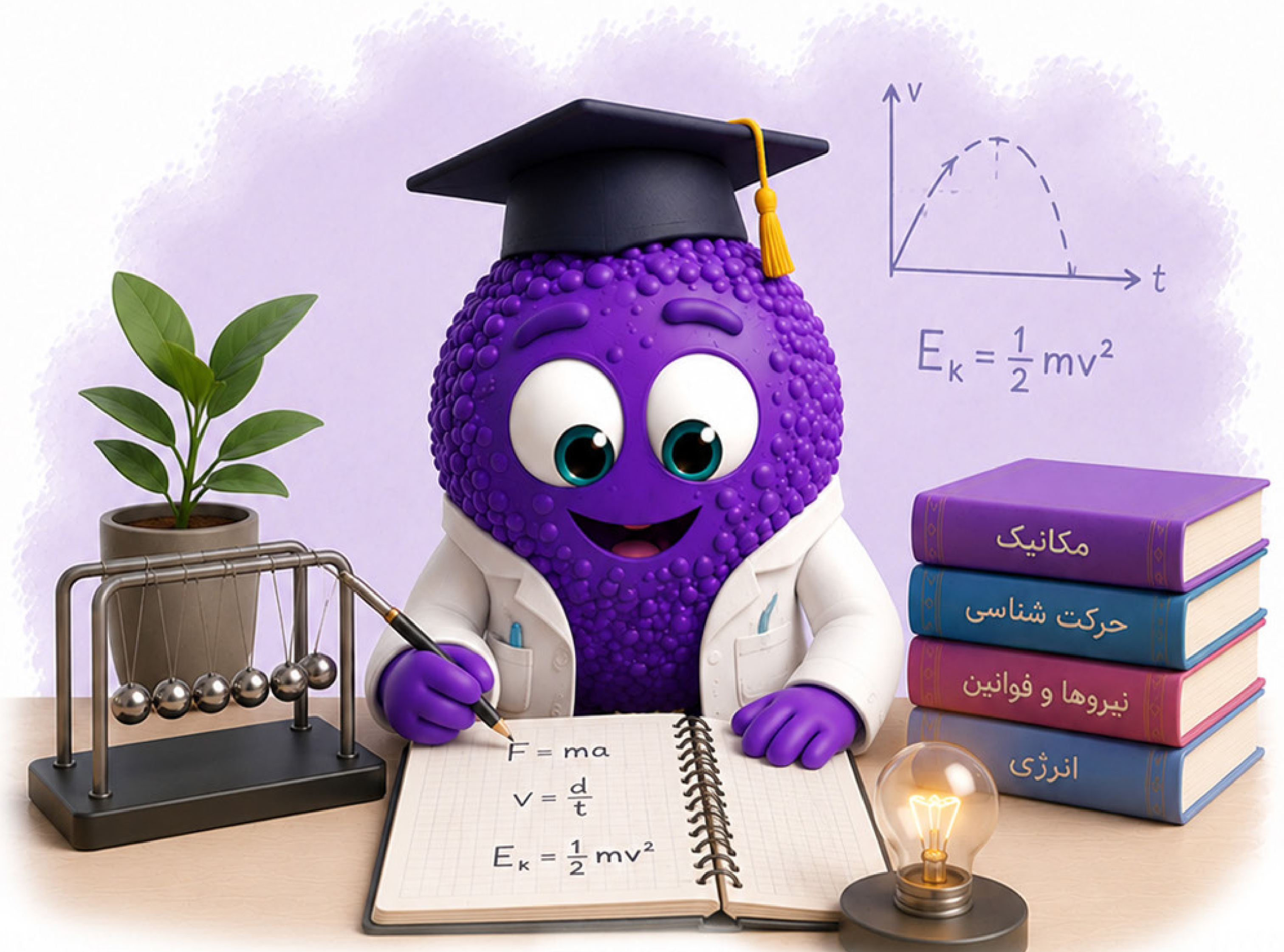
گروه آموزشی مشاوره‌ای نوتروفیل

درس

آمار و احتمال یازدهم ـ فصل ۳

نوتروبیست




نوتروفیل خونه رتبه برترها
قبولی های کنکور ۱۴۰۴
.تک رقمی نوتروفیل
رتبه ۸
ایمان نیک‌نام جهرمی
.دو رقمی‌های نوتروفیل
رتبه ۱۴
امیرمحمدکیانی
رتبه ۱۶
آریا قهرمانی
رتبه ۲۰
سینا راضی
رتبه ۳۲
امیرمحمد رضائی
رتبه ۵۵
مهسا سیاوشی
رتبه ۵۹
ایمان انفرادی
رتبه ۶۱
بهار هلالی
رتبه ۷۵
محمدصالح عارفی
رتبه ۸۰
محمدمهدی شریفی
.سه رقمی و چهار رقمی
های نوتروفیل
رتبه ۱۴۷
محدثه حیدری
رتبه ۱۶۰
اشکان کوثری
رتبه ۱۶۹
هانیه خواجه
رتبه ۲۲۲
امیرمحمد شکوهی
رتبه ۲۵۹
ابوالفضل ناصریان
رتبه ۲۷۱
فاطمه سادات موسوی
رتبه ۳۰۸
سید علی اکرمی
رتبه ۳۴۱
حمیدرضا بشیری
رتبه ۴۳۲
سید محمدصادق حسینی
رتبه ۴۷۳
زهرا بابائی
رتبه ۵۱۴
محمدپارسا عبداله آبادی
رتبه ۵۳۲
فاطمه شاهسوند
رتبه ۵۳۷
ریحانه حیدری
رتبه ۵۳۹
نجمه کیخا
رتبه ۵۴۶
حسین تفضلی نژاد
رتبه ۵۵۷
محمدصالح زارعی
رتبه ۵۷۰
زهرا ولی نژاد
رتبه ۶۰۶
سجاد محمودی زاده
رتبه ۶۶۱
فاطمه اصغری
رتبه ۶۶۷
سیاوش مصطفایی
رتبه ۶۷۲
محمدماهان عنبرستانی
رتبه ۶۹۱
بهار ضرغامی
رتبه ۷۱۴
محمد یزدیان
رتبه ۷۸۱
احسان قنبری
رتبه ۷۸۶
نیما غفاری
رتبه ۸۰۳
مائده رنجبر
رتبه ۸۰۴
آرمین رضایی
رتبه ۸۹۳
فاطمه مشاوری نجف آبادی
رتبه ۹۰۹
کیمیا فدائی
رتبه ۹۴۷
صفورا بقائی
رتبه ۱۰۵۲
پویان فریور افشار
رتبه ۱۰۵۸
الینا جلالی فر
رتبه ۱۱۲۲
علی طاهر زاده
رتبه ۱۱۲۷
زهرا بابائی
رتبه ۱۲۳۴
مطهره توحیدی
رتبه ۱۲۳۶
مبینا ایزدی
رتبه ۱۲۸۴
بهار امیری
رتبه ۱۲۸۴
فاطمه معین زاده
رتبه ۱۳۵۰
علی زینلی
رتبه ۱۳۷۲
پارسا رضایی
رتبه ۱۴۲۴
سید امیرحسین حسینی
رتبه ۱۴۸۳
سینا خاوری خراسانی
رتبه ۱۴۹۳
محمد مهدی خرم زاده
رتبه ۱۵۰۳
فاطمه رحیم زاده
رتبه ۱۵۳۴
فاطیما عبیری
رتبه ۱۶۲۸
امیرمحمد فکور حقیقی
رتبه ۱۶۳۹
ابوالفضل نیرومند
رتبه ۱۶۷۸
سجاد ینکی
رتبه ۱۶۹۶
ندا ملکشاهی
رتبه ۱۷۳۱
محمدرضا محسنی
رتبه ۱۷۵۴
هلیا حاجیلوئی
رتبه ۱۹۶۶
مهسا رضایی مقدم
رتبه ۲۰۱۵
علی شیرزاد
رتبه ۲۵۵۹
سارا حمزه
رتبه ۲۶۲۵
زهره مجمعی
رتبه ۲۷۱۱
محمد غلامی
رتبه ۲۷۵۱
فهمیه سیدآبادی
رتبه ۲۷۸۱
سعید شبانی
رتبه ۲۷۹۴
مریم بادلی
رتبه ۲۸۱۰
هدیه رحیمی
رتبه ۲۸۸۱
پارسا جمال امیدی
رتبه ۳۱۳۳
صبا شایع ثانی
رتبه ۳۲۴۴
هلیا سجادی
رتبه ۳۳۴۳
سینا ارزمانی

فصل سوم: آمار توصیفی

☑ **توصیف و نمایش داده‌ها:** بعد از گردآوری داده‌ها باید به تنظیم، دسته‌بندی و خلاصه کردن آن‌ها پرداخت. برای این منظور دو روش وجود دارد: جدول فراوانی، رسم نمودار.

☑ **جدول فراوانی:** برای تشکیل دادن جدول فراوانی برای داده‌های جمع‌آوری شده، مراحل زیر را انجام می‌دهیم:

۱ـ بیش‌ترین و کم‌ترین داده را پیدا می‌کنیم (max و min). از تفاضل آن‌ها دامنه‌ی تغییرات داده‌ها را پیدا می‌کنیم که با R نمایش داده می‌شود: $R = Max - Min$

۲ـ اکنون دامنه تغییرات را به تعداد دلخواهی تقسیم می‌کنیم. تقسیم‌بندی قاعده خاصی ندارد. به این عمل دسته‌بندی داده‌ها می‌گوییم.

اگر تعداد دسته‌ها را با k نشان دهیم از تقسیم R بر k عددی پیدا می‌شود که طول دسته نامیده می‌شود و با C نمایش داده می‌شود. $C = \frac{R}{k}$

۳ـ پس از دسته‌بندی داده‌ها جدولی با ستون‌های مختلف رسم می‌کنیم:

درصد فراوانی نسبی	فراوانی نسبی	فراوانی مطلق	مرکز دسته	دسته‌ها

مرکز دسته (x_i): برای پیدا کردن مرکز دسته ابتدا و انتهای دسته را با هم جمع کرده سپس بر دو تقسیم می‌کنیم

فراوانی مطلق: تعداد داده‌هایی که در هر دسته قرار می‌گیرد را فراوانی مطلق نامیده و با f_i نشان می‌دهیم.

فراوانی نسبی: اگر فراوانی مطلق هر دسته را به مجموع تمام داده‌ها تقسیم کنیم فراوانی نسبی هر دسته حاصل می‌شود و با F_i نشان می‌دهیم.

درصد فراوانی نسبی: کافیست اعداد ستون چهارم را در ۱۰۰ ضرب کنیم.

مثال: جدول فراوانی داده‌های زیر را با ۵ دسته تشکیل دهید. ۱۰,۱۰,۱۲,۱۳,۱۵,۱۵,۱۷,۱۸,۲۰,۲۰,۲۲,۲۲
۲۳,۲۳,۲۵,۲۶,۲۸,۲۸,۲۸,۲۹,۳۰

$$Max = ۳۰ \quad , \quad min = ۱۰ \quad \Rightarrow \quad R = ۳۰ - ۱۰ = ۲۰ \quad \Rightarrow \quad C = \frac{۲۰}{۵} = ۴$$

راه‌حل:

درصد فراوانی نسبی	فراوانی نسبی	فراوانی مطلق	مرکز دسته	دسته
$\frac{۴}{۲۰} \times ۱۰۰ = ۲۰\%$	$\frac{۴}{۲۰}$	4	12	۱۰ − ۱۴
$\frac{۳}{۲۰} \times ۱۰۰ = ۱۵\%$	$\frac{۳}{۲۰}$	3	16	۱۴ − ۱۸
$\frac{۳}{۲۰} \times ۱۰۰ = ۱۵\%$	$\frac{۳}{۲۰}$	3	20	۱۸ − ۲۲
$\frac{۵}{۲۰} \times ۱۰۰ = ۲۵\%$	$\frac{۵}{۲۰}$	5	24	۲۲ − ۲۶
$\frac{۵}{۲۰} \times ۱۰۰ = ۲۵\%$	$\frac{۵}{۲۰}$	5	28	۲۶ − ۳۰
		$n = ۲۰$		

← دقت شود که نقاط انتهایی دسته‌ها به‌جز دسته آخر؛ جزو دسته‌ها به‌حساب نمی‌آیند.

تذکر: مفهوم درصد فراوانی نسبی: درصد فراوانی نسبی نشان می‌دهد که چند درصد داده‌ها در دسته‌ی موردنظر قرار دارند. برای نمونه دو مثال قبل ۲۰٪ داده‌ها در دسته اول و ۲۵٪ داده‌ها در دسته‌ی چهارم قرار دارند.

☑ رسم نمودار «نمودارهای میله‌ای، مستطیلی، دایره‌ای»:

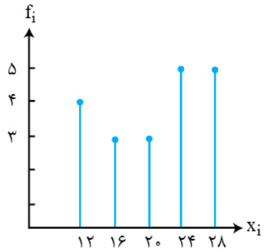
الف) نمودار میله‌ای
- محور x ها: مرکز دسته‌ها (x_i)
- محور y ها: فراوانی مطلق

ب) نمودار مستطیلی: «بافت نگاشت»
- محور x ها: حدود دسته‌ها
- محور y ها: فراوانی مطلق

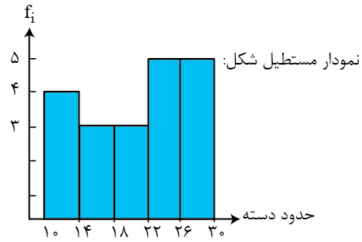
پ) نمودار دایره‌ای:
- زاویه دسته iام: $\frac{f_i}{n} \times ۳۶۰$
- فراوانی مطلق دسته iام: f_i

مثال: هرسه نمودار مثال قبل را رسم کنید:

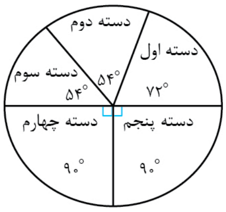
راه‌حل: نمودار میله‌ای:



نمودار مستطیل‌شکل:



نمودار دایره‌ای: ابتدا زاویه هر دسته را می‌یابیم:



$$a_1 = \frac{f_1}{n} \times 360 = \frac{4}{20} \times 360 = 72^\circ$$

$$a_2 = \frac{f_2}{n} \times 360 = \frac{3}{20} \times 360 = 54^\circ$$

$$a_3 = 54^\circ \qquad a_4 = \frac{5}{20} \times 360 = 90^\circ \qquad a_5 = 90^\circ$$

تذکر: هر ویژگی از اشیاء یا اشخاص که در اعضای جامعه یکسان نیست و معمولاً از یک عضو به عضو دیگر تغییر می‌کند را متغیر می‌نامیم.

انواع متغیرها: گسسته، پیوسته، کیفی ⇐ گسسته: مثل تعداد آدم‌ها. آدم نصفه نداریم ← پس ← عدد صحیح پیوسته: مثل وزن $49/87$ kg ← پس ← الزاماً عدد صحیح نیست.

مثال: جملات زیر را کامل کنید:

الف) برای متغیرهای پیوسته از نمودار .... استفاده می‌شود.

ب) برای متغیرهای گسسته از نمودارهای .... و .... استفاده می‌شود.

پ) برای متغیرهای کیفی از نمودارهای .... و .... استفاده می‌شود.

راه‌حل: **الف)** بافت نگاشت **ب)** دایره‌ای و میله‌ای **پ)** دایره‌ای و میله‌ای

☑ **معیارهای گرایش به مرکز:** در بخش اول فصل با دسته‌بندی داده‌ها و رسم نمودارها آشنا شدیم. اما هیچ کمیتی که بتواند به‌طور عددی یک جامعه را به ما معرفی کند ارائه نکردیم. در این بخش به معرفی کمیت‌هایی که به تمرکز داده‌ها و بخش مرکزی داده‌ها اشاره می‌دارند، خواهیم پرداخت.

☑ **مد (نما):** داده‌ای که بیش‌ترین فراوانی «تکرار» را داشته باشد. ممکن است مد یکتا نباشد، یعنی یک جامعه آماری می‌تواند دو یا چند مد داشته باشد.

مثال: در جدول فراوانی زیر مقدار مد کدام است؟

x_i	۱	۲	۳	۴
f_i	۸	۱۰	۳	۱۶

راه‌حل: عدد ۴ بیش‌ترین فراوانی را دارد، پس مد است.

☑ **میانگین:** اگر $x_1, x_2, \dots, x_n$ داده‌های آماری باشند آن‌گاه میانگین آن‌ها را با $\bar{x}$ نشان می‌دهند. و همچنین:

$$\bar{x} = \frac{x_1 + x_2 + \cdots + x_n}{n} = \frac{\sum_{i=1}^{n}}{n} = \frac{\text{مجموع اعداد}}{\text{تعداد اعداد}}$$

☑ **میانگین موزون داده‌ها:** اگر در داده‌های آماری، همه یا بعضی از داده‌ها دارای فراوانی «تکرار» باشند آن‌گاه از میانگین موزون استفاده می‌کنیم. در این روش باید هر داده را در فراوانی‌اش ضرب کنیم، سپس آن‌ها را با هم جمع کنیم و بر مجموع فراوانی‌ها تقسیم کنیم.

$$\bar{x}_w = \frac{x_1 f_1 + x_2 f_2 + \cdots + x_n f_n}{f_1 + f_2 + \cdots + f_n} \Rightarrow \bar{x}_w = \frac{\sum_{i=1}^{n} x_i f_i}{\sum_{i=1}^{n} f_i}$$

مثال: نمرات ریاضی، فیزیک، شیمی و عربی یک دانش‌آموز به‌ترتیب ۱۴ و ۱۵ و ۱۸ و ۲۰ هستند. اگر ضرایب این چهار دوست به‌ترتیب ۴، ۳ و ۲ و ۱ باشد مطلوب است میانگین موزون «معدل» نمرات دانش‌آموز؟

راه‌حل:

$$\bar{x}_w=\frac{۱۴\times۴+۱۵\times۳+۱۸\times۲+۲۰\times۱}{۴+۳+۲+۱}=\frac{۱۵۷}{۱۰}=۱۵/۷$$

☑ **میانه:** پس از مرتب کردن داده‌های آماری (به‌صورت صعودی)، میانه داده‌ای است که از نیمی از داده‌ها بیش‌تر و از نیم دیگر آن‌ها کم‌تر باشد. برای یافتن میانه ابتدا داده‌ها را از کوچک به بزرگ مرتب می‌کنیم. اگر تعداد داده‌ها فرد باشد، عدد وسط میانه است و اگر تعداد داده‌ها زوج باشد، میانگین دو عدد وسط میانه است. میانه را با $Q_۲$ نشان می‌دهیم.

اگر از داده‌های قبل از میانه مجدداً میانه بگیریم، چارک اول داده‌ها $(Q_۱)$ و اگر از داده‌های بعد از میانه مجدداً میانه بگیریم چارک سوم داده‌ها $(Q_۳)$ را حساب کرده‌ایم.

مثال: در داده‌های زیر مطلوب است میانه، $Q_۱$ و $Q_۳$؟ $۱,۲,۴,۴,۸,۸,۹,۱۰,۱۰,۲۰,۳۰,۴۰$

$Q_۲=\frac{۸+۹}{۲}=۸/۵\rightarrow$ (میانه میانگین دو داده وسط است، یعنی داده ششم و هفتم)

راه‌حل: داده‌ها مرتب هستند خداروشکر! ☺

$۱,۲,۴,۴,۸,۸\rightarrow Q_۱=\frac{۴+۴}{۲}=۴$ $\qquad$ $۹,۱۰,۱۰,۲۰,۳۰,۴۰\rightarrow Q_۳=\frac{۱۰+۲۰}{۲}=۱۵$

تذکر: حواستون باشه که $Q_۱$ و $Q_۲$ و $Q_۳$ می‌تونه یکی از داده‌ها باشند یا نباشند! ☺

☑ **معیارهای پراکندگی:** هرگاه بخواهیم میزان دوری یا نزدیکی داده‌ها به یکدیگر را بررسی کنیم از شاخص‌های پراکندگی استفاده می‌کنیم، هر چقدر مقدار عددی این شاخص‌ها کم‌تر باشد بیانگر این موضوع است که داده‌ها به هم نزدیک‌ترند. پس اگر همه داده‌ها با هم برابر بودند مقدار شاخص‌های پراکندگی باید صفر باشند. مهم‌ترین شاخص‌های پراکندگی عبارتند از: انحراف معیار، واریانس، ضریب تغییرات

☑ **انحراف معیار:** یک معیار سنجش برای میزان پراکندگی داده‌ها حول میانگینشان است. اگر انحراف معیار را به توان ۲ برسانیم واریانس به‌دست می‌آید. انحراف معیار را با σ و واریانس را با $\sigma^۲$ نمایش می‌دهیم.

اگر n داده به‌صورت $x_۱,x_۲,\dots,x_n$ با میانگین $\bar{x}$ داشته باشیم، واریانس و انحراف معیار از رابطه زیر محاسبه می‌شود.

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

$$\text{انحراف معیار}=\sqrt{\text{واریانس}}=\sigma$$

تذکر:
- σ کوچک‌تر ⇔ پراکندگی حول میانگین کم‌تر ⇔ داده‌ها به هم نزدیک‌تر
- σ بزرگ‌تر ⇔ پراکندگی حول میانگین بیش‌تر ⇔ داده‌ها از هم دورتر

مثال: برای داده‌های روبه‌رو مطلوب است واریانس و انحراف معیار: داده‌ها: $۲,۳,۵,۷,۸$

راه‌حل:

$$\bar{x}=\frac{۲+۳+۵+۷+۸}{۵}=\frac{۲۵}{۵}=۵ \qquad \sigma^۲=\frac{(۲-۵)^۲+(۳-۵)^۲+(۵-۵)^۲+(۷-۵)^۲+(۸-۵)^۲}{۵}$$

$$\Rightarrow\sigma^۲=۵/۲\quad\rightarrow\quad\sigma=\sqrt{۵/۲}$$

☑ **ضریب تغییرات:** از تقسیم انحراف معیار (G) بر میانگین داده‌ها ($\bar{x}$) عددی به‌دست می‌آید که ضریب تغییرات نام دارد و با CV نمایش داده می‌شوند. $CV=\frac{G}{\bar{x}}$

مثال: اگر ضریب تغییرات ده داده برابر ۲ و میانگین آن‌ها ۴ باشد واریانس داده‌ها چقدر است؟

راه‌حل: $CV=۲ \quad n=۱۰ \quad \bar{x}=۴ \quad \Rightarrow CV=\frac{\sigma}{\bar{x}} \Rightarrow ۲=\frac{\sigma}{۴} \Rightarrow \sigma=۸ \Rightarrow \sigma^۲=۶۴$

تذکر ۱: اگر داده‌ها k برابر شوند. آن‌گاه تمام شاخص‌های مرکزی و انحراف معیار k برابر شده. واریانس $k^۲$ برابر می‌شود ولی ضریب تغییرات ثابت می‌ماند. «با شرط $k>۰$»

تذکر ۲: اگر داده‌ها با k جمع شوند آن‌گاه تمامی شاخص‌های مرکزی با k جمع می‌شوند اما انحراف معیار و واریانس تغییری نمی‌کنند. ضریب تغییرات نیز به شکل مقابل تغییر می‌کند: $CV_{\text{جدید}}=\frac{\sigma}{\bar{x}+k}$

مثال: اگر ۲۰ داده آماری را دو برابر کرده پس ۷ واحد از هرکدام کم کنیم ضریب تغییرات ۱/۵ برابر می‌شود. مجموع داده‌های قبلی چقدر بوده است؟

راه‌حل: اگر داده‌ها را ۲ برابر کنیم هم میانگین و هم انحراف معیار ۲ برابر می‌شوند. و اگر داده‌ها ۷ واحد کم کنیم فقط از میانگین ۷ واحد کم می‌شود. $\sigma_{\text{جدید}}=۲\sigma_{\text{قبلی}} \qquad \Rightarrow \bar{x}_{\text{جدید}}=۲\bar{x}_{\text{قبلی}}-۷$

حالا وقت آن است که فرمول ضریب تغییرات را بنویسیم:

$$\frac{CV_{\text{جدید}}}{CV_{\text{قبلی}}}=\frac{\frac{۲\sigma}{۲\bar{x}-۷}}{\frac{\sigma}{\bar{x}}}=\frac{۳}{۲}\Rightarrow \bar{x}=۱۰/۵$$

☑ **نمودار جعبه‌ای:** برای رسم این نمودار از ۵ مقدار استفاده می‌کنیم:

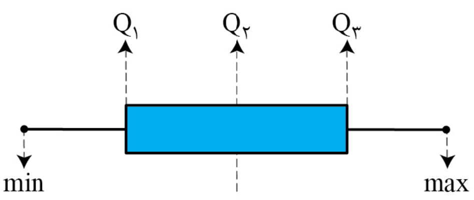
۱ـ min

۲ـ max

۳ـ $Q_۱$

۴ـ $Q_۲$

۵ـ $Q_۳$

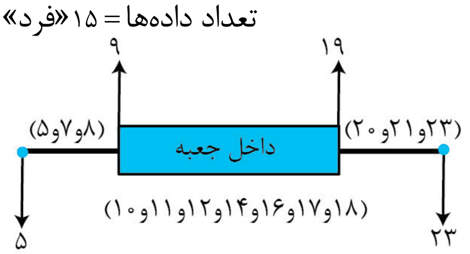


مثال: نمودار جعبه‌ای داده‌های روبه‌رو را رسم کنید: ۵,۷,۸,۹,۱۰,۱۱,۱۲,۱۴,۱۶,۱۷,۱۸,۱۹,۲۰,۲۱,۲۳

راه‌حل: ابتدا ۵ مقدار موردنظر را طبق نکته قبل می‌یابیم: $max=۲۳ \qquad min=۵$

تعداد داده‌ها = ۱۵ «فرد» $Q_۲=۱۴$ داده‌های قبل از میانه: $۵,۷,۸,\boxed{۹},۱۰,۱۱,۱۲ \to Q_۱=۹$

داده‌های بعد از میانه: $۱۶,۱۷,۱۸,\boxed{۱۹},۲۰,۲۱,۲۳ \to Q_۳=۱۹$



☑ **نرخ بیکاری:**

تعریف جامعه بیکار: به افراد ۱۰ساله و یا بالاتر از ۱۰ ساله‌ای گفته می‌شود که ۳ شرط زیر را توأم با هم داشته باشند

✓ در هفته مشخص حتی یک ساعت هم کار نکرده‌اند.

✓ آمادگی انجام کار دارند.

✓ در هفته مشخص و ۳ هفته قبل از آن جویای کار بوده‌اند.

تعریف جامعه شاغل: به افراد ۱۰ساله‌ای که در طول هفته مشخص حداقل یک ساعت کار کرده باشند.

جمعیت فعال = جمعیت بیکار + جمعیت شاغل