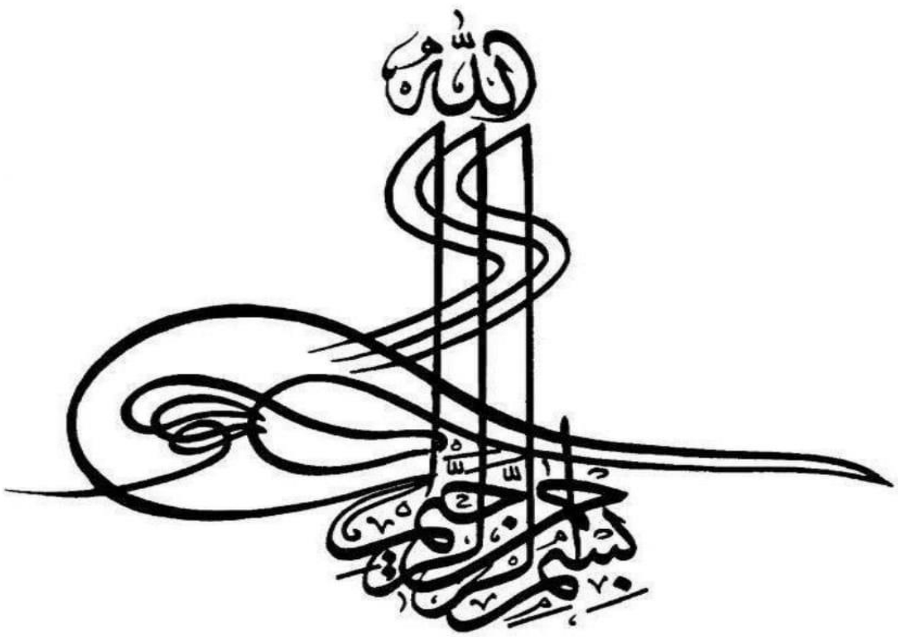
استاد : حبیب هاشمی

مبحث : جزوه آمار کنکور رشته ریاضی

استاد : حبیب هاشمی

مبحث : جزوه آمار کنکور رشته ریاضی

حبیب هاشمی ۰۹۱۲۰۹۱۸۷۰۱



حبیب هاشمی ۰۹۱۲۰۹۱۸۷۰۱

درسنامه کنکوری آمار و مدل سازی

جهت استفاده دانش آموزان
"ریاضی، انسانی و تجربی و به ویژه کنکوری‌های هر سه رشته"

- **طبقه بندی سوالات به صورت موضوعی**
- **پاسخ کاملا تشریحی**
- **تمرین‌های برای آمادگی**

۱۳۹۵

استاد : حبیب هاشمی

مبحث : جزوه آمار کنکور رشته ریاضی

فهرست مطالب

عنوان **صفحه**

حبیب هاشمی ۰۹۱۲۰۹۱۸۷۰۱

حبیب هاشمی ۰۹۱۲۰۹۱۸۷۰۱

مقدمه

اثری که در اختیار شماست مجموعه‌ای از مطالب آمار و مدل سازی، مثال‌های متنوع و تست‌های سراسری می‌باشد؛ از این رو شما را در جمع‌بندی مطالبی که تاکنون مطالعه کرده اید یاری می‌رساند. با توجه به لزوم تشریح و توضیح بیشتر این درس برای رشته‌های ریاضی، تجربی و انسانی سعی شده است که پاسخ مثال ها به اندازه کافی تشریحی باشد و نکات لازم هر تست گفته شود. در این درسنامه تست‌های سراسری نیز گنجانده شده است تا سلیقه طراحان کنکور برجسته و مشخص شود. امیدواریم این درسنامه برایتان مفید باشد و بتوانید از آن بهره مندی لازم را ببرید.

موفق باشید.

۱ اندازه‌گیری و مدل سازی

۱-۱ اندازه‌گیری

در رسیدن به اطلاعات عددی با معیار مناسب برای انجام بررسی آماری، نسبت دادن عدد به موضوع مورد مطالعه را اندازه گیری گوییم.

اولین قدم در رسیدن به اطلاعات عددی برای انجام بررسی آماری یک یا چند موضوع اندازه‌گیری است.

به عنوان نمونه برای رسیدن به اطلاعات عددی در بررسی پیشرفت سالانه تولید گندم در کشور، مقدار گندم تولید شده در سال اندازه گیری شده و به صورت عدد مشخص می‌شود.

۱-۲ مدل سازی ریاضی

بیان مسئله به زبان ریاضی را مدل سازی ریاضی گوییم.

ویژگی های یک مدل سازی با ارزش:

۱-ساده و ابتدایی　　　۲- به پدیده مورد نظر نزدیک تر بودن

مثال۱: نسبت دادن عدد به موضوع موردمطالعه برای رسیدن به اطلاعات عددی با معیاری مناسب کدام است؟

۱) نمونه‌گیری　　√ ۲) اندازه‌گیری　　۳) مدل سازی ریاضی　　۴) سرشماری

مثال۲: اولین قدم دررسیدن به اطلاعات عددی ....است.

۱) نمونه‌گیری　　√ ۲) اندازه‌گیری　　۳) مدل سازی ریاضی　　۴) سرشماری

مثال۳: کدام نوع مدل سازی ریاضی باارزش‌تر است؟(سراسر ی انسانی۸۷)

۱) خطای اندازه‌گیری برابر　　۲) نتیجه حاصل همان پدیده‌ی موردنظر

۳) فقط مفاهیم ریاضی ساده‌تر　　√ ۴) مفاهیم ریاضی ساده‌تر – نتیجه به پدیده موردنظر نزدیک‌تر

مثال۴: بیان مسئله به زبان ریاضی را .... گویند.

۱)نمونه‌گیری　　۲) اندازه‌گیری　　√ ۳) مدل سازی ریاضی　　۴) سرشماری

۳-۱ خطای اندازه‌گیری

تفاضل مقدار واقعی و مقدار اندازه‌گیری شده را خطای اندازه‌گیری گوییم و با E نشان می‌دهند.

***مقدار مثبت خطای اندازه‌گیری از واحد اندازه‌گیری کوچک‌تر است.**

***خطای اندازه‌گیری مثبت یا منفی است و هیچ‌گاه صفر نیست.**

***درمحاسبات خطای اندازه گیری، می‌توان از E با توان‌های بزرگ تر از یک چشم پوشی کرد.**

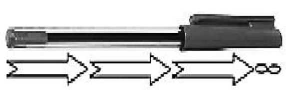
***فرم کلی مدل اندازه‌گیری به‌صورت زیر است:**

(E) خطای اندازه‌گیری + مقدار اندازه‌گیری شده = مقدار واقعی

مثال۵: اگر مقدار اندازه‌گیری کوچک‌تر از مقدار واقعی باشد، خطای اندازه گیری............ است؟

۱) منفی ۲√) مثبت ۳) صفر ۴) بزرگ‌تر از واحد

مثال۶: در اندازه‌گیری زیر واحد بزرگ A و واحد کوچک B است. مدل ریاضی طول خودکار کدام است؟



۱) $l=۳B+۲A$ ۲) $l=۳B+۲A+E$

۳) $l=۳A+۲B$ ۴√) $l=۳A+۲B+E$

مثال۷: در مثال۶ کدام گزینه صحیح است؟

۱) $|E|<A+B$ ۲) $|E|<A$ ۳√) $|E|<B$ ۴) $|E|<AB$

حل: چون واحد B کوچک‌تر از واحد A است گزینه۳صحیح است.

مثال۸: فرض کنید مدل طول مدادی $L=۴+E_۱$ و مدل شعاع قاعده آن $R=۲+E_۲$ است مدلی برای حجم مداد بیابید.

حل:

ارتفاع × مساحت قاعده (دایره) = حجم مداد (استوانه) $\Rightarrow V=\pi R^۲.L$

$$V=\pi(۴+E_۱)^۲(۲+E_۲)=\pi\left(۱۶+۸E_۱+E_۱^۲\right)(۲+E_۲)$$

$$=\pi\left(۳۲+۱۶E_۲+۱۶E_۱+۸\underbrace{E_۱E_۲}_{۰}+۲\underbrace{E_۱^۲}_{۰}+\underbrace{E_۱^۲E_۲}_{۰}\right)$$

$$\Longrightarrow V=۳۲\pi+\underbrace{۱۶\pi E_۱+۱۶\pi E_۲}_{E}\Longrightarrow V=۳۲\pi+E$$

مثال۹: مدل طول و عرض مستطیلی به ترتیب $E_1 + 8$ و $E_2 + 3$ است. مدل مساحت مستطیل کدام است؟

۱) $24 + E_1E_2$ ۲) $24 + E_1 + E_2$ ۳) $24 + 3E_1 + 8E_2$ √ ۴) $24 + 8E_1 + 3E_2$

حل: گزینه ۳

$$(8 + E_1)(3 + E_2) = 24 + 8E_2 + 3E_1 + \underbrace{E_1E_2}_{\cdot} = 24 + 8E_2 + 3E_1$$

۲ جامعه و نمونه

۱-۲ جامعه آماری

مجموعه افراد یا اشیایی که موضوع یا موضوعاتی را روی آن‌ها مطالعه می‌کنیم را جامعه آماری گوییم.

صفت: به کمیت یا کیفیتی که متعلق به عناصر جامعه آماری باشد صفت گوییم و بر دو نوع است:

۱)**صفت ثابت**: همه ی عناصر جامعه آن را دارا باشند مانند کارمندان شرکت نفت

۲)**صفت متغیر**: در یک فرد به فرد دیگر تغییر می‌کند مانند گروه خونی

نکته: اعضای جامعه آماری حداقل یک صفت مشترک دارند.

مشاهده آماری: جمع آوری اطلاعات مربوط به صفات متغیر دریک جامعه آماری را مشاهده آماری می‌گویند.

۲-۲ روش‌های مطالعه آمار

مطالعه آماری موضوع یا موضوعات به دو روش ***سرشماری*** *و* ***نمونه‌گیری*** انجام می‌گیرد.

۱-۲-۲ سرشماری

مطالعه یا بررسی تمام اعضای جامعه آماری را روش سرشماری گویند.

مهمترین مشکلات سرشماری:

** بالا بودن هزینه*

** وقت‌گیر بودن*

** در دسترس نبود ن اعضای جامعه*

** از بین رفتن اعضای جامعه در بعضی مطالعات*

۲-۲-۲ نمونه گیری

مطالعه یا بررسی اعضایی از جامعه را به عنوان نمونه، روش نمونه گیری گویند.

نمونه: زیرمجموعه‌ای از جامعه آماری را نمونه گویند.

تعداد اعضای جامعه و نمونه را به ترتیب اندازه جامعه و نمونه گویند.

نکته: عمل نمونه گیری مهمترین بخش آمار است، نمونه‌ای می‌تواند نمایانگر خصوصیات جامعه آماری باشد که:

۱) به اندازه کافی بزرگ باشد.

۲) اعضایش به‌طور تصادفی انتخاب شوند.

هرچه تنوع موضوع در جامعه بیشتر باشد اندازه نمونه بزرگتر انتخاب می‌شود.

*انتخاب اعضای نمونه نباید از قانونی پیروی کند.

*معمولاً اندازه نمونه حداقل ۱۰ درصد اندازه جامعه است.

مثال۱: افراد یا اشیایی که موضوع یا موضوعاتی روی آن‌ها مطالعه می‌شود، کدام است؟

۱)نمونه ۲)سرشماری ۳)جامعه√ ۴)نمونه‌گیری

مثال۲: مطالعه و بررسی آماری تمام اعضای جامعه کدام است؟

۱)نمونه ۲)سرشماری√ ۳)جامعه ۴)نمونه‌گیری

مثال۳: از بین رفتن اعضای جامعه در بعضی موضوعات مربوط به کدام روش آماری است؟

۱)نمونه گیری√ ۲) سرشماری ۳)اندازه گیری ۴)مدل سازی

مثال۴: برای کدام موضوع در کلاس اندازه نمونه بزرگ‌تر است؟

۱)قد ۲)سن ۳)هوش ۴)معدل√

مثال۵: کدام نمونه به‌خوبی بیانگر خصوصیات جامعه نیست؟

۱)به اندازه کافی بزرگ باشد. ۲) اعضایش قانونی انتخاب شوند.√

۳)اعضایش تصادفی انتخاب شوند. ۴)اندازه‌اش بیشتراز ۱۰درصدجامعه باشد.

نمونه تصادفی: نمونه‌ای است که اعضایش به‌طور تصادفی انتخاب شوند و روش انتخاب اعضایش از دو ویژگی زیر پیروی کند:

۱) امکان انتخاب تمام اعضای جامعه در نمونه وجود داشته باشد.

۲) شانس انتخاب تمام اعضای جامعه برای نمونه یکسان باشد.

۳-۲-۲ روش اعداد تصادفی

یکی از روش‌های ساده و پراستفاده برای انتخاب نمونه تصادفی ساده، روش اعداد تصادفی است.

عدد تصادفی: عددی کوچک‌تر از یک و بزرگ‌تر از صفر با سه رقم اعشار را عدد تصادفی گویند و برای نوشتن آن می‌توان از ماشین‌حساب استفاده کرد. مانند عدد ۰/۵۶۴

مراحل روش اعداد تصادفی برای انتخاب نمونه:

* اعضای جامعه را شماره‌گذاری می‌کنیم.(شماره‌گذاری می‌تواند از هر عدد طبیعی شروع شود.)

*به تعداد اندازه نمونه عدد تصادفی مشخص می‌کنیم.

*شماره‌های اعضای نمونه را به‌صورت زیر به دست می‌آوریم:

*اولین شماره جامعه+ [عدد تصادفی × اندازه جامعه] = شماره نمونه

([ ] نماد جزء صحیح است.)

نکته: درصورتی‌که شماره‌های مساوی برای شماره نمونه به دست بیاید عدد تصادفی را عوض کنید.

مثال۶: در مورد روش انتخاب نمونه تصادفی کدام گزینه صحیح **نیست؟**

۱)امکان انتخاب تمام اعضای جامعه در نمونه وجود داشته باشد.

۳) اندازه نمونه به‌اندازه کافی بزرگ باشد.

۲) شانس انتخاب تمام اعضای جامعه برای نمونه یکسان باشد.

۴√) انتخاب اعضای نمونه از قانونی پیروی کند.

مثال۷: به کمک اعداد تصادفی ۰/۴۲۱ و ۰/۸۴۲ دو شماره از بین اعداد طبیعی ۱ تا ۱۵ بیابید؟

$[0/421 \times 15] + 1 = [6/315] + 1 = 6 + 1 = 7$

$[0/842 \times 15] + 1 = [12/63] + 1 = 12 + 1 = 13$

مثال۸: با استفاده از عدد تصادفی۰/۳۴۲، عدد طبیعی بین ۱۵ تا ۷۸ کدام است؟

۱) ۲۱ ۲) ۲۲ ۳) ۳۶ √ ۴) ۳۷

حل:

$64 = 1 + 15 - 78 =$ اندازه جامعه

$[0/342 \times 64] + 15 = [21/888] + 15 = 21 + 15 = 36$

مثال۹:در یک لیست ۳۰ نفری، می‌خواهیم ازشماره ی۲۰ تا ۳۰ یک نفر را انتخاب کنیم اگر ماشین حساب، عدد تصادفی۰/۲۹۷راتولید کند، کدام شماره انتخاب می‌شود؟

۱) ۲۱ ۲) ۲۲ ۳) ۲۳ √ ۴) ۲۴

حل:

$11 = 1 + 20 - 30 =$ اندازه جامعه

$[0/297 \times 11] + 1 = [3/267] + 1 = 3 + 1 = 4$

بنابراین۴امین نفر با شروع از شماره ی ۲۰می شود شماره ی۲۳.

داده: نتیجه بررسی و مطالعه‌ی نمونه را داده گویند.

هدف اصلی علم آمار تبدیل داده‌ها به اطلاعات است.

۴-۲-۲روش‌های جمع‌آوری داده‌ها

روش جمع‌آوری داده‌ها در آمار با توجه به موضوع موردمطالعه عبارت اند از:

- داده‌های از پیش تهیه‌شده
- پرسش(مصاحبه یا پرسشنامه کتبی)
- مشاهده و ثبت وقایع
- آزمایش

نکته: تفاوت روش پرسش کتبی(پرسشنامه) و پرسش شفاهی(مصاحبه) در ایـن است که در پرسشنامه سؤالات برای همه یکسان است. اما در مصاحبه با توجه به اینکه مصاحبه‌گر ممکن است در مورد سؤالاتی توضیح بیشـتری دهـد، سـؤالات برای همه یکسان نیست.

۵-۲-۲ نکات طراحی پرسشنامه

- سازمان‌دهی محتوای پرسشنامه
- در نظر گرفتن هدف بررسی
- تهیه فهرستی از عناوین مربوط به اطلاعاتی که باید جمع‌آوری شود.
- عدم جمع‌آوری داده‌ها و اطلاعات اضافی
- استفاده از سؤالات ساده و واضح به طوری که برداشت متفاوتی از پرسش ها نشود.
- استفاده از سؤالاتی که دارای جواب یک‌کلمه‌ای یا عددی هستند.
- عدم استفاده از سؤالات هدایت‌کننده
- استفاده از سؤالات چندگزینه‌ای

مثال۱۰: نتیجه مطالعه و بررسی نمونه را................. گویند.

۱) سرشماری √ ۲) داده ۳) نمونه‌گیری ۴)اندازه‌گیری

مثال۱۱: -برای موضوع *استفاده از چتر در هوای بارانی* کدام روش جمع‌آوری داده‌ها مناسب است؟

۱)داده‌های از پیش تهیه‌شده ۲) پرسش ۳) آزمایش √۴) مشاهده و ثبت وقایع

مثال۱۲: روش جمع‌آوری *داده‌های از پیش تهیه‌شده* برای کدام موضوع مناسب است؟

√۱) بررسی تعداد تصادفات رانندگی در یک بزرگراه

۲) تأثیر گوش دادن موسیقی کلاسیک بر یادگیری

۳) تأثیر رژیم گرفتن در کاهش وزن

۴) تعیین رنگ موردعلاقه مردم

مثال۱۳: جمع‌آوری داده‌ها به کدام روش موردقبول نیست؟(سراسر تجربی ۹۱)

۱)مصاحبه √ ۲) پرسش هدایت‌شونده ۳) آزمایش ۴) مشاهده و ثبت وقایع

مثال۱۴: در طراحی پرسشنامه کدام گزینه صحیح نیست؟

۱)سازمان‌دهی محتوا √ ۲)استفاده از سؤالات هدایت‌شونده

۳) استفاده از سؤالات ساده و واضح ۴) استفاده از سؤالات چندگزینه‌ای

مثال۱۵: در کدام مورد عمل سرشماری انجام نشده است؟(سراسری انسانی)

۱) تمام افراد جامعه مورد مطالعه قرار گیرد. ۲) نمونه همان جامعه آماری است.

۳) اندازه ی نمونه برابر اندازه ی جامعه است. √ ۴) نمونه، زیر مجموعه ی جامعه ی آماری است.

مثال۱۶: هدف اصلی علم آمار کدام است؟

۱) تبدیل اطلاعات به یکدیگر ۲) تبدیل اطلاعات به داده ها

√۳)تبدیل داده ها به اطلاعات ۴)تبدیل داده ها به یکدیگر

مثال۱۷:کدام طریق برای جمع آوری داده ها مناسب نیست؟(سراسری تجربی ۹۰ خارج از کشور)

۱)مصاحبه √۲)الگوی خاص ۳)مشاهده ۴)آزمایش

۳ متغیرهای تصادفی

۱-۳ متغیر تصادفی

موضوع یا موضوعاتی را که در نمونه یا جامعه موردمطالعه قرار می‌گیرند، متغیر تصادفی گوییم.

> ***متغیر تصادفی*** *دارای حالت‌ها یا مقادیری است که به‌طور* ***تصادفی*** *در بین اعضای جامعه یا نمونه* ***تغییر*** *می‌کنند.*

متغیر تصادفی به دو دسته **کمی** و **کیفی** تقسیم می‌شوند:

متغیر تصادفی کمی: متغیری که قابل‌اندازه‌گیری است و نتیجه آن عدد هست.

متغیر تصادفی کیفی: متغیری که قابل‌اندازه‌گیری نیست و دارای مقولات یا حالات مختلف است.

مثال۱: کدام متغیر تصادفی کمی است؟

۱)جنسیت افراد ۲) وضعیت تأهل افراد √ ۳) نمره درس دانش آموزان ۴)گروه خونی افراد

مثال۲: کدام متغیر تصادفی کیفی است؟

۱)میزان بارندگی در شهر ۲)درآمد افراد شاغل ۳)قد دانش آموزان √ ۴)نوع تلفن افراد

متغیر تصادفی کمی در بین اعضای نمونه یا جامعه قابل مقایسه است.

مثال۳: با کدام متغیر **نمی‌توان** دانش آموزان کلاس را با هم مقایسه کرد؟

۱)نمره درس √ ۲)گروه خونی ۳) وزن ۴)تعداد اعضای خانواده

۱-۱-۳ انواع متغیر تصادفی

متغیر تصادفی کمی دارای دو نوع پیوسته و گسسته است. همچنین متغیر تصادفی کیفی نیز دارای دو نوع ترتیبی و اسمی است.

متغیر کمی پیوسته: متغیر کمی که مقادیرش بین دو مقدار دلخواه تغییر می‌کند و قابل‌شمارش نیستند.

متغیر کمی گسسته: متغیر کمی که مقادیرش اعداد صحیح است و قابل‌شمارش هستند.

متغیرهای تصادفی کمی گسسته از نوع تعداد می‌باشند.

تذکر: متغیرهای تصادفی که می‌توان تعداد آن‌ها به‌صورت نیم هم شمرد جز متغیرهای تصادفی گسسته هستند مانند تعداد طبقات ساختمان که در صورت ناتمام بودن یک طبقه به‌صورت ۱/۵، ۲/۵، ۳/۵... طبقه شمرده می‌شود.

متغیر کیفی ترتیبی: متغیر کیفی که حالاتش دارای ترتیب هستند. مانند مقاطع تحصیلی(ابتدایی، متوسطه اول، متوسطه دوم، دانشگاهی) و میزان تحصیلات (کاردانی، کارشناسی، کارشناسی ارشد و دکتری)

متغیر کیفی اسمی: متغیر کیفی که حالاتش دارای ترتیب نیستند. مانند جنسیت(زن و مرد) و گروه خونی (A,B,AB,O)

مثال۴: نوع متغیرهای زیر را مشخص کنید؟

الف)میزان آلودگی هوای شهر تهران — کمی پیوسته

ب) زمان انتظار بیمار در مطب پزشک — کمی پیوسته

پ) تعداد غایبین کلاس‌های دبیرستان — کمی گسسته

ت)تعداد مکالمات تلفن افراد — کمی گسسته

ث) ارزشیابی توصیفی دانش آموزان(ضعیف، متوسط، خوب،عالی) — کیفی ترتیبی

ح) مراحل رشد انسان(نوزاد، کودک، نونهال، نوجوان، جوان، میان‌سال، پیر) — کیفی ترتیبی

خ)وضعیت تأهل افراد(مجرد، متأهل) — کیفی اسمی

ج) گروه خونی افراد (A , B , AB , O) — کیفی اسمی

مثال۵: کدام‌یک از متغیرهای زیرکمی پیوسته است.

۱) تعداد نامه‌های یک صندوق ۲) جنسیت افراد √ ۳) میزان بارندگی ۴) گروه خونی

مثال۶: متغیر « نوع تلفن مورداستفاده شهروندان » .... است.

۱) کمی گسسته √ ۲) کیفی اسمی ۳) کیفی ترتیبی ۴) کمی پیوسته

مثال۷: تعداد پاکت‌های موجود در یک صندوق پست یک متغیر...........است؟

۱)کیفی- ترتیبی ۲) کیفی اسمی ۳) کمی پیوسته √ ۴) کمی گسسته

مثال۸: مقطع تحصیلات افراد یک شهر، کدام نوع متغیر است؟(سرا سری انسانی۸۸)

√ ۱)کیفی ترتیبی ۲) کیفی اسمی ۳) کمی پیوسته ۴) کمی گسسته

مثال۹: قطر تنه ی درختان یک باغ کدام نوع متغیر است؟(سراسری انسانی)

۱)کیفی ترتیبی ۲) کیفی اسمی √ ۳) کمی پیوسته ۴) کمی گسسته

۴ دسته‌بندی داده‌ها و جدول فراوانی

۱- ۴ اهداف دسته‌بندی داده‌ها

۱) خلاصه کردن داده‌ها برای تبدیل سریع داده‌ها به اطلاعات

۲) تبدیل داده‌های کمی به کیفی(یک‌کاسه کردن) جهت راحت توصیف کردن داده‌ها

۴-۱-۱ دسته‌بندی داده‌های گسسته

در بعضی از موضوعات که در آن‌ها از داده‌های گسسته استفاده می‌شود با یک‌کاسه کردن داده‌ها بر اساس دسته‌بندی داده‌ها می‌توان اطلاعات را به‌صورت توصیفی نشان داد.

مثال۱: تعداد خانوار ۳۰ دانش‌آموز به‌صورت زیر است:

۳-۵-۴-۶-۵-۴-۳-۷-۹-۸-۷-۶-۸-۵-۷-۱۰-۵-۴-۳-۹-۹-۶-۵-۶-۵-۴-۸-۶-۴-۶-۷

داده را در سه دسته ۳ تا ۵ نفر، ۶ تا ۸ نفر و بیشتر از ۸ نفر به ترتیب به سه نوع خانوار کم‌جمعیت، جمعیت متوسط و پرجمعیت دسته‌بندی می‌کنیم.

نوع خانوار	دسته‌ها
کم‌جمعیت	۳ تا ۵ نفر
جمعیت متوسط	۶ تا۸ نفر
پرجمعیت	بیشتر از ۸ نفر

فراوانی مطلق داده: تعداد دفعات تکرار داده x_i را فراوانی مطلق داده گویند و با f_i نشان می‌دهند.

مجموع فراوانی داده‌ها با تعداد داده‌ها برابر است. به عبارتی $\sum f_i = n$

۴-۱-۲ جدول فراوانی داده‌های گسسته

بعد از یک‌کاسه کردن داده‌های گسسته تعداد داده را در هر دسته شمرده و به‌عنوان فراوانی مطلق در نظر می‌گیریم و جدول فراوانی را برای آن‌ها تشکیل می‌دهیم.

مثال۲: در مثال ۱ جدول فراوانی را تشکیل دهید.

فراوانی مطلق	نوع خانوار	دسته‌ها
۱۴	کم‌جمعیت	۳ تا ۵ نفر
۱۳	جمعیت متوسط	۶ تا۸ نفر
۳	پرجمعیت	بیشتر از ۸ نفر

۳-۱-۴ دسته‌بندی داده‌های پیوسته

مراحل دسته‌بندی داده‌ها را با یک مثال مطرح می‌کنیم:

مثال۳: داده‌های زیر مربوط به نمرات پایانی اول درس آمار و مدل سازی یک کلاس ۳۰ نفر است:

۱۴	۱۲.۵	۱۲	۱۶	۱۴	۹	۱۶	۱۵	۷	۱۳
۱۸	۱۵	۱۴	۱۸	۱۳	۱۷	۱۹	۱۰	۶	۱۲
۱۵	۱۴	۱۵	۱۶	۱۳	۱۷	۱۶	۱۲	۱۱	۱۴

۱) **محاسبه دامنه تغییرات داده‌ها:** حداکثر اختلاف داده‌ها را دامنه تغییرات داده گوییم و به‌صورت زیر محاسبه می‌شود:

$$R = b - a$$

که در آن b بزرگ‌ترین و a کوچک‌ترین داده می‌باشند. در مثال ۳ داریم:

$$R=b-a=۱۹-۶=۱۳$$

نکته: هرچه دامنه تغیرات بزرگتر باشد نشان دهنده این است که اختلاف در جامعه بیشتر است.

اگر داده ها را a برابر کنیم دامنه تغییرات $\|a\|$برابر می‌شود.
اگرداده ها را b واحد کم(اضافه) کنیم دامنه تغییرات تغییر نمی کند.

۲) **مشخص کردن تعداد دسته‌ها:** تعداد دسته‌ها را می‌توان به‌طور اختیاری با توجه به تعداد داده‌ها بین ۵ تا ۲۰ دسته تعیین کرد. تعداد دسته همیشه مشخص می‌شود.

تذکر: می‌توان تعداد دسته‌ها را نیز با استفاده از فرمول $k = [\sqrt{n}]$ که در آن k تعداد دسته‌ها، n تعداد داده‌ها و نماد [ ] جز صحیح است، محاسبه کرد.

در مثال ۳ تعداد دسته‌ها را ۵ در نظر می‌گیریم یعنی:

$$k = [\sqrt{۳۰}] = [۵/۴۷۷] = ۵$$

۳) **تعیین طول دسته‌ها:** طول دسته‌ها به صورت‌های زیر تعریف می‌شود :

۱) تفاضل کران‌های پایین(کران‌های بالا) متوالی دسته‌ها

۲) تفاضل نماینده(مرکز)های متوالی دسته‌ها

۳) تفاضل کران بالا و کران پایین هر دسته

برای تعیین طول دسته از رابطه $\mathbf{c} = \frac{\mathbf{R}}{\mathbf{K}}$ می‌توان استفاده کرد. در محاسبه طول دسته ممکن است مقدار آن عددی صحیح مثبت نباشد که با توجه به‌دقت ثبت داده‌ها (تعداد ارقام اعشاری داده‌ها) مقدار طول دسته(c) را به عدد بالاتر با همان دقت ثبت داده‌ها تغییر می‌دهیم. در مثال۳ داریم: $c = \frac{۱۳}{۵} = ۲/۶ = ۳$

۴) **تعیین کران‌های دستهiام:** کوچک‌ترین و بزرگ‌ترین مقدار دستهi ام به ترتیب کران پایین(a_i) و کران بالای(b_i) دستهiام گوییم.

برای تعیین کران‌های دسته‌ها ابتدا کران پایین دسته اول را به‌صورت زیر محاسبه می‌کنیم.

دسته اول$[a\ ,a+c)$

دسته دوم$[a+c\ ,a+۲c)$

⋮

دسته iام$[a+(i-۱)\ ,a+ic)$

دسته آخر$[b-c, b]$

مقدار طول دسته به‌دست‌آمده را در تعداد دسته‌ها ضرب می‌کنیم و آن را R' می‌نامیم. با محاسبه نصف تفاضل $R'-R$ مقداری به دست می‌آید که در محاسبه کران‌های از آن استفاده می‌کنیم. در مثال۳داریم:

$$\frac{R'-R}{۲} = \frac{۳\times۵-۱۳}{۲} = \frac{۱۵-۱۳}{۲} = \frac{۲}{۲} = ۱$$

حال مقدار به‌دست‌آمده را از کوچک‌ترین داده کم می‌کنیم تا کران پایین دسته اول مشخص شود. در مثال۳داریم:

کران پایین دسته اول: $۶-۱=۵$

در موردی که نسبت دامنه به تعداد دسته ها (طول دسته)مقداری صحیح مثبت هست، مقدار$R'-R$ صفر به دست می‌آید. در این صورت کران پایین دسته اول و کوچک ترین داده برابرند. البته در غیر اینصورت می‌توان مقدار طول دسته را با تغییر به عدد صحیح مثبت به این حالت تبدیل کرد. در این صورت دسته ها را به صورت
$[a_i\ ,b_i)$ نمایش می‌دهیم که در آن a_i کران پایین و b_i کران بالای دسته iام است.

حال با اضافه کردن مقدار طول دسته (c) به کران پایین دسته اول، کران پایین دسته دوم مشخص می‌شود. به همین ترتیب کران پایین دسته‌های بعدی نیز مشخص می‌شوند. در مثال۳داریم:

کران پایین دسته دوم: $۵+۳=۸$

کران پایین دسته سوم: ۱۱ = ۳ + ۸

کران پایین دسته چهارم :۱۴ = ۳ + ۱۱

کران پایین دسته پنجم: ۱۷ = ۳ + ۱۴

بعد از تعیین کران‌های پایین دسته‌ها، برای تعیین کران بالای دسته‌ها می‌توان با اضافه کردن طول دسته به کران‌های پایین، کران‌های بالا را مشخص کرد. یا به عبارتی کران پایین هر دسته با کران بالای دسته قبلش برابر است بنابراین در مثال۳ کران‌ها بالا به‌صورت زیر محاسبه می‌شوند:

کران بالای دسته اول = کران پایین دسته دوم= ۸

کران بالای دسته دوم = کران پایین دسته سوم = ۱۱

کران بالای دسته سوم = کران پایین دسته چهارم = ۱۴

کران بالای دسته چهارم =کران پایین دسته پنجم = ۱۷

کران بالای دسته پنجم =کران پایین دسته پنجم + طول دسته= ۲۰ = ۳ + ۱۷

در آخر با مقایسه کران بالای دسته‌ی آخر و بزرگ‌ترین داده صحت دسته‌بندی داده‌ها را بررسی می‌کنیم. در این مقایسه باید کران بالای دسته‌ی آخر کمتر از بزرگ‌ترین داده نباشد. همان‌طور که معلوم است کران بالای هر دسته با کران پایین دسته بعدی برابر است که نشان‌دهنده پیوستگی داده‌ها است. بنابراین دسته‌ها در مثال۳ به‌صورت زیر هستند:

دسته‌ها
۵-۸
۸-۱۱
۱۱-۱۴
۱۴-۱۷
۱۷-۲۰

نکته: رابطه بین تعداد دسته‌ها، طول دسته و دامنه تغییرات :

$$c = \frac{R}{k} \ , \ R = ck \quad , \quad k = \frac{R}{c}$$

مثال۴: نمرات دانش آموزان یک کلاس در درس ریاضی به‌صورت زیر است. تعداد دسته‌های لازم برای تشکیل جدول توزیع فراوانی کدام است؟(طول هر دسته ۴ فرض شود.)

۴,۱۷,۲۰,۱۴,۱۲,۹,۱۳,۱۸,۱۵,۱۵,۱۴,۴,۳,۵,۵,۵,۲۰

۱) ۴.۵ ۲) ۴ ۳) ۵ √ ۴) ۴/۲۵

$k = \frac{R}{C}, \ a = ۳ \ , b = ۲۰ \ , R = ۲۰ - ۳ = ۱۷$

$\rightarrow K = \frac{۱۷}{۴} = ۴.۲۵ \rightarrow \boxed{K = ۵}$

مثال۵: طول دسته کدام نیست؟

۱)تفاضل کران‌های پایین(بالا) متوالی دسته‌ها ۲) تفاضل کران‌های بالا و پایین هر دسته

۳) تفاضل مرکز(نماینده) متوالی دسته‌ها √ ۴) تفاضل دامنه و تعداد دسته‌ها

۵)مرکز(نماینده) دسته‌ی iام: نصف مجموع کران‌های دستهiام یعنی، $x_i = \frac{a_i + b_i}{۲}$

نکته:در جدول فراوانی، مراکز دسته تشکیل دنباله حسابی با قدرنسبت طول دسته می‌دهندبنابراین رابطه زیر را می‌توان بین دو مرکز دسته بنویسیم:

$$x_j = x_i + (j - i) \times c$$

مثال۶: در مثال ۳ داریم:

مرکز دسته x_i	دسته ها
$\frac{۵ + ۸}{۲} = ۶/۵$	۵ − ۸
۹/۵	۸ − ۱۱
۱۲/۵	۱۱ − ۱۴
۱۵/۵	۱۴ − ۱۷
۱۸/۵	۱۷ − ۲۰

برای دودسته غیر متوالی i و j داریم:

$c = \frac{a_i - a_j}{i - j} \quad , \quad c = \frac{b_i - b_j}{i - j} \quad , \quad c = \frac{x_i - x_j}{i - j} \quad , \quad a_i = x_i - \frac{c}{۲} \quad , \quad b_i = x_i + \frac{c}{۲}$

رابطه بین کران‌های دستهiام، طول دسته ها(c) و کران پایین دسته اول(کوچک ترین داده $a_۱$): $[a_۱ + (i - ۱) \times c \ , \ a_۱ + i \times c)$

مثال۷: در ۵۶ داده آماری، بزرگ‌ترین و کوچک‌ترین آن‌ها به ترتیب ۸۶ و۶۵ است. این داده‌ها به ۷ طبقه دسته‌بندی‌شده‌اند. اگر داده‌هایی که در یک دسته قرار دارند یکسان در نظر گرفته شود، مقدار مشترک(مرکز) آن‌ها در دسته پنجم کدام است؟(سرا سری انسانی۸۸)

۱)۷۷ ۲)۷۷/۵ ۳)۷۸ ۴)۷۸/۵ √

حل:

$$\begin{cases} R = Max - Min = ۸۶ - ۶۵ = ۲۱ \\ \text{طول دسته} = \dfrac{R}{\text{تعداد دسته}} = \dfrac{۲۱}{۷} = ۳ \end{cases}$$

پس دسته‌ها به‌صورت (۶۸،۶۵] و (۷۱،۶۸] و (۷۴،۷۱] و (۷۷،۷۴] و دسته پنجم (۸۰،۷۷] هستند و درنتیجه مقدار مشترک(مرکز) دسته پنجم $۷۸/۵ = \frac{۷۷+۸۰}{۲}$ است.

مثال۸: دسته سوم یک جدول توزیع فراوانی به‌صورت(۵۲,۵۷] است. کران بالای دسته هفتم، کران پایین دسته اول، مرکز دسته ششم را به دست آورید.

$$c = \frac{b_۷ - b_۳}{۷-۳} \Rightarrow ۵ = \frac{b_۷ - ۵۷}{۴} \Rightarrow b_۷ - ۵۷ = ۲۰ \Rightarrow b_۷ = ۷۷$$

$$c = \frac{a_۱ - a_۳}{۱-۳} \Rightarrow ۵ = \frac{a_۱ - ۵۲}{-۲} \Rightarrow a_۱ - ۵۲ = -۱۰ \Rightarrow a_۱ = ۴۲$$

$$c = \frac{x_۶ - x_۳}{۶-۳} \Rightarrow ۵ = \frac{x_۶ - ۵۴/۵}{۳} \Rightarrow x_۶ - ۵۴/۵ = ۱۵ \Rightarrow x_۶ = ۶۹/۵$$

مثال۹: یک سری داده‌های آماری را دسته‌بندی کرده‌ایم، به‌طوری‌که مرکز دسته دوم ۶ و مرکز دسته دهم ۳۴ است. با فرض برابر بودن طول همه‌ی دسته‌ها، کران بالای دسته ششم کدام است؟

۱) ۲۱/۵ ۲) ۲۱/۷۵ √ ۳) ۲۳ ۴) ۲۰

حل:

$$c = \frac{x_{۱۰} - x_۲}{۱۰-۲} = \frac{۳۴-۶}{۸} = \frac{۲۸}{۸} = ۳/۵ \Longrightarrow [۴/۲۵ \text{ و } ۷/۷۵]$$

$۶-۱/۷۵ \quad ۶ \quad ۶+۱/۷۵$

مرکز دسته دوم

$$c = \frac{b_۶ - b_۲}{۶-۲} \Rightarrow ۳/۵ = \frac{b_۶ - ۷/۷۵}{۴} \Rightarrow b_۶ - ۷/۷۵ = ۱۴ \Rightarrow b_۶ = ۲۱/۷۵$$

مثال۱۰: در یک دسته‌بندی داده‌ها، کران پایین دسته دوم و مرکز دسته هشتم به ترتیب ۷, ۳۳ است. اگر این داده‌ها در ۱۲ طبقه دسته‌بندی کرده باشیم، کران بالای دسته آخر چقدر است؟

۱) ۴۷ ۲) ۵۷ ۳) ۵۱ √ ۴) ۴۳

حل:

کران پایین دسته هشتم $= \frac{c}{۲} -$ مرکز دسته هشتم $\Rightarrow$ کران پایین دسته هشتم $= ۳۳ - \frac{c}{۲}$

$$c = \frac{a_۸ - a_۲}{۸ - ۲} \Rightarrow c = \frac{(۳۳ - \frac{c}{۲}) - ۷}{۶} \Rightarrow c = ۴$$

$$c = \frac{b_{۱۲} - b_۲}{۱۲ - ۲} \Rightarrow ۴ = \frac{b_{۱۲} - ۱۱}{۱۰} \Rightarrow b_{۱۲} = ۵۱$$

مثال۱۱: تعدادی داده را در ۴ دسته، دسته‌بندی کرده‌ایم اگر ۱۹، ۳۱ به ترتیب مراکز دسته‌های اول و آخر باشند تمامی دسته‌ها و مراکز آن‌ها را تعیین کنید.

$$c = \frac{x_۱ - x_۴}{۱ - ۴} = \frac{۱۹ - ۳۱}{-۳} = \frac{-۱۲}{-۳} = ۴$$

$$a_۱ = x_۱ - \frac{c}{۲} = ۱۹ - \frac{۴}{۲} = ۱۷$$

مرکز دسته	دسته‌ها $[a_i, b_i)$
۱۹	[۱۷،۲۱)
۲۳	[۲۱،۲۵)
۲۷	[۲۵،۲۹)
۳۱	[۲۹،۳۳]

۴-۲ انواع فراوانی

برای تشکیل جدول فراوانی داده‌های دسته‌بندی‌شده فراوانی‌هایی را به آن اضافه می‌کنیم این فراوانی‌ها در توصیف داده‌ها به ما کمک می‌کنند.

۴-۲-۱ فراوانی مطلق دسته iام

تعداد داده‌ها در دسته iام را فراوانی مطلق آن گوییم و با f_i نشان می‌دهیم.

نکته: برای اینکه سریع‌تر و راحت‌تر فراوانی مطلق دسته‌ها را مشخص کنیم بهتر است داده‌ها را به‌طور غیر نزولی مرتب شوند.

مثال۱۲: در مثال۳ فراوانی مطلق دسته‌ها را به جدول دسته‌بندی و مرکز دسته‌ها اضافه می‌کنیم.

۱۵و۱۵و۱۵و۱۵و۱۴و۱۴و۱۴و۱۴و۱۴و۱۳ و۱۳ و ۱۳ و ۱۲و ۱۲ و ۱۲ و ۱۲ و ۱۱ و۱۰ و ۹ و ۷ و ۶
۱۹ و ۱۸ و ۱۸ و ۱۷ و ۱۷ و ۱۶ و ۱۶ و ۱۶ و ۱۶

دسته‌ها	مرکز دسته x_i	فراوانی مطلق f_i
۵-۸	$\frac{۵+۸}{۲} = ۶/۵$	۲
۸-۱۱	۹/۵	۲
۱۱-۱۴	۱۲/۵	۸
۱۴-۱۷	۱۵/۵	۱۳
۱۷-۲۰	۱۸/۵	۵
جمع		۳۰

مجموع فراوانی‌های مطلق دسته‌ها با تعداد داده‌ها برابر است.

$$\sum f_i = n$$

۲-۲-۴ فراوانی نسبی دسته iام

نسبت فراوانی مطلق دسته iام به مجموع فراوانی‌های مطلق(تعداد داده‌ها) را فراوانی نسبی دسته iام گوییم و با $\frac{f_i}{n}$ نشان می‌دهیم.

نکته: مجموع فراوانی‌های نسبی دسته‌ها برابر یک است.

۳-۲-۴ فراوانی تجمعی دسته iام

تعداد داده‌های کوچک‌تر از کران بالای دسته iام را فراوانی تجمعی دسته iام گوییم. البته در مورد دسته آخر وقتی که کران بالای آن بزرگترین داده است(مقدار طول دسته صحیح ومثبت است) تعداد داده‌های کوچکتر ومساوی کران بالای آن را فراوانی تجمعی دسته آخر گوییم.

*فراوانی تجمعی با فراوانی مطلق دسته اول برابر است.

*فراوانی تجمعی دسته آخر با تعداد داده‌ها برابر است.

*فراوانی تجمعی هر دسته (دسته دوم و بعد از آن) با مجموع فراوانی مطلق آن دسته با فراوانی‌های مطلق دسته‌ی قبلش برابر است. یعنی $\sum_{k=۱}^{i} f_i$: فراوانی تجمعی دسته iام

مثال۱۳: فراوانی تجمعی هر دسته برابر تعداد داده هایی است که مقدار آنها از .... آن دسته .... است.

۱) کران پایین – کوچک تر ۲) کران بالا - کوچک تر √

۳) کران پایین- بزرگ تر ۴) کران بالا - بزرگ تر

۴-۲-۴ درصد فراوانی نسبی(درصد)دسته iام

ضرب فراوانی نسبی دسته iام در ۱۰۰ را درصد فراوانی نسبی(درصد) دسته iام گوییم. یعنی،

$$\frac{f_i}{n} \times ۱۰۰$$

نکته: مجموع درصد فراوانی نسبی دسته‌ها برابر ۱۰۰ است.

مثال۱۴: برای داده‌های دسته‌بندی‌شده مثال ۳ فراوانی‌ها را به جدول اضافه می‌کنیم.

دسته‌ها	مرکز دسته x_i	فراوانی مطلق f_i	فراوانی نسبی $\frac{f_i}{n}$	فراوانی تجمعی $\sum_{k=1}^{i} f_i$	درصد فراوانی نسبی $\frac{f_i}{n} \times ۱۰۰$
۸-۵	۶/۵	۲	$\frac{۲}{۳۰} \approx ۰/۰۷$	۲	$۰/۰۷ \times ۱۰۰ = ۷$
۱۱-۸	۹/۵	۲	$\frac{۲}{۳۰} \approx ۰/۰۷$	۴	$۰/۰۷ \times ۱۰۰ = ۷$
۱۴-۱۱	۱۲/۵	۸	$\frac{۸}{۳۰} \approx ۰/۲۶$	۱۲	$۰/۲۶ \times ۱۰۰ = ۲۶$
۱۷-۱۴	۱۵/۵	۱۳	$\frac{۱۳}{۳۰} \approx ۰/۴۳$	۲۵	$۰/۴۳ \times ۱۰۰ = ۴۳$
۲۰-۱۷	۱۸/۵	۵	$\frac{۵}{۳} \approx ۰/۱۷$	۳۰	$۰/۱۷ \times ۱۰۰ = ۱۷$
جمع		n =۳۰	۱		۱۰۰

مثال۱۵: داده‌های زیر وزن ۲۵ دانش آموز یک کلاس هست. جدول فراوانی آن ها در ۵ دسته تشکیل دهید.

۵۶ – ۴۵ – ۶۰ – ۶۵- ۴۱ – ۳۵ – ۴۳ – ۷۵- ۷۹- ۸۱- ۴۵- ۶۵- ۷۵- ۴۶- ۴۰ – ۷۸

۳۵ – ۸۸ – ۳۸ – ۶۴ – ۸۸ – ۷۵ ۶۰ – ۴۵ – ۵۰

ابتدا داد ه ها را مرتب می‌کنیم:

۳۵ — ۳۵ — ۳۸ — ۴۰ — ۴۱ — ۴۳ — ۴۵ — ۴۵ — ۴۵ — ۴۶ — ۵۰ — ۵۶ — ۶۰

۶۰ — ۶۴ — ۶۵ — ۶۵ — ۷۵ — ۷۵ — ۷۵ — ۷۸ — ۷۹ — ۸۱ — ۸۸ — ۸۹

دسته‌ها	مرکز دسته x_i	فراوانی مطلق f_i	فراوانی نسبی $\frac{f_i}{n}$	فراوانی تجمعی $\sum_{k=1}^{i} f_i$	درصد فراوانی نسبی $\frac{f_i}{n}\times ۱۰۰$
[۵ , ۴۶)	۴۰/۵	۹	$۹/۲۵ = ۰/۳۶$	۹	۳۶
[۴۶ , ۵۷)	۵۱/۵	۳	$۳/۲۵ = ۰/۱۲$	۱۲	۱۲
[۵۷ , ۶۸)	۶۲/۵	۵	$۵/۲۵ = ۰/۲۰$	۱۷	۲۰
[۶۸ , ۷۹)	۷۳/۵	۴	$۴/۲۵ = ۰/۱۶$	۲۱	۱۶
[۷۹ , ۹۰]	۸۴/۵	۴	$۴/۲۵ = ۰/۱۶$	۲۵	۱۶
جمع		**n =۲۵**	**۱**		**۱۰۰**

فراوانی تجمعی دسته‌ی (i - ۱) ام − فراوانی تجمعی دسته‌ی iام = فراوانی مطلق دسته i ام

مثال۱۶: در جدول زیر مقدار x را طوری بیابید تا فراوانی نسبی دسته سوم برابر ۰/۲ باشد؟

دسته ها	۴/۵ − ۹ / ۵	۹/۵ − ۱۴ / ۵	۱۴/۵ − ۱۹ / ۵	۱۹/۵ − ۲۴/۵
فراوانی	۷	x	۴	۶

$$\text{فراوانی نسبی دسته دوم} = \frac{\text{فراوانی مطلق دسته دوم}}{\text{کل فراوانی ها}} \Rightarrow \frac{۲}{۱۰} = \frac{۴}{۷+x+۴+۶} \to x = ۳$$

مثال۱۷: در جدول داده‌های زیر اگر درصد فراوانی نسبی دسته وسط برابر ۳۰ باشد فراوانی دسته ۱۲-۱۰ کدام است؟ (سراسر ی تجربی)

۱۴(۱ ۱۶ (۲ ۱۸ (۳√ ۲۰(۴

مرکز دسته‌ها	۷	۹	۱۱	۱۳	۱۵
فراوانی	۹	۱۵	x	۱۰	۸

حل:

$$\text{درصد فراوانی نسبی دسته سوم} = \frac{f_i}{N} \times ۱۰۰ \to$$

$$۳۰ = \frac{x}{۹+۱۵+x+۱۰+۱۸} \times ۱۰۰$$

$$\frac{۳۰}{۱} = \frac{۱۰۰x}{۴۲ + x} \to \frac{۳}{۱} = \frac{۱۰x}{۴۲ + x} \to ۱۲۶ + ۳x = ۱۰x \to ۱۲۶ = ۷x \to x = \frac{۱۲۶}{۷} = ۱۸$$

مثال۱۸: اندازه قد ۱۲۰ دانش‌آموز، در جدول زیر دسته‌بندی‌شده است. فراوانی دسته‌ی چهارم کدام است؟(سرا سری تجربی)

مرکز دسته	۱۵۵	۱۵۸	۱۶۱	۱۶۴	۱۶۷	۱۷۰
درصد فراوانی نسبی	۱۰	۱۵	۱۸	x	۲۰	۱۲

۱) ۲۰ ۲) ۲۴ ۳) ۲۵ ۴) ۳۰ √

حل:

$$۱۰ + ۱۵ + ۱۸ + x + ۲۰ + ۱۲ = ۱۰۰ \Rightarrow x = ۲۵$$

درصد فراوانی نسبی دسته چهارم $= ۲۵$

$$\frac{\text{فراوانی مطلق دسته چهارم}}{\text{کل فراوانی ها}} \times ۱۰۰ \Rightarrow ۲۵ = \frac{f_i}{۱۲۰} \times ۱۰۰ \Rightarrow ۲۵ = \frac{۱۰۰f_i}{۱۲۰} \to \boxed{f_i = ۳۰}$$

مثال۱۹: ۸۰ داده آماری در ۷ دسته دسته‌بندی‌شده ، فراوانی تجمعی در دسته‌های سوم و چهارم به ترتیب ۴۸ و ۵۶ است. فراوانی نسبی دسته‌ی چهارم کدام است؟

۱) ۰/۸ ۲) ۰/۱ √ ۳) ۰/۱۲ ۴) ۰/۱۵

حل:

فراوانی تجمعی دسته‌ی سوم- فراوانی تجمعی دسته چهارم= فراوانی مطلق دسته چهارم

$$۸ = ۵۶ - ۴۸ = \text{فراوانی مطلق دسته چهارم}$$

$$\text{فراوانی نسبی دسته چهارم} = \frac{\text{فراوانی مطلق دسته چهارم}}{\text{کل فراوانی ها}} = \frac{۸}{۸۰} = ۰/۱$$

مثال۲۰: فراوانی نسبی دسته‌ای برابر ۰/۶ است. اگر از این دسته سه عضو برداریم و در دسته‌ی دیگری قرار دهیم، فراوانی نسبی به ۰/۴ کاهش می‌یابد فراوانی مطلق دسته‌ی اولیه و تعداد داده‌ها را حساب کنید.

$$\frac{f_i}{n} = ۰/۶ \Rightarrow f_i = ۰/۶n \quad (۱)$$

حال اگر از آن دسته سه عضو برداریم و در دسته‌ی دیگری قرار دهیم و فراوانی نسبی به ۰/۴ کاهش یابد داریم:

$$\frac{f_i - ۳}{n} = ۰/۴ \to f_i = ۰/۴n + ۳ \quad (۲)$$

$$(۱),(۲) \to \begin{cases} f_i = ۰/۶n \\ f_i = ۰/۴n + ۳ \end{cases} \Rightarrow ۰/۶n = ۰/۴n + ۳$$

$$\Rightarrow ۰/۲n = ۳ \Rightarrow n = ۱۵, f_i = ۰/۶ \times ۱۵ = ۹$$

مثال۲۱: هشتاد داده‌ی آماری در ۷دسته طبقه‌بندی‌شده‌اند. اگر ۲۰ داده‌ی جدید به این جدول افزوده شود، فراوانی نسبی دسته‌ی وسط تغییر نمی‌کند. نسبت افزایش داده‌های دسته‌ی مذکور به فراوانی مطلق این دسته در داده‌های قبل کدام است؟(سرا سری ریاضی۹۰)

۱) $\frac{۳}{۸}$ **۲)** $\frac{۱}{۵}$ **۳)** $\frac{۱}{۴}$ √ **۴)** $\frac{۱}{۸}$

حل: اگر فراوانی دسته چهارم(وسط) داده‌های قبلی را f و فراوانی دسته چهارم(وسط) داده‌های جدید را F درنظر بگیریم. چون فراوانی نسبی در دو حالت تغییر نمی‌کند، داریم:

$$\frac{f}{۸۰} = \frac{F}{۱۰۰} \Rightarrow \frac{F}{f} = \frac{۱۰۰}{۸۰} = \frac{۵}{۴}$$

بنابراین:

$$\frac{\text{افزایش داده دسته‌های چهارم}}{\text{فراوانی مطلق دسته چهارم قبلی}} = \frac{F - f}{f} = \frac{F}{f} - ۱ = \frac{۵}{۴} - ۱ = \frac{۱}{۴}$$

مثال۲۲: در جدول فراوانی تجمعی داده‌های دسته‌بندی‌شده اگر درصد فراوانی نسبی دسته وسط ۲۴ باشد، فراوانی دسته چهارم کدام است؟ (سرا سری ریاضی۸۵)

مرکز دسته	۱۳	۱۵	۱۷	۱۹	۲۱
فراوانی تجمعی	۵	۱۴	a	۴۱	۵۰

۱) ۱۴ ۲) ۱۵ √ ۳) ۱۶ ۴) ۱۷

حل:

$\text{فراوانی تجمعی دسته آخر} = n \Rightarrow n = ۵۰$ (تعداد داده‌ها)

$$۲۴ = \frac{f_۳}{۵۰} \times ۱۰۰ \Rightarrow f_۳ = ۱۲$$

$$f_۳ = a - ۱۴ \Rightarrow ۱۲ = a - ۱۴ \Rightarrow a = ۲۶$$

$$f_۴ = ۴۱ - a = ۴۱ - ۲۶ = ۱۵$$

مثال۲۳: داده‌های جدول زیر، داده‌های آماری پیوسته است. چند درصد داده‌ها، در فاصله‌ی [۱۸/۵ , ۲۱/۵) قرار دارند؟(سرا سری تجربی۸۸)

مرکزدسته	۱۴	۱۷	۲۰	۲۳	۲۶
فراوانی تجمعی	۵	۱۳	۲۵	۳۴	۴۰

۱)۲۰ ۲)۲۵ ۳)۳۰ √ ۴)۴۰

حل:مرکز دسته [۱۸/۵ , ۲۱/۵) برابر است با $۲۰ = \frac{۱۸/۵ + ۲۱/۵}{۲}$ پس فراوانی دسته سوم:

$f_۳ = ۲۵ - ۱۳ = ۱۲$

درصد فراوانی نسبی دسته سوم: $\frac{f_۳}{n} \times ۱۰۰ = \frac{۱۲}{۴۰} \times ۱۰۰ = ۳۰$

تذکر : تعداد داده‌ها با فراوانی تجمعی دسته آخر(پنجم) برابر است یعنی، $n = ۴۰$

مثال۲۴: در دسته بندی داده‌های آماری، مناسب ترین مقداری که می‌توانیم به هر یک از افراد یک دسته نسبت دهیم کدام است؟(سراسری انسانی)

۱)مرکز دسته √ ۲) کران پایین دسته ۳)میانگین مقادیر دسته ۴)کران بالا دسته

مثال۲۵: در دسته بندی ۱۳۵ داده ی آماری در ۱۵ دسته، کران‌های دسته ی چهارم به صورت [۷۴ , ۷۷) است. اگر این داده ها در ۹ دسته، دسته بندی شوند، کران پایین دسته ی آخر، کدام است؟(سراسری انسانی)

۱)۹۵ ۲)۹۸ ۳)۱۰۲ ۴)۱۰۵ √

حل:

در دسته بندی قدیم:

طول دسته $= ۷۷ - ۷۴ = ۳, k = ۱۵$

$\rightarrow R = c \times k = ۳ \times ۱۵ = ۲۵$ دامنه تغییرات

کران بالای دسته چهارم $b_۴ = a + ۴c \rightarrow ۷۷ = a + ۴ \times ۳$

$\rightarrow a = ۷۷ - ۱۲ = ۶۵$ کوچکترین داده

در دسته بندی جدید:

$k = ۹ \; , \; R = ۴۵ \Rightarrow c = \frac{R}{k} = \frac{۴۵}{۹} = ۵$

تعداد دسته ها   طول دسته جدید

$b = a + ۹c = ۶۵ + ۹ \times ۵ = ۱۱۰$

$\Rightarrow a_۹ = b - c = ۱۱۰ - ۵ = ۱۰۵$

کران پایین دسته آخر (دسته نهم) = ۱۰۵بزرگ ترین داده

مثال۲۶: مقادیر ۱۲۰ داده ی آماری، در بازه [۵۹ , ۲۳] می‌باشند. این داده ها در ۹ طبقه، دسته بندی شده اند. اگر مجموع فراوانی‌های دو دسته ی آخر ۱۵ باشد، چند درصد داده ها کمتر از ۵۱، هستند؟ (سراسری انسانی۹۵)

۱) ۸۲/۵ ۲) ۸۷/۵√ ۳) ۹۰ ۴) ۹۲/۵

حل: تعداد دسته ها تا کران بالای ۵۱: $c = \frac{۵۹-۲۳}{۹} = \frac{۳۶}{۹} = ۴ \to k = \frac{۵۱-۲۳}{۴} = ۷$

بنابراین دودسته آخر دسته‌های هشتم و نهم هستند. پس تعداد داده‌های کمتر از ۵۱ با مجموع فراوانی این ۷ دسته برابر است یعنی ۱۰۵ = ۱۵ − ۱۲۰.

درصد داده‌های کمتر از ۵۱: $\frac{۱۰۵}{۱۲۰} \times ۱۰۰ = ۸۷/۵$

مثال۲۷: اگر داده‌های آماری ۱,۲,۵,۷دارای دامنه تغییرات R باشند و به داده‌های فوق اعداد ۰ و ۶ را اضافه می‌کنیم دامنه تغیرات اعداد جدید کدام است؟

۱) R ۲) $۱ + R$√ ۳) $R - ۱$ ۴) ۲R

حل: دامنه تغییرات داه‌های قدیم $۶ = ۱ - ۷ = R$ و دامنه تغییرات داده‌های جدید $۷ = ۰ - ۷$بنابراین دامنه تغییرات داده‌های جدید برابر است با $R + ۱$.

مثال۲۸: کوچکترین و بزرگترین داده‌های آماری ۱۷/۲ و ۲۲/۶ هستند اگر کران پایین دسته دوم ۱۷/۸ باشد مطلوبست:

الف) دامنه تغییرات

حل:

$$R = ۲۲/۶ - ۱۷/۲ = ۵/۴$$

ب) طول طبقات

حل:

$c =$ کران پایین دسته ی دوم − کران پایین دسته ی اول $= ۱۷/۸ - ۱۷/۲ = ۰/۶$

پ) تعداد طبقا ت

حل:

$$n = \frac{R}{C} = \frac{۵/۴}{۰/۶} = ۹$$

ت) کران پایین دسته‌ی سوم

حل:

$$c = \frac{\text{کران پایین دسته سوم} - \text{کران پایین دسته اول}}{۳ - ۱} \Rightarrow$$

$$۰/۶ = \frac{x - ۱۷/۲}{۲} \rightarrow x = ۲ \times ۰/۶ + ۱۷/۲ = ۱۸/۴$$

ث) کران بالای دسته‌ی ششم

حل:

$$c = \frac{\text{کران بالای دسته ششم} - \text{کران بالای دسته اول}}{۶ - ۱} \Rightarrow ۰/۶ = \frac{x - ۱۷/۸}{۵} \rightarrow x = ۲۰/۸$$

ج) مرکز دسته‌ی پنجم

حل:

$$c = \frac{\text{مرکز دسته پنجم} - \text{مرکز دسته اول}}{۵ - ۱} \Rightarrow ۰/۶ = \frac{x - ۱۷/۵}{۴} \rightarrow x = ۱۹/۹$$

د) مرکزدسته‌ی آخر(سراسری تجربی ۸۶ خارج ازکشور)

حل:

$$c = \frac{\text{مرکز دسته نهم} - \text{مرکز دسته اول}}{۹ - ۱} \Rightarrow ۰/۶ = \frac{x - ۱۷/۵}{۸} \rightarrow x = ۲۲/۳$$

مثال۲۹: در جدول زیر فراوانی نسبی دسته دوم برابر ۰/۲۵ است. فراوانی نسبی دسته سوم کدام است؟

x_I	۶	۱۰	۱۴	۱۸
f_i	۲	۵	x	۵

۱) ۰/۴ ✓ ۲) ۰/۴۵ ۳) ۰/۳۵ ۴) ۰/۲۵

حل:

$$\text{فراوانی نسبی دسته دوم} = \frac{\text{فراوانی مطلق دسته دوم}}{\text{کل فراوانی}} \Rightarrow ۰/۲۵ = \frac{۵}{۲ + ۵ + x + ۵}$$

$$\rightarrow \frac{۲۵}{۱۰۰} = \frac{۵}{۱۲ + x} \rightarrow ۱۲ + x = ۲۰ \rightarrow x = ۸$$

$$\text{فراوانی نسبی دسته سوم} = \frac{\text{فراوانی مطلق دسته سوم}}{\text{کل فراوانی}} = \frac{۸}{۲۰} = ۰/۴$$

مثال۳۰: دانش آموزان یک مدرسه با سال تولد یکسان را وزن کشی کرده و عدد صحیح وزن آنان را یادداشت کرده ایم چند درصد آن ها وزن کمتر از ۵۰ دارند؟ (سراسرس تجربی ۸۸ خارج از کشور)

وزن	۴۶	۴۷	۴۸	۴۹	۵۰	۵۱
تعداد	۸	۹	۱۲	۱۵	۶	۵

۱) ۷۲ ۲) ۷۵ ۳) ۷۸ ۴) ۸۰ √

حل:

وزن	۴۶	۴۷	۴۸	۴۹	۵۰	۵۱
درصد فراوانی نسبی	$\frac{۸}{۵۵} \times ۱۰۰$	$\frac{۹}{۵۵} \times ۱۰۰$	$\frac{۱۲}{۵۵} \times ۱۰۰$	$\frac{۱۵}{۵۵} \times ۱۰۰$	$\frac{۶}{۵۵} \times ۱۰۰$	$\frac{۵}{۵۵} \times ۱۰۰$

$$\frac{۸+۹+۱۲+۱۵}{۵۵} \times ۱۰۰ = ۸۰$$

مثال۳۱: کوچکترین و بزرگترین داده‌های آماری ۳۱ و ۵۲ می‌باشند این داده ها در ۷ دسته ، دسته بندی شده اند. ۳۷ درصد داده ها کمتر از ۴۰ و ۴۸ در صد آنها بیشتر یا مساوی ۴۳ می‌باشد، اگر فراوانی کل ۸۰ باشد فراوانی دسته وسط کدام است؟

۱) ۹ ۲) ۱۲ √ ۳) ۱۵ ۴) ۱۶

حل:

حدود طبقات	[۳۱,۳۴)	[۳۴,۳۷)	[۳۷,۴۰)	[۴۰,۴۳)	[۴۳,۴۶)	[۴۶,۴۹)	[۴۹,۵۲)
	۳۷ درصد کمتر از ۴۰			دسته وسط	۴۸ درصد بیشتر از ۴۳		

دسته چهارم $= [a+۳c, a+۴c) = [۳۱+۹, ۳۱+۱۲) = [۴۰,۴۳)$

درصد فراوانی دسته وسط(چهارم) $= ۱۰۰ - (۳۷+۴۸) = ۱۵$

$$۱۵ = \frac{\text{فراوانی مطلق دسته وسط}}{۸۰} \times ۱۰۰ \Rightarrow ۱۵ = \frac{\text{فراوانی}}{۸۰} \times ۱۰۰ \Rightarrow \text{فراوانی} = ۱۲$$

مثال۳۲: اطلاعات مربوط به دو دسته ی اول در دسته بندی تعدادی داده آماری که در دسته هایی با طول‌های مساوی دسته بندی شده اند به صورت زیر است. با توجه به جدول کران بالای دسته چهارم کدام است؟

۱) ۱۶ ۲) ۱۱ √ ۳) ۱۵ ۴) ۱۸

مرکز دسته	دسته ها
۴	$[a,b)$
d	$[۵,c)$

حل:کران بالای دسته اول برابر کران پایین دسته ی دوم است. پس $b=۵$ در دسته ی اول داریم:

$$\frac{a+b}{۲}=۴\rightarrow\frac{a+۵}{۲}=۴\rightarrow a=۳$$

بنابراین دسته ی اول به صورت $[۳,۵)$ است در نتیجه طول دسته برابر۲ است.

$$\text{طول دسته}=\frac{\text{کران بالای دسته ی چهارم}-\text{کران بالای دسته ی اول}}{۴-۱}\rightarrow ۲=\frac{x-۵}{۳}\rightarrow x=۱۱$$

۵ نمودارها (شاخص‌های هندسی)

نمودارها یکی دیگر از راه‌های خلاصه کردن داده‌ها به‌صورت هندسی برای رسیدن به اطلاعات هستند. استخراج اطلاعات از داده ها با نمودارها خیلی سریع تر و راحت تر است. نمودارها بر اساس نوع داده ها به چند دسته تقسیم می‌شوند.

۵-۱ نمودار میله ای(ستونی)

این نمودار برای داده‌های کیفی و گسسته مناسب است اما می‌توان برای داده‌های دسته بندی شده هم از آن استفاده کرد. در این نمودار برچسب محور افقی(x ها) حالات داده ها(داده‌های کیفی)، مقادیر داده ها(داده‌های گسسته) یا مرکز دسته ها(داده‌های دسته بندی شده) و برچسب محور عمودی(y ها) فراوانی، فراوانی نسبی یا درصد فراوانی نسبی است.

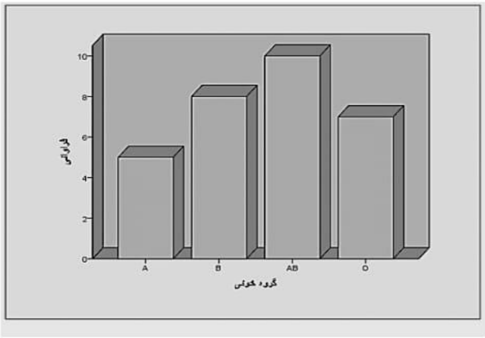
مثال۱: نمودار داده‌های زیر را رسم کنید.

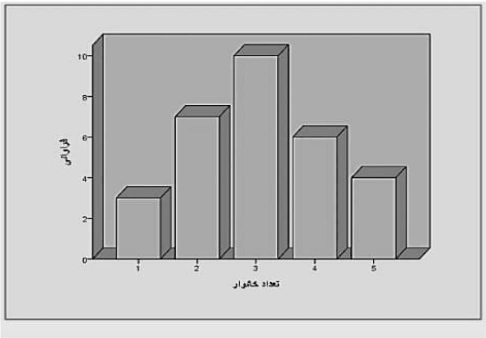
الف) داده‌های کیفی

فراوانی	گروه خونی
۵	A
۸	B
۱۰	AB
۷	O

ب) داده‌های گسسته

فراوانی	تعداد خانوار
۳	۱
۷	۲
۱۰	۳
۶	۴
۴	۵





مثال۲: برای داده‌های کیفی و کمی گسسته کدام نمودار مناسب است؟

۱√)میله ای ۲)مستطیلی ۳)چند بر فراوانی ۴)همه موارد

۲-۵ نمودار مستطیلی

برای داده‌های پیوسته به صورت دسته بندی شده رسم می‌شود. به عبارتی :

نمودار مستطیلی نمایش هندسی داده‌های دسته بندی شده است

مثال۳: نمایش هندسی دسته بندی داده ها کدام است؟

۱)میله ای √ ۲)مستطیلی ۳)چند بر فراوانی ۴)دایره ای

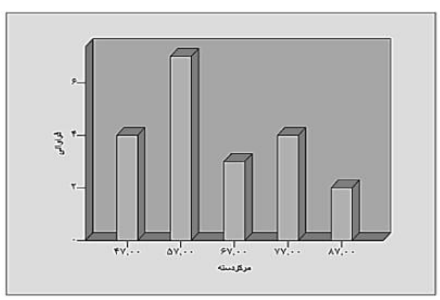
برای رسم نمودار مستطیلی کران‌های دسته ها را روی محور افقی(xها) وفراوانی مطلق، نسبی ودرصد فراوانی نسبی روی محور عمودی(yها) قرار می‌گیرد. برای هر دسته مستطیلی به ضلع طول دسته روی محور افقی و ضلع دیگر آن در امتداد محور عمودی به اندازه فراوانی مطلق، نسبی یا درصد فراوانی نسبی آن دسته رسم می‌شود.

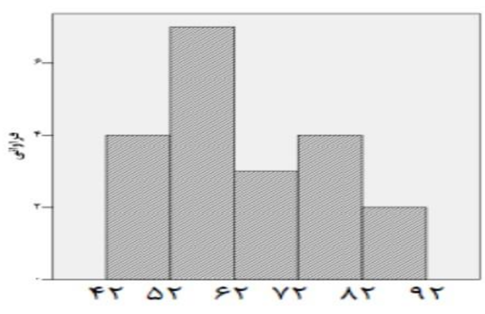
مثال۴: در نمودار مستطیلی برچسب محور طول ها(xها) کدام است؟

۱)مرکز دسته √ ۲)کران‌های دسته ۳)فراوانی ۳)درصد

مثال۵: با توجه به جدول زیر نمودار میله‌ای و مستطیلی را رسم کنید:

مرکز دسته	فراوانی	دسته ها
۴۷	۴	۴۲-۵۲
۵۷	۷	۵۲-۶۲
۶۷	۳	۶۲-۷۲
۷۷	۴	۷۲-۸۲
۸۷	۲	۸۲-۹۲





مجموع مساحت مستطیل‌های نمودار مستطیلی برابر است با حاصلضرب طول دسته ها در تعداد داده ها(مجموع فراوانی‌های مطلق)

$S = c \times \sum f_i = c \times n$ *مجموع مساحت مستطیل ها*

نام دیگر نمودار مستطیلی، هیستوگرام است. در نمودار مستطیلی هر چه مساحت بیشتر باشد یعنی تعداد داده بیشتری در آن دسته قرار دارند.

مثال۶: اگر مجموع مساحت مستطیل ها در نمودارمستطیلی برابر ۸۰ و فراوانی نسبی و مطلق دسته سوم به ترتیب ۰/۴ و ۸ باشد، طول دسته چقدر است؟

۱) ۴/۸ ۲) ۴/۲ ۳) ۵ ۴) ۴ √

حل:

$$\begin{cases} \dfrac{f_۳}{n} = ۰/۴ & \text{فراوانی نسبی دسته سوم} \\ f_۳ = ۸ & \text{فراوانی مطلق دسته سوم} \end{cases} \Rightarrow n = \dfrac{۸}{\frac{۰}{۴}} = ۲۰ \Rightarrow ۸۰ = c \times ۲۰$$

$$\Rightarrow c = \frac{۸۰}{۲۰} = ۴$$

۳-۵ نمودار چند بر فراوانی

نمودار چندبر فراوانی نیز برای داده‌های پیوسته با توزیع فراوانی داده ها مناسب است. در این نمودار نقاطی را در صفحه‌ی مختصات پیدا می‌کنیم و به هم وصل می‌کنیم که طول آنها مرکز دسته ها و عرض آنها فراوانی همان دسته می‌باشد.

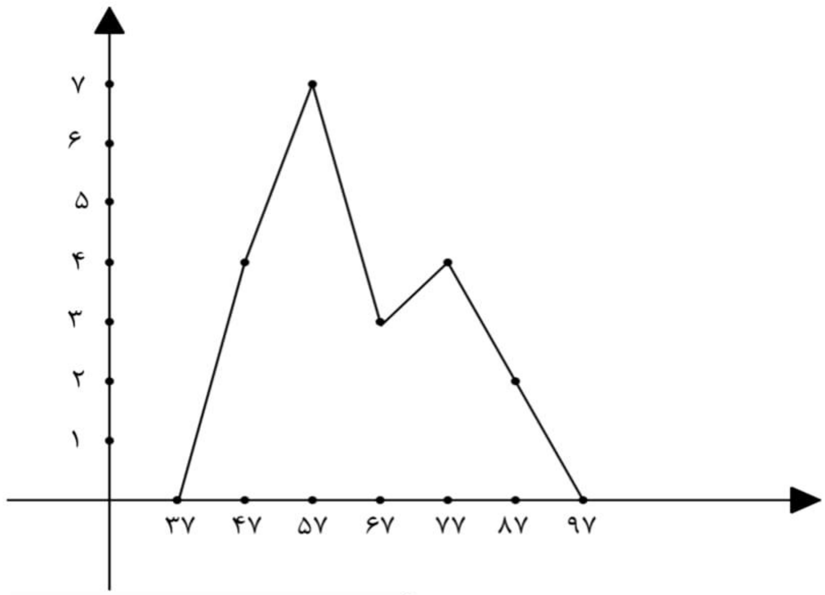
مثال۷: طول نقاط در نمودار چند بر فراوانی کدام است؟

۱)کران پایین دسته ۲)کران بالای دسته ۳√)مرکز دسته ۴)فراوانی

نکته: در نمودار چندبر فراوانی دو دسته با فراوانی‌های صفر به ابتدا و انتهای دسته ها اضافه می‌کنیم تا شکل بسته شود که در این صورت مساحت زیر آن با محور xها (S') با مجموع مساحت مستطیل‌های نمودار مستطیلی (S) برابر است.

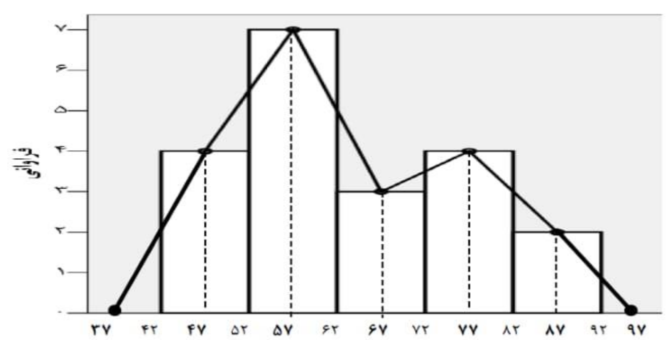
$$\acute{S} = S = c \times \sum f_i = c \times n$$

مثال۸: نمودار چند بر فراوانی مثال۵ را رسم کنید.



$\underbrace{(\text{مرکز دسته آخر} + c)\ \text{کران بالای دسته آخر} + \frac{c}{۲}}_{۸۷+۱۰=۹۷}$ $\underbrace{(\text{مرکز دسته اول} - c)\ \text{کران پایین دسته اول} - \frac{c}{۲}}_{۴۷-۱۰=۳۷}$

نکته: نمودار چند بر فراوانی را می‌توان روی نمودار مستطیلی رسم کرد. برای مثال ۵ داریم:



مثال۹: در داده‌های آماری دسته بندی شده، مساحت نمودار مستطیلی آن را S و سطح زیر نمودار چندبر فراوانی را که دو سر آن بر روی محور افقی باشد S' می‌نامیم نسبت $\frac{S}{S'}$ چگونه است؟ (سراسری تجربی ۸۹ خارج از کشور)

۱)کوچکتر از یک ۲) بزرگتر از یک ۳) برابر یک ۴) اظهارنظر نمی توان کرد.

مثال۱۰: در دسته بندی داده‌های آماری پیوسته، دسته اول ۱۳-۱۰ تعداد دسته ها ۴ تا است. اگر مساحت زیر نمودار چندبر فراوانی با مساحت زیر نمودار مستطیلی برابر باشد طول اولین و آخرین نقطه‌ی نمودار چندبر کدام است؟

۱) ۱۱/۵, ۱۴/۵ ۲) ۷,۲۵ ۳) ۱۰,۲۲ ۴√) ۲۳/۵, ۸/۵

حل:

$c=۳ \Rightarrow$ ۱۳-۱۰ ۱۶-۱۳ ۱۹-۱۶ ۲۲-۱۹

مرکز دسته ها ۱۱/۵ ۱۴/۵ ۱۷/۵ ۲۰/۵

۸/۵= ۱۱/۵-۳ ۲۳/۵ =۳ +۲۰/۵

مثال۱۱: در جدول فراوانی داده‌های پیوسته و دسته بندی شده، دو نقطه (۲۱,۴۲) و (۲۴,۵۱) متوالیا از نمودار فراوانی تجمعی است. کدام نقطه در رسم چندبر فراوانی به کار می‌رود؟ (سراسری ریاضی ۸۸)

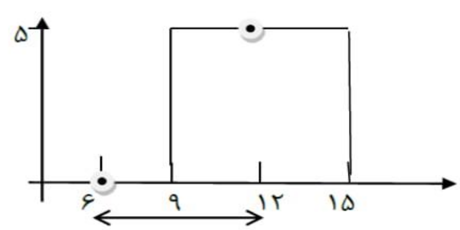
۱) (۲۱,۵۱) ۲) (۲۲/۵ ,۹) ۳) (۲۲/۵ ,۴۲) ۴√) (۲۴,۹)

حل: طول نقطه ها مرکز دسته و عرض نقطه فراوانی تجمعی است. پس مرکز دو دسته متوالی ۲۱ و ۲۴ است و فراوانی تجمعی آنها ۴۲ و ۵۱ است. فراوانی مطلق دسته دوم برابر است با ۹= ۴۲-۵۱ و نقطه (۹, ۲۴) در نمودار چندبرفراوانی می‌آید.

مثال۱۲: هرگاه دسته اول یک جدول فراوانی بصورت (۹,۱۵] و فراوانی آن ۵ باشد، طول‌های نقاط اول و دوم نمودار چندبر فراوانی تکمیل شده چه اعدادی هستند؟

۱√) ۶,۱۲ ۲) ۰,۵ ۳) ۶,۵ ۴) ۱۲,۰

حل:



طول دسته $c=۱۵-۹=۶$

مثال۱۳: اگر مجموع مساحت مستطیل ها در نمودارمستطیلی برابر۸۰ و فراوانی نسبی و مطلق دسته سوم به ترتیب ۰/۴ و ۸ باشد، طول دسته چقدر است؟

۱) ۴/۸ ۲) ۴/۲ ۳) ۵ ۴√) ۴

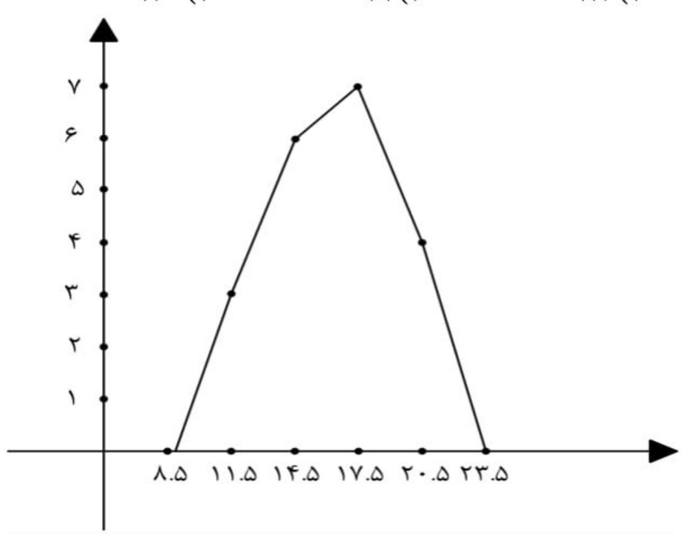
حل:

فراوانی نسبی دسته سوم $= \frac{f_3}{n} \Rightarrow 0/4 = \frac{8}{n} \Rightarrow n = 20$ تعداد داده ها

$$S = c \times n \Rightarrow 80 = c \times 20 \Rightarrow c = \frac{80}{20} = 4$$

مثال۱۴: نمودار چندبر فراوانی زیر داده شده است. مجموع مساحت‌های مستطیل‌های نمودار مستطیلی کدام است؟ (مثال کتاب)

۱) ۶۰√ ۲) ۴۸ ۳) ۷۲ ۴) ۸۰



حل:

$$S = c \times \sum f_i = 3(3 + 6 + 7 + 4) = 60$$

مثال۱۵: فراوانی تجمعی یک سری داده آماری پیوسته، به صورت زیر است اگر درصد فراوانی نسبی دسته وسط ۱۰ باشد، مساحت زیر نمودار چندبر فراوانی این داده ها کدام است؟

مرکز دسته ها	۱	۳	۵	۷	۹
فراوانی تجمعی	۲	۶	۸	۱۷	a

۱) ۸۰ ۲) ۲۰ ۳) ۳۰ ۴) ۴۰√

حل:

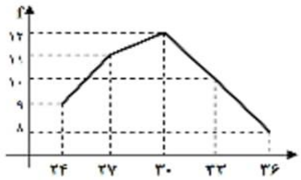
درصد فراوانی دسته وسط $= \dfrac{\text{فراوانی مطلق دسته وسط}}{\text{کل فراوانی ها}} \times 100 \rightarrow 10 = \dfrac{8-6}{a} \times 100 \rightarrow a = 20$

مجموع فراوانی ها = آخرین عدد در فراوانی تجمعی $= \sum f_i = a$

$S = C.\sum f_i = ۲ \times ۲۰ = ۴۰$

مثال۱۶: به داده‌های آماری با نمودار زیر دو داده ۲۹ و۳۲ افزوده می‌شود، درصد فراوانی نسبی در دسته وسط جدید کدام است؟(ریاضی۹۴)

۱) ۲۳ ۲) ۲۴ ۳) ۲۵√ ۴) ۲۶



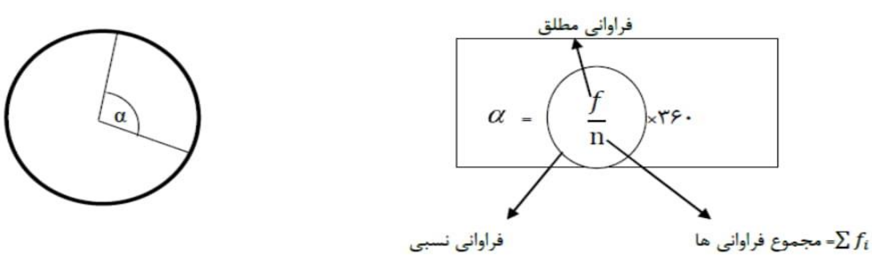
حل: طول دسته برابر است با $۲۷ - ۲۴ = ۳$ c= بنابراین دسته ها به صورت زیر هستند، که داده ۲۹ در دسته سوم(دسته وسط) وداده ۳۲ در دسته چهارم قرار می‌گیرد. اضافه کردن این دو داده هیچ تغییری در دامنه تغییرات و تعداد دسته ایجاد نمی کند و درنتیجه طول دسته ها هم ثابت است.

دسته ها	۲۲/۵-۲۵/۵	۲۵/۵-۲۸/۵	۲۸/۵-۳۱/۵	۳۱/۵-۳۴/۵	۳۴/۵-۳۷/۵
فراوانی	۹	۱۱	۱۲+۱	۱۰+۱	۸

درصد فراوانی دسته وسط $\frac{f_۳}{\sum f_i} \times ۱۰۰ = \frac{۱۳}{۹+۱۱+۱۳+۱۱+۸} \times ۱۰۰ = \frac{۱۳}{۵۲} \times ۱۰۰ = ۲۵$

۴-۵ نمودار دایره ای:

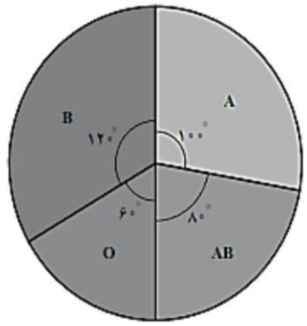
نمودار دایره ای برای داده‌های کیفی مناسب است اما می‌توان برای داده ای دسته بندی شده هم از آن استفاده کرد. برای رسم نمودار دایره ای به هر حالت متغیر کیفی یا دسته قطاعی یا زاویه مرکزی از دایره را متناسب با فراوانی ان به صورت زیر بدست می آوریم.



گروه خونی	A	B	AB	O

مثال۱۷: داده‌های زیر مربوط به گروه خونی ۱۸ دانش آموز یک کلاس است. نمودار دایره ای را برای آن رسم کنید.

۳	۴	۶	۵	فراوانی



$$\alpha_A=\frac{۵}{۱۸}\times ۳۶۰=۱۰۰^\circ$$
$$\alpha_B=\frac{۶}{۱۸}\times ۳۶۰=۱۲۰^\circ$$
$$\alpha_{AB}=\frac{۴}{۱۸}\times ۳۶۰=۸۰^\circ$$
$$\alpha_O=\frac{۳}{۱۸}\times ۳۶۰=۶۰^\circ$$

*اگر فراوانی داده ها را چند برابر کنیم زاویه‌های مرکزی در نمودار دایره ای تغییری نمی کنند.

*ترتیب کنارهم قرار گرفتن نواحی (قطاع ها)در نمودار دایره ای در تفسیر داده ها تأثیری ندارد.

*تغییر شعاع دایره در نمودار دایره ای تغییری در تفسیر داده ها ندارد.

مثال۱۸: در یک شرکت دارویی جدول توزیع کارکنان را با نمودار دایره‌ای نشان می‌دهیم. زاویه مربوط به کارکنان ارشد، چند درجه است؟(سراسری تجربی۹۳)

۱) ۸۴° ۲) ۹۲° ۳) ۹۶° √ ۴) ۱۰۵°

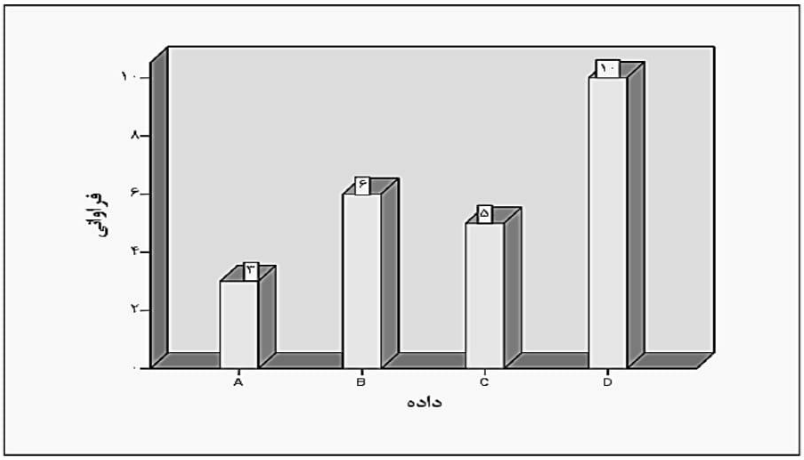
دکتری	ارشد	کارشناسی	کاردانی	دیپلم	نوع مدرک
۳۰	۱۲۰	۱۸۰	۹۰	۳۰	تعداد

حل:

$$\alpha=\frac{۱۲۰}{۳۰+۹۰+۱۸۰+۱۲۰+۳۰}\times ۳۶۰=\frac{۱۲۰}{۴۵۰}\times ۳۶۰=۹۶$$

مثال۱۹: اگر نمودار میله‌ای یک سری داده‌های آماری بصورت زیر باشد، در نمودار دایره‌ای، زاویه مربوط به D چند درجه از زاویه مربوط به B بیشتر است؟

۱) ۹۰ ۲) ۳۰ ۳) ۴۵ ۴) ۶۰ √

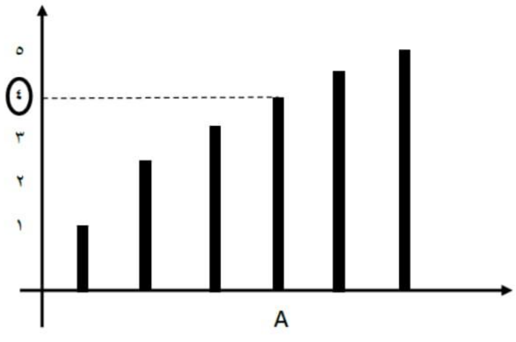


حل: $\alpha_D = \frac{۱۰}{۳+۶+۵+۱۰} \times ۳۶۰ = \frac{۱۰}{۲۴} \times ۳۶۰ = ۱۵۰$

$\alpha_B = \frac{۶}{۳+۶+۵+۱۰} \times ۳۶۰ = \frac{۶}{۲۴} \times ۳۶۰ = ۹۰ \rightarrow \alpha_D - \alpha_B = ۱۵۰ - ۹۰ = ۶۰$

مثال۲۰: در مقایسه سطح زیر کشت غله‌ای در شش استان نمودار میله‌ای مقابل رسم شده است. در نمودار دایره‌ای زاویه مرکزی متناظر استان A چند درجه است؟ (قسمت غیر صحیح هر میله ۰/۵ است)(سراسری تجربی ۹۰)

۱) ۶۴ ۲) ۷۲ √ ۳) ۸۰ ۴) ۹



حل:

$\alpha_A = \frac{۴}{۱+۲.۵+۳+۴+۴.۵+۵} \times ۳۶۰ \quad \alpha_A = \frac{۴}{۲۰} \times ۳۶۰ = ۷۲$

مثال۲۱: داده‌های آماری در ۹ طبقه دسته بندی شده اند. فراوانی تجمعی نسبی در دو دسته‌ی چهارم و پنجم به ترتیب ۰/۲۸ و ۰/۴۰ است. در نمودار دایره ای زاویه مربوط به دسته‌ی پنجم چند درجه است؟ (سراسری ریاضی ۸۶)

۱) ۴۰/۵ ۲) ۴۱/۴ ۳) ۴۲/۶ ۴) ۴۳/۲ √

حل:

فراوانی تجمعی نسبی دسته چهارم- فراوانی تجمعی نسبی دسته پنجم = فراوانی نسبی دسته پنجم

$$\frac{f}{n} = ۰/۴۰ - ۰/۲۸ = ۰/۱۲ \Rightarrow \alpha = \frac{f}{n} \times ۳۶۰^\circ = ۰/۱۲ \times ۳۶۰^\circ = ۴۳/۲$$

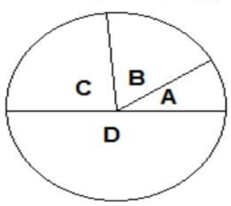
مثال۲۲: در نمودار دایره ای زاویه مربوط به یک دسته ۷۲° است. فراوانی نسبی این دسته کدام است؟

۱) ۰/۱ ۲) ۰/۴ ۳) ۰/۳ ۴) ۰/۲ √

حل:

$$\alpha = \frac{f}{n} \times ۳۶۰^\circ \rightarrow ۷۲^\circ = \frac{f}{n} \times ۳۶۰^\circ \rightarrow \frac{f}{n} = ۰/۲$$

مثال۲۳: در نمودار دایره ای مقابل تعداد افرادی که دردسته‌های D,C,B قرار دارند به ترتیب ۲, ۳, ۶ برابر تعداد افرادی است که در A قرار دارند زاویه A کدام است؟



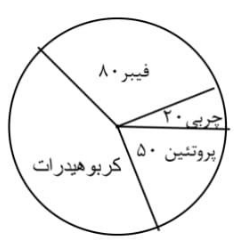
۱) ۲۰° ۲) ۳۰° √ ۳) ۴۵° ۴) ۶۰°

حل:

$$A + B + C + D = ۳۶۰^\circ$$
$$\rightarrow A + ۲A + ۳A + ۶A = ۳۶۰^\circ \rightarrow A = ۳۰^\circ$$

مثال۲۴: نمودار دایره ای زیر سهم وزنی- ترکیبات تشکیل دهنده یک بسته غذایی کنسرو شده را نشان می‌دهد. چند گرم کربوهیدرات در بسته‌ی ۴۸۰ گرمی از این محصول وجود دارد؟

۱) ۲۴۰ ۲) ۲۵۰ ۳) ۳۰۰ ۴) ۲۸۰ √



حل:

$$\alpha_{کربوهیدرات} = ۳۶۰^\circ - (۸۰^\circ + ۲۰^\circ + ۵۰^\circ) = ۲۱۰^\circ, \alpha = \frac{f_i}{n} \times ۳۶۰^\circ$$

$$۲۱۰^\circ = \frac{f_i}{۴۸۰} \times ۳۶۰^\circ \rightarrow f_i = \frac{۲۱۰ \times ۸}{۶} = ۲۸۰$$

مثال۲۵: شرکتی ۱۶۰ کارمند دارد که مدرک تحصیلی آنان با ۶ کد متمایز مشخص شده اند در نمودار دایره ای، زاویه‌ی مرکزی هر گروه با واحد درجه مطابق جدول زیر است. تعداد کارکنان با کد ۴ کدام است؟ (سراسری تجربی ۹۰ خارج از کشور)

کد	۱	۲	۳	۴	۵	۶
زاویه مرکزی	۲۷	۴۵	۹۹	a	۵۴	۱۸

۱) ۵۲ √ ۲) ۵۴ ۳) ۵۶ ۴) ۵۸

حل:

$$۷ + ۴۵ + ۹۹ + \alpha + ۵۴ + ۱۸ = ۳۶۰ \rightarrow \alpha = ۱۱۷, \alpha = \frac{f_i}{n} \times ۳۶۰^\circ$$

$$۱۱۷ = \frac{f}{۱۶۰} \times ۳۶۰ \rightarrow f = \frac{۱۱۷ \times ۱۶۰}{۳۶۰} = ۵۲$$

مثال۲۶: در جدول زیر زاویه مرکزی مربوط به دسته‌ای به مرکز ۱۶ در نمـودار دایـره‌ای برابـر ۲۴° است. چند درصد داده ها از ۱۳/۵ کمتر است؟

۱) ۷۵٪ ۲) ۶۰٪ ۳) ۶۶/۶٪ ۴) ۷۶/۶٪ √

دسته ها	۳/۵ -۸/۵	۸/۵ -۱۳/۵	۱۳/۵-۱۸/۵	۱۸/۵-۲۳/۵
فراوانی	۹	۱۴	x	۵
	تعداد داده که‌های از ۱۳.۵ کمتر است.			

حل:

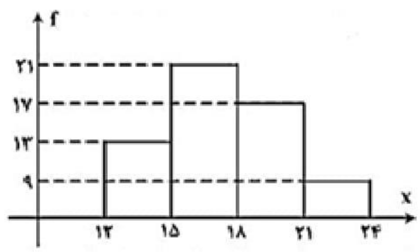
$$\alpha = \frac{f_i}{n} \times ۳۶۰^\circ \Rightarrow ۲۴ = \frac{x}{۹ + ۱۴ + x + ۵} \times ۳۶۰ \Rightarrow x = ۲$$

تعداد کل داده ها $n = \sum f_i = ۹ + ۱۴ + ۲ + ۵ = ۳۰$

$\frac{۹+۱۴}{۳۰} \times ۱۰۰ = ۷۶/۶\%$ درصد داده هایی که از ۱۳/۵ کمترند

مثال۲۷: از داده‌های آماری با نمودار مستطیلی زیر، سه داده ۱۴ و۱۶ و ۱۶ حذف شده است. در نمودار دایره ای داده‌های جدید، بزرگ ترین زاویه مرکزی نظیر دسته ها، چند درجه است؟ (سراسری تجربی۹۴)

۱) ۹۰ ۲) ۱۰۵ ۳) ۱۲۰ √ ۴) ۱۳۵



حل: داده ۱۴ از دسته اول کم می‌شود یعنی فراوانی آن به ۱۲ و دو داده ۱۶ از دسته دوم کم می‌شود یعنی فراوانی آن به ۱۹ تغییر می‌یابد. بنابراین بزرگترین زاویه مربوط به دسته دوم با فراوانی ۱۹ است. از طرفی تعداد داده ها با مجموع فراوانی ها یعنی

$۹+۱۲+۱۷+۱۹ = ۵۷$ برابر است پس، $\alpha = \frac{۱۹}{۵۷} \times ۳۶۰ = ۱۲۰$

مثال ۲۸ :نمودار دایره ای برای کدام متغیر مناسب است و اندازه زاویه مرکزی هر قسمت متناسب با کدام است(سراسری انسانی۹۵)؟

۱)کیفی– فراوانی نسبی √ ۲) کیفی– فراوانی تجمعی

۳) گسسته – فراوانی مطلق ۴) گسسته– فراوانی تجمعی

۵-۵ نمودار ساقه وبرگ

این نمودار الهام گرفته از درخت است. به این صورت که چون در درخت هر ساقه شامل برگ هایی است، داده‌های کمی را می‌توان به دو قسمت یکی برگ که همان اولین رقم سمت راست (یکان یا رقم اعشار) و دیگری ساقه که سایر ارقام است تقسیم کرد. جلوی هر ساقه، برگ‌های آن را به ترتیب از کوچک به بزرگ می‌نویسیم. این نمودار تمام داده ها را در خودش دارد.

نکته : برگ ها همیشه یک رقم و تکراری می‌توانند باشند(از ۰ تا ۹)

نکته: داده ها در این نمودار قابل بازیافت هستند.

مثال ۲۹: در داده‌های آماری زیر نمودار ساقه و برگ را رسم کنید.

۱۰،۵۰،۶۵،۳۳،۴۸،۵،۱۱،۲۳،۳۷،۲۶،۲۶،۳۲،۱۷،۷،۱۳،۱۹،۲۹،۴۳،۲۱،۲۲

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

با رسم نمودار ساقه وبرگ، داده ها مرتب می‌شوند.

مثال ۳۰: در نمودار ساقه و برگ زیر به ترتیب چند عدد می‌تواند به جای x و y قرار می‌گیرد؟

ساقه	برگ
۳	۰ ۰ ۱ x ۶ ۷ ۹
۴	۱ ۱ ۱ ۲ ۳ ۴ ۷
۵	۰ ۱ ۲ ۵ ۶ y
۶	۰ ۲ ۳

۱) ۴٫۳   ۲) ۴٫۶ √   ۳) ۶٫۴   ۴) ۶٫۱

حل:

$$X = ۱,۲,۳,۴,۵,۶ \qquad Y = ۶,۷,۸,۹$$

مثال۳۱: داده‌های آماری، به صورت نمودار ساقه وبرگ زیر داده شده است. چند درصد این داده ها در بازه ی(۳۴٫۴۵] است؟(سراسری انسانی۹۴)

۱) ۳۲   ۲) ۳۲.۵   ۳) ۳۶ √   ۴) ۳۷.۵

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

حل:تعداد داده با تعداد برگ ها برابر ۲۵ است. و داده هایی که در دسته (۴۵ , ۳۴] قرار می‌گیرند با شمارش داده ها برابر ۹ است. بنابراین درصد داده ها در دسته(۴۵ , ۳۴]برابر است با $\frac{۹}{۲۵} \times ۱۰۰ = ۳۶٪$ پس گزینه ۳ صحیح است.

مثال۳۲: توزیع گروه خونی تعدادی از افراد در جدول زیر آمده است. درصد مساحت مربوط به گروه خونی O در نمودار دایره ای کدام است؟

O	AB	B	A
۱۲	۱۰	۱۴	۲۴

۱) ۱۵   ۲) ۲۰ √   ۳) ۲۵   ۴) ۴۰

حل:

$$\alpha = \frac{۱۲}{۶۰} \times ۳۶۰ = ۷۲ \rightarrow \frac{۷۲}{۳۶۰} \times ۱۰۰ = ۲۰$$

۶ شاخص‌های مرکزی

مقداری که مرکز داده ها را مشخص می‌کند شاخص مرکزی گویند. شاخص مرکزی نشان دهنده تمرکز داده است.

شاخص های مرکزی سه نوع اند: مد، میانه و میانگین

۱-۶ مد(نما) Mo:

داده ای که بیشترین فراوانی (تعداد دفعات تکرار) را دارد مد گوییم و با Mo نشان می‌دهیم.

نکته: در رأی گیری انتخاباتی اساس تصمیم گیری مد است. به عبارتی نامزدی که بیشترین تعداد رأی را دارد انتخاب می‌شود.

نکته: بیشترین استفاده مد برای داده‌های کیفی است.

مثال۱: مد داده‌های زیر را بیابید.

الف) ۵ – ۲ – ۳ – ۵ – ۲ – ۱ – ۴ – ۲ – ۶ – ۲

ب)

داده ها	۱۲	۱۴	۱۶	۱۸	۲۰
فراوانی	۲	۳	۲	۲	۱

حل: در قسمت الف چون داده ۲ بیشترین فراوانی دارد پس مد یعنی ۲ = Mo

در قسمت ب مد ۱۴ است چون بیشترین فراوانی را دارد یعنی ۱۴ = Mo

مد ممکن است منحصر به فرد نباشد.

نکته: جامعه چند مدی، جامعه ای است که چند داده از آن دارای بیشترین فراوانی یکسان هستند و علت آن همگن(یک دست) نبودن جامعه است.

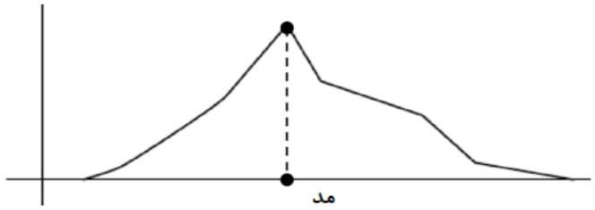
در جامعه‌های چند مد شاخص مرکزی مد معتبر نیست.

مثال۲: در جامعه غیر یک دست کدام شاخص معتبر نیست؟

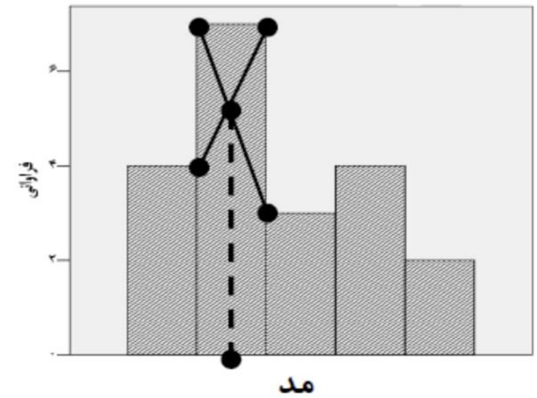
۱)مد √ ۲)میانه ۳)میانگین ۴)چارک

نکته: برای تعیین محل تقریبی مد از روی نمودار چندبرفراوانی از بالاترین نقطه خطی عمود بر محور افقی(xها) رسم می‌کنیم.

مثال۳: محل تقریبی مد در نمودار زیر را مشخص کنید :



نکته: برای تعیین تقریبی مد از روی نمودار مستطیلی به صورت شکل زیر عمل می‌کنیم.



مد تقریبی در داده‌های دسته بندی شده مرکز دسته ای است که بیشترین فراوانی را دارد.

۲-۶ میانه Md:

پس ازمرتب کردن داده ها، مقداری را که تعداد داده ها ی بعد از آن با تعداد داده‌های قبل از آن برابر است میانه می‌نامیم.

میانه برای داده‌های کیفی قابل محاسبه نیست و فقط برای داده‌های کمی محاسبه می‌شود.

مثال۴: کدام شاخص برای داده‌های کیفی قابل محاسبه است؟

۱)میانگین ۲)میانه ۳)مد√ ۴)چارک

۱-۲-۶ روش محاسبه میانه

۱) داده ها را به طور غیر نزولی مرتب می‌کنیم(تکرار داده ها نوشته می‌شود).

۲) بر اساس تعداد داده ها (n):

الف) اگر تعداد داده ها فرد باشد: داده ی وسط میانه است وشماره آن برابر است با $\frac{n}{2} + 0/5$.

ب) اگر تعداد داده ها زوج باشد: نصف مجموع دو داده وسط میانه است و شماره‌های آنها برابر است با $\frac{n}{2}$ و $\frac{n}{2} + 1$.

مثال۵: میانه و مد داده‌های زیر را بیابید.

الف) ۵ − ۲ − ۷ − ۶ − ۵ − ۴ − ۱

حل:

داده‌های مرتب شده : ۱ − ۲ − ۴ − ۵ − ۵ − ۶ − ۷

شماره داده وسط : $\frac{7}{2} + 0.5 = 4$ و $Mo = 5$

بنابراین داده چهارم میانه است یعنی $Md = 5$.

ب) ۴ − ۸ − ۴ − ۵ − ۲ − ۱

حل:

داده‌های مرتب شده : ۱ − ۲ − ۴ − ۴ − ۵ − ۸

شماره داده‌های وسط : $\frac{6}{2} = 3$ و $\frac{6}{2} + 1 = 4$, $Mo = 4$

بنابراین نصف مجموع داده‌های سوم وچهارم میانه است یعنی $Md = \frac{4+4}{2} = 4$

ج)

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

حل:

داده ها در نمودار ساقه وبرگ مرتب اند.

شماره داده‌های وسط: $\frac{16}{2} = 8$ و $\frac{16}{2} + 1 = 9$ $Mo = 26$

بنابراین نصف مجموع داده‌های هشتم و نهم میانه است یعنی $Md = \frac{23+25}{2} = 24$.

اگر تعداد داده ها زوج باشد و دو داده ی وسط مساوی نباشند میانه مقداریست غیر از داده ها.

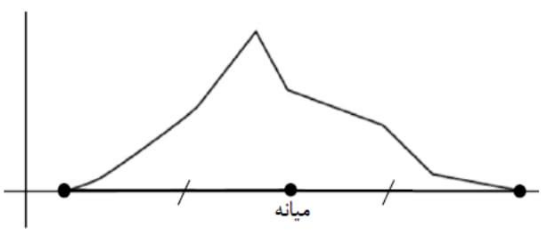
مثال۶: در کدام مورد میانه مقداریست غیر از داده ها؟

۱)تعداد داده ها زوج و دو داده وسط مساوی ✓۲) تعداد داده ها زوج و دو داده وسط نامساوی

۳) تعداد داده ها فرد ۴) همه موارد

وسط دو نقطه انتهایی نمودار چند برفراوانی روی محور افقی(xها) محل تقریبی میانــه داده هــا است.

مثال۷: محل تقریبی میانه در نمودار زیر را مشخص کنید.



۳-۶ چارک ها

با تعیین میانه، داده ها به دو نیمه ی اول و دوم با تعداد داده‌های مساوی (۵۰درصد داده ها در هر نیمه) تقسیم می‌شوند. میانه نیمه اول را **چارک اول**، میانه نیمه دوم را **چارک ســوم** و میانه را **چارک دوم** گوییم وبا نمادهای Q_1، Q_2 و Q_3 نشان می‌دهیم. با تعیین چارک ها تعداد داده ها به چهار قسمت مساوی تقسیم می‌شوند.

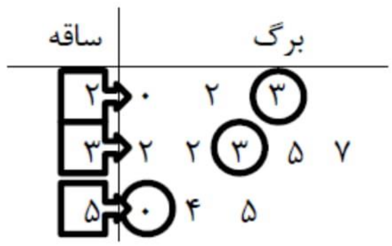
مثال۸: داده‌های ۲۹, ۱۴, ۲۰, ۹, ۱۳, ۲۱, ۲۶, ۱۱, ۳, ۱۴ مفروضند، مطلوبست میانه، چــارک اول و چارک سوم:

$$\text{میانه} = \frac{14+14}{2} = 14$$

$$\underbrace{3, 9, \overset{Q_1}{\overbrace{11}}, 13, 14}_{\text{نیمه اول}}, \underbrace{14, 20, \overset{Q_2}{\overbrace{21}}, 26, 29}_{\text{نیمه دوم}}$$

مثال۹: با توجه به داده‌های نمودار ساقه وبرگ مقدار $Q_3 - Q_1$کدام است؟

۱)۱۷ ۲)۲۳ ۳)۲۷ √ ۴)۱۰



حل:گزینه ۳ صحیح است چون چارک اول ۲۳ و چارک سوم ۵۰ بنابراین ۲۷ = ۲۳ − ۵۰

تقریباً ۲۵ درصد $\left(\frac{۱}{۴}\right)$ داده ها کوچکتر و ۷۵ درصد$\left(\frac{۳}{۴}\right)$ داده ها بزرگتر از چارک اول اند.

تقریباً ۵۰ درصد$\left(\frac{۱}{۲}\right)$ داده ها کوچکتر و ۵۰ درصد$\left(\frac{۱}{۲}\right)$ داده ها بزرگتر از چارک دوم (میانه) اند.

تقریباً ۷۵ درصد $\left(\frac{۳}{۴}\right)$ داده ها کوچکتر و ۲۵ درصد$\left(\frac{۱}{۴}\right)$ داده ها بزرگتر از چارک سوم اند.

مثال۱۰: تقریباً چند درصد داده ها از چارک سوم کمتراند؟

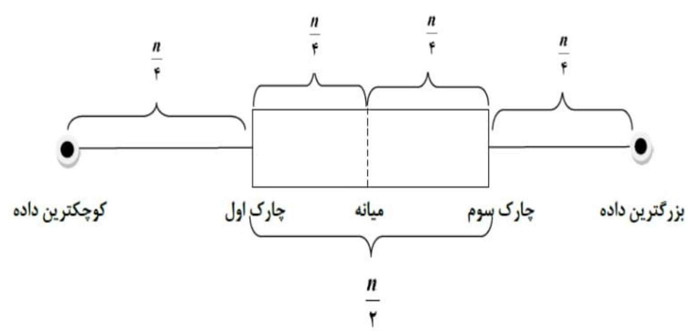
۱)۲۵ ۲)۵۰ ۳) ۷۵√ ۴)۱۰۰

مثال۱۱: تقریباً ۲۵ درصد داده ها از کدام مقدار بزرگتر است؟

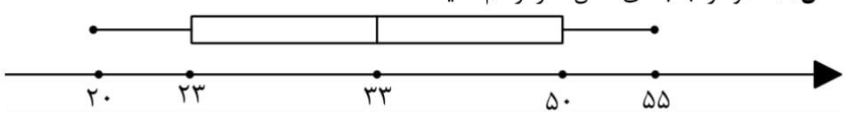
۱)چارک اول ۲)چارک دوم ۳) چارک سوم √ ۴) مد

۱-۳-۶ نمودار جعبه ای

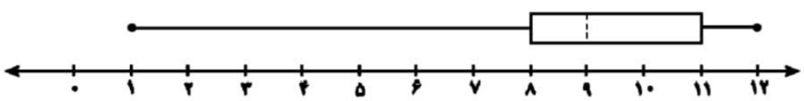
نمودار اکتشافی مکان داده ها که بر اساس پنج مقدار ۱-کوچکترین داده ۲- چارک اول ۳- میانه (چارک دوم) ۴- چارک سوم ۵- بزرگترین داده شامل سه قسمت، جعبه، دنباله راست ودنباله چپ به صورت زیر رسم می‌شود.



مثال۱۲: نمودار جعبه ای مثال ۹ را رسم کنید.



مثال۱۳ : در نمودار زیر بلند تر بودن دنباله چپ نسبت به دنباله راست نشان دهنده چیست؟



۱)تعداد بیشتر داده‌های ۲) پراکندگی بیشتر داده ها √

۳) تمرکز بیشتر داده ها ۴) درصد بیشتر داده ها

۶-۴ میانگین

یکی دیگر از شاخص‌های مرکزی داده ها میانگین است. در این بخش سه نوع میانگین ۱)ساده ۲)وزندار ۳)از روی داده‌های دسته بندی شده را مورد بررسی قرار می‌دهیم.

۶-۴-۱ میانگین ساده

نسبت مجموع داده ها به تعداد داده ها را میانگین ساده گوییم. n داده $x_1 + x_2 + \cdots + x_n$ را در نظر بگیرید، میانگین را با $\bar{x}$ نشان داده و به صورت $\bar{X} = \frac{x_1+x_2+\cdots+x_n}{n} = \frac{\sum_{i=1}^{n} x_i}{n}$ تعریف می‌کنیم که در آن نماد سیگما $\sum$ به معنی مجموع است.

مثال۱۴: میانگین داده‌های ۱۲،۹،۷،۳، a برابر ۸ است. مقدار a کدام است؟

۱)۸ ۲)۹ √ ۳)۱۰ ۴)۱۱

حل:

$$\bar{x} = \frac{3+7+9+12+a}{5} \Rightarrow 8 = \frac{31+a}{5} \Rightarrow 31 + a = 40 \Rightarrow a = 9$$

۶-۴-۲ مجموع داده ها

مجموع داده ها نیز یکی از مقادیری است که می‌توان عملکرد گروه ها را با آن مقایسه کرد. با داشتن میانگین می‌توان مجموع داده ها را به صورت $\sum x_i = n\bar{x}$ به دست آورد.

مثال۱۵: اگر میانگین داده‌های آماری ۱۵،۲۰ ,x_9 ,... ,x_1 برابر۲۳ باشد، میانگین داده‌های ۶،۱۸ , x_9 ,... ,x_1 کدام است؟

۱)۲۲ √ ۲)۲۴ ۳)۲۵ ۴)۱۲

حل:

$$۲۳ = \frac{x_۱+\cdots+x_۹+۱۵+۲۰}{۱۱} \to \sum x_i + ۳۵ = ۲۵۳ \to \sum_{i=۱}^{۹} x_i = ۲۱۸$$

$$\bar{x} = \frac{x_۱+\cdots+x_۹+۶+۱۸}{۱۱} = \frac{۲۱۸+۲۴}{۱۱} = ۲۲$$

مثال۱۶: میانگین ۱۰ داده آماری برابر ۱۳/۲ است اگر داده‌های ۵،۷ را از بین آنها حذف کنیم، میانگین داده‌های جدید کدام است؟

۱) ۱۴/۵ ۲) ۱۴/۸ ۳) ۱۵ √ ۴) ۱۵/۲

حل:

$$\bar{x} = \frac{x_۱+\cdots+x_{۱۰}}{۱۰} \to ۱۳.۲ = \frac{x_۱+\cdots+x_{۱۰}}{۱۰} \to x_۱+\cdots+x_{۱۰} = ۱۳۲\ ,\quad \begin{cases} x_۹ = ۵ \\ x_{۱۰} = ۷ \end{cases}$$

$$\Rightarrow x_۱+\cdots+x_۸ = ۱۳۲-(۵+۷) = ۱۲۰ \qquad \text{جدید}\ \bar{x} = \frac{۱۲۰}{۸} = ۱۵$$

اگر میانگین داده های $x_۱, \dots, x_n$ برابر $\bar{x}$ باشد و تعدادی داده به آنها اضافه یا تعدادی از داده ها را حذف کنیم که میانگین داده‌های اضافه یا حذف شده نیز $\bar{x}$ باشد در این صورت میانگین داده‌های جدید نیز $\bar{x}$ خواهد بود.

نکته: میانگین در جامعه آماری منحصربفرد است همواره در داده‌های آمار داریم:

$$x_{Min} \le \bar{x} \le x_{Max}$$

مثال۱۷: میانگین ۸ داده برابر ۱۶ است. اگر دو داده ۱۵ و ۱۷ را به آنها اضافه کنیم میانگین داده‌های جدید کدام است؟

۱) ۱۰ ۲) ۱۵ ۳) ۱۶ √ ۴) ۱۷

حل: چون میانگین دو داده ۱۵ و۱۷ برابر ۱۶ است بنابراین میانگین داده‌های جدید ۱۶ است.

اگر داده‌های آماری تشکیل دنباله‌ی عددی (با هر قدرنسبتی) دهند برای بدست آوردن میانگین داده ها به صورت زیر عمل می‌کنیم:

$$\bar{x} = \frac{\text{جمله اول}+\text{جمله آخر}}{۲}$$

مثال۱۸: اگر میانگین داده‌های آماری $x_۱, \dots, x_n$ برابر $\bar{x}$ باشد، میانگین داده‌های آماری به‌صورت $۳x_۱+۱, ۳x_۲+۳, \dots ۳x_n+۲n-۱$ کدام است؟

۱) $۳\bar{x}$ ۲) $۳\bar{x}+n$ √ ۳) $۳\bar{x}+n+۱$ ۴) $۳\bar{x}+n-۱$

حل:

$$\frac{(۳x_۱+۳x_۲+\cdots+۳x_n)}{n} + \frac{۱+۳+\cdots+(۲n-۱)}{n} = ۳\bar{x} + \frac{۱+(۲n-۱)}{۲} = ۳\bar{x}+n$$

مثال۱۹: اگر میانگین داده‌های آماری $x_{۱۰}, \dots, x_{۱}$ برابر ۱۲ باشد، میانگین داده‌های آماری بصورت $x_{۱} + ۳, x_{۲} + ۶, \dots, x_{۱۰} + ۳۰$ کدام است؟

۱) ۲۸/۵ √ ۲) ۴۷ ۳) ۲۳/۵ ۴) ۲۱

حل:

$$\frac{(x_{۱} + x_{۲} + \cdots + x_{n})}{n} + \frac{۳ + ۶ + \cdots + ۳۰}{n} = \bar{x} + \frac{۳ + ۳۰}{۲} = ۱۲ + ۱۶/۵ = ۲۸/۵$$

مثال۲۰: اگر داده ها ۷ عدد متوالی صحیح باشند، تفاضل میانگین از میانه کدام است.

۱) ۱ ۲) صفر √ ۳) ۱- ۴) معلوم نیست.

حل: فرض می‌کنیم اولین عدد a است بنابراین داده ها به صورت زیر هستند:

$$a, a+۱, a+۲, a+۳, a+۴, a+۵, a+۶$$

با توجه به داده ها میانه برابر است با a+۳ و چون داده ها تشکیل دنباله عددی می‌دهند داریم :

$$\bar{x} = \frac{a + a + ۶}{۲} = \frac{۲a + ۶}{۲} = a + ۳$$

پس تفاضل میانگین از میانه برابر صفر است.

اگر $\bar{x}$ میانگین داده‌های $x_{۱}, \dots, x_{n}$ باشد آن گاه مجموع اختلافات داده ها از میانگین برابر است با صفر.

$$\sum_{i=۱}^{n}(x_{i} - \bar{x}) = (x_{۱} - \bar{x}) + (x_{۲} - \bar{x}) + \cdots + (x_{n} - \bar{x}) = ۰$$

این ویژگی در حالتی که توزیع داده ها متقارن است برای میانه و مد نیز برقرار است. اما در حالتی که توزیع داده ها نامتقارن است، فقط برای میانگین برقرار است.

مثال۲۱: اگر $\bar{x}$، میانگین داده های $x_{۱}, \dots, x_{n}$ باشد، مقدار $\sum(x_{i} - \bar{x})$ کدام است؟

۱) $n \times \bar{x}$ ۲) n ۳) صفر √ ۴) یک

اگر $\bar{x}$ ، میانگین داده‌های $x_{۱}, \dots, x_{n}$ باشد آن گاه میانگین داده‌های

$$ax_{۱} + b, \dots, ax_{n} + b$$

که با نماد $\overline{ax + b}$ نشان می‌دهیم عبارت است از : $\overline{ax + b} = a\bar{x} + b$

هر بلایی سر داده ها بیاد سر میانگین هم میاد.

مثال۲۲: میانگین داده‌های $x_{۱}, \dots, x_{n}$ برابر ۸ است میانگین داده های $۲x_{۱} - ۳, ۲x_{۲} - ۳, \dots, ۲x_{۳} - ۳$ کدام است؟

۱) ۸ ۲) ۱۶ ۳) ۵ ۴) ۱۳ √

حل:گزینه ۴ درست است چون: $\overline{2x-3}=2\bar{x}-3=2\times 8-3=13$

مثال۲۳: اگر میانگین داده‌های آماری $x_1, \dots, x_n$ برابر $\bar{x}$ باشد, میانگین داده‌های آماری به‌صورت $\frac{x_1-3}{3}, \frac{x_2-3}{3}, \dots, \frac{x_n-3}{3}$ کدام است؟

√ ۱) $\frac{1}{3}\bar{x}-1$ ۲) $\bar{x}$ ۳) $3\bar{x}$ ۴) صفر

حل: $\overline{\left(\frac{x-3}{3}\right)}=\overline{\frac{1}{3}x-1}=\frac{1}{3}\bar{x}-1$

مثال۲۴: تمام داده‌های نمودار ساقه و برگ زیر را ۳ برابر کرده، سپس ۴۰ واحد از آنها کم می‌کنیم. میانگین داده‌های جدید کدام است؟(سراسری ریاضی،۹۲)

۱) ۲۴۰ √ ۲) ۲۴۵ ۳) ۲۵۰ ۴) ۲۵۵

ساقه	برگ
۸	۰ ۱ ۵
۹	۲ ۴ ۶ ۷
۱۰	۰ ۰ ۳ ۴ ۸

حل:

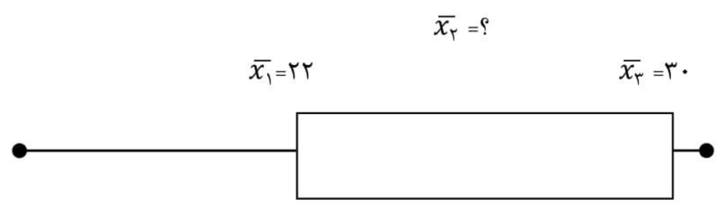
$$\bar{x}=\frac{80+81+85+92+94+96+97+100+100+103+104+108}{12}=\frac{1140}{12}=95$$

$$\overline{3x-40}=3\bar{x}-40=3\times 95-40=245$$

در نمودار جعبه ای اگر میانگین سه قسمت نمودار (دنباله چپ، جعبه و دنباله راست) به ترتیب $\overline{x_1}$, $\overline{x_2}$, $\overline{x_3}$ باشند میانگین کل داده ها برابراست با: $\bar{x}=\frac{\overline{x_1}+2\overline{x_2}+\overline{x_3}}{4}$

مثال۲۵: در نمودار جعبه‌ای ۳۶ داده آماری، میانگین داده‌ها دو طرف جعبه بطور جداگانه به ترتیب ۲۲ و ۳۰ می‌باشد. اگر میانگین تمام داده‌ها ۲۷/۵ باشد آنگاه میانگین داده‌های داخل جعبه کدام است؟(سراسری ریاضی،۹۰).

۱) ۲۸ ۲) ۲۸/۵ √ ۳) ۲۹ ۴) ۲۹/۵



حل:

$$27.5=\frac{22+2\overline{x_2}+30}{4}\to 52+2\overline{x_2}=110\to \overline{x_2}=\frac{110-52}{2}=29$$

مثال۲۶: در داده‌های آماری با نمودار ساقه و برگ، داده‌های کمتر ازچارک اول و بیشتر از چارک سوم را حذف می‌کنیم میانگین داده‌های باقی مانده کدام است؟(سراسری ریاضی، ۸۷)

√ ۱) ۴۲/۷ ۲) ۴۲/۹ ۳) ۴۳/۲ ۴) ۴۳/۴

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

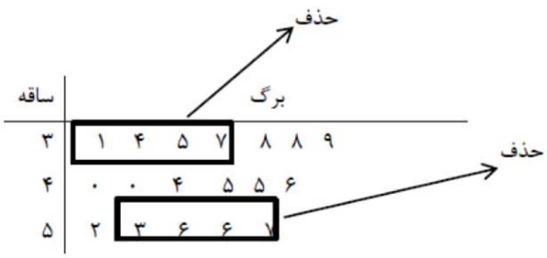
حل:

گام ۱: تعداد داده‌ها را می‌شماریم که ۱۸ تا است پس میانه، نصف مجموع داده‌های نهم و دهم است یعنی : $۴۲=\frac{۴۰+۴۴}{۲}$

گام ۲: تعداد داده‌های قبل از میانه را می‌شماریم که ۹ تا است پس چارک اول داده ی پنجم است یعنی :$Q_۱$ =۳۸

گام ۳: تعداد داده‌های بعد از میانه را می‌شماریم که ۹ تا است پس چارک سوم داده‌ی پنجم بعد از میانه است. یعنی: $Q_۳$ =۵۲

گام ۴: داده‌های قبل ۳۸ وبعد ۵۲ را حذف می‌کنیم.



گام۵: حالا میانگین ده داده باقیمانده را حساب می‌کنیم.

$$\bar{x}=\frac{۳۸+۳۸+۳۹+۴۰+۴۰+۴۴+۴۵+۴۵+۴۶+۵۲}{۱۰}=\frac{۴۲۷}{۱۰}=۴۲.۷$$

مثال۲۷: در داده‌های ۲۵, ۲۰, ۲۱, ۲۶, ۱۲, ۱۴, ۱۵, ۲۴, ۲۰, ۱۶, ۱۴, ۱۸ میانگین داده‌های بزرگتر از چارک اول و کوچکترین از چارک سوم» کدام است؟

۱) ۱۸/۵ √ ۲) ۱۸/۳۳ ۳) ۱۸/۶۶ ۴) ۱۸/۷۵

حل:

$$12 ، 14 ، \boxed{14 ، 15} ، 16 ، \boxed{18 ، 20} ، 20 ، \boxed{21 ، 24} ، 25 ، 26$$

$$Q_1 = \frac{14+15}{2} = 14/5 \qquad Q_2 = \frac{18+20}{2} = 19 \qquad Q_3 = \frac{21+24}{2} = 22/5$$

داده‌های بزرگتر از چارک اول و کوچکتر از چارک سوم: ۱۵,۱۶,۱۸,۲۰,۲۰,۲۱

$$\bar{x} = \frac{15+16+18+20+20+21}{6} = \frac{110}{6} = 18/33$$

مثال۲۸: داده‌های آماری به صورت ساقه وبرگ نشان داده شده اند. در نمودار جعبه ای، تفاضل میانه از میانگین داده‌های داخل جعبه کدام است؟(سراسری۹۴ تجربی)

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

۱)صفر ✓ ۲)۰/۵ ۳)۱ ۴)۱/۵

حل: تعداد داده برابر تعداد برگ ها یعنی ۲۵ است.پس میانه داده ی سیزدهم یعنی ۶۲ است. تعداد داده‌های کوچکتر از میانه ۱۲ است. پس چارک اول نصف مجموع داده‌های ششم و هفتم است. یعنی$55 = \frac{54+56}{2} = Q_1$. همچنین تعداد داده‌های بزرگتر از میانه ۱۲ است. پس چارک سوم نصف مجموع داده‌های ششم و هفتم بعد از میانه است. یعنی$71 = \frac{71+71}{2} = Q_3$.

ساقه	برگ
$\boxed{5}$	۰ ۱ ۱ ۲ ۴ $\boxed{4\ 6}$ ۷ ۹ ۹
$\boxed{6}$	۰ ۰ $\boxed{2}$ ۳ ۳ ۵ ۵ ۶
$\boxed{7}$	$\boxed{1\ 1}$ ۲ ۲ ۴ ۷ ۸

داده‌های داخل جعبه :

۵۶ − ۵۷ − ۵۹ − ۵۹ − ۶۰ − ۶۰ − ۶۲ − ۶۳ − ۶۳ − ۶۵ − ۶۵ − ۶۶ − ۷۱

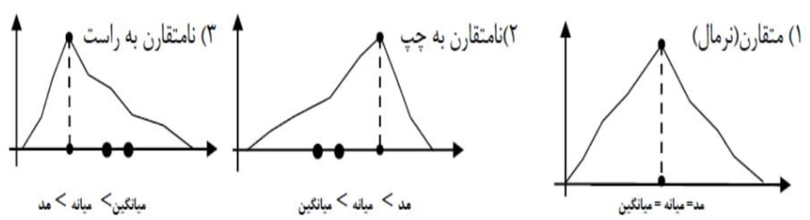
$$\bar{x} = \frac{56 + 57 + 2(59) + 2(60) + 62 + 2(63) + 2(65) + 66 + 71}{13} = 62$$

بنابراین تفاضل میانه از میانگین داده‌های داخل جعبه صفر است.

۵-۶ مقایسه شاخص‌های مرکزی(میانگین، میانه و مد)

یکی از مهم ترین کاربردهای شاخص‌های مرکزی مقایسه آنها در تعیین وضعیت توزیع داده است.

در مقایسه میانگین، میانه ومد سه حالت برای توزیع داده ها به صورت زیر داریم:



تعداد داده ها در دو طرف مد برابرند	تعداد داده های کوچکتر از مد بیشتر از تعداد داده ها بزرگتر از مد است	تعداد داده های بزرگتر از مد بیشتر از تعداد داده ها کوچکتر از مد است

مثال۲۹: در کدام حالت توزیع داده ها نامتقارن به راست است؟

۱) میانه > مد > میانگین ۲) مد > میانه > میانگین

۳√) میانگین> میانه > مد ۴) مد > میانگین > میانه

در دو حالت توزیع نامتقارن داده ها، میانگین به خوبی نمایانگر تمرکز داده ها نیست و بهتر است از میانه برای وضعیت داده ها استفاده کرد.

مثال۳۰ : در یک امتحان ریاضی ۱۵ نفر نمرات زیر را کسب نمودند:

۰,۱,۳,۱۴,۱۴,۱۵,۱۶,۱۶,۱۷,۱۷,۱۸,۱۸,۱۸,۱۹,۲۰

کدام شاخص نمایانگر بهتری از وضعیت نمرات دانش آموزان است؟

۱)مد ۲)میانه √ ۳)میانگین ۴)چارک

حل: با توجه به داده ها مد برابر است با ۱۸، میانه ۱۶ و میانگین،

$$\bar{x} = \frac{0+1+3+2(14)+15+2(16)+2(17)+3(18)+19+20}{15} = \frac{206}{15} = 13.73$$

بنابراین، مد > میانه > میانگین پس توزیع داده ها نامتقارن به راست و میانه نمایانگر بهتری از وضعیت نمرات است.

مثال۳۱: اگر مد داده ها کوچکتر از میانه باشد، تعداد داده های............... از مد بیشتر از تعداد داده های................. از مد است.

۱)کوچکتر –بزرگتر √ ۲)بزرگتر – کوچکتر ۳)کوچکتر – کوچکتر ۴)بزرگتر - بزرگتر

مثال۳۲: در داده های: x , ۶۰ , ۸۰ , ۱۱۰ , ۷۰ مد، میانه و میانگین برابرند، x کدام است؟

۱) ۶۰ ۲) ۷۰ √ ۳) ۸۰ ۴) ۱۱۰

حل: چون مد داریم پس x یکی از داده ها و مد است. از طرفی مد معمولاً کوچکترین یا بزرگترین داده نیست(شاخص مرکزی است). بنابراین x یکی از دو مقدار ۷۰ یا ۸۰ است.اگر ۷۰ باشد مد ومیانه برابر ۷۰ هستند ولی میانگین برابراست با

$$\bar{x} = \frac{۶۰ + ۷۰ + ۷۰ + ۸۰ + ۱۱۰}{۵} = \frac{۳۹۰}{۵} = ۷۸$$

بنابراین x برابر است با ۸۰ یعنی گزینه۳.

۶-۶ میانگین وزندار:

میانگین داده هایی که دارای وزن (ارزش)های متفاوتی هستند، با داده‌های بدون وزن فرق می‌کند. در محاسبه میانگین بایستی این وزن ها را دخالت داد. بدین ترتیب میانگین چنین داده هایی را میانگین وزندار گویند.

اگر داده های $x_1, \dots, x_n$ به ترتیب دارای وزن‌های $w_1, \dots, w_n$ باشند میانگین وزندار به صورت $\overline{x_w} = \frac{w_1x_1+w_2x_2+\cdots+w_nx_n}{w_1+w_2+\cdots+w_n} = \frac{\sum w_ix_i}{\sum w_i}$ است.

مثال۳۳: نمرات ۴ درس همراه با واحد آن ها به صورت زیر است. اگر میانگین نمرات ۱۲ باشد x کدام است؟

نمره درس	۱۲	۱۸	x	۱۵
واحد درس	۴	۱	۳	۲

۱) ۸ √ ۲) ۱۰ ۳) ۱۲ ۴) ۱۳

حل:

اگر نمره درس را x_i و واحد درس w_i در نظر بگیریم داریم:

$$۱۲ = \frac{۴ \times ۱۲ + ۱ \times ۱۸ + ۳ \times x + ۲ \times ۱۵}{۴ + ۱ + ۳ + ۲} = \frac{۹۶ + ۳x}{۱۰}$$

$$\rightarrow ۳x = ۱۲۰ - ۹۶ = ۲۴ \rightarrow x = \frac{۲۴}{۳} = ۸$$

مثال۳۴: نمره‌ی کل آزمون عمومی یک داوطلب مطابق جدول زیر ۵۸ درصد است نمره آزمون انگلیسی او چند درصد است؟(سراسری تجربی۸۱)

زبان انگلیسی	معارف	عربی	ادبیات	درس
؟	۷۰	۵۲	۶۵	درصد
۲	۳	۲	۴	ضریب

۱) ۳۱   ۲) ۳۲ √   ۳) ۳۳   ۴) ۳۴

حل:

$$\overline{x_w} = \frac{۶۵\times۴+۵۲\times۲+۷۰\times۳+x\times۲}{۴+۲+۳+۲} \to ۵۸ = \frac{۵۷۲+۲x}{۱۱} \to x = ۳۲$$

مثال۳۵: در یک روز از ۲۲ نفر مشتری محصول یک کارخانه تولید مواد غذایی، ۵ نفر محصول را در بسته بندی ۱ تایی، ۵ نفر در بسته بندی ۲تایی، ۳ نفر در بسته بندی ۳تایی، ۶ نفر در بسته بندی ۴تایی و ۳ نفر در بسته بندی ۶ تایی خریدند. میانگین تعداد محصول در هر بسته خریداری شده چقدر است؟

حل: اگر تعداد محصول در هر بسته بندی را x_i و تعداد مشتریان هر محصول را w_i در نظر بگیریم به عبارتی

x_i	۱	۲	۳	۴	۶
w_i	۵	۵	۳	۶	۳

$$\overline{x_w} = \frac{۱\times۵+۲\times۵+۳\times۳+۴\times۶+۶\times۳}{۲۲} = \frac{۶۶}{۲۲} = ۳$$

۷ شاخص‌های پراکندگی

تفاوت بین داده ها از لحاظ مقداری یکی دیگر از ویژگی‌های داده ها برای توصیف آن ها و مقایسه چند دسته داده است. در این فصل مقادیری برای داده ها محاسبه می‌کنیم که این ویژگی را در خود خلاصه می‌کنند.
مقداری که تفاوت داده ها را نسبت به هم و یا نسبت به مرکز نشان می‌دهد شاخص پراکندگی گوییم.

شاخص های پراکندگی پنج نوع اند: دامنه تغییرات، انحراف از میانگین، واریانس، انحراف از معیار و ضریب تغییرات.

۱-۷دامنه تغییرات (R)

حداکثر اختلاف داده ها را که برابر است با تفاضل بزرگترین داده از کوچکترین داده، دامنه تغییرات داده ها گویند و به صورت زیر نشان می‌دهیم:

$$R = X_{Max} - X_{Min}$$

مثال۱: با مقایسه دامنه تغییرات داده‌های زیر پراکندگی داده ها را بررسی کنید.

الف) ۲۰، ۱۵، ۱۵، ۱۵، ۱۴، ۱۴، ۱۴، ۱۴، ۱۳/۵، ۱۳، ۱۲/۵، ۱۲، ۱۲، ۱۲، ۴، ۲

ب) ۱۹، ۱۷، ۱۶، ۱۴، ۱۲، ۱۰، ۷، ۵، ۱

حل: الف) R = ۲۰ − ۲ = ۱۸   ب) R = ۱۹ − ۱ = ۱۸

دامنه دو دسته داده برابر است ولی پراکندگی داده ها یکی نیست. داده‌های دسته الف به جز داده‌های تفاوتی یک یا نیم واحدی دارند ولی داده‌های دسته ب تفاوتی بیش از دو واحد دارند.
به داده‌های ۲۰ ,۴,۲ در دسته داده‌های الف، داده‌های دور افتاده گویند.

نکته: *یکی از معایب دامنه تغییرات این است که وقتی داده دور افتاده داریم به خوبی پراکندگی را نشان نمی دهد. در این حالت بهتر است برای محاسبه دامنه بجای بزرگترین داده و کوچکترین داده به ترتیب از چارک سوم و اول استفاده کنیم و به آن دامنه چارکی گوییم.*

در مثال ۱ چارک اول و سوم به ترتیب ۱۲ و ۱۵ می‌باشد. بنابراین دامنه چارکی برابر است با ۳ = ۱۲ − ۱۵ که به خوبی پراکندگی داده‌های دسته الف را نشان می‌دهد.

اگر دامنه‌ی تغییرات داده‌های آماری $x_1, \dots, x_n$ برابر(R) باشد، آنگاه دامنه تغییرات داده‌های آماری $ax_1 \pm b, ax_2 \pm b, \dots, ax_n \pm b$ برابر($|a|R$) خواهد بود.
به عبارتی با اضافه کردن یک مقدار غیر صفر به داده ها دامنه تغییرات ثابت می‌ماند ولی با ضرب یک مقدار مخالف صفر در داده ها، دامنه تغییرات در مقدار مثبت آن مقدار ضرب می‌شود.

مثال۲: دامنه ی تغییرات داده‌های آماری ۵, ۵, ۵, ۵, ۵ کدام است؟ $R = ۵ - ۵ = ۰$

اگر شاخص پراکندگی داده‌های $x_1, \dots, x_n$ برابر ۰ باشد، آنگاه $x_1 = x_2 = \cdots = x_n$ و برعکس.

مثال۳: اگر دامنه‌ی تغییرات داده‌های آماری $x_1, x_2, \dots, x_n$ برابر صفر باشد، میانگین داده‌های آماری بصورت $2x_1+1, 2x_2+1, \dots, 2x_n+1$ کدام است؟

۱) $\frac{x_1+x_n}{2}$ ۲) $2(x_1, x_2, \dots, x_n)+1$ ۳) x_n ۴) $2x_1+1$ √

حل: چون دامنه‌ی تغییرات صفر است داده‌ها با هم برابرند

$$x_1 = x_2 = \cdots = x_n \to 2x_n + 1 = 2x_n + 1 = \cdots = 2x_n + 1$$
$$\overline{2x + 1} = 2\bar{x} + 1 = 2x_1 + 1$$

مثال۴: نمرات آزمون مهارت فنی دو کارگر A و B به صورت زیر است:

A: ۱۵ ،۱۴ ،۱۵ ،۱۶ ،۱۷ ،۱۹

B : ۱۶،۱۴ ،۱۷ ،۱۴ ،۱۷ ،۱۸

دقت عملکرد کدام بیشتر است؟(ریاضی۹۳)

۱) A ۲) B √ ۳) یکسان ۴) غیر پیش بینی

حل: دامنه تغییرات نمرات دو نفر را مقایسه می‌کنیم، هر کدام کوچکتر است پراکندگی کمتر ودقت عملکرد بیشتر است.

$$\left.\begin{aligned} R_A &= 19 - 14 = 5 \\ R_B &= 18 - 14 = 4 \end{aligned}\right\} \Rightarrow R_B < R_A$$

مثال۵: در یک دسته بندی داده‌ها، مرکز سه دسته متوالی بترتیب ۶/۸, ۷/۲, ۷/۶ و تعداد دسته‌ها برابر ۱۲ است، دامنه تغییرات کدام است؟

۱) ۵/۴ ۲) ۴/۸ √ ۳) ۷/۲ ۴) ۴/۶

حل: $n = 12 \to R = kc = 12 \times 0/4 = 4/8$

مثال۶: دامنه تغییرات داده‌های $3, -a^2, 4 + a^2, 1, 2$ سه برابر میانه است، میانه داده‌های $a^2, 2a^4, 3a^6, 4a^8$ کدام است؟

۱) ۲/۵ √ ۲) ۵ ۳) ۱۰ ۴) ۱۲/۵

حل: $-a^۲, ۱,۲,۳, a^۲+۴ \rightarrow$ میانه $M=۲$

$R = a^۲ + ۴ - (-a^۲) = ۲a^۲ + ۴ \rightarrow ۲a^۲ + ۴ = ۳ \times ۲ \rightarrow a^۲ = ۱ \rightarrow a = \pm ۱$

$(\pm۱)^۲, ۲(\pm۱)^۴, ۳(\pm۱)^۶, ۴(\pm۱)^۸ =$ داده جدیدهای

داده‌های جدید : ۱, ۲, ۳, ۴ $\rightarrow$ میانه $=$ ۲/۵

مثال۷: در نمودار جعبه‌ای داده‌های آماری ۱۴, ۱۶, ۹, ۱۱, ۱۸, ۲۰, ۱۹, ۱۵, ۱۳, ۸, ۷ دامنه تغییرات اعداد داخل جعبه کدام است؟

۱) ۵ √ ۲) ۸ ۳) ۱۰ ۴) ۱۱

حل: داده‌های مرتب شده:

۷ ، ۸ ، ۹ ، ۱۱ ، ۱۳ ، ۱۴ ، ۱۵ ، ۱۶ ، ۱۸ ، ۱۹ ، ۲۰

چارک اول میانه چارک سوم

داده‌های داخل جعبه: ۱۶، ۱۵، ۱۴، ۱۳، ۱۱

۱۶ - ۱۱ = ۵ :دامنه

انحراف از میانگین:

تفاضل هر داده از میانگین را انحراف از میانگین آن داده گویند.

اگر $x_۱, x_۲, \ldots, x_n$ داده‌ها باشند آنگاه انحراف از میانگین داده ها به صورت زیر است:

$$(x_۱ - \bar{x}), (x_۲ - \bar{x}), \ldots, (x_n - \bar{x})$$

مثال۸: انحراف از میانگین داده‌های روبرو را بدست آورید. ۸, ۳, ۷, ۴, ۹, ۵

$$\bar{x} = \frac{۵+۹+۴+۷+۳+۸}{۶} = \frac{۳۶}{۶} = ۶$$

انحرافات از میانگین : ۶-۵, ۶-۹, ۶-۴, ۶-۷, ۶-۳, ۶-۸

انحراف از میانگین : ۱-, ۳, ۲-, ۱+, ۳-, ۲

مجموع انحراف از میانگین داده‌ها همیشه صفر است.

در مثال قبل داریم: $-۱ + ۳ - ۲ + ۱ - ۳ + ۲ = ۰$

۷-۲واریانس(پراش)

میانگین مجذور(مربع) انحراف از میانگین داده ها را واریانس گویند و با $\sigma^۲$(مجذور سیگما) نشان می‌دهند. به عبارتی:

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

مثال۹: واریانس داده‌های ۴, ۴, ۵, ۵, ۵, ۳, ۲ کدام است؟

۱) $\frac{۵}{۷}$ ۲) $\frac{۶}{۷}$ ۳) $\frac{۸}{۷}$ √ ۴) $\frac{۹}{۷}$

حل:

$$\bar{x} = \frac{۲+۳+۵+۵+۵+۴+۴}{۷} = \frac{۲۸}{۷} = ۴$$

$$\sigma^۲ = \frac{(۲-۴)^۲+(۳-۴)^۲+(۵-۴)^۲+(۵-۴)^۲+(۵-۴)^۲+(۴-۴)^۲+(۴-۴)^۲}{۷}$$

$$\to \frac{۴+۱+۱+۱+۱}{۷} = \frac{۸}{۷}$$

مثال۱۰: در ۵ داده‌ی آماری انحراف از میانگین داده‌ها برابر a , ۲, ۰, ۱-, ۳- می‌باشد. واریانس این داده‌ها کدام است؟

۱) ۲/۱ ۲) ۲/۴ ۳) ۳/۲ ۴) ۳/۶ √

حل: مجموع انحراف از میانگین داده‌ها برابر صفر است.

$$-۳ + (-۱) + ۰ + ۲ + a = ۰ \to a = ۲$$

$$\sigma^۲ = \frac{\sum (x_i - \bar{x})^۲}{n} = \frac{(-۳)^۲ + (-۱)^۲ + ۰^۲ + ۲^۲ + ۲^۲}{۵}$$

$$\to \frac{۹+۱+۴+۴}{۵} = \frac{۱۸}{۵} = ۳/۶$$

دستور دیگری برای محاسبه واریانس:

$$\sigma^۲ = \text{(میانگین مجذورداده‌ها)} - \text{(مجذور میانگین داده‌ها)}$$

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

توجه کنید در $x_۱^۲ + x_۲^۲ + \cdots + x_n^۲$ ابتدا مقادیر x_i ها را به توان ۲ می‌رسانیم و سپس آنها را با هم جمع می‌کنیم و در $\bar{x}^۲$ میانگین داده‌ها را به توان ۲ می‌رسانیم.

نکته: زمای که مجموع مربعات انحراف از میانگین برابر فراوانی کل است واریانس داده ها برابر یک است.

مثال۱۱: واریانس داده‌های ۹, ۷, ۶, ۵, ۱ را بدست آورید؟

$$\bar{x}\ \frac{۲۸}{۵} = \frac{۱+۵+۶+۷+۹}{۵} = =۵/۶$$

$$x_۱^۲ + x_۲^۲ + \cdots + x_۵^۲ = ۱^۲+۵^۲+۶^۲+۷^۲+۹^۲ = ۱۹۲$$

$$\sigma^۲ = \frac{۱۹۲}{۵} - (۵/۶)^۲ = ۳۸/۴ - ۳۱/۳۶ = ۷/۰۴$$

مثال۱۲: مجموع مجذورات ۱۰ داده‌ی آماری برابر ۳۴۰ و میانگین این داده‌ها برابر ۵ است، واریانس این داده‌ها کدام است؟

۱) ۹ √ ۲) ۱۶ ۳) ۴ ۴) ۳

حل: $\sigma^۲ = \frac{۳۴۰}{۱۰} - ۵^۲ = ۹$

***نکته:** مشکل واریانس این است که واحد اندازه گیری واریانس مربع واحد اندازه گیری داده ها است. به عبارتی واریانس از جنس داده ها نیست. به این معنی که اگر داده ها بر حسب متر باشند(از جنس طول) آنگاه واریانس برحسب مترمربع(از جنس مساحت) است.*

۳-۷انحراف معیار

جذر واریانس را انحراف معیار گویند. و با نماد σ (سیگما) نشان می‌دهند.

مثال۱۳: در ۱۰ داده آماری مجموع مجذورات داده‌ها برابر ۸۵ و مجموع داده‌ها برابر ۲۵ است، انحراف معیار کدام است؟

۱) ۲/۵ ۲) ۱/۷۵ ۳) ۱/۲۵ ۴) ۱/۵ √

حل:

$\bar{x} = \frac{۲۵}{۱۰} = ۲/۵$

$\sigma^۲ = \frac{۸۵}{۱۰} - (۲/۵)۲ = ۸/۵ - ۶/۲۵ = ۲/۲۵ \to \sigma = \sqrt{۲/۲۵} = ۱/۵$

تمرین۱: مجموع مربعات ۱۰ داده آماری برابر ۰/۹۶۴ و میانگین آنها ۳ می‌باشد. انحراف معیار این داده‌ها را بدست آورید.

جواب: ۰/ ۸

مثال۱۴: اگر میانگین داده‌های $x_۱, x_۲, \dots, x_۵$ برابر ۵ باشد و واریانس برابریک باشد حاصل $x_۱^۲ + x_۲^۲ + \cdots + x_۵^۲$ کدام است؟

۱) ۱۲۰ ۲) ۱۲۵ ۳) ۱۳۰ √ ۴) ۳۸۰

حل:

$\sigma^۲ = \frac{\sum x_i^۲}{n} - \bar{x}^۲$

$۱ = \frac{\sum x_i^۲}{۵} - ۵^۲ \to ۱ = \frac{\sum x_i^۲}{۵} - ۲۵ \to ۲۶ = \frac{\sum x_i^۲}{۵} \to \sum x_i^۲ = ۱۳۰$

مثال۱۵: میانگین ۱۰ داده آماری مساوی ۲ و انحراف معیار آنها مساوی ۰/۳ است. اگر همه‌ی داده‌ها را به توان ۲ برسانیم میانگین داده‌های جدید را بدست آورید؟

$\bar{x} = ۲, \sigma = ۰/۳$

$\sigma^۲ = \frac{x_۱^۲ + x_۲^۲ + \cdots + x_{۱۰}^۲}{۱۰} - \bar{x}^۲ \to (۰.۳)^۲ = \frac{x_۱^۲ + x_۲^۲ + \cdots + x_{۱۰}^۲}{۱۰} - ۲^۲$

$\to \frac{x_۱^۲ + x_۲^۲ + \dots + x_{۱۰}^۲}{۱۰} = ۴.۰۹$

تمرین۲: الف) واریانس و انحراف معیار داده‌های ۷, ۵, ۳,۱ را بدست آورید.

جواب: واریانس: ۵ , انحراف معیار : $\sqrt{5}$

ب) داده‌ها را با ۵ جمع کنید سپس واریانس و انحراف معیار را بدست آورید.

جواب: واریانس : ۵ , انحراف معیار: $\sqrt{5}$

پ) از داده‌ها ۳ واحد کم کنید سپس واریانس و انحراف معیار را بدست آورید.

جواب: واریانس : ۵ , انحراف معیار : $\sqrt{5}$

نتیجه: *در حالت کلی ثابت می‌شود که اگر داده‌ها را با عددی جمع یا تفریق کنیم واریانس و انحراف معیار تغییری نمی کند:*

$$\sigma^2_{x\pm a} = \sigma^2_x \quad , \quad \sigma_{x\pm a} = \sigma_x$$

ت) داده‌ها را در ۳ ضرب کنید سپس واریانس و انحراف معیار را بدست آورید.

جواب: واریانس : $5\times3^2 = 45$, انحراف معیار : $3\times\sqrt{5}$

ث) داده‌ها را در ۴- ضرب کنید سپس واریانس و انحراف معیار را بدست آورید.

جواب: واریانس : $5\times(-4)^2 = 80$, انحراف معیار : $\sqrt{5}\,|4| = 4\sqrt{5}$

نتیجه: در حالت کلی ثابت می‌شود که اگر داده‌ها را در عددی ضرب کنیم واریانس در مجذور آن عدد و انحراف معیار در قدر مطلق آن عدد ضرب می‌شود:

$$\sigma^2_{ax} = a^2\sigma^2_x \quad , \quad \sigma_{ax} = |a|\,\sigma_x$$

مثال۱۶: انحراف معیار داده‌های $x_1, x_2, ..., x_{10}$ مساوی ۵ است. انحراف معیار داده‌های $-2x_1+5$, $-2x_2+5$, ..., $-2x_{10}+5$ کدام است؟

۱) ۵-　　۲) ۱۰-　　۳) ۱۰ √　　۴) ۵

حل: $\sigma_{جدید} = |-2| \times 5 = 10$

مثال۱۷: انحراف معیار داده‌های $x_1, x_2, ..., x_n$ مساوی ۴ است. واریانس داده‌های $3x_1-2$, $3x_2-2$, ..., $3x_n-2$ کدام است؟

۱) ۱۶　　۲) ۳۶　　۳) ۴۸　　۴) ۱۴۴√

حل: $\sigma = 4 \rightarrow \sigma^2 = 16 \rightarrow \sigma^2_{جدید} = 3^2 \times 16 = 144$

مثال۱۸: داده‌های آماری با میانگین ۸ و واریانس ۲/۲۵ موجود است، تمام داده‌ها را دو برابر می‌کنیم تا داده‌های جدید حاصل شود. انحراف معیار داده‌های جدید کدام است؟

۱) ۳√　　۲) ۳/۵　　۳) ۲/۲۵　　۴) ۱/۵

حل:

روش اول: $\sigma^{۲} = ۲/۲۵$ قدیم $\rightarrow$ $\sigma^{۲} = ۴ \times ۲/۲۵ = ۹$ جدید $\rightarrow$ $\sigma = \sqrt{۹} = ۳$ جدید

روش دوم: $\sigma^{۲} = ۲/۲۵ \rightarrow$ $\sigma = ۱/۵$ قدیم $\Rightarrow$ $\sigma = ۲ \times ۱/۵ = ۳$ جدید

مثال۱۹: اگر انحراف معیار قیمت‌های چند کالا در سال گذشته برابر۳۰ و امسال ٪۱۰ به قیمت‌ها افزوده شده باشد واریانس قیمت‌های امسال آن کالاها کدام است؟

۱) ۱۰۸۹ ✓ ۲) ۹۰۰ ۳) ۳۳ ۴) ۳۰

حل: فرض می‌کنیم قیمت هر کالا x باشد. اضافه شدن ۱۰ درصد قیمت هر کالا به خودش به صورت زیر است:

$$x + ۰.۱x = ۱.۱x$$

به عبارتی قیمت ها در ۱/۱ ضرب می‌شوند. بنابراین داریم:

$\sigma = ۳۰$ قدیم $\rightarrow \sigma$ جدید $= ۱.۱ \times ۳۰ = ۳۳ \rightarrow \sigma^{۲} = ۳۳^{۲} = ۱۰۸۹$

مثال۲۰: انحراف معیار داده‌های x +۱ و x+۵ و x+۳ و x+۳ کدام است؟

۱) $۲\sqrt{۲}$ ۲) $\frac{\sqrt{۲}}{۲}$ ۳) $\sqrt{۲}$ ✓ ۴) ۲

حل: مقدار $x + ۳$را از داده ها کم می‌کنیم تا داده های۲−,۰,۰,۲ به دست آیند. انحراف معیار داده‌های جدید با انحراف معیار داده‌های قبلی برابر است.

$$\bar{x} = \frac{۰ + ۰ + ۲ + (-۲)}{۴} = ۰ \rightarrow \sigma = \sqrt{\frac{۰^{۲} + ۰^{۲} + ۲^{۲} + (-۲)^{۲}}{۴}} = \sqrt{۲}$$

مثال۲۱: در نمودار ساقه وبرگ داده‌های آماری روبرو واریانس داده‌های بین چارک اول و چارک سوم کدام است؟(سراسری انسانی۹۵)

۱) ۱۷/۲۴

۲) ۱۷/۸۲

۳) ۱۸/۰۲

۴) ۱۸/۴۴ ✓

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

حل: داده ها را از روی نمودارمی نویسیم:

$$۳۲,۳۳,۳۴,\boxed{۳۴,۳۶},۳۹,۴۰,۴۱,\boxed{۴۳},۴۵,۴۵,۴۷,\boxed{۵۱,۵۲},۵۴,۵۷,۵۸$$

$Q_{۱} = \frac{۳۴+۳۶}{۲} = ۳۵$ میانه $Q_{۳} = \frac{۵۱+۵۲}{۲} = ۵۱.۵$

داده‌های بین چارک اول وسوم: ۳۶، ۳۹، ۴۰، ۴۱، ۴۳، ۴۵، ۴۵، ۴۷، ۵۱

۴۳ را از داده ها کم می‌کنیم : ‎−۷، −۴، −۳، −۲، ۰، ۲، ۲، ۴، ۸

$$\bar{x} = \frac{-۷ - ۴ - ۳ - ۲ + ۰ + ۲ + ۲ + ۴ + ۸}{۹} = ۰$$

$$\sigma^{۲} = \frac{(-۷)^{۲} + (-۴)^{۲} + (-۳)^{۲} + (-۲)^{۲} + (۰)^{۲} + (۲)^{۲} + (۲)^{۲} + (۴)^{۲} + (۸)^{۲}}{۹} - .۲$$

$$\sigma^{۲} = \frac{۱۶۶}{۹} = ۱۸.۴۴$$

تمرین ۳ : اگر داده‌های آماری ۱۱, ۱۵, ۱۷, ۱۶, ۱۴, ۹, ۱۱, ۱۲, ۱۵, ۱۸, ۱۴, را با نمودار جعبه‌ای نشان دهیم، انحراف معیار داده‌های داخل جعبه کدام است؟ (سراسری تجربی، ۸۸)

۱) ۱/۱ ۲) ۱/۲ ۳) ۱/۲۵ ۴) ۱/۳

تمرین۴: در داده‌های ۲۵, ۱۹, ۱۸, ۱۷, ۲۲, ۲۱, ۱۵, ۱۸, ۲۰, ۲۳, ۲۴ واریانس داده‌های بزرگتر از مُد و کوچکتر از چارک سوم کدام است؟

۱) ۱/۱۱ ۲) ۲/۲۵ ۳) ۲/۵ ۴) ۱/۲۵

نکته: اگر n داده‌ی آماری تشکیل دنباله‌ی عددی با قدر نسبتd بدهند واریانس از فرمول زیر بدست می‌آید:

$$\sigma^{۲} = \frac{n^{۲}-۱}{۱۲} d^{۲}$$

مثال۲۲: انحراف معیار داده‌های ۱۳۸۷, ۱۳۸۶, ۱۳۸۵, ۱۳۸۴, ۱۳۸۳ کدام است؟

۱) $\sqrt{۲}$ ✓ ۲) ۲ ۳) ۱ ۴) ۴

حل: ۵ = n داده تشکیل دنباله عددی با قدر نسبت۱ = d می‌دهند بنابراین داریم:

$$\sigma^{۲} = \frac{۵^{۲}-۱}{۱۲}(۱)^{۲} = ۲ \rightarrow \sigma = \sqrt{۲}$$

تمرین۵: پنج داده‌ی آماری تشکیل تصاعد عددی با قدر نسبت ۶ می‌دهند، واریانس این داده‌ها کدام است؟

۱)۶۴ ۲) ۶۸ ۳) ۷۲ ۴) ۷۸

نکته: اگر واریانس داده‌ها صفر باشد داده‌ها باهم برابرند و بالعکس.

مثال۲۳: اگر میانگین داده‌های d,c,b,a برابر ۵ و واریانس آنها صفر باشد، میانگین داده‌های a+b , c+d, d, c,b,a کدام است؟

۱) ۱۰ ۲) ۴ ۳) ۵ ۴) $\frac{۲۰}{۳}$ ✓

حل: چون در نمونه اول واریانس صفر است پس= $a = b$ ← $\bar{x} = \frac{۳\times۵+۲\times۱۰}{۶} = \frac{۴۰}{۶} = \frac{۲۰}{۳}$

$c = d = ۵$

نکته: *برای محاسبه واریانس با اضافه یا کم کردن چند داده ابتدا مجموع داده ها و مجموع مجذور داده ها را بدست آورده سپس با اضافه و کم کردن داده ها به آن ها بعد ازمحاسبه میانگین، واریانس را محاسبه می‌کنیم. فرمول‌های محاسبه مجموعه ها عبارت اند از:*

$$\sum x_i^{۲} = n(\sigma^{۲} + \bar{x}^{۲}) \ , \ \sum x_i = n\bar{x}$$

مثال۲۴: میانگین و واریانس ۸ داده آماری به ترتیب برابر ۱۵ و ۴ است. اگر داده‌های ۱۲ و ۱۸ به آنها اضافه شود واریانس ۱۰ داده جدید کدام است؟ (سراسری تجربی، ۸۴)

۱) ۴ ۲) ۴/۵ ۳) ۴/۸ ۴) ۵ √

در این سؤال چون میانگین دو داده ۱۲ و ۱۸ برابر با ۱۵ است بنابراین میانگین داده‌های جدید تغییری نکرده و همان ۱۵ است.

حل:

$$\text{قدیم} \sum x_i^{۲} = n(\sigma^{۲} + \bar{x}^{۲}) = ۸(۴ + ۱۵^{۲}) = ۱۸۳۲$$

$$\text{جدید} = \sum x_i^{۲} = ۱۸۳۲ + ۱۲^{۲} + ۱۸^{۲} = ۲۳۰۰ \to \sigma^{۲} = \frac{۲۳۰۰}{۱۰} - ۱۵^{۲} = ۲۳۰ - ۲۲۵ = ۵$$

مثال۲۵: در ۲۵ داده آماری میانگین و انحراف معیار بترتیب ۳۰ و ۸ می‌باشد، اگر داده‌های ناجور ۱۰, ۱۵, ۴۵, ۵۰ از بین آنها حذف شوند، واریانس داده‌های باقیمانده کدام است؟(سراسری تجربی ۹۳)

۱) ۱۴/۷۲ ۲) ۱۴/۸۱ ۳) ۱۵/۳۳ ۴) ۱۶/۶۶ √

حل: چون میانگین داده‌های ناجور برابر ۳۰ همان میانگین قبلی است پس میانگین داده‌های جدید تغییر نمی کند.

$$\text{قدیم} \sum x_i^{۲} = n(\sigma^{۲} + \bar{x}^{۲}) = ۲۵(۶۴ + ۳۰^{۲}) = ۲۴۱۰۰$$

$$\text{جدید} = \sum x_i^{۲} = ۲۴۱۰۰ - (۱۰^{۲} + ۱۵^{۲} + ۴۵^{۲} + ۵۰^{۲}) = ۲۴۱۰۰ - ۴۸۵۰ = ۱۹۲۵۰$$

$$\text{واریانس جدید} = \sigma^{۲} = \frac{۱۹۲۵۰}{۲۱} - ۳۰^{۲} = ۹۱۶.۶۶ - ۹۰۰ = ۱۶.۶۶$$

تمرین۶ : میانگین و واریانس ۸ داده‌ی آماری ۱۵ و ۴ است. اگر داده‌های ۲۲ و ۱۸ به آنها اضافه شود واریانس ۱۰ داده‌ی جدید را به دست آورید.

۴-۷واریانس در جدول فراوانی

برای محاسبه واریانس در جدول فراوانی داده ها از فرمول زیر استفاده می‌کنیم:

$$\sigma^{۲} = \frac{\sum f_i (x_i - \bar{x})^{۲}}{n} = \frac{\sum f_i x_i^{۲}}{n} - \bar{x}^{۲}$$

که در آن؛

f_i : فراوانی داده(دسته) i ام

x_i : داده(مرکز دسته)i ام

میانگین : $\bar{x} = \frac{\sum f_i x_i}{n}$

n : تعداد داده ها (مجموع فراوانی)

مثال۲۶: واریانس جدول زیر کدام است؟

۱) ۴ ۲) ۶ ۳) ۳ ۴) ۲ √

x_i	۲	۴	۶
f_i	۲	۴	۲

حل: ابتدا میانگین را محاسبه می‌کنیم:

$$\bar{x} = \frac{۲\times۲+۴\times۴+۶\times۲}{۲+۴+۲} = ۴$$

$$\sigma^۲ = \frac{۲(۲-۴)^۲+۴(۴-۴)^۲+۲(۶-۴)^۲}{۲+۴+۲} = \frac{۸+۰+۸}{۸} = ۲$$

***نکته**: در محاسبه واریانس از روی جدول دسته بندی داده ها با تعداد دسته‌های فرد معمولاً به جای مرکز دسته ها از انحراف مراکز دسته ها از مرکز دسته وسط استفاده می‌شود.*

مثال۲۷: در جدول زیر، واریانس داده‌ها کدام است؟(سراسری تجربی ۹۰ خارج از کشور)

۱) ۱۱/۷۲ ۲) ۱۱/۹۶ ۳) ۱۲/۲۴ √ ۴) ۱۲/۳۶

مرکز دسته	۱۲	۱۵	۱۸	۲۱	۲۴
فراوانی	۴	۳	۹	۷	۲

حل:

از همه داده‌ها ۱۸ واحد کم می‌کنیم

$x_i - ۱۸$	-۶	-۳	۰	۳	۶
فراوانی	۴	۳	۹	۷	۲

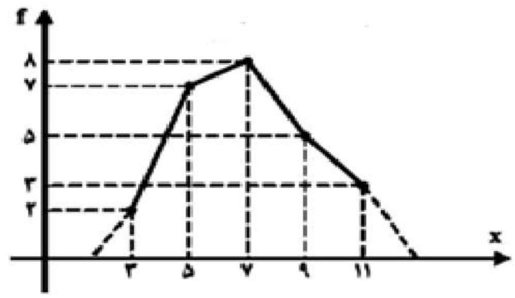
$$\bar{x} = \frac{-۲۴-۹+۰+۲۱+۱۲}{۲۵} = ۰$$

$$\to \sigma^۲ = \frac{۴(-۶-۰)^۲+۳(-۳-۰)^۲+۹(۰-۰)^۲+۷(۳-۰)^۲+۲(۶)^۲}{۲۵}$$

$$\sigma^۲ = \frac{۱۴۴+۲۷+۶۳+۷۲}{۲۵} = \frac{۳۰۶}{۲۵} = ۱۲.۴۴$$

مثال۲۸: با توجه به نمودار چندبر فراوانی زیر، واریانس کل داده ها، کدام است؟(سراسری ریاضی ۹۵)

۱) ۴/۵ ۲) ۴/۸ ۳) ۵/۱۲ √ ۴) ۴/۹۲



حل:

$x_i - ۷$	۴	۲	۰	۲-	۴-
f_i	۳	۵	۸	۷	۲

$$\bar{x} = \frac{۲(-۴)+۷(-۲)+۰+۵(۲)+۳(۴)}{۲+۷+۸+۵+۳} = ۰$$

$$s^۲ = \frac{۲(-۴)^۲+۷(-۲)^۲+۰+۵(۲)^۲+۳(۴)^۲}{۲+۷+۸+۵+۳} = \frac{۳۲+۲۸+۲۰+۴۸}{۲۵} \rightarrow$$

$$\frac{۱۲۸}{۲۵} = ۵ / ۱۲$$

تمرین ۷: واریانس داده‌های جدول زیر کدام است؟

۱) ۲ ۲) ۲/۲۵ ۳) ۲/۵ ۴) ۲/۷۵

حدود دسته	۲-۰	۴-۲	۶-۴	۸-۶
فراوانی	۱	۲	۹	۴

تذکر: اگر حدود دسته‌ها را به ما بدهند ابتدا مرکز دسته‌ها را بدست می‌آوریم.

مثال۲۹: جدول زیر مقادیر انحراف از میانگین ۱۲ داده آماری دسته بندی شده را مشخص می‌کند انحراف معیار داده‌ها کدام است؟

۱) $\frac{۷}{\sqrt{۶}}$ √ ۲) $\frac{\sqrt{۲۶}}{۲}$ ۳) $\frac{۷}{\sqrt{۳}}$ ۴) $\sqrt{۲۶}$

انحراف از میانگین	۳-	۲-	a	۱	۸
فراوانی مطلق	۲	۳	x	۴	۱

حل: چون ۱۲ داده آماری داریم جمع فراوانی‌های مطلق برابر ۱۲ است

$۲+۳+x+۴+۱ = ۱۲ \rightarrow x = ۲$

مجموع انحراف از میانگین داده‌ها همواره صفر است.

$$(-۳)(۲)+(-۲)(۳) + a\ (۲)+(۱)(۴)+۸(۱) = ۰ \rightarrow a = ۰$$

با توجه به جدول مقادیر انحراف از میانگین را داریم.

$$\sigma^{۲} = \frac{۲(-۳)^۲+۳(-۲)^۲+۲(۰)^۲+۴(۱)^۲+۱(۸)^۲}{۱۲} = \frac{۹۸}{۱۲} = \frac{۴۹}{۶}\ \sigma = \sqrt{\frac{۴۹}{۶}} = \frac{۷}{\sqrt{۶}} \rightarrow$$

۵-۷ ضریب تغییرات:(C.V)

آخرین شاخص پراکندگی ضریب تغییرات است. یک از مشکلات استفاده از شاخص‌های پراکندگی بیان شده در مقایسه پراکندگی دو یا چند دسته داده با واحدهای‌های اندازه گیری مختلف است بعنوان مثال برای مقایسه پراکندگی وزن(برحسب کیلوگرم) و قد(برحسب سانتی متر) دانش آموزان کلاس استفاده از شاخص‌های دامنه، واریانس و انحراف معیار مناسب نیست. برای رفع این مشکل از شاخص پراکندگی باید استفاده کرد که مستقل از واحد اندازه گیری داده ها باشد.

نسبت انحراف معیار به میانگین را ضریب تغییرات می‌نامند، که بدون واحد است:

$$CV = \frac{\sigma}{\bar{x}}$$

ضریب تغییرات عبارت است از میزان پراکندگی به ازای یک واحد از میانگین

> **تذکر:** $\overline{X}$ نباید صفر شود و چون میزان پراکندگی مورد نظر است باید نا منفی باشد، از این رو ضریب تغییرات را درباره‌ی داده‌های مثبت مانند قد, وزن, سن و میزان محصول و غیره بکار می‌برند.

مثال ۳۰: میانگین طول اضلاع مربع هایی ۱۵ واحد با ضریب تغییرات ۰/۲ محاسبه شده است. میانگین مساحت این مربع ها کدام است؟(سراسری تجربی۹۵ و ریاضی۹۱)

۱) ۲۲۹ ۲) ۲۳۲ ۳) ۲۳۴ √ ۴) ۲۳۶

حل: طول اضلاع مربع ها را داده ها در نظر می‌گیریم. رابطه میانگین طول اضلاع و میانگین مساحت ها(مربع طول اضلاع) به صورت زیر است:

$$\sigma^۲ = \overline{x^۲} - \bar{x}^۲ \Rightarrow \overline{x^۲} = \sigma^۲ + \bar{x}^۲$$

از طرفی:

$$CV = \frac{\sigma}{\bar{x}} \rightarrow ۰.۲ = \frac{\sigma}{۱۵} \rightarrow \sigma = (۰.۲ \times ۱۵)^۲ = ۳$$

بنابراین:

$$\overline{x^۲} = ۳^۲ + ۱۵^۲ = ۹ + ۲۲۵ = ۲۳۴$$

مثال ۳۱: میانگین اضلاع مربع هایی برابر ۸ و میانگین مساحت آنها ۶۵/۴۴ می‌باشد. ضریب تغییرات در طول اضلاع این مربع ها ,کدام است ؟ (سراسری تجربی ۹۴ خارج از کشور)

۱) ۰/۱۲ ۲) ۰/۱۵ √ ۳) ۰/۲ ۴) ۰/۲۵

حل: طول اضلاع مربع ها را داده ها در نظر می‌گیریم.

$$\sigma^۲ = \overline{x^۲} - \bar{x}^۲ = ۶۵.۴۴ - ۸^۲ = ۱.۴۴ \to \sigma = ۱.۲$$

$$\mathrm{CV} = \frac{\sigma}{\bar{x}} = \frac{۱.۲}{۸} = ۰.۱۵$$

مثال۳۲: اگر ضریب تغییرات داده‌های آماری $x_۱, x_۲, ..., x_n$ برابر ۰/۲ و میانگین این داده‌ها ۱۰ باشد، واریانس داده‌های $-۲x_۱, -۲x_۲, ..., -۲x_n$ کدام است؟

۱) ۴ ۲) ۸ ۳) ۱۶ √ ۴) ۴-

$$\mathrm{CV} = \frac{\sigma}{\bar{x}} \quad \frac{۲}{۱۰} \Longrightarrow \sigma \quad \frac{\sigma}{۱۰}۲ \Rightarrow \sigma^۲ = ۲^۲ = ۴ \quad \text{جدید } \sigma^۲ = (-۲)^۲ \times ۴ = ۱۶$$

مثال۳۳: ضریب تغییرات یک سری از داده‌های آماری نصف انحراف معیار است. اگر به هریک از داده‌های آماری ۳ واحد اضافه کنیم میانگین، چه عددی خواهد شد؟

۱) ۲ ۲) ۲/۵ ۳) ۴ ۴) ۵ √

$$\mathrm{CV} = \frac{\sigma}{\bar{x}} =_{\sigma} \frac{۱}{۲} \to \frac{\sigma}{\bar{x}} \to \bar{X} = ۲ \xrightarrow{\text{به تمام داده ها ۳ واحد می‌اضافه کنیم}} \bar{x} = ۲ + ۳ = ۵$$

> *اگرداده‌های آماری را در عدد مثبتی ضرب شوند، CV ثابت می‌ماند:*
>
> $$\text{جدید } CV = \frac{|a|\sigma}{a\bar{x}} = \frac{a\sigma}{a\bar{x}} = \frac{\sigma}{\bar{x}} = CV \text{ قدیم}$$
>
> *اگر عدد مثبتی به داده‌ها اضافه شود انحراف معیار تغییری نمی کند ولی میانگین افزایش می‌یابد؛ بنابر این CV کاهش می‌یابد و برعکس.*

مثال۳۴: داده‌های ۵، ۴، ۳، ۲، ۱ $= x_i$ مفروض است. ضریب تغییرات داده‌های $u_i = ۱۲x_i + ۶$ کدام است؟(سراسری ریاضی۹۵)

۱) ۰/۴ √ ۲) ۰/۴۸ ۳) ۰/۵۲ ۴) ۰/۶

حل: میانگین داده‌های x_i برابر ۳ است. بنابراین: $\bar{u} = ۱۲\bar{x} + ۶ = ۱۲ \times ۳ + ۶ = ۴۲$. با توجه به داده ها،

$$\sigma_x = \sqrt{\frac{(۱-۳)^۲ + (۲-۳)^۲ + ۰ + (۴-۳)^۲ + (۵-۳)^۲}{۵}} = \sqrt{\frac{۱۰}{۵}} = \sqrt{۲} \cong ۱.۴۱$$

با توجه به نکات بالا،

$$CV_u = \frac{\sigma_u}{\bar{u}} = \frac{۱۲\sigma_x}{۴۲} = \frac{۱۲ \times ۱.۴۱}{۴۲} = ۰.۴$$

مثال۳۵: در ۱۵۰ داده‌ی آماری با میانگین ۱۲, به دو برابر هریک از داده‌ها ۳ واحد اضافه می‌کنیم تا داده‌های جدید حاصل شوند، ضریب تغییرات داده‌های جدید چند برابرضریب تغییرات داده‌های قبلی است؟ (سراسری ریاضی، ۹۲).

۱) $\frac{۷}{۹}$ ۲) $\frac{۵}{۶}$ ۳) $\frac{۷}{۸}$ ۴) $\frac{۸}{۹}$ √

حل:

$$CV_۱ = \frac{\sigma}{\bar{x}} \xrightarrow{\text{به دو برابر هر یک از داده ها ۳ واحد می‌اضافه کنیم.}} CV_۲ = \frac{۲\sigma}{۲\bar{x} + ۳}$$

$$\frac{CV_۲}{CV_۱} = \frac{\frac{۲\sigma}{(۲ \times ۱۲ + ۳)}}{\frac{\sigma}{۱۲}} = \frac{۸}{۹}$$

مثال ۳۶: اگر ٪ ۲۰ نمره‌ی هر دانش آموز به نمره‌ی او اضافه شود، ضریب تغییرات نمره‌ها جدید چگونگی تغییرمی کند ؟

۱)بیشتر می‌شود. ۲)کمتر می‌شود ۳)ثابت می‌ماند √ ۴) چیزی نمی توان گفت

حل:

$$x_۱, x_۲, \dots, x_n \xrightarrow{\text{٪۲۰ به هر نمره اضافه شود}} x_۱ + ۰/۲x_۱, x_۲ + ۰/۲x_۲, \dots, x_n + ۰/۲x_n$$

درنتیجه نمرات جدید : $۱/۲x_۱ , ۱/۲x_۲,\dots,۱/۲x_n$

یعنی نمرات در ۱/۲ ضرب می‌شوند بنابراین ضریب تغییرات ثابت می‌ماند.

مثال۳۷: در ۶۰ داده‌ی آماری، میانگین ۳ و انحراف معیار ۱/۲ است اگر به تمام داده‌ها ۹ واحد اضافه شود ضریب تغییرات داده‌های جدید کدام است؟(سراسری تجربی ۸۵)

۱) ۰/۱ √ ۲) ۰/۲ ۳) ۰/۳ ۴) ۰/۴

حل:اگر ۹ واحد به هر کدام از داده‌ها اضافه کنیم انحراف معیار تغییر نمی کند ولی میانگین را باید بعلاوه ۹ کرد.

$$\bar{X} = ۳+۹ = ۱۲ \rightarrow CV = \frac{۱/۲}{۱۲} = ۰/۱$$

مثال۳۸: در داده‌های آماری با میانگین $\bar{X}$ و انحراف معیار σ، اگر به هر یک از داده‌ها مقدار $\bar{X}$ را اضافه کنیم تا داده‌های جدید حاصل شود، ضریب تغییرات داده‌های جدید چند برابر ضریب تغییرات داده‌های قبلی است؟ (سراسری تجربی ۸۶)

۱) $\frac{۱}{۴}$ ۲) $\frac{۱}{۲}$ √ ۳) ۱ ۴) ۲

حل:

$$CV_۱=\frac{\sigma}{\bar{x}}\xrightarrow{\text{به داده ها } \bar{x} \text{ می‌اضافه شود}}CV_۲=\frac{\sigma}{\bar{x}+\bar{x}}=\frac{\sigma}{۲\bar{x}}\rightarrow\frac{CV_۲}{CV_۱}=\frac{\frac{\sigma}{۲\bar{x}}}{\frac{\sigma}{\bar{x}}}=\frac{۱}{۲}$$

۶-۷مقایسه ی دقت یا پراکندگی دو گروه

برای مقایسه‌ی دو گروه ابتدا میانگین داده‌ها را بدست می‌آوریم ؛که به یکی از دو حالت زیر برخورد می‌کنیم:

حالت ۱: اگر میانگین دو گروه با هم برابر باشد واریانس را بدست می‌آوریم. هرچه واریانس بیشتر باشد پراکندگی بیشتر است؛ پس دقت کمتر است.

حالت۲: اگر میانگین دو گروه با هم برابر نباشد ضریب تغییرات را بدست می‌آوریم. هرچه ضریب تغییرات بیشتر باشد پراکندگی بیشتر است پس دقت کمتر است.

مثال۳۹: دو نفر در یک آزمایشگاه در ۵ روز متوالی همزمان شروع به کار کردند. امتیازات دقت کاری آنها مطابق جدول زیر است، دقت کاری کدام بیشتر است؟

۱√)نفر اول ۲) نفر دوم ۳) یکسان ۴) نیاز به اطلاعات بیشتری دارد

نفر اول	۷	۹	۸	۹	۷
نفر دوم	۱۰	۸	۶	۷	۹

حل: $\overline{X_۱}=\overline{X_۲}=۸$

چون میانگین‌ها برابرند پس واریانس را حساب می‌کنیم.

$$\sigma_۱{}^۲=\frac{(۷-۸)^۲+(۹-۸)^۲+(۸-۸)^۲+(۹-۸)^۲+(۷-۸)^۲}{۵}=\frac{۴}{۵}=۰.۸$$

$$\sigma_2{}^2=\frac{(۱۰-۸)^۲+(۸-۸)^۲+(۶-۸)^۲+(۷-۸)^۲+(۹-۸)^۲}{۵}=\frac{۱۰}{۵}=۲$$

دقت کاری نفر اول بیشتر است(هرچه واریانس کمتر باشد دقت بیشتر و پراکندگی کمتر است)

مثال۴۰: دو سری داده آماری ۵, ۴, ۳, ۲, ۱ : A و ۶۵, ۶۴, ۶۳, ۶۲, ۶۱ = B مفروضند، پراکندگی نسبی کدام سری داده‌ی آماری بیشتر است؟

۱√)داده‌های A ۲)داده‌های B

۳)داده‌های A وB پراکندگی نسبی یکسانی دارند. ۴)نمی توان مقایسه کرد

چون میانگین‌ها با هم برابرنیستند، از ضریب تغییرات کمک می‌گیریم.

داده‌های B = ۶۳+ داده‌های A

$$\begin{cases} A:1,2,3,4,5 & \begin{cases}\bar{x}_A = 3 \\ \sigma_A = \sqrt{2}\end{cases} \Rightarrow CV_A = \dfrac{\sqrt{2}}{3} \\ B:61,62,63,64,65 & \begin{cases}\bar{x}_B = 63 \\ \sigma_B = \sqrt{2}\end{cases} \Rightarrow CV_B = \dfrac{\sqrt{2}}{63}\end{cases} \Rightarrow CV_A > CV_B$$

منابع:

۱- بخشعلی زاده، شهرناز؛ پاشا، عین الله؛ رستگار، آرش؛ آمار و مدل سازی.۹۴. چاپ و نشر کتاب‌های درسی ایران.

۲- آذر ، عادل؛ مومنی، منصور ؛ آمار و کاربرد آن در مدیریت(تحلیل آماری).۹۴. انتشارات سازمان مطالعه و تدوین کتب علوم انسانی.

۳- موسی پور ، داوود؛ زارعی،علی؛ کلیات آمار.۹۵. ناشرحافظ پژوه.

۴- رنجبران، هادی؛ آمار واحتمال وکاربرد آن در مدیریت وحسابداری.۹۴. ناشر اثبات.