

آمار و احتمال

پایه یازدهم « رشته ی ریاضی فیزیک »

فصل ۳ : آمار توصیفی

تهیه کننده : جابر عامری

دبیر ریاضی شهرستان های اهواز و باوی



www.mathtower.ir

@amerimath



درس اول : توصیف و نمایش داده ها

همانطور که می دانید، در یک فعالیت آماری بعد از گردآوری داده ها، تنظیم و رده بندی و خلاصه کردن آنها، جهت نتیجه گیری و تعیین شاخص ها ، لازم است. در این درس به طبقه بندی داده ها و رسم نمودار های آماری می پردازیم.

دسته بندی اطلاعات

می دانیم که واقعیت هایی درباره های یک شیء یا فرد که در محاسبه ، برنامه ریزی و پیش بینی به کار می روند را داده (اطلاعات) می نامند. واضح است که باید داده های جمع آوری شده در هر زمینه ای که باشند را دسته بندی نمود تا از این طریق بتوان به راحتی داده ها را توصیف و از آنها نتیجه گیری کرد، یکی از روش های دسته بندی داده ها تشکیل **جدول فراوانی** موسوم به توزیع فراوانی است در واقع یکی از کارآمدترین روشها برای خلاصه کردن و سازمان بندی کردن اطلاعات توزیع فراوانی می باشد. که در اینجا به توضیح آن می پردازیم.

جدول فراوانی

در یک مجموعه ی داده های آماری به تعداد دفعاتی که هر داده تکرار می شود، **فراوانی مطلق** یا به اختصار فراوانی آن داده گویند.

مثال: نمرات مسئولیت پذیری بیست مدیر به شرح زیر است، فراوانی این داده ها را مشخص کنید.

x_i : ۱۲ و ۱۵ و ۱۷ و ۱۲ و ۱۲ و ۵ و ۶ و ۱۲ و ۶ و ۱۰ و ۵ و ۱۵ و ۱۲ و ۶ و ۱۵ و ۱۰ و ۱۲ و ۱۵

$F(x=12) = 6$ فراوانی ۱۲ برابر ۶ است. $F(x=5) = 4$ فراوانی ۵ برابر ۴ است.

$F(x=15) = 4$ فراوانی ۱۵ برابر ۴ است. $F(x=6) = 3$ فراوانی ۶ برابر ۳ است.

$F(x=17) = 1$ فراوانی ۱۷ برابر ۱ است. $F(x=10) = 2$ فراوانی ۱۰ برابر ۲ است.

وبه طور خلاصه داریم.

| | | | | | | | |
|-----------------|---|---|----|----|----|----|-----|
| نمره (x_i) | ۵ | ۶ | ۱۰ | ۱۲ | ۱۵ | ۱۷ | جمع |
| فراوانی (F_i) | ۴ | ۳ | ۲ | ۶ | ۴ | ۱ | ۲۰ |

نتیجه: واضح است که جمع فراوانی های مطلق با اندازه نمونه یا جامعه برابر است.



روشهای دسته بندی داده ها

هر ویژگی از اشیاء یا اشخاص که در اعضای جامعه یکسان نیست و معمولاً از یک عضو به عضو دیگر تغییر می کند را متغیر می نامند. نماد یا عددی که به آن ویژگی یک عضو نسبت داده می شود را مقدار متغیر (مشاهده) می گویند. متغیرها براساس اینکه مقادیر عددی می گیرند یا خیر، دارای دو نوع اصلی کمی و کیفی می باشد،

متغیر کمی: متغیرهایی هستند که مقادیر عددی می گیرند و برای آنها عملیات ریاضی از قبیل جمع، تفریق و معدل گیری قابل انجام است.

مثال: سن، نمره، تعداد افراد خانواده، طول مکالمه ی تلفنی و

متغیر کیفی: متغیرهایی هستند که صرفاً برای دسته بندی افراد یا اشیاء در گروه ها به کار می روند و لزوماً مقدار عددی نمی گیرند.

مثال: جنسیت (مرد و زن)، کیفیت کالا (مرغوب، متوسط، نامرغوب)

هدف علم آمار نتیجه گیری از تعداد زیادی داده است که از جامعه یا نمونه بدست می آیند. تنظیم جدول آماری جهت توصیف نتایج مفید است. دسته بندی داده ها با توجه به نوع متغیر آنها به صورت های مختلف انجام می شود.

الف) دسته بندی کیفی: این روش برای دسته بندی داده هایی استفاده می شود که دارای متغیر کیفی می باشند.

مثال: گروه خونی ۲۵ دانش آموز یک کلاس به شرح زیر است.

A,O,AB,O,B,A,O,B,B,O,B,AB,A,AB,O,A,A,B,A,B,B,O,O,AB,O

حال می توان این مشاهدات را به صورت زیر دسته بندی کرد.

| ردیف | دسته (گروه خونی) | خط نشان (T) | فراوانی (F) |
|------|------------------|---------------------|-------------|
| ۱ | A | / | ۶ |
| ۲ | B | // | ۷ |
| ۳ | AB | | ۴ |
| ۴ | O | /// | ۸ |
| جمع | - | - | ۲۵ |



مثال: جدول داده شده را کامل کنید.

| ردیف (i) | دسته ها | فراوانی مطلق | فراوانی نسبی | درصد فراوانی نسبی | زاویه ی قطاع | مرکز دسته ها |
|----------|---------|--------------|--------------|-------------------|--------------|--------------|
| ۱ | ۵-۹ | ۳ | ۰/۱۲ | ۱۲ | ۴۳/۲ | ۷ |
| ۲ | ۹-۱۳ | ۵ | ۰/۲۰ | ۲۰ | ۷۲ | ۱۱ |
| ۳ | ۱۳-۱۷ | ۷ | ۰/۲۸ | ۲۸ | ۱۰۰/۸ | ۱۵ |
| ۴ | ۱۷-۲۱ | ۶ | ۰/۲۴ | ۲۴ | ۸۶/۴ | ۱۹ |
| ۵ | ۲۱-۲۵ | ۳ | ۰/۱۲ | ۱۲ | ۴۳/۲ | ۲۳ |
| ۶ | ۲۵-۲۹ | ۱ | ۰/۰۴ | ۴ | ۱۴/۴ | ۲۷ |
| جمع | - | ۲۵ | ۱ | ۱۰۰ | ۳۶۰ | - |

تمرین ۳: جدول زیر را کامل کنید.

| ردیف (i) | دسته ها | فراوانی مطلق | فراوانی نسبی | درصد فراوانی نسبی | زاویه ی قطاع | مرکز دسته ها |
|----------|---------|--------------|--------------|-------------------|--------------|--------------|
| ۱ | ۱-۵ | ۳ | | | | |
| ۲ | ۵-۹ | ۷ | | | | |
| ۳ | ۹-۱۳ | ۶ | | | | |
| ۴ | ۱۳-۱۷ | ۴ | | | | |
| جمع | - | | | | | |

نتیجه:

۱: تفاضل مراکز دو دسته متوالی با طول دسته ها برابر است. $(C = x_{i+1} - x_i)$

۲: مجموع فراوانی های نسبی تمام دسته ها برابر یک است.

$$(f_1 + f_2 + f_3 + \dots + f_n = \sum f_i = 1)$$

۳: مجموع درصد فراوانی های نسبی تمام دسته ها برابر ۱۰۰ است.

$$(P_1 + P_2 + P_3 + \dots + P_n = \sum P_i = 100)$$

۴: مجموع زاویه های قطاع های تمام دسته ها برابر ۳۶۰ است.

$$(\alpha_1 + \alpha_2 + \alpha_3 + \dots + \alpha_n = \sum \alpha_i = 360)$$



تمرین برای حل :

۴: در جدول زیر جای خالی را کامل کنید.

| دسته ها | فراوانی مطلق | فراوانی نسبی | مرکز دسته ها | زاویه ی قطاع |
|---------|--------------|--------------|--------------|--------------|
| ۲-۵ | ۲ | ۰/۲ | ۳/۵ | |
| ۵-۸ | | | ۶/۵ | ۱۰۸ |
| | ۵ | ۰/۵ | | ۱۸۰ |
| جمع | ۱۰ | | - | |

۵: داده های زیر یک نمونه ی ۲۵ تایی از نمرات دانش آموزان پایه ی دهم یک منطقه، در مسابقه ی علمی

۳۰ نمره ای می باشند.

۲۲/۵ و ۲۱ و ۲۱/۵ و ۱۷/۵ و ۱۶ و ۱۶/۷۵ و ۱۲/۵ و ۱۴ و ۱۸ و ۱۴/۷۵ و ۱۳/۵ و ۱۸

۲۰ و ۱۸ و ۱۶ و ۱۷ و ۱۴ و ۱۳/۵ و ۱۹ و ۱۹/۷۵ و ۲۱ و ۲۰/۵ و ۱۵ و ۱۵/۵ و ۱۶/۵

جدول فراوانی زیر را برای این داده ها ، تکمیل کنید.

| دسته ها | فراوانی مطلق | فراوانی نسبی | درصد فراوانی نسبی | زاویه ی مرکزی |
|---------|--------------|--------------|-------------------|---------------|
| | | | | |
| | | | | |
| | | | | |
| جمع | | | | |

۶: از ۲۵ دانش آموز پایه ی یازدهم پیرامون رنگ مورد علاقه برای خودرو پرسیده شد و جدول فراوانی زیر

تنظیم گردید. این جدول را تکمیل کنید.

| دسته ها | فراوانی مطلق | فراوانی نسبی | درصد فراوانی نسبی | زاویه ی مرکزی |
|---------|--------------|--------------|-------------------|---------------|
| سفید | ۶ | | | |
| نقره ای | ۸ | | | |
| یشمی | ۷ | | | |
| آبی | ۴ | | | |
| جمع | ۲۵ | | | |

۷: در یک جدول فراوانی، اگر فراوانی نسبی مربوط به گروه خونی B برابر ۰/۴ باشد و مجموع فراوانی های

همه ی گروه ها ی خونی برابر ۲۰ باشد. فراوانی گروه خونی B را تعیین کنید.

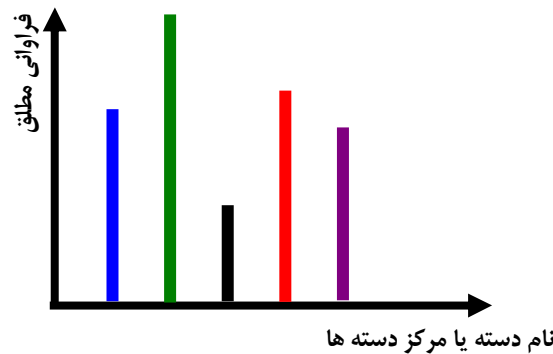
نمودارهای آماری

نمایش هندسی داده های دسته بندی شده را نمودار (شاخص هندسی) می گویند. نمودارها وسیله ای سودمند برای به تصویر درآوردن و تجسم جامعه یا نمونه می باشند، بطوری که تفهیم نتایج به کمک آنها در کمترین زمان و با ساده ترین بیان صورت می گیرد.

نمودارهای آماری با توجه به نوع متغیرهای بکار رفته در داده ها دارای سه دسته اصلی زیر می باشند. به جدول زیر دقت شود.

| نام نمودارهای مناسب | نوع متغیرهای بکار رفته در داده های آماری |
|-----------------------------------------------|------------------------------------------|
| نمودار میله ای ، نمودار دایره ای | کیفی |
| نمودار میله ای ، نمودار دایره ای | کمی گسسته |
| نمودار بافت نگاشت (مستطیلی) ، نمودار چند ضلعی | کمی پیوسته |

نمودار میله ای (ستونی): در این نمودار مهم این است که گروه های متفاوت جامعه یا نمونه از نظر فراوانی مقایسه شوند. محور افقی این نمودار شامل دسته ها (نام متغیرها) و محور قائم آن شامل فراوانی مطلق (یا فراوانی نسبی و گاهی درصد فراوانی نسبی) است.

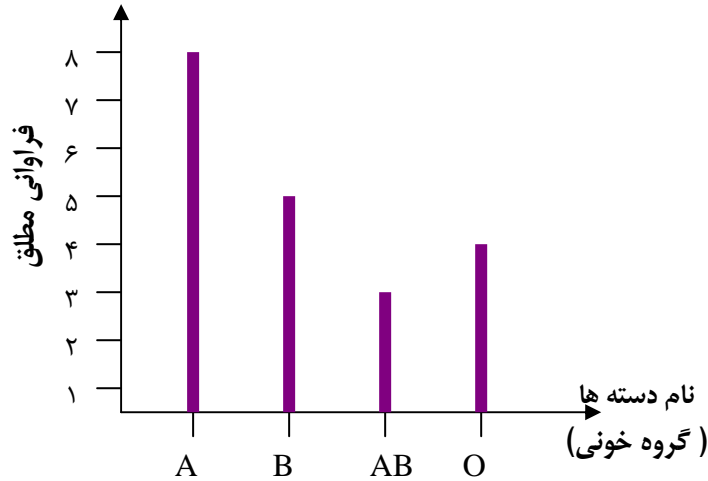


مثال: اطلاعات آماری زیر مربوط به گروه خونی ۲۰ دانش آموز است، پس از تنظیم جدول فراوانی نمودار میله ای رسم کنید.

AB, A, AB, A, A, A, B, A, A, B, O, B, O, A, A, B, B, O, AB, O

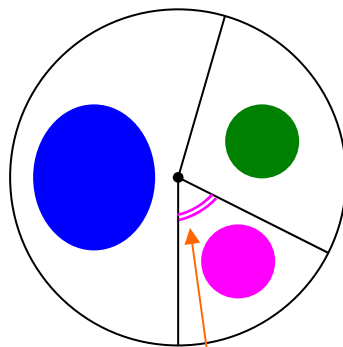
حل:

| دسته ها | A | B | AB | O | جمع |
|---------|---|---|----|---|-----|
| فراوانی | ۸ | ۵ | ۳ | ۴ | ۲۰ |



نمودار دایره ای: این نمودار نیز برای مقایسه گروه های متفاوت جامعه بکار می‌رود، برای رسم این نمودار

زاویه هایی مرکزی (زاویه های قطاع ها) روی دایره طوری مشخص می‌شوند که اندازه‌ی هر زاویه بر حسب درجه به کمک جدول فراوانی به دست آمده باشد.



مثال: در جدول زیر مدرک تحصیلی کارکنان یک اداره نوشته شده است. متناظر با این جدول یک نمودار

دایره ای رسم کنید.

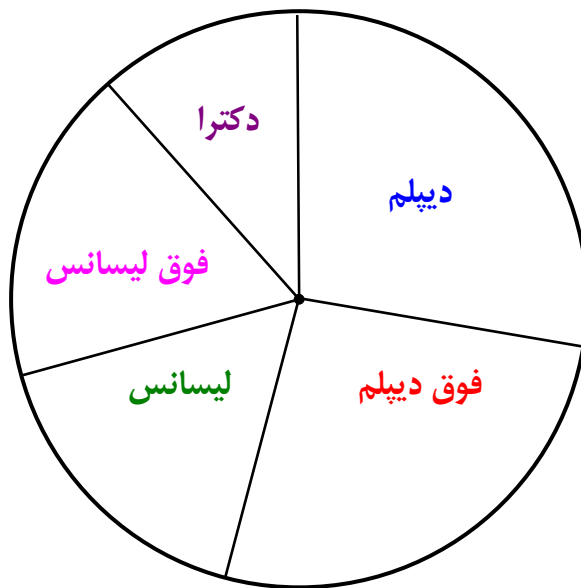
| مدرک تحصیلی | دیپلم | فوق دیپلم | لیسانس | فوق لیسانس | دکتر |
|-----------------|-------|-----------|--------|------------|------|
| فراوانی (تعداد) | ۱۰ | ۹ | ۶ | ۷ | ۴ |

حل: ابتدا زاویه‌ی قطاع مربوط به هر دسته را تعیین و سپس روی دایره، زاویه های قطاع متناظر را تعیین

می‌کنیم.



| نوع بیمه | فراوانی | زاویه مرکزی (بر حسب درجه) |
|------------|---------|----------------------------|
| دیپلم | ۱۰ | ۱۰۰ |
| فوق دیپلم | ۹ | ۹۰ |
| لیسانس | ۶ | ۶۰ |
| فوق لیسانس | ۷ | ۷۰ |
| دکتر | ۴ | ۴۰ |
| جمع | ۳۶ | ۳۶۰ |



تمرین ۸: در یک شرکت دارویی جدول توزیع کارکنان را با نمودار دایره ای نشان می دهیم. زاویه ی قطاع مربوط به کارکنان ارشد، چند درجه است؟

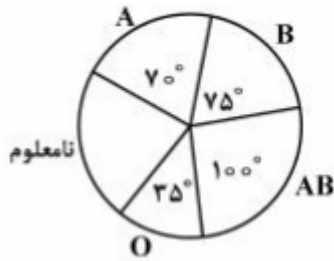
| نوع مدرک | دیپلم | کاردانی | کارشناسی | ارشد | دکتر |
|----------|-------|---------|----------|------|------|
| تعداد | ۳۰ | ۹۰ | ۱۸۰ | ۱۲۰ | ۳۰ |

۱۰۵° (۴) ۹۶° (۳) ۹۲° (۲) ۸۴° (۱)

حل :

$$n = 30 + 90 + 180 + 120 + 30 = 450$$

$$\alpha_i = \frac{f_i}{n} \times 360 = \frac{120}{450} \times 360 = 12 \times 8 = 96$$



تمرین ۹: نمودار دایره ای روبرو، متناسب با تعداد کارکنان سازمانی با گروه خونی متمایز است. گروه خونی ۳۲ نفر از آنان تعیین نشده است. چند نفر آنها، دارای نوع خون B هستند؟

- ۴۰ (۴)
- ۳۶ (۳)
- ۳۰ (۲)
- ۲۵ (۱)

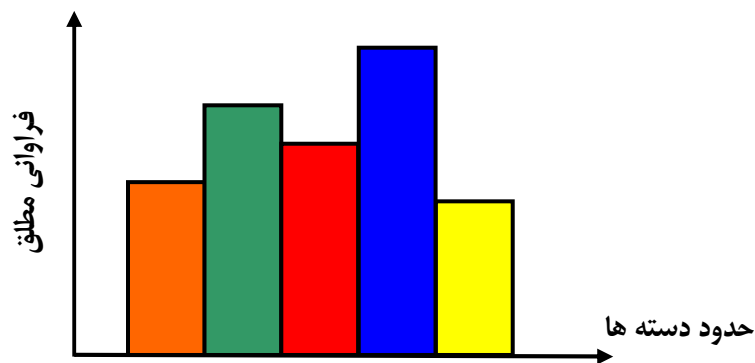
حل: ابتدا زاویه‌ی مربوط به دسته‌ی «نامعلوم» را مشخص می‌کنیم.

$$70 + 75 + 100 + 35 + \alpha = 360 \rightarrow \alpha = 80$$

حال از تناسب استفاده می‌کنیم. بدین شکل که برای دسته‌ی نامعلوم ۸۰ درجه برای ۳۲ نفر است. لذا برای دسته‌ی B زاویه‌ی ۷۵ درجه برای چند نفر است؟

$$\frac{32}{80} = \frac{x}{75} \rightarrow x = \frac{32 \times 75}{80} = 30$$

نمودار بافت نگاشت (مستطیلی): در این نمودار مستطیل‌هایی رسم می‌شود که قاعده‌ی آنها روی محور افقی و برابر طول هر یک از دسته‌ها می‌باشد. بطوری که ارتفاع مستطیل‌ها به موازات محور قائم و متناسب با فراوانی مطلق یا درصد فراوانی نسبی باشد.

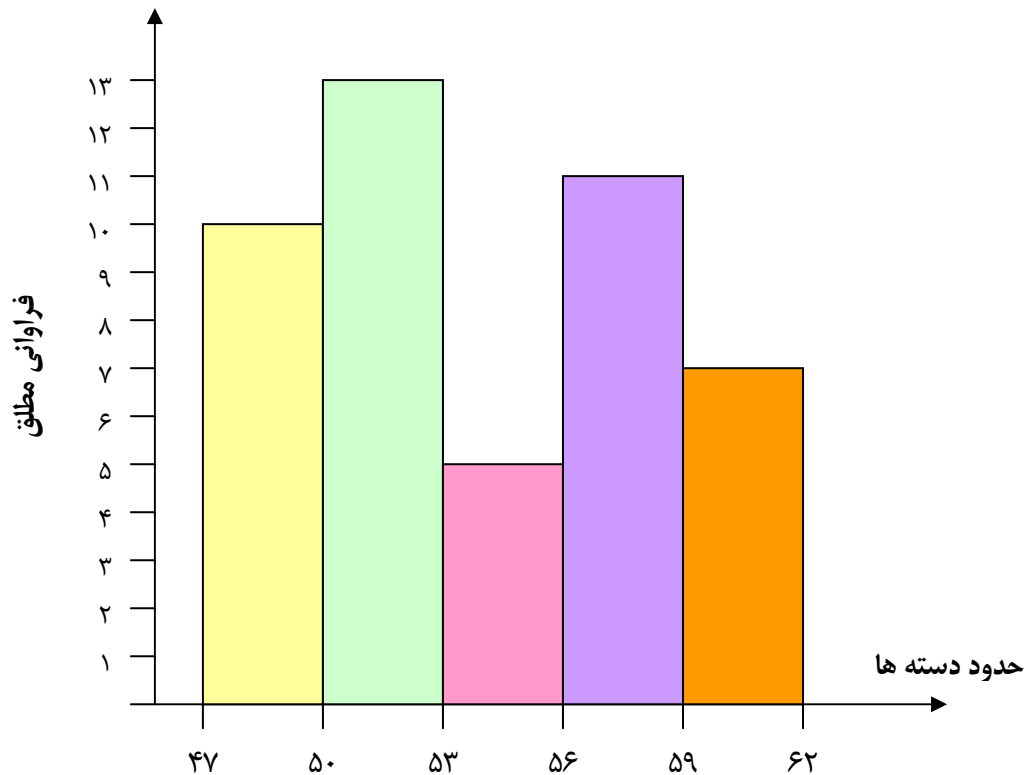


مثال: با توجه به جدول فراوانی زیر یک نمودار مستطیلی رسم کنید.

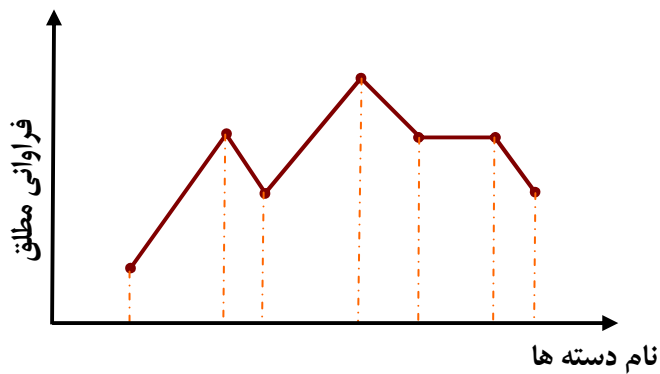
| | | | | | |
|--------------|-------|-------|-------|-------|-------|
| حدود دسته‌ها | ۴۷-۵۰ | ۵۰-۵۳ | ۵۳-۵۶ | ۵۶-۵۹ | ۵۹-۶۲ |
| فراوانی | ۱۰ | ۱۳ | ۵ | ۱۱ | ۷ |



حل:



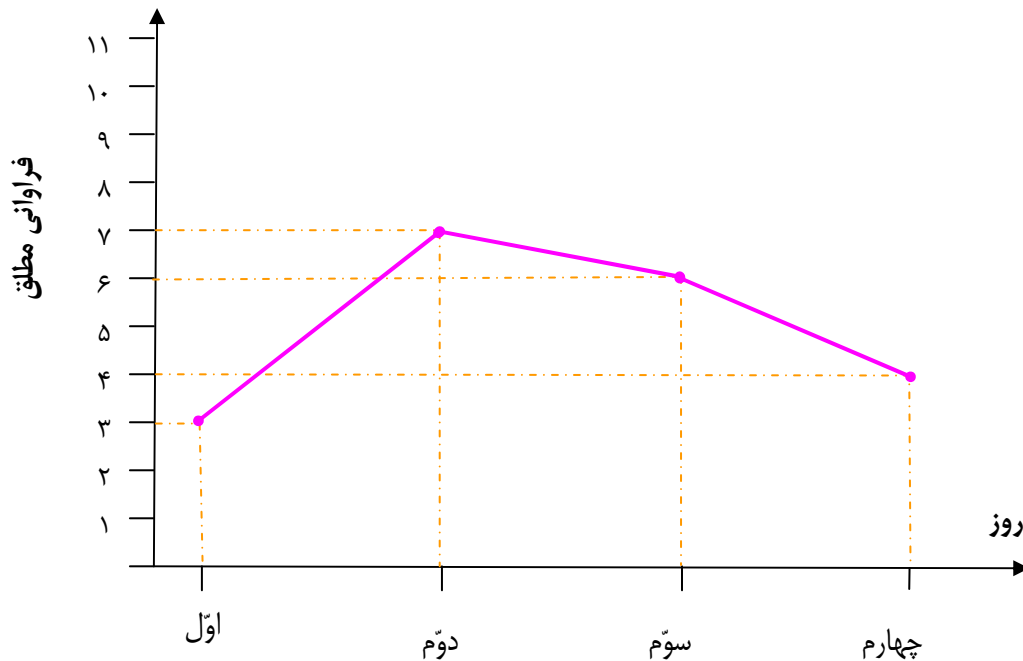
نمودار چندضلعی (چندبر): نمودار چندضلعی جهت نمایش تغییرات داده‌هایی بکار می‌رود که دارای متغیر کمی پیوسته (اغلب زمان) می‌باشند. برای رسم آن روی محور افقی مرکز دسته ها و روی محور قائم فراوانی مطلق یا در صد فراوانی نسبی منظور می‌شوند.



مثال: جدول زیر تعداد مقالات ارائه شده در یک کنفرانس ریاضی در طول روزهای برگزاری آن است. با توجه به این جدول، نمودار چند ضلعی رسم کنید.

| دسته ها | روز اوّل | روز دوّم | روز سوّم | روز چهارم |
|--------------|----------|----------|----------|-----------|
| فراوانی مطلق | ۳ | ۷ | ۶ | ۴ |

حل:



تمرین ۱۰: با توجه به مثال قبل به سئوالات زیر پاسخ دهید.

الف: کمترین و بیشترین مقالات در چه روز هایی ارائه شده اند؟

ب: چند درصد مقالات در روز سوم ارائه شده اند؟

تمرین برای حل:

۱۱: رنگ چشم ۱۲۸ فرد به شرح زیر است.

۶۴ نفر قهوه ای ، ۲۳ نفر آبی ، ۳۶ نفر سبز ، ۵ نفر سایر رنگ ها

الف: کدام نمودار برای این داده ها مناسب است. (میله ای ، دایره ای ، هر دو)

ب: تعیین کنید که چند درصد این افراد رنگ چشم آبی دارند؟

۱۲: جملات زیر را کامل کنید.

الف: برای متغیرهای پیوسته از نمودار های و استفاده می شود.

ب: برای متغیرهای گسسته از نمودارهای و استفاده می شود.

پ: برای متغیرهای کیفی از نمودارهای و استفاده می شود.

تهیه کننده: جابر عامری عضو گروه ریاضی دوره ی دوم متوسطه استان خوزستان

**درس دوم : معیارهای گرایش به مرکز**

بدیهی است که یکی از اهداف یک فعالیت آماری، نتیجه گیری پیرامون نمونه یا جامعه است. بعد از دسته بندی داده ها و رسم نمودار های مناسب، گاهی محاسبه ی معیارهای گرایش به مرکز داده ضروری است. در این درس با مهمترین این معیارها آشنا می شویم. اما قبل از ورود به بحث نماد جمع و ویژگی های آن را معرفی می کنیم.

معرفی نماد جمع و ویژگی های آن

اگر $x_1, x_2, x_3, \dots, x_n$ مجموعه ای از n عدد باشند، در این صورت عبارت

$$x_1 + x_2 + x_3 + \dots + x_n$$

را به شکل $\sum_{i=1}^n x_i$ نمایش می دهند. به عبارتی دیگر

$$\sum_{i=1}^n x_i = x_1 + x_2 + x_3 + \dots + x_n$$

مثال : اگر $x_1 = 2$ و $x_2 = 3$ و $x_3 = 5$ باشد، مقدار $\sum_{i=1}^3 x_i$ را تعیین کنید.

حل :

$$\sum_{i=1}^3 x_i = x_1 + x_2 + x_3 = 2 + 3 + 5 = 10$$

مثال : اگر $x_1 = 2$ و $x_2 = 3$ و $x_3 = 5$ باشد، مقدار $\sum_{i=1}^3 4x_i$ را تعیین کنید.

حل :

$$\sum_{i=1}^3 4x_i = 4x_1 + 4x_2 + 4x_3 = 4(x_1 + x_2 + x_3) = 4(2 + 3 + 5) = 40$$

مثال : اگر $x_1 = 2$ و $x_2 = 3$ و $x_3 = 5$ و $y_1 = 7$ و $y_2 = 1$ و $y_3 = 0$ باشد، مقدار $\sum_{i=1}^3 (x_i + y_i)$ را

تعیین کنید.

حل :

$$\sum_{i=1}^3 (x_i + y_i) = (x_1 + y_1) + (x_2 + y_2) + (x_3 + y_3) = (2 + 7) + (3 + 1) + (5 + 0) = 18$$



خواص سیگما

اگر $(x_1, x_2, x_3, \dots, x_n)$ و $(y_1, y_2, y_3, \dots, y_n)$ دو مجموعه‌ی عددی و k یک عدد حقیقی باشد. در این صورت خواص زیر را می توان برای سیگما ثابت نمود.

الف)
$$\sum_{i=1}^n kx_i = k \sum_{i=1}^n x_i$$

ب)
$$\sum_{i=1}^n (x_i + y_i) = \sum_{i=1}^n x_i + \sum_{i=1}^n y_i$$

ج)
$$\sum_{i=1}^n k = nk$$

تمرین ۱: هر یک از ویژگی های فوق را ثابت کنید.

معیارهای مرکزی

هر عدد که معرف مرکز مجموعه‌ی داده ها باشد را معیار مرکزی یا شاخص مرکزی (پارامتر مرکزی) می نامند. به کمک معیار های مرکزی می توان موقعیت کلی داده ها را تعیین کرد، لذا برای مقایسه‌ی دو یا چند جامعه یک روش مناسب، محاسبه‌ی معیار های مرکزی آنها است. معیار های مرکزی دارای سه نوع مهم هستند. این سه نوع عبارتند از ، میانگین (\bar{x}) ، میانه (\tilde{x}) و مد یا نما (\hat{x}) می باشند که میانگین مهمترین معیار مرکزی محسوب می شود.

میانگین

در یک مجموعه‌ی داده های آماری، عدد متمایل به وسط آنها را میانگین (میانگین حسابی) می نامند. میانگین یا متوسط داده ها، با خارج قسمت مجموع اندازه‌ی داده ها بر تعداد آنها برابر است.

$$\bar{x} = \frac{x_1 + x_2 + x_3 + \dots + x_n}{n} = \frac{\sum x_i}{n}$$

مثال: میانگین داده های زیر را بدست آورید.

$$x_i : 5, 8, 10, 12, 14, 17$$



حل:

$$\bar{x} = \frac{\sum x_i}{n} = \frac{5 + 8 + 10 + 12 + 14 + 17}{6} = \frac{66}{6} = 11$$

تمرین ۲: میانگین نمرات یک کلاس ۱۶ نفری در درس آمار و احتمال برابر ۱۲/۵ است. جمع نمرات دانش آموزان این کلاس را بدست آورید.

تمرین ۳: میانگین داده های زیر برابر ۲۲ است. مقدار a را بدست آورید.

$$15 \text{ و } 20 \text{ و } 21 \text{ و } 24 \text{ و } 25 \text{ و } a \text{ و } 22 \text{ و } 30 \text{ و } 25 \text{ و } 20$$

میانگین وزنی

گاهی به دلایلی برای داده های یک مجموعه آماری ضریب های (تعداد تکرار) متفاوتی در نظر گرفته می شود. مانند:

۱: ضریب های نمرات امتحانات مستمر و پایانی در محاسبه نمره ی سالانه

۲: ضریب های دروس در کنکور

۳: واحد های دروس

در این صورت تعداد تکرار هر داده را وزن آن داده می نامند. اگر داده های $(x_1 \text{ و } x_2 \text{ و } x_3 \text{ و } \dots \text{ و } x_n)$ به ترتیب دارای تعداد تکرار یا ضریب $(w_1 \text{ و } w_2 \text{ و } w_3 \text{ و } \dots \text{ و } w_n)$ باشند، می توان میانگین را به روش زیر محاسبه نمود.

| | | | | | |
|-----------------|-------|-------|-------|-------|-------|
| داده ها (x_i) | x_1 | x_2 | x_3 | | x_n |
| ضریب (w_i) | w_1 | w_2 | w_3 | | w_n |

$$\bar{x} = \frac{w_1x_1 + w_2x_2 + \dots + w_nx_n}{w_1 + w_2 + \dots + w_n} = \frac{\sum w_i x_i}{\sum w_i}$$

مثال: میانگین داده های مربوط به جدول زیر را محاسبه کنید.

| | | | | |
|-----------------|----|----|----|----|
| داده ها (x_i) | ۱۴ | ۱۷ | ۱۶ | ۱۵ |
| ضریب (w_i) | ۴ | ۲ | ۳ | ۱ |

حل:

$$\bar{x} = \frac{\sum w_i x_i}{\sum w_i} = \frac{4(14) + 2(17) + 3(16) + 1(15)}{4 + 2 + 3 + 1} = \frac{56 + 34 + 48 + 15}{10} = \frac{153}{10} = 15.3$$

مثال: میانگین داده های زیر را بدست آورید.

۶۲ و ۶۵ و ۶۵ و ۶۵ و ۶۲ و ۶۲ و ۶۰ و ۶۲ و ۵۸ و ۶۰ و ۶۲ و ۶۰ و ۵۸ و ۷۲

حل:

| x_i | F_i | $F_i \cdot x_i$ |
|-------|-------|-----------------|
| ۷۲ | ۱ | ۷۲ |
| ۵۸ | ۲ | ۱۱۶ |
| ۶۰ | ۳ | ۱۸۰ |
| ۶۲ | ۵ | ۳۱۰ |
| ۶۵ | ۴ | ۲۶۰ |
| جمع | ۱۵ | ۹۳۸ |

$$\bar{x} = \frac{\sum F_i x_i}{n} = \frac{938}{15} \approx 62.53$$

نتیجه: در یک جدول فراوانی اگر x_i مرکز دسته باشد. می توان میانگین داده ها را به کمک فرمول زیر

محاسبه کرد.

مثال: با توجه به جدول زیر میانگین حسابی را تعیین کنید.

| حدود دسته | ۵ - ۱۵ | ۱۵ - ۲۵ | ۲۵ - ۳۵ | ۳۵ - ۴۵ |
|-----------|--------|---------|---------|---------|
| فراوانی | ۳ | ۴ | ۱ | ۲ |

حل:

| دسته ها | F_i | x_i | $F_i \cdot x_i$ |
|---------|-------|-------|-----------------|
| ۵ - ۱۵ | ۳ | ۱۰ | ۳۰ |
| ۱۵ - ۲۵ | ۴ | ۲۰ | ۸۰ |
| ۲۵ - ۳۵ | ۱ | ۳۰ | ۳۰ |
| ۳۵ - ۴۵ | ۲ | ۴۰ | ۸۰ |
| جمع | ۱۰ | - | ۲۲۰ |

$$\bar{x} = \frac{\sum F_i x_i}{n} = \frac{220}{10} = 22$$



مثال: اگر میانگین داده ها در جدول زیر برابر ۴ باشد، فراوانی مطلق دسته‌ی آخر را به دست آورید.

| حدود دسته | ۰ - ۲ | ۲ - ۴ | ۴ - ۶ | ۶ - ۸ |
|-----------|-------|-------|-------|-------|
| فراوانی | ۵ | ۷ | ۴ | x |

حل:

| مرکز دسته | ۱ | ۳ | ۵ | ۷ |
|-----------|---|---|---|---|
| فراوانی | ۵ | ۷ | ۴ | x |

$$\bar{x} = \frac{\sum F_i x_i}{\sum F_i} \rightarrow 4 = \frac{5 + 21 + 20 + 7x}{5 + 7 + 4 + x} \rightarrow 4(16 + x) = 46 + 7x \rightarrow 3x = 18 \rightarrow x = 6$$

توجه: اگر دو یا چند گروه از یک مجموعه‌ی داده‌های آماری با اندازه‌های متفاوت دارای میانگین

مشخصی باشند. میانگین کل که به میانگین مرکب معروف است، را می‌توان به روش زیر محاسبه نمود.

$$\bar{z} = \frac{m\bar{x} + n\bar{y}}{m + n}$$

که در آن اندازه و میانگین هر گروه مطابق جدول زیر است.

| - | اندازه | میانگین |
|----------|--------|-----------|
| گروه اول | m | \bar{x} |
| گروه دوم | n | \bar{y} |

البته این فرمول فوق برای بیش از دو گروه از یک جامعه قابل تعمیم است.

مثال: میانگین ۴ داده‌ی آماری برابر ۷ و میانگین ۶ داده‌ی دیگر ۵ است. میانگین کل داده‌ها را بدست آورید.

حل:

$$\bar{z} = \frac{m\bar{x} + n\bar{y}}{m + n} = \frac{4(7) + 6(5)}{4 + 6} = \frac{28 + 30}{10} = \frac{58}{10} = 5.8$$

تمرین ۴: اگر دو گروه دارای اندازه‌های برابر باشند، میانگین مرکب را از چه روشی می‌توان محاسبه نمود؟



تمرین برای حل :

۵: در جدول زیر نمره های یک دانش آموز در درس حسابان داده شده است، میانگین وزنی (نمره ی سالانه) وی را محاسبه کنید.

| - | مستمر اوّل | پایانی اوّل | مستمر دوّم | پایانی دوّم |
|------|------------|-------------|------------|-------------|
| نمره | ۱۰ | ۹ | ۱۲ | ۱۴ |
| ضریب | ۱ | ۲ | ۱ | ۶ |

۶: دانش آموزی درکنکور سراسری شرکت می کند و نتیجه ی کارنامه ی آزمون آن به شرح زیر است.

| مواد امتحانی | ریاضیات | فیزیک | شیمی | زبان انگلیسی | ادبیات و زبان فارسی | دین و زندگی |
|--------------|---------|-------|------|--------------|---------------------|-------------|
| درصد | ۷۴ | ۶۵ | ۸۲ | ۵۵ | ۹۵ | ۱۰۰ |
| ضریب دروس | ۴ | ۳ | ۲ | ۱ | ۲ | ۳ |

متوسط درصد مواد امتحانی این دانش آموز را محاسبه کنید.

۷: سازمانی دارای سه بخش کاری است، تعداد پرسنل و متوسط دریافتی هر بخش بصورت زیر است.

میانگین دریافتی کل پرسنل را محاسبه کنید.

| بخش | اداری | آموزشی | فرهنگی | جمع |
|---------------|-------|--------|--------|-----|
| تعداد پرسنل | ۲۰ | ۳۰ | ۵۰ | ۱۰۰ |
| متوسط دریافتی | ۵۰ | ۴۰ | ۶۰ | - |

۸: دستمزد روزانه ی ۶ نفر از کارگران یک کارگاه هر یک ۷۵ هزار تومان و دستمزد روزانه ی ۲ نفر دیگر

هریک ۶۰ هزار تومان است. میانگین دستمزد کل کارگران را حساب کنید.

۹: میانگین داده های جدول مقابل ۱۴ است، مقدار a را بیابید.

| داده | ۱۰ | a | ۱۲ | ۱۶ |
|---------|----|-----|----|----|
| فراوانی | ۱ | ۲ | ۱ | ۶ |

۱۰: دانش آموزی در کنکور سراسری شرکت می کند و نتیجه ی کارنامه ی آزمون آن به شرح زیر است.

| مواد امتحانی | ریاضیات | فیزیک | شیمی | زبان انگلیسی | ادبیات و زبان فارسی | دین و زندگی |
|--------------|---------|-------|------|--------------|---------------------|-------------|
| درصد | ۵۳ | ? | ۶۷ | ۳۴ | ۸۰ | ۶۷ |
| ضریب دروس | ۴ | ۳ | ۱ | ۱ | ۴ | ۳ |

اگر معدل موزون درصد این دانش آموز ۶۳ باشد، درصد این دانش آموز در درس فیزیک را تعیین کنید.



۱۱: میانگین ۵ داده‌ی آماری ۱۷ است. اگر دو عدد ۱۶ و ۱۱ را به داده‌های قبلی اضافه کنیم. میانگین جدید چه عددی خواهد شد؟

۱۲: ثابت کنید که مجموع تفاضل هر یک از داده‌های یک مجموعه آماری از میانگین آنها برابر صفر است.

$$\sum_{i=1}^n (x_i - \bar{x}) = 0$$

خواص میانگین

اگر $(x_1, x_2, x_3, \dots, x_n)$ و $(y_1, y_2, y_3, \dots, y_n)$ دو مجموعه‌ی عددی و k یک عدد حقیقی باشد. در این صورت خواص زیر را می‌توان برای میانگین بررسی کرد.

الف: میانگین حاصل جمع داده‌های یک مجموعه‌ی آماری با یک عدد ثابت با حاصل جمع آن عدد و میانگین آن داده‌ها برابر است.

$$z_i = x_i + k \rightarrow \bar{z} = \bar{x} + k$$

ب: میانگین حاصل ضرب داده‌های یک مجموعه‌ی آماری در یک عدد ثابت با حاصل ضرب آن عدد و میانگین آن داده‌ها برابر است.

$$z_i = k \cdot x_i \rightarrow \bar{z} = k \cdot \bar{x}$$

ج: میانگین حاصل جمع داده‌های متناظر از دو مجموعه‌ی داده‌های آماری با حاصل جمع میانگین‌های آنها برابر است.

$$z_i = x_i + y_i \rightarrow \bar{z} = \bar{x} + \bar{y}$$

تمرین برای حل:

۱۳: هر یک از خواص فوق را اثبات کنید.

۱۴: ثابت کنید که میانگین داده‌های مساوی، برابر هر یک از آنها است.

۱۵: الف: اگر $(x_1, x_2, x_3, \dots, x_n)$ یک مجموعه از داده‌های آماری باشند و $y_i = ax_i + b$. ثابت کنید که $\bar{y} = a\bar{x} + b$

ب: هرگاه میانگین داده‌های آماری $(x_1, x_2, x_3, \dots, x_n)$ برابر ۱۷ باشد. میانگین داده‌های $(1 + 2x_1, 1 + 2x_2, 1 + 2x_3, \dots, 1 + 2x_n)$ را بدست آورید.



تذکر: یکی از کاربرد های خواص میانگین، در محاسبه ی سریع میانگین مورد استفاده قرار می گیرد. در این روش یک داده به دلخواه (بهتر است داده ی متوسط) را انتخاب کرده و از تمام داده ها کم می کنیم. سپس میانگین داده های جدید را محاسبه می کنیم. با اضافه کردن داده ای که قبلاً از تمام داده ها کم کرده بودیم به میانگین بدست آمده، میانگین واقعی داده ها بدست می آید. به مثال زیر توجه کنید.

مثال: نمرات ریاضی ۴۰ دانش آموز یک کلاس در جدول زیر آمده است. میانگین وزنی نمرات، کدام است؟
(کنکور ۹۸ ریاضی)

| | | | | | | |
|-----|----|----|----|----|----|----|
| x | ۱۰ | ۱۲ | ۱۴ | ۱۵ | ۱۷ | ۱۸ |
| f | ۵ | ۸ | ۷ | ۱۰ | ۶ | ۴ |

$$14/75 (4) \qquad 14/4 (3) \qquad 14/25 (2) \qquad 14/2 (1)$$

حل: از تمام داده ها ۱۴ واحد کم می کنیم و میانگین داده های جدید را به دست می آوریم و در آخر به میانگین به دست آمده ۱۴ واحد را اضافه می کنیم.

| | | | | | | |
|--------------|----|----|---|----|---|---|
| $y = x - 14$ | -۴ | -۲ | ۰ | ۱ | ۳ | ۴ |
| f | ۵ | ۸ | ۷ | ۱۰ | ۶ | ۴ |

| | | | | | | | |
|--------------|-----|-----|---|----|----|----|-----|
| y | -۴ | -۲ | ۰ | ۱ | ۳ | ۴ | جمع |
| f | ۵ | ۸ | ۷ | ۱۰ | ۶ | ۴ | ۴۰ |
| $f \times y$ | -۲۰ | -۱۶ | ۰ | ۱۰ | ۱۸ | ۱۶ | ۸ |

$$\bar{y} = \frac{\sum (f_i \times y_i)}{\sum f_i} = \frac{8}{40} = 0.2$$

$$\bar{y} = \bar{x} - 14 \rightarrow \bar{x} = \bar{y} + 14 = 0.2 + 14 = 14.2$$

تمرین برای حل:

۱۶: به روش فوق میانگین داده های متناظر با جدول زیر را محاسبه کنید.

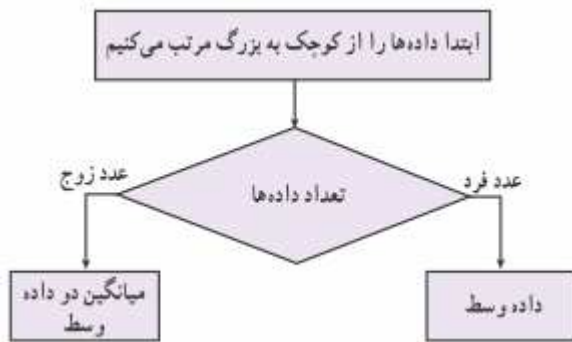
| | | | | |
|-----|----|----|----|----|
| x | ۱۳ | ۱۷ | ۱۶ | ۱۵ |
| f | ۵ | ۳ | ۴ | ۸ |



میانه

در یک مجموعه از داده های آماری، میانه داده ای است که نصف داده از آن بیشتر و نصف داده ها از آن کمتر است. به عبارت دیگر در یک مجموعه ی داده های آماری که به صورت غیر نزولی (از کوچک به بزرگ) مرتب شده باشند، عدد وسط این داده ها را میانه می نامند. لذا برای محاسبه ی میانه ابتدا داده ها را به صورت

غیر نزولی مرتب می کنیم، آنگاه



الف) اگر تعداد داده ها فرد باشد، داده ی وسط میانه است.

ب) اگر تعداد داده ها زوج باشد، میانگین دو داده ی وسط میانه است.

مثال: میانه ی هریک از مجموعه های زیر را بیابید.

الف) ۵ و ۸ و ۵ و ۱۰ و ۹ و ۸ و ۷ و ۵ و ۳

ب) ۶۰ و ۲۲ و ۲۱ و ۲۰ و ۱۹ و ۱۷ و ۱۷ و ۱۶ و ۱۵ و ۱۳

حل: ابتدا داده ها را از کوچک به بزرگ مرتب می کنیم.

الف:

۳ و ۵ و ۵ و ۷ و ۸ و ۸ و ۹ و ۱۰



چون تعداد داده ها فرد است، پس داده ی وسط میانه است. لذا میانه برابر $\tilde{x} = 7$ می باشد.

ب:

۱۳ و ۱۵ و ۱۶ و ۱۷ و ۱۷ و ۱۹ و ۲۰ و ۲۱ و ۲۲ و ۶۰



چون تعداد داده ها زوج است، میانگین دو داده ی وسط میانه است. لذا میانه برابر $\tilde{x} = \frac{17 + 19}{2} = 18$ می باشد.

می باشد.

نتیجه: اگر تمام داده های یک مجموعه ی داده های آماری برابر باشند، میانه نیز برابر هر یک از آنها است.



توجه: اگر در یک مجموعه‌ی داده‌های آماری داده‌ی دور افتاده‌ای وجود داشته باشند. چون میانه تحت تأثیر داده‌های دور افتاده قرار نمی‌گیرد، نسبت به میانگین، معیار مناسبتری محسوب می‌شود. داده‌ی دور افتاده، داده‌ای است که نسبت به سایر داده‌ها تفاوت بسیار دارد.

چارک‌ها

در یک مجموعه‌ی داده‌های آماری که داده‌های آن به صورت غیر نزولی مرتب شده باشند، عدد وسط این داده‌ها را میانه یا چارک دوم می‌نامند و آنرا با Q_2 نمایش می‌دهند. از طرفی میانه‌ی نیمه‌ی اول داده‌ها را چارک اول (Q_1) و میانه‌ی نیمه‌ی دوم آنها را چارک سوم (Q_3) می‌نامند.

مثال: چارک‌های اول تا سوم داده‌های زیر را تعیین کنید.

۱۹ و ۳۱ و ۲۵ و ۱۸ و ۳۲ و ۴۳ و ۴۱ و ۳۴ و ۱۸ و ۲۷ و ۱۴ و ۲۳ و ۱۵ و ۱۰ و ۱۲

حل:

| | | | | | | | |
|----------------|----------------|----------------|----------------|----------------|----------------|----------------|-----------|
| <u>۱۰ و ۱۲</u> | <u>۱۴ و ۱۵</u> | <u>۱۸ و ۱۸</u> | <u>۱۹ و ۲۳</u> | <u>۲۵ و ۲۷</u> | <u>۳۱ و ۳۲</u> | <u>۳۴ و ۴۱</u> | <u>۴۳</u> |
| ↑ | ↑ | ↑ | ↑ | ↑ | ↑ | ↑ | ↑ |
| Q_1 | Q_2 | Q_3 | | | | | |

لذا چارک اول $Q_1 = 15$ و چارک دوم (میانه) $Q_2 = 23$ و چارک سوم $Q_3 = 32$

توجه: برای تعیین چارک‌ها، ابتدا میانه را محاسبه می‌کنیم ولی برای تعیین چارک‌های اول و سوم، میانه شرکت داده نمی‌شود.

تمرین ۱۷: چارک‌های اول تا سوم داده‌های زیر را محاسبه کنید.

۴۱ و ۱۷ و ۱۶ و ۱۹ و ۱۳ و ۱۴ و ۱۸ و ۱۶ و ۱۵ و ۱۷



۲۱: توضیح دهید اگر تمام داده های یک مجموعه‌ی آماری در عدد ثابت ۲ ضرب شوند، میانه و مد آنها چه تغییری می‌کنند؟

۲۲: یک کارخانه تولید خودرو برای رنگ اتومبیل از مشتریان نظرخواهی کرد و جدول زیر را از یک نمونه بدست آورد. با توجه به این جدول مد داده ها را تعیین کنید.

| رنگ مورد علاقه مشتری | سبز | سفید | قرمز | آبی | زرد |
|----------------------|-----|------|------|-----|-----|
| تعداد (نفر) | ۱۵ | ۴۲ | ۱۳ | ۱۷ | ۵ |

۲۳: ثابت کنید که اگر داده های یک مجموعه‌ی داده های آماری تشکیل دنباله‌ی حسابی می‌دهند، میانگین

آنها برابر $\frac{a+b}{2}$ است که در آن a جمله‌ی اول و b جمله‌ی آخر دنباله است.

۲۴: در هر مورد تعیین کنید که کدام معیار گرایش به مرکز، مناسبتر است.

الف: نتایج رأی گیری در انتخاب ریاست جمهوری

ب: تعیین میزان دمای یک شهر در یک هفته

ج: تعیین طول عمر لامپ های کم مصرف

د: تعیین حد وسط حقوق کارکنان چند اداره با احتساب کلیه‌ی حقوق ها

۲۵: داده های زیر را در نظر بگیرید.

۶ و ۹ و ۱۱ و ۱۳ و ۵ و ۴ و ۶ و ۱۰

الف: میانگین، میانه و مد را محاسبه کنید.

ب: یک داده از این داده ها را طوری تغییر دهید که مد و میانه ثابت بماند.

ج: دو داده‌ی دیگر به این داده ها، طوری اضافه کنید که میانگین تغییر نکند.

د: دو داده‌ی دیگر به این داده ها، طوری اضافه کنید که میانگین و میانه ثابت بمانند ولی مد تغییر کند.

۲۶: داده های زیر را طوری تغییر دهید که میانه ثابت بماند ولی میانگین دو واحد اضافه شود.

۱۰ و ۱۹ و ۱۲ و ۱۷ و ۲۲



۲۷: در هر مورد جای خالی را کامل کنید.

الف: اگر در یک مجموعه‌ی داده‌های آماری، داده‌ی دورافتاده وجود داشته باشد، نسبت به میانگین،

معیار بهتری محسوب می‌شود.

ب: اگر در مجموعه‌ی داده‌های آماری، تمام داده‌ها باشند. این داده‌ها مُد ندارند.

تهیه کننده: جابر عامری

عضو گروه ریاضی دوره ی دوم متوسطه

استان خوزستان

**درس سوّم : معیارهای پراکندگی**

در فعالیت های آماری، گاهی لازم می شود که میزان پراکندگی (دوری و نزدیکی) داده را نسبت به همدیگر یا نسبت به میانگین محاسبه شود. برای اینکار از معیارهای پراکندگی استفاده می شود. در این درس با مهمترین این معیارها آشنا می شویم.

معیارهای پراکندگی

هر عدد که میزان پراکندگی داده ها نسبت به همدیگر یا نسبت به میانگین را نشان دهد، را معیار پراکندگی یا شاخص پراکندگی (پارامتر پراکندگی) می نامند. مهمترین معیارهای پراکندگی عبارتند از ، دامنه ی تغییرات ، واریانس ، انحراف معیار و ضریب تغییرات می باشند.

از بین این معیارها، دامنه ی تغییرات پراکندگی داده ها را نسبت به همدیگر و بقیه، پراکندگی را نسبت به میانگین که مهمترین معیار مرکزی می باشد، نشان می دهند و توجه داشته باشید که انحراف معیار به جهت ویژگی هایی که دارد ، مهمترین معیار محسوب می شود.

دامنه ی تغییرات

در یک مجموعه ی داده های آماری تفاضل کمترین داده از بیشترین آنها را دامنه ی تغییرات می نامند. به عبارتی دیگر دامنه ی تغییرات، طول بازه ای است که داده ها در آن قرار دارند. در این صورت، اگر a کوچکترین و b بزرگترین داده ی یک مجموعه ی داده های آماری باشند. دامنه ی تغییرات به شکل زیر است.

$$R = b - a$$

دامنه ی تغییرات را می توان به بیشترین اختلاف بین داده های یک مجموعه ی داده های آماری تعبیر کرد.

مثال: دامنه ی تغییرات داده های زیر را حساب کنید و تعبیر آن را بنویسید.

$$۱۳ و ۸ و ۱۴ و ۱۵ و ۱۰$$

حل:

$$R = b - a = ۱۵ - ۸ = ۷$$

تعبیر: بیشترین اختلاف بین داده های این مجموعه برابر ۷ است.



نتیجه:

- ۱: بزرگی دامنه‌ی تغییرات نشان دهنده‌ی تفاوت زیاد در جامعه است، هر چه قدر این دامنه بیشتر باشد، تفاوت بین داده‌ها زیاد است و هر چه قدر این دامنه کمتر باشد، داده‌ها به هم نزدیکترند. اگر دامنه‌ی تغییرات صفر باشد، تمام داده‌ها برابر هستند و جامعه همگون است.
- ۲: دامنه‌ی تغییرات ضعیف‌ترین شاخص پراکندگی است و معرف خوبی برای پراکندگی داده‌ها نمی‌باشد، زیرا برای محاسبه‌ی آن فقط از بزرگترین و کوچکترین داده استفاده می‌شود و تعداد یا مقادیر بقیه‌ی داده‌ها تأثیری بر مقدار آن ندارند.

واریانس و انحراف معیار

میانگین توان دوم تفاضل داده‌ها از میانگین آنها را واریانس (پراش) می‌نامند. در این صورت

$$\sigma^2 = \frac{\sum (x_i - \bar{x})^2}{n}$$

توجه داشته باشید که واریانس در ضمن داشتن اهمیت زیاد دارای دو اشکال عمده است.

الف: تحت تأثیر داده‌های بزرگ قرار می‌گیرد.

ب: واحد اندازه‌گیری واریانس مجذور واحد اصلی متغیر داده‌ها است. مثلاً اگر واحد اندازه‌گیری داده‌ها سانتی متر باشد، واحد اندازه‌گیری واریانس سانتی متر مربع خواهد بود.

برای رفع این دو اشکال از انحراف معیار استفاده می‌شود. ریشه‌ی دوم واریانس را انحراف معیار (انحراف استاندارد) می‌نامند.

$$\sigma = \sqrt{\sigma^2}$$

بنابراین طبق این تعریف بدیهی است که انحراف معیار دارای همان واحدی خواهد بود که داده‌ها بر حسب آن محاسبه شده‌اند.

مثال: واریانس و انحراف معیار داده‌های زیر را بدست آورید.

۹ و ۴ و ۷ و ۳ و ۲



حل:

| x_i | $x_i - \bar{x}$ | $(x_i - \bar{x})^2$ |
|----------|-----------------|---------------------|
| ۲ | -۳ | ۹ |
| ۳ | -۲ | ۴ |
| ۷ | ۲ | ۴ |
| ۴ | -۱ | ۱ |
| ۹ | ۴ | ۱۶ |
| جمع = ۲۵ | - | ۳۴ |

$$\bar{x} = \frac{\sum x_i}{n} = \frac{۲۵}{۵} = ۵$$

میانگین

$$\sigma^2 = \frac{\sum (x_i - \bar{x})^2}{n} = \frac{۳۴}{۵} = ۶/۸$$

واریانس

$$\sigma = \sqrt{\sigma^2} = \sqrt{۶/۸} \approx ۲/۶$$

انحراف معیار

توجه: اگر مجموعه‌ی داده‌های آماری دارای تکرار باشند، می‌توان به کمک فرمول زیر واریانس را محاسبه کرد.

$$\sigma^2 = \frac{\sum F_i (x_i - \bar{x})^2}{n}$$

همچنین اگر x_i مرکز دسته‌ها در یک جدول فراوانی باشد، می‌توان به کمک این فرمول واریانس داده‌ها را نیز به دست آورد.

مثال: واریانس و انحراف معیار داده‌های زیر را بدست آورید.

۱۲ و ۱۵ و ۱۲ و ۱۲ و ۱۲ و ۱۲ و ۱۱ و ۱۱ و ۱۱ و ۱۰ و ۱۰ و ۸ و ۷ و ۷ و ۷

حل:



| x_i | F_i | $F_i x_i$ | $x_i - \bar{x}$ | $(x_i - \bar{x})^2$ | $F_i (x_i - \bar{x})^2$ |
|-------|-------|-----------|-----------------|---------------------|-------------------------|
| ۷ | ۳ | ۲۱ | -۳ | ۹ | ۲۷ |
| ۸ | ۱ | ۸ | -۲ | ۴ | ۴ |
| ۱۰ | ۲ | ۲۰ | ۰ | ۰ | ۰ |
| ۱۱ | ۳ | ۳۳ | ۱ | ۱ | ۳ |
| ۱۲ | ۵ | ۶۰ | ۲ | ۴ | ۲۰ |
| ۱۵ | ۱ | ۱۵ | ۵ | ۲۵ | ۲۵ |
| جمع | ۱۵ | ۱۵۷ | - | - | ۷۹ |

$$\bar{x} = \frac{\sum F_i x_i}{n} = \frac{157}{15} = 10.46 \approx 10$$

$$\sigma^2 = \frac{\sum F_i (x_i - \bar{x})^2}{n} = \frac{79}{15} = 5.26$$

$$\sigma = \sqrt{\sigma^2} = \sqrt{5.26} \approx 2.29$$

تمرین برای حل :

۱: سن ۲۰ نفر شرکت کننده در یک کلاس نهضت سواد آموزی به شرح زیر است.

۱۲ و ۲۱ و ۱۹ و ۲۴ و ۲۱ و ۲۳ و ۲۷ و ۱۳ و ۲۳ و ۱۷ و ۲۵ و ۱۲ و ۲۷ و ۲۲ و ۲۲ و ۲۲ و ۲۵ و ۱۷ و ۱۳ و ۱۸ و ۲۷

الف: دامنه‌ی تغییرات داده‌ها را به دست آورید.

ب: جدول زیر را با توجه به این داده‌ها تکمیل نموده و سپس انحراف معیار آنها را به دست آورید.

| دسته‌ها | F_i | x_i | $F_i x_i$ | $x_i - \bar{x}$ | $(x_i - \bar{x})^2$ | $F_i (x_i - \bar{x})^2$ |
|---------|-------|-------|-----------|-----------------|---------------------|-------------------------|
| ۱۰-۱۵ | | | | | | |
| ۱۵-۲۰ | | | | | | |
| ۲۰-۲۵ | | | | | | |
| ۲۵-۳۰ | | | | | | |
| جمع | | | | | | |

۲: ثابت کنید که واریانس داده‌های مساوی برابر صفر است.



خواص واریانس

اگر $(x_1, x_2, x_3, \dots, x_n)$ یک مجموعه‌ی عددی و k یک عدد حقیقی باشد. در این صورت خواص زیر را می توان برای واریانس آنها بررسی کرد.

۱: واریانس حاصل جمع داده ها با یک عدد ثابت ، با واریانس آن داده ها برابر است.

$$y_i = x_i + k \rightarrow \sigma_y^2 = \sigma_x^2$$

۲: واریانس حاصل ضرب داده ها در یک عدد ثابت ، با حاصل ضرب مربع آن عدد در واریانس آن داده ها برابر است.

$$y_i = k.x_i \rightarrow \sigma_y^2 = k^2 . \sigma_x^2$$

تمرین ۲: هر یک از خواص فوق را اثبات کنید.

خواص انحراف معیار

مشابه آنچه که برای واریانس داشتیم. خواص زیر را برای انحراف معیار نیز می توان بیان کرد.

۱: انحراف معیار حاصل جمع داده ها با یک عدد ثابت با انحراف معیار آن داده ها برابر است.

$$y_i = x_i + k \rightarrow \sigma_y = \sigma_x$$

۲: انحراف معیار حاصل ضرب داده ها در یک عدد ثابت، با حاصل ضرب قدر مطلق آن عدد در انحراف معیار آن داده ها برابر است.

$$y_i = k.x_i \rightarrow \sigma_y = |k| . \sigma_x$$

تمرین ۳: هر یک از خواص فوق را اثبات کنید.

تمرین ۴: الف: اگر $(x_1, x_2, x_3, \dots, x_n)$ یک مجموعه از داده های آماری باشند و $y_i = ax_i + b$.

$$\text{ثابت کنید که } \sigma_y^2 = a^2 . \sigma_x^2$$

ب: هرگاه واریانس داده های آماری $(x_1, x_2, x_3, \dots, x_n)$ برابر ۱۶ باشد. انحراف معیار داده های

$(1 + 2x_1, 1 + 2x_2, 1 + 2x_3, \dots, 1 + 2x_n)$ را محاسبه کنید.



تمرین ۵: اگر واریانس داده های یک مجموعه ی داده های آماری برابر ۱۸ و میانگین آنها ۸ باشد و تمام داده ها را دو برابر کنیم. میانگین و واریانس داده های جدید را تعیین کنید.

روش دیگر برای محاسبه ی واریانس

با توجه به تعریف واریانس و خواص سیگما می توان نتیجه گرفت که :

$$\sigma^2 = \frac{\sum x_i^2}{n} - \bar{x}^2$$

برای اثبات این موضوع می توان به شکل زیر عمل کرد.

$$\sigma^2 = \frac{\sum (x_i - \bar{x})^2}{n}$$

$$\rightarrow \sigma^2 = \frac{\sum (x_i^2 - 2\bar{x}x_i + \bar{x}^2)}{n} = \frac{\sum x_i^2}{n} + \frac{\sum (-2\bar{x}x_i)}{n} + \frac{\sum \bar{x}^2}{n}$$

$$\rightarrow \sigma^2 = \frac{\sum x_i^2}{n} + \frac{\sum (-2\bar{x}x_i)}{n} + \frac{\sum \bar{x}^2}{n} = \frac{\sum x_i^2}{n} - 2\bar{x} \frac{\sum x_i}{n} + \frac{n\bar{x}^2}{n}$$

$$\rightarrow \sigma^2 = \frac{\sum x_i^2}{n} - 2\bar{x} \cdot \bar{x} + \bar{x}^2 = \frac{\sum x_i^2}{n} - 2\bar{x}^2 + \bar{x}^2 = \frac{\sum x_i^2}{n} - \bar{x}^2$$

نتیجه: اگر داده های آماری دارای تکرار باشند. می توان به شکل زیر نیز عمل کرد.

$$\sigma^2 = \frac{\sum F_i x_i^2}{n} - \bar{x}^2$$

مثال: واریانس و انحراف معیار داده های زیر را به روش فوق محاسبه کنید.

۷ و ۹ و ۶ و ۵ و ۳



حل :

| x_i | x_i^2 |
|----------|---------|
| ۳ | ۹ |
| ۵ | ۲۵ |
| ۶ | ۳۶ |
| ۹ | ۸۱ |
| ۷ | ۴۹ |
| جمع = ۳۰ | ۲۰۰ |

$$\bar{x} = \frac{\sum x_i}{n} = \frac{30}{5} = 6$$

$$\sigma^2 = \frac{\sum x_i^2}{n} - \bar{x}^2 = \frac{200}{5} - (6)^2 = 40 - 36 = 4$$

$$\sigma = \sqrt{4} = 2$$

تمرین برای حل :

۶: اگر برای یک جامعه داده های آماری داشته باشیم.

$$\sum x_i^2 = 580 \text{ و } \sum x_i = 100 \text{ و } n = 20$$

واریانس و انحراف معیار داده ها را تعیین کنید.

واریانس مرکب^۱

اگر دو یا چند گروه از یک مجموعه داده های آماری با اندازه های متفاوت دارای واریانس های متفاوتی باشند، واریانس کل که به واریانس مرکب معروف است را می توان از روش زیر محاسبه کرد.

$$\sigma_t^2 = \frac{m(\bar{x}^2 + \sigma_x^2) + n(\bar{y}^2 + \sigma_y^2)}{m + n} - \bar{z}^2$$

¹. فقط برای مطالعه



با فرض اینکه این دو گروه که دارای میانگین و واریانس مطابق جدول زیر می باشند.

| - | اندازه | میانگین | واریانس |
|----------|--------|-----------|--------------|
| گروه اول | m | \bar{x} | σ_x^2 |
| گروه دوم | n | \bar{y} | σ_y^2 |

همچنین \bar{z} میانگین مرکب دو گروه است.

توجه داشته باشید که

الف: با ریشه گیری از فرمول فوق انحراف معیار مرکب نیز بدست می آید.

ب: این فرمول، برای بیش از دو گروه از یک جامعه نیز قابل تعمیم است.

مثال: اطلاعات زیر مربوط به اجرای یک آزمون هوش در دو کلاس است.

| - | اندازه | میانگین | انحراف معیار |
|----------|--------|---------|--------------|
| کلاس اول | ۲۵ | ۱۰۳ | ۹ |
| کلاس دوم | ۲۴ | ۱۰۴ | ۱۲ |

واریانس و انحراف معیار مرکب را محاسبه کنید.

حل:

$$\bar{z} = \frac{m\bar{x} + n\bar{y}}{m+n} = \frac{25(103) + 24(104)}{25+24} = \frac{5071}{49} \approx 103/5$$

$$\sigma_z^2 = \frac{m(\bar{x}^2 + \sigma_x^2) + n(\bar{y}^2 + \sigma_y^2)}{m+n} - \bar{z}^2$$

$$= \frac{25(103^2 + 9^2) + 24(104^2 + 12^2)}{25+24} - (103/5)^2 = \frac{267250 + 263040}{49} - 10712/25 = 110$$

$$\sigma_z = \sqrt{\sigma_z^2} = \sqrt{110} \approx 10/48$$

تمرین ۷: اگر اندازه های دو گروه برابر باشند، واریانس مرکب از چه رابطه ای به دست می آید؟



ضریب پراکندگی

خارج قسمت انحراف معیار داده های یک مجموعه ی داده های آماری بر میانگین آنها را ضریب پراکندگی (ضریب تغییرات) گویند.

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

توجه کنید که طبق تعریف ، ضریب پراکندگی داده ها فاقد واحد اندازه گیری می باشد، لذا اغلب در موارد زیر استفاده می شود.

الف : برای مقایسه ی دو یا چند جامعه ی آماری که واحد اندازه گیری داده های آنها متفاوت باشد.

مثلاً: مقایسه ی سود دو شرکت که سود یکی بر حسب ریال و سود دیگری بر حسب دلار باشد.

ب : برای مقایسه ی دو یا چند جامعه ی آماری که واریانس های آنها برابر باشند ولی میانگین های متفاوت دارند.

مثلاً: با توجه به اطلاعات زیر دو شرکت x و y دارای انحراف معیار برابر هستند ولی میانگین مساوی دارند، لذا شرکت y از پراکندگی شرکت بهتری در مقایسه با شرکت x محسوب می شود.

| - | میانگین | انحراف معیار | ضریب پراکندگی |
|----------|---------|--------------|---------------|
| شرکت x | ۱۰ | ۲ | ۰/۲ |
| شرکت y | ۲۰ | ۲ | ۰/۱ |

مثال : ضریب پراکندگی داده های زیر را محاسبه کنید.

۴ و ۹ و ۳ و ۸ و ۶

| x_i | $x_i - \bar{x}$ | $(x_i - \bar{x})^2$ |
|----------|-----------------|---------------------|
| ۴ | -۲ | ۴ |
| ۹ | ۳ | ۹ |
| ۳ | -۳ | ۹ |
| ۸ | ۲ | ۴ |
| ۶ | ۰ | ۰ |
| جمع = ۳۰ | --- | ۲۶ |

$$\bar{x} = \frac{\sum x_i}{n} = \frac{30}{5} = 6$$



$$\sigma^2 = \frac{\sum (x_i - \bar{x})^2}{n} = \frac{26}{5} = 5.2$$

$$\sigma = \sqrt{5.2} \approx 2.28$$

$$CV = \frac{\sigma}{\bar{x}} = \frac{2.28}{6} \approx 0.38$$

نتیجه: اگر تمام داده های آماری برابر باشند، ضریب تغییرات آنها صفر است.

مثال: فرض کنیم وزن دو نوزاد ۱/۵ و ۲/۵ کیلوگرم و وزن دو فرد چهل ساله ۸۰ و ۸۱ کیلوگرم است.

الف: تفاوت وزن دو نوزاد چقدر است.

ب: تفاوت وزن دو فرد چهل ساله چقدر است.

پ: انحراف معیارهای وزن هر دو دسته را به دست آورید.

ت: فکر می کنید تفاوت وزن ها در کدام دسته زیاده تر به نظر می رسد؟

ث: ضریب تغییرات هر دو دسته را به دست آورید. چه نتیجه ای می گیرید؟

حل:

$$R_x = 2/5 - 1/5 = 1 \quad R_y = 81 - 80 = 1$$

| x_i | $(x_i - \bar{x})^2$ |
|-------------------------------------|-------------------------------------------------|
| ۱/۵ | $(1/5 - 2)^2 = 0.25$ |
| ۲/۵ | $(2/5 - 2)^2 = 0.25$ |
| $\bar{x} = \frac{1/5 + 2/5}{2} = 2$ | $\sigma_x = \sqrt{\frac{0.25 + 0.25}{2}} = 0.5$ |

ضریب تغییرات

$$C.V_x = \frac{\sigma_x}{\bar{x}} = \frac{0.5}{2} = 0.25$$

| y_i | $(y_i - \bar{y})^2$ |
|--------------------------------------|-------------------------------------------------|
| ۸۰ | $(80 - 80.5)^2 = 0.25$ |
| ۸۱ | $(81 - 80.5)^2 = 0.25$ |
| $\bar{y} = \frac{80 + 81}{2} = 80.5$ | $\sigma_y = \sqrt{\frac{0.25 + 0.25}{2}} = 0.5$ |

ضریب تغییرات

$$C.V_y = \frac{\sigma_y}{\bar{y}} = \frac{0.5}{80.5} = 0.006$$

چون انحراف معیار هر دو گروه برابر است. نمی توان گفت که تفاوت وزن ها در کدام دسته بیشتر است. لذا

بهتر است از ضریب تغییرات استفاده کرد. گروهی که ضریب تغییرات آن کمتر است، تفاوت وزن های کمتر

دارد.



تمرین برای حل :

۸: برای داده های مقابل ضریب تغییرات را محاسبه کنید. ۱۲ و ۱۵ و ۳ و ۶ و ۹

۹: ضریب تغییرات داده های مربوط به جدول زیر را تعیین کنید.

| | | | | |
|---------|---|---|---|---|
| داده ها | ۱ | ۲ | ۴ | ۵ |
| فراوانی | ۵ | ۱ | ۲ | ۴ |

۱۰: اگر ضریب تغییرات ۱۰ داده برابر ۲ و میانگین آنها برابر ۴ باشد، واریانس داده ها را به دست آورید.

۱۱: اگر $n=10$ و $\sum x_i = 60$ و $\sum x_i^2 = 400$ ضریب پراکندگی داده ها را تعیین کنید.

۱۲: نشان دهید که اگر تمام داده های یک مجموعه ی داده های آماری را در عدد مثبت k ضرب کنیم، ضریب تغییرات آنها تغییری نمی کند.

۱۳: توضیح دهید که اگر تمام داده های یک مجموعه ی داده های آماری را با عدد مثبت k جمع کنیم، ضریب تغییرات آنها چه تغییری می کند؟

۱۴: داده های یک مجموعه ی داده های آماری بر حسب متر اندازه گیری شده اند. واحد اندازه گیری معیارهای میانگین، واریانس، انحراف معیار و ضریب تغییرات را مشخص کنید.

۱۵: اگر همه ی داده های یک مجموعه ی داده های آماری مساوی باشند. درباره ی میانگین و واریانس آن مجموعه چه می توان گفت؟

۱۶: اگر تمام داده های یک مجموعه ی آماری مساوی و میانگین آنها ۵ باشد. مقدار هر یک از شاخص های زیر را در صورت وجود بنویسید. الف : مد ب : میانه ج : انحراف معیار

۱۷: میانگین یک مجموعه ی داده های آماری ۵ و واریانس آنها ۱۲ است. در صورتی که تمام داده ها را سه برابر کنیم، میانگین و واریانس داده های جدید را بنویسید.

۱۸: امتیازات مهارت کاری دو نفر از پرسنل یک آزمایشگاه در ۵ روز کاری به صورت زیر است. تعیین کنید که دقت عمل کدام یک بیشتر است.

A : ۲۰ و ۲۳ و ۲۴ و ۲۶ و ۲۷ و B : ۲۱ و ۲۲ و ۲۳ و ۲۶ و ۲۸



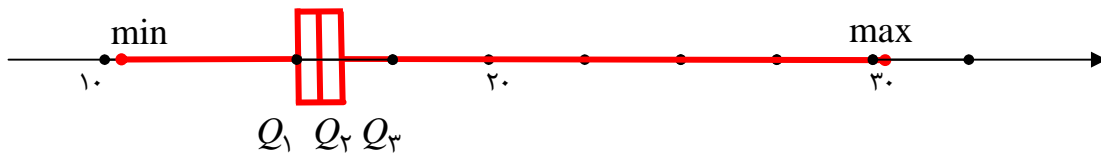
مثال: داده های زیر نرخ بیکاری یک کشور در ده سال گذشته است.

۱۱/۵ و ۱۱/۳ و ۱۰/۵ و ۱۰/۴ و ۱۱/۹ و ۱۳/۵ و ۱۲/۳ و ۱۲/۲ و ۱۰/۲ و ۳۰/۱

نمودار جعبه ای متناظر با این داده ها را رسم کنید.

$11/5, 11/3, 10/5, 10/4, 11/9, 13/5, 12/3, 12/2, 10/2, 30/1$ و $11/9, 12/2, 12/3, 13/5, 30/1$

$min = 10/2$ و $Q_1 = 10/5$ و $Q_2 = \frac{11/5 + 11/9}{2} = 11/7$ و $Q_3 = 12/3$ و $max = 30/1$



مثال: در بررسی سن بازیکنان تیم ملی، کدام نمودار مناسبتر است.

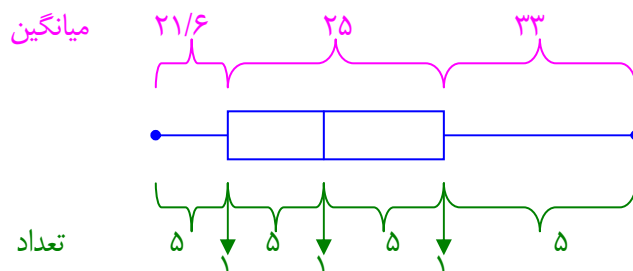
- الف) دایره ای
- ب) میله ای
- ج) جعبه ای
- د) بافت نگاشت

حل: با توجه به اینکه نمودار جعبه ای، پراکندگی سن بازیکنان در اطراف میانه را به خوبی نشان می دهد، نمودار جعبه ای مناسبتر است.

مثال: در نمودار جعبه ای ۲۳ داده ای آماری، میانگین دنباله های سمت چپ و سمت راست به ترتیب ۲۱/۶ و ۳۳ و میانگین داده های داخل و روی جعبه ۲۵ می باشد. میانگین کل این داده ها کدام است؟

- ۱) ۲۵/۸
- ۲) ۲۶
- ۳) ۲۶/۱
- ۴) ۲۶/۲

حل: با کمی دقت می توان برای مسئله نمودار زیر را رسم نمود.



در این صورت میانگین کل به صورت زیر است.

$\bar{x} = \frac{5(21/6) + 13(25) + 5(33)}{23} = \frac{598}{23} = 26$



تمرین برای حل :

۱۹ : داده های زیر را در نظر بگیرید.

۱۹ و ۳۱ و ۲۵ و ۱۸ و ۳۲ و ۳۳ و ۴۱ و ۳۴ و ۱۶ و ۲۷ و ۱۴ و ۲۳ و ۱۵ و ۱۰ و ۱۲

الف : چارک های اول تا سوم داده های زیر را تعیین کنید.

ب : دامنه ی میان چارکی را محاسبه کنید.

ج : نمودار جعبه ای ترسیم کنید.

۲۰ : داده های زیر را در نظر بگیرید.

۷ و ۹ و ۱۱ و ۱۲ و ۵ و ۴ و ۶ و ۱۰

الف : داده ها را طوری تغییر دهید که میانگین ۲ واحد اضافه شود ولی انحراف معیار تغییر نکند.

ب : داده ها را طوری تغییر دهید که میانگین و میانه ۳ برابر شوند.

ج : دو داده ی دیگر به این داده ها طوری اضافه کنید میانگین و میانه تغییر نکنند.

۲۱ : در هر مورد جای خالی را کامل کنید.

الف : تفاضل بیشترین و کمترین داده را گویند.

ب : در نمودار جعبه ای ۵۰ درصد داده ها قبل از و ۵۰ درصد داده ها بعد از قرار دارند.

پ : در نمودار جعبه ای یک مجموعه ی داده های آماری درصد داده ها قبل از چارک اول و درصد

داده ها قبل از چارک دوم و درصد داده ها قبل از چارک سوم قرار می گیرند.

ت : معیارهای پراکندگی که میانگین در آنها نقش دارد و هستند.

ث : اگر داده های آماری k برابر شوند، ضریب تغییرات (عدد k مثبت است).

ج : واحد اندازه گیری واریانس ، واحد اندازه گیری داده ها است.

ح : هر قدر انحراف معیار کمتر باشد، میزان داده ها، کمتر است.

تهیه کننده : جابر عامری عضو گروه ریاضی دوره ی دوم متوسطه استان خوزستان