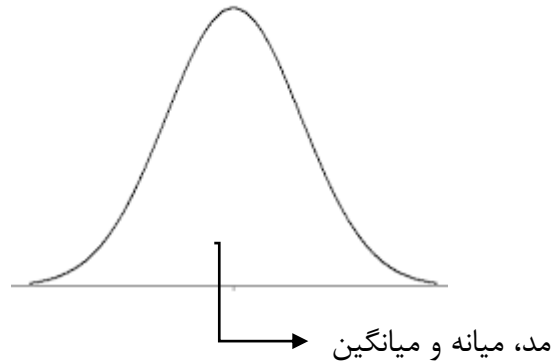
$$\sum (x - 15) = 20 \quad \sum (x - 13) = 2 \quad \sum (x - 14) = 0 \quad \sum (x - 18) = 13$$

الف) ۱۳ ب) ۱۸ ج) ۱۵ د) ۱۴

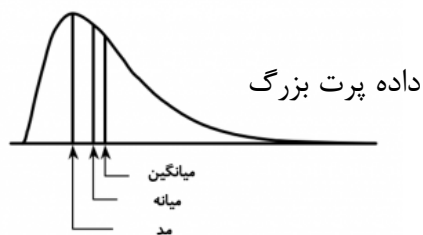
پاسخ: گزینه د. همانطور که قبل تر گفته شد، مجموع انحراف داده‌ها از میانگین همواره صفر است، بنابراین صرفاً مجموع انحراف داده‌ها از عدد ۱۴ صفر است، پس عدد ۱۴ میانگین است.

چولگی (کجی)

شاخص چولگی تحت تاثیر داده‌های پرت قرار می‌گیرد و برای سنجش تقارن توزیع استفاده می‌شوند.
- میانگین، میانه و نما در منحنی متقارن و نرمال، با هم مساوی هستند و هر سه در وسط منحنی قرار دارند. مانند شکل زیر است:



- میانگین و مد و میانه در منحنی با چولگی (کجی) مثبت یا راست، میانگین بزرگتر از میانه و میانه بزرگتر از مد است. مانند شکل زیر:



- میانگین و مد و میانه در منحنی با چولگی (کجی) منفی یا چپ، مد بزرگتر از میانه و میانه بزرگتر از میانگین است. مانند شکل زیر:



محاسبه چولگی (کجی).....

$$sk = \frac{\bar{x} - m_o}{s}$$

$\bar{x}$ = میانگین

m_o = مد

s = انحراف معیار

نکته ۷. اگر کجی بین -۱ و ۱ باشد، به این معنی است که منحنی تقریباً نرمال است

نکته ۸. در فصل بعد در مورد گشتاورها به شکل کامل صحبت خواهد شد، اما در اینجا نکته مهم این است که کجی (چولگی)، گشتاور سوم محسوب می‌شود.

چندک‌ها

چندک‌ها از خانواده میانه هستند و مانند آن شاخص ترتیبی دارند. معروف‌ترین چندک‌ها، چهارک، دهک، صدک است.

Md		
Q ₁	Q ₂	Q ₃
P ₂₅	D ₅	P ₇₅
	P ₅₀	

محاسبه چندک‌ها.....

محاسبه چهارک‌ها و چندک‌ها مانند محاسبه میانه است. برای مثال محاسبه چهارک سوم:

$$Q_3 = L + \left[\frac{\frac{3n}{4} - cf}{F} \right] i$$

برای محاسبه انواع صدک و دهک نیز به همین شکل برای مثال محاسبه صدک ۲۶ ام مساوی است:

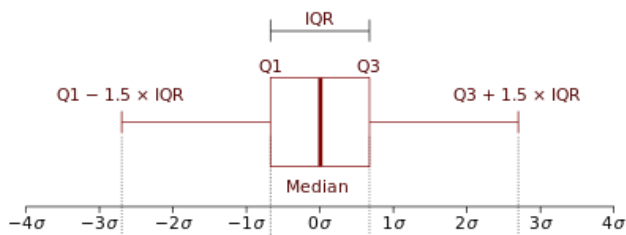
$$P_{26} = L + \left[\frac{\frac{26n}{100} - cf}{F} \right] i$$

نمودار جعبه ای (Box plot)

این نمودار با نمرات کمی و پیوسته کار می‌کند و کاربرد آن یافتن داده‌های پرت است.

برای یافتن داده پرت، می‌توانیم از مسیر زیر استفاده کنیم.

۱,۵ برابر دامنه چارکی را از چارک اول کم و ۱,۵ برابر دامنه چارکی به چارک سوم اضافه کنیم. دامنه نرمال بودن بدست می‌آید، داده‌هایی که بعد از این دامنه هستند، پرت محسوب می‌شوند.

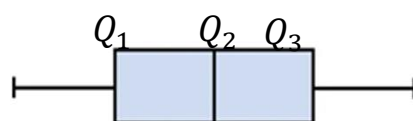


نکته ۹. نحوه محاسبه دامنه چارکی:

$$IQR = Q_3 - Q_1$$

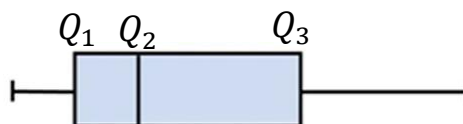
نکته ۱۰. کاربرد دیگر نمودار جعبه‌ای، در تشخیص چولگی راست و چپ است

در منحنی نرمال فاصله بین چارک اول تا دوم برابر با فاصله چارک دوم تا سوم است مانند شکل زیر:



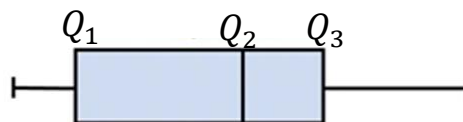
$$Q_2 - Q_1 = Q_3 - Q_2$$

در کجی مثبت (چولگی به راست)، فاصله بین چارک اول تا دوم کمتر از فاصله چارک دوم تا سوم است مانند شکل زیر:



$$Q_2 - Q_1 < Q_3 - Q_2$$

در کجی منفی (چولگی به چپ)، فاصله بین چارک اول تا دوم بیشتر از فاصله چارک دوم تا سوم است مانند شکل زیر:



$$Q_2 - Q_1 > Q_3 - Q_2$$

نکته ۱۱. شاخص‌های مرکزی به تمام عملیات ریاضی (جمع و تفریق، ضرب و تقسیم) حساس هستند. یعنی اگر همه داده‌ها با عددی جمع، تفریق و یا ضرب و تقسیم بشوند، میانگین، میانه و مد نیز در همان اعداد جمع، تفریق و یا ضرب و تقسیم می‌شوند.

مثال ۱۱. اگر در کنکور بیشتر سوالات سخت باشند، آنگاه:

الف) میانگین بزرگتر میانه (ب) میانه مساوی با مد (ج) مد بزرگتر از میانه (د) میانه مساوی با میانگین
پاسخ گزینه الف. اگر امتحانی بسیار مشکل باشد، یعنی بیشتر افراد نتوانسته‌اند به سوال پاسخ دهند، پس منحنی به سمت راست کج می‌شود و بیشتر تمرکز (مد) در قسمت منفی است. پس در این حالت میانگین بزرگتر از میانه و میانه بزرگتر از مد است.

مثال ۱۲. اگر مد و میانگین به ترتیب ۴۳ و ۴۰ باشند و انحراف استاندارد برابر با ۱ باشد، آنگاه میزان کجی:

الف) +۱ (ب) -۱ (ج) ۰ (د) -۳

$$sk = \frac{\bar{x} - m_o}{s} = \frac{40 - 43}{1} = -3$$

پاسخ گزینه الف. طبق فرمول کجی داریم:

مثال ۱۳. اگر در یک دنباله داده‌ها، مقدار چارک اول برابر ۲۰۰ و مقدار چارک سوم ۳۰۰ باشد، آنگاه کدام یک از داده‌های زیر پرت محسوب می‌شود:

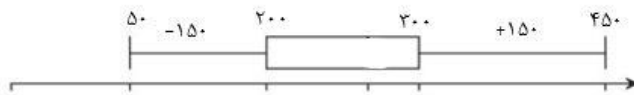
الف) ۱۵۰ (ب) ۷۵ (ج) ۵۰۰ (د) ۳۵۰

پاسخ: گزینه ج. طبق فرمول در ادامه، اعداد بین ۵۰ تا ۴۵۰، عدد پرت محسوب نمی‌شوند.

$$IQR = 300 - 200 = 100$$

$$Q_1 - (1.5 \times IQR) = 200 - (1.5 \times 100) = 50$$

$$Q_3 + (1.5 \times IQR) = 300 + (1.5 \times 100) = 450$$



مثال ۱۴. اگر چارک اول ۱۰۰، چارک دوم ۱۲۰ و چارک سوم ۱۵۰ باشد، آنگاه:

الف) میانگین بزرگتر میانه (ب) میانه مساوی با مد (ج) مد بزرگتر از میانه (د) میانه مساوی با میانگین

پاسخ گزینه الف. طبق فرمول زیر فاصله بین چارک اول تا دوم کمتر از فاصله چارک دوم تا سوم است، پس کجی مثبت (چولگی به راست) است و در کجی مثبت، میانگین بزرگتر از میانه و میانه بزرگتر از مد است.

$$Q_3 - Q_2 = 150 - 120 = 30 \quad Q_2 - Q_1 = 120 - 100 = 20$$

$$Q_2 - Q_1 < Q_3 - Q_2$$

مثال ۱۵. اگر میانگین نمره درسی ریاضی ۶۴ باشد و از هر نفر ۲ نمره کم شود، در این صورت میانگین جدید کدام است؟

الف) ۶۲ (ب) ۸ (ج) ۶۴ (د) ۱۲۸

پاسخ گزینه د. شاخص‌های مرکزی به تمامی اعمال ریاضی حساس هستند، پس اگر از نمرات ۲ نمره کم بشود، از میانگین نیز ۲ نمره کم خواهد شد.

سوالات طبقه بندی شده فصل دوم آمار

۱- کدام عبارت نمایانگر چولگی مثبت است؟ (مشاوره-آزاد ۸۵)

- | | | | |
|----------------------|--------------------|--------------------|--------------------|
| (۱) میانگین = $46/2$ | (۲) میانگین = 58 | (۳) میانگین = 28 | (۴) میانگین = 28 |
| میان = $42/1$ | میان = $52/1$ | میان = 30 | میان = 28 |
| نما = $47/3$ | نما = 51 | نما = 29 | نما = $29/1$ |

۲- چنانچه بخواهیم توزیع درآمد خانواده را در یک مجموعه بررسی کنیم؛ استفاده از چه روش آماری را توصیه می کنید؟ (روانشناسی بالینی وزارت بهداشت - ۸۵)

- (۱) میانگین (۲) نما (۳) میان (۴) دامنه
- ۳- برای نشان دادن شاخص مرکزی یک سری از نمرات در چه صورتی میان به میانگین ترجیح داده می شود؟ (روانشناسی بالینی وزارت بهداشت - دکتری ۸۵)

- (۱) اگر پراکندگی نرمال یا تقریباً نرمال باشد.
- (۲) پراکندگی دارای کجی زیاد باشد.
- (۳) پراکندگی براساس مقیاس نسبی باشد.
- (۴) نمرات از یک نمونه بزرگ و معرف جامعه به دست آمده باشد.

۴- نمره درس ریاضی ۶ دانش آموز که به صورت تصادفی از یک کلاس انتخاب شده اند به این قرار است:
۲۰ - ۱۸ - ۱۷ - ۱۶ - ۱۵ و صفر برای قضاوت کلی در مورد این کلاس کدام اندازه گرایش به مرکز مناسبتر است؟ (سراسری ۸۶)

- (۱) کجی (۲) میان (۳) میانگین (۴) نما
- ۵- اگر توزیع نمونه های یک آزمون منفی باشد، اندازه میانگین توزیع: (مشاوره - آزاد ۸۶)
- (۱) کوچکتر از میان است (۲) بزرگتر از میان است (۳) کوچکتر از کجی است (۴) بزرگتر از نما است
- ۶- در کدام رابطه منحنی دارای کجی مثبت خواهد بود؟

(روانشناسی، عمومی، تربیتی صنعتی و سازمانی - آزاد ۸۶)

- (۱) $M=Mo=Mn$ (۲) $Mo>Mn>M$ (۳) $M<Mo<Mn$ (۴) $M>Mn>Mo$

۷- برای نشان دادن تمایل مرکزی توزیع، نمرات معمولاً زمانی میان به میانگین ترجیح داده میشود که: (روانشناسی بالینی وزارت بهداشت-۸۶)

- (۱) توزیع نرمال و یا تقریباً نرمال باشد
- (۲) توزیع انحراف قابل ملاحظه ای داشته باشد.
- (۳) نمرات براساس مقیاس نسبیتی اندازه گیری شده باشد.
- (۴) توزیع تنها یک نما داشته باشد.

۸- نما و میانه ی نمرات یک آزمون، کدام شاخص از گزینه های زیر هستند:

- (۱) واریانس (۲) تمایل مرکزی (۳) تفاوت آماری (۴) پراکندگی

۹- در یک آزمون روان شناختی این هفت عدد (۶-۷-۸-۹-۱۰-۱۵-۱۵) به دست آمده است میانگین میانه و نما را به ترتیب مشخص کنید: (روانشناسی بالینی وزارت بهداشت-۸۶)

میانگین	میانه	نما
(۱) ۹	۱۵	۹
(۲) ۹	۹	۱۵
(۳) ۱۰	۹	۱۵
(۴) ۱۰	۱۵	۱۵

۱۰- میانگین و میانه یک جامعه به ترتیب ۳۰ و ۵۰ است. فرض کنید توزیع جامعه تقریباً نرمال باشد، نما را محاسبه کنید. (روانشناسی عمومی - آزاد ۸۷)

۴۰ (۱)	۹۰ (۲)	۶۵ (۳)	۲۵ (۴)
--------	--------	--------	--------

۱۱- کدام یک از عبارتهای زیر از ویژگی‌های میانگین است؟ (مشاوره - آزاد ۸۷)

$$\sum (x - \bar{x})^2 = \sum (x - A)^2 \quad (۱)$$

$$\sum (x - \bar{x})^2 > \sum (x - A)^2 \quad (۲)$$

$$\sum (x - \bar{x})^2 < \sum (x - Md)^2 \quad (۳)$$

$$\sum (x - \bar{x})^2 \leq \sum (x - A)^2 \quad (۴)$$

۱۲- میانگین نمرات آمار سه کلاس به شرح زیر است: کلاس اول با تعداد ۴۰ نفر ۱۰ کلاس دوم با تعداد ۵۰ نفر، ۱۲ و کلاس سوم با تعداد ۳۰ نفر، ۱۵ است. میانگین کل نمرات آمار سه کلاس کدام یک از موارد زیر است؟ (مشاوره - آزاد ۸۷)

۱۲/۳۳ (۱)	۱۳ (۲)	۱۲/۰۸ (۳)	۱۳/۲۵ (۴)
-----------	--------	-----------	-----------

۱۳- مجموع مجذور انحرافات از میانگین کدام است؟ (روانشناسی عمومی - آزاد ۸۷)

(۱) حداقل است	(۲) یک است.	(۳) حداکثر است	(۴) صفر است.
---------------	-------------	----------------	--------------

۱۴- میانگین صفت متغیر x در جامعه‌ای $\bar{x} = 125$ است. اگر فراوانی‌ها را بر عدد ۱۰ تقسیم کنیم میانگین چقدر خواهد شد؟ (روانشناسی، عمومی، صنعتی و سازمانی - آزاد ۸۸)

۱۲/۵۱	۱۲۵ (۲)	۱۳۵ (۳)	۱۱۵ (۴)
-------	---------	---------	---------

۱۵- جامعه‌ای با توزیع نرمال دارای میانگین ۱۰۰ و میانه ۱۲۰ است، نما چقدر است؟

(روانشناسی، عمومی، صنعتی و سازمانی - آزاد ۸۸)

۱۱۰ (۱)	۱۲۰ (۲)	۱۴۰ (۳)	۱۶۰ (۴)
---------	---------	---------	---------

۱۶- چنانچه قصد داشته باشیم برای توزیع نمرات، زیر مناسب‌ترین شاخص مرکزی را انتخاب کنیم کدام یک از شاخص‌های زیر را انتخاب خواهید کرد؟ (۸-۱۲۴-۹-۱۲۳-۱۳۰-۱۲۵) (روانشناسی، عمومی، صنعتی و سازمانی - آزاد ۸۸)

(۱) میانگین	(۲) نما	(۳) میانگین هندسی	(۴) میانه
-------------	---------	-------------------	-----------

۱۷- کدام یک از گزینه‌های زیر شاخص‌های گرایش «مرکزی را نشان می‌دهد؟ (تربیتی - آزاد ۸۸)

(۱) نما، میانگین، واریانس	(۲) واریانس، میانگین، انحراف استاندارد
(۳) انحراف استاندارد، میانه، میانگین	(۴) نما، میانه، میانگین

۱۸- در یک منحنی بهنجار: (تربیتی - آزاد ۸۸)

- (۱) نما، بزرگ ترین عدد و سپس به ترتیب میانه و میانگین قرار می گیرند.
- (۲) میانه، میانگین و نما بر هم منطبق است.
- (۳) میانه، بزرگ ترین عدد و سپس نما و میانگین قرار می گیرند.
- (۴) میانگین بزرگترین عدد و سپس میانه و نما قرار می گیرند.

۱۹- میانگین نمره ۱۵ با فراوانی ۸ و نمره ۲۵ با فراوانی ۱۲ برابر است با: (مشاوره - آزاد ۸۸)

- (۱) ۲۰ (۲) ۱۸/۵ (۳) ۲۲/۵ (۴) ۲۱

۲۰- در توزیع های متقارن کدام یک از حالات زیر درست است؟ (مشاوره - آزاد ۸۸)

(۱) $\bar{x} > Mo > Md$

(۲) $\bar{x} < Mo < Md$

(۳) $\bar{x} = Mo = Md$

(۴) $Md > \bar{x} > Mo$

۲۱- باثبات ترین شاخص گرایش مرکزی کدام است؟ (سراسری ۸۹)

- (۱) نما (۲) میانه (۳) میانگین (۴) انحراف معیار

۲۲- کدام مورد بیانگر کجی مثبت می باشد؟ (سراسری ۸۹)

- (۱) میانگین = میانه = نما (۲) نما < میانه < میانگین (۳) میانگین < نما < میانه (۴) میانگین < میانه < نما

۲۳- اگر مجموع انحراف داده ها از عدد ۷ برابر صفر باشد، میانگین داده ها:

(وزارت بهداشت - ۸۹)

- (۱) ۴۹ (۲) $\sqrt{7}$ (۳) ۷ (۴) ۴/۵

۲۴- کدام شاخص گرایش مرکزی را برای تفسیر اندازه های روبرو انتخاب می کنید؟ $x_i = 3, 8, 7, 7, 13, 11, 18, 22, 4$

- (۱) میانگین (۲) نما (۳) میانگین وزنی (۴) میانه

۲۵- اگر در یک توزیع، فراوانی ها را در عدد ثابتی ضرب کنیم؛ چه تغییر در میانگین اتفاق می افتد؟

- (۱) تغییر نمی کند (۲) تغییر می کند (۳) افزایش می یابد (۴) کاهش می یابد.

۲۶- کدام یک از توزیع ها احتمالاً کجی دارد؟

- (۱) نمرات یک آزمون آسان (۲) تعداد کتاب های کتابخانه های شهر (۳) میزان علاقه به کتاب های کتابخانه (۴) میزان مهارت مراجعین در استفاده از نرم افزارها

۲۷- میانگین اعداد ۱ تا ۱۰۰ کدام است؟

- (۱) $\frac{101}{100}$ (۲) $\frac{500}{100}$ (۳) $\frac{5500}{100}$ (۴) $\frac{5050}{100}$

۲۸- در چه شرایطی استفاده از میانگین توزیع نامناسب است؟

- (۱) توزیع متقارن ولی از هر دو طرف کشیده باشد.
- (۲) داده ها فاصله ای باشند.
- (۳) توزیع دارای کجی زیادی باشد.
- (۴) داده ها نسبی باشند.

- ۲۹- درجه پستی و بلندی نمودار توزیع فراوانی را مقادیر مختلف کدام مشخصه آماری تعیین می‌کند؟
(۱) میانگین (۲) کجی (۳) انحراف استاندارد (۴) مد یا نما
- ۳۰- اگر توزیع داده‌ها با میانگین ۱۵ دارای کجی مثبت باشد. کدام یک از حالت‌ها برای دیگر شاخص‌های مرکزی درست است؟
(۱) نما کوچک‌تر از ۱۵ (۲) نما بزرگ‌تر از ۱۵ (۳) میانه بزرگ‌تر از ۱۵ (۴) میانه و نما هردو برابر ۱۵

کلید سوالات طبقه بندی شده فصل دوم آمار

سوال	پاسخ	سوال	پاسخ
۱۶	۴	۱	۲
۱۷	۴	۲	۲
۱۸	۲	۳	۲
۱۹	۴	۴	۲
۲۰	۳	۵	۱
۲۱	۳	۶	۴
۲۲	۴	۷	۲
۲۳	۳	۸	۲
۲۴	۴	۹	۳
۲۵	۱	۱۰	۲
۲۶	۱	۱۱	۴
۲۷	۴	۱۲	۳
۲۸	۳	۱۳	۱
۲۹	۴	۱۴	۲
۳۰	۱	۱۵	۴