۷- شاخص های پراکندگی

تفاوت بین داده ها از لحاظ مقداری یکی دیگر از ویژگی های داده ها برای توصیف آن ها و مقایسه چند دسته داده است. در این فصل مقادیری برای داده ها محاسبه می کنیم که این ویژگی را در خود خلاصه می کنند.

مقداری که تفاوت داده ها را نسبت به هم و یا نسبت به مرکز نشان می دهد شاخص پراکندگی گوییم.

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

۱.۱ دامنه 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 می دهیم:

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

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

الف) ۲۰، ۱۵، ۱۵، ۱۵، ۱۴، 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. داده های دسته الف به جز داده های تفاوتی یک یا نیم واحدی دارند ولی داده های دسته ب تفاوتی بیش از دو واحد دارند. به داده های ۴,۲, ۲۰ در دسته داده های الف، داده های 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 می باشد. بنابراین دامنه چارکی برابر است با ۳ = ۱۲ − ۱۵ که به خوبی پراکندگی داده های دسته الف را نشان می دهد.

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 می ماند ولی با ضرب یک مقدار مخالف صفر در داده ها، دامنه تغییرات در مقدار مثبت آن مقدار ضرب می شود.

مثال۲: دامنه ی تغییرات داده‌های آماری ۵, 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 می کنیم، هر کدام کوچکتر است پراکندگی کمتر ودقت عملکرد بیشتر است.

$$\left.\begin{array}{l} R_A = ۱۹ - ۱۴ = ۵ \\ R_B = ۱۸ - ۱۴ = ۴ \end{array}\right\} \Rightarrow R_B < R_A$$

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: $n = ۱۲ \rightarrow R = kc = ۱۲ \times ۰/۴ = ۴/۸$

مثال۶: دامنه تغییرات داده‌های $۳, -a^۲, ۴ + a^۲, ۱,۲$ سه برابر میانه است، میانه داده‌های $a^۲, ۲a^۴, ۳a^۶, ۴a^۸$ کدام است؟

√۱) ۲/۵   ۲) ۵   ۳)۱۰   ۴) ۱۲/۵

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

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

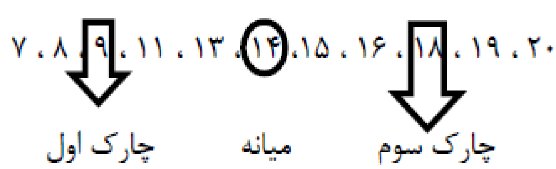
داده های جدید $= (\pm ۱)^۲, ۲(\pm ۱)^۴, ۳(\pm ۱)^۶, ۴(\pm ۱)^۸$

داده‌های جدید : ۱ ,۲ ,۳ ,۴ → 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: داده های مرتب شده:



داده های داخل جعبه: ۱۱ ،۱۳ ،۱۴ ،۱۵ ،۱۶

دامنه: ۱۶ - ۱۱ = ۵

انحراف از میانگین:

تفاضل هر داده از میانگین را انحراف از میانگین آن داده گویند. اگر $x_۱, x_۲, \dots, x_n$ داده‌ها باشند آنگاه انحراف از میانگین داده ها به صورت زیر است:

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

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

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

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

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

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 می دهند. 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 داده های 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: $\bar{x}=\frac{۲+۳+۵+۵+۵+۴+۴}{۷}=\frac{۲۸}{۷}=۴$

$$\sigma^۲=\frac{(۲-۴)^۲+(۳-۴)^۲+(۵-۴)^۲+(۵-۴)^۲+(۵-۴)^۲+(۴-۴)^۲+(۴-۴)^۲}{۷}=\frac{۴+۱+۱+۱+۱}{۷}=\frac{۸}{۷}$$

مثال۱۰: در ۵ داده‌ی آماری انحراف از میانگین داده‌ها برابر a , ۲, ۰, ۱-, ۳- می باشد. 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. $-۳+(-۱)+۰+۲+a=۰\to a=۲$

$$\sigma^۲=\frac{\sum(x_i-\bar{x})^۲}{n}=\frac{(-۳)^۲+(-۱)^۲+۰^۲+۲^۲+۲^۲}{۵}=\frac{۹+۱+۴+۴}{۵}=\frac{۱۸}{۵}=۳/۶$$

دستور دیگری برای محاسبه واریانس:

$\sigma^{2} =$ *(میانگین مجذورداده‌ها) − (مجذور میانگین داده‌ها)*

$$\sigma^{2} = \frac{(x_1)^{2}+(x_2)^{2}+\dots+(x_n)^{2}}{n} - (\bar{x})^{2} = \frac{\sum_{i=1}^{n}(x_i)^{2}}{n} - (\bar{x})^{2}$$

توجه کنید در $x_1^2 + x_2^2 + \cdots + x_n^2$ ابتدا مقادیر x_i ها را به توان ۲ می رسانیم و سپس آنها را با هم جمع می کنیم و در $\bar{x}^{2}$ میانگین داده‌ها را به توان ۲ می رسانیم.

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؟

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

$$x_1^{۲} + x_2^{۲} + \cdots + x_5^{۲} = ۱^{۲}+۵^{۲}+۶^{۲}+۷^{۲}+۹^{۲} = ۱۹۲$$

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

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.*

۷.۳انحراف معیار

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

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 ۹۰/۰۶۴ و میانگین آنها ۳ می باشد. انحراف معیار این داده‌ها را بدست آورید.

جواب: ۰/۸٪

مثال۱۴: اگر میانگین داده‌های $x_۱, x_۲, \ldots, x_۵$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؟

$\bar{x} = ۲, \sigma = ۰/۳$

$$\sigma^۲ = \frac{x_۱^۲+x_۲^۲+\ldots+x_{۱۰}^۲}{۱۰} - \bar{x}^۲ \Rightarrow (۰/۳)^۲ = \frac{x_۱^۲+x_۲^۲+\ldots+x_{۱۰}^۲}{۱۰} - ۲^۲ \Rightarrow \frac{x_۱^۲+x_۲^۲+\ldots+x_{۱۰}^۲}{۱۰} = ۴/۰۹$$

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

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

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

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

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

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

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

$$\sigma_{x\pm a}^۲ = \sigma_x^۲ \quad , \quad \sigma_{x\pm a} = \sigma_x$$

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

جواب: واریانس : ۵×$۳^۲$= ۴۵ , انحراف معیار : $\sqrt{۵}$ ×۳

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

جواب: واریانس : ۵×$(-۴)^۲$ = ۸۰ , انحراف معیار : $\sqrt{۵}$ |۴| = $\sqrt{۵}$ ۴

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

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

مثال۱۶: انحراف معیار داده‌های $x_1, x_2, ..., x_{10}$ مساوی ۵ است. انحراف معیار داده‌های $-۲x_1+۵$, $-۲x_2+۵$, ..., $-۲x_{10}+۵$ کدام است؟

۱) ۵-　　۲) ۱۰-　　۳) ۱۰ √　　۴) ۵

حل: جدید$\sigma^2 = |-۲| \times ۵ = ۱۰$

مثال۱۷: انحراف معیار داده‌های $x_1, x_2, ..., x_n$ مساوی ۴ است. واریانس داده های $۳x_1-۲$, $۳x_2-۲$, ..., $۳x_n-۲$ کدام است؟

۱) ۱۶　　۲) ۳۶　　۳) ۴۸　　۴) ۱۴۴√

حل: $\sigma = ۴ \rightarrow \sigma^2 = ۱۶ \rightarrow$ جدید$\sigma^2 = ۳^2 \times ۱۶ = ۱۴۴$

مثال۱۸: داده‌های آماری با میانگین ۸ و واریانس ۲/۲۵ موجود است، تمام داده‌ها را دو برابر می کنیم 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: قدیم $\sigma^2 = ۲/۲۵ \rightarrow$ جدید σ^2 =۴× ۲/۲۵ =۹ → جدید $\sigma = \sqrt{۹}$ =۳

روش دوم: $\sigma^2 = ۲/۲۵ \rightarrow$ قدیم $\sigma = ۱/۵ =>$ جدید σ =۲×۱/۵ =۳

مثال۱۹: اگر انحراف معیار قیمت‌های چند کالا در سال گذشته برابر ۳۰ و امسال ٪۱۰ به قیمت‌ها افزوده شده باشد واریانس قیمت-های امسال آن کالاها کدام است؟

۱) ۱۰۸۹ ✓   ۲) ۹۰۰   ۳) ۳۳   ۴) ۳۰

حل: فرض می کنیم 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 می شوند. بنابراین داریم:

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

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 می کنیم تا داده های $-۲, ۲, ۰, ۰$ به دست آیند. انحراف معیار داده های جدید با انحراف معیار داده های 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 داده های آماری روبرو واریانس داده های بین چارک اول و چارک سوم کدام است؟(سراسری انسانی۹۵)

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

۱) ۱۷/۲۴

۲) ۱۷/۸۲

۳) ۱۸/۰۲

۴) ۱۸/۴۴ ✓

حل: داده ه را از روی نمودارمی نویسیم:

۳۲، ۳۳، ۳۴، ۳۴، ۳۶، ۳۹، ۴۰، ۴۱، ۴۳، ۴۵، ۴۵، ۴۷، ۵۱، ۵۲، ۵۴، ۵۷، ۵۸

$Q_۱ = \frac{۳۴+۳۶}{۲} = ۳۵$ میانه $Q_۳ = \frac{۵۱+۵۲}{۲} = ۵۱.۵$

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

۴۳ را از داده ها کم می کنیم : ۸، ۴، ۲، ۲، ۰، ۲-، ۳-، ۴-، ۷-

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

$$\sigma^۲ = \frac{(-۷)^۲+(-۴)^۲+(-۳)^۲+(-۲)^۲+(۰)^۲+(۲)^۲+(۲)^۲+(۴)^۲+(۸)^۲}{۹} - ۰^۲ = \frac{۱۶۶}{۹} = ۱۸.۴۴$$

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 نسبت d بدهند واریانس از فرمول زیر بدست می آید:*
>
> $$\sigma^۲ = \frac{n^۲-۱}{۱۲} d^۲$$

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

✓ ۱) $\sqrt{۲}$ ۲) ۲ ۳) ۱ ۴) ۴

حل: n = ۵ داده تشکیل دنباله عددی با قدر نسبت$d = ۱$ می دهند بنابراین داریم: $\sigma^۲ = \frac{۵^۲-۱}{۱۲}(۱)^۲ = ۲ \to \sigma = \sqrt{۲}$

تمرین۵: پنج داده‌ی آماری تشکیل تصاعد عددی با قدر نسبت ۶ می دهند، 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$a = b = c = d = ۵ \leftarrow \bar{x} = \frac{۳\times۵+۲\times۱۰}{۶} = \frac{۴۰}{۶} = \frac{۲۰}{۳}$

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 می کنیم. فرمول های 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 و۱۸ برابر با ۱۵ است بنابراین میانگین داده های جدید تغییری نکرده و همان ۱۵ است.

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

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 می باشد، 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 داده های ناجور برابر ۳۰ همان میانگین قبلی است پس میانگین داده های جدید تغییر نمی کند.

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

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

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.

۷.۴واریانس در جدول فراوانی

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

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

که در آن ؛

f_i : فراوانی داده(دسته) i ام

x_i : داده(مرکز دسته)i ام

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

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

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

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

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

مرکزدسته	۱۲	۱۵	۱۸	۲۱	۲۴
فراوانی	۴	۳	۹	۷	۲

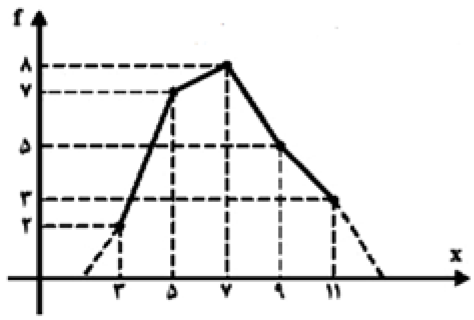
حل: از همه داده‌ها ۱۸ واحد کم می کنیم

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

$$\bar{x}=\frac{-۲۴-۹+۰+۲۱+۱۲}{۲۵}=۰\rightarrow \sigma^۲=\frac{۴(-۶-۰)^۲+۳(-۳-۰)^۲+۹(۰-۰)^۲+۷(۳-۰)^۲+۲(۶)^۲}{۲۵}$$
$$=\frac{۱۴۴+۲۷+۶۳+۷۲}{۲۵}=\frac{۳۰۶}{۲۵}=۱۲.۴۴$$

مثال۲۸: با توجه به مودار 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	۲	۷	۸	۵	۳

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

$$s^۲=\frac{۲(-۴)^۲+۷(-۲)^۲+۰+۵(۲)^۲+۳(۴)^۲}{۲+۷+۸+۵+۳}=\frac{۳۲+۲۸+۲۰+۴۸}{۲۵}=\frac{۱۲۸}{۲۵}=۵/۱۲$$

تمرین ۷: واریانس داده‌های جدول زیر کدام است؟

۱)۲ ۲) ۲/۲۵ ۳) ۲/۵ ۴) ۲/۷۵

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

تذکر: *اگر حدود دسته‌ها را به ما بدهند ابتدا مرکز دسته‌ها را بدست می آوریم.*

مثال۲۹: جدول زیر مقادیر انحراف از میانگین ۱۲ داده آماری دسته بندی شده را مشخص می کند 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.

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

۷.۵ضریب تغییرات:(C.V)

آخرین شاخص پراکندگی ضریب تغییرات است. یک از مشکلات استفاده از شاخص های پراکندگی بیان شده در مقایسه پراکندگی دو یا چند دسته داده با واحدهای های اندازه گیری مختلف است بعنوان مثال برای مقایسه پراکندگی وزن(برحسب کیلوگرم) و قد(برحسب سانتی متر) دانش آموزان کلاس استفاده از شاخص های 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 می نامند که بدون واحد است:

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

تذکر: *$\bar{X}$ 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 می برند.*

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 می گیریم. 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:

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

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 می باشد. 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 می گیریم.

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

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

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

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

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

$$CV = \frac{\sigma}{\bar{x}} \rightarrow \frac{۱}{۲} = \sigma \; \frac{\sigma}{\bar{x}} \rightarrow \bar{X} = ۲ \xrightarrow{\text{به تمام داده ها ۳ واحد اضافه می کنیم}} \bar{x} = ۲ + ۳ = ۵$$

- *اگرداده‌های آماری را در عدد مثبتی ضرب شوند، CV ثابت می ماند.*

$$\text{جدید } CV = \frac{|a|\sigma}{a\bar{x}} = \frac{a\sigma}{a\bar{x}} = \frac{\sigma}{\bar{x}} = CV \text{ قدیم}$$

- *اگر عدد مثبتی به داده‌ها اضافه شود انحراف معیار تغییری نمی کند ولی میانگین افزایش می یابد، بنابر این CV کاهش می یابد و برعکس.*

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

✓ ۱) ۰/۴　　　۲) ۰/۴۸　　　۳) ۰/۵۲　　　۴) ۰/۶

حل: میانگین داده های 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 می کنیم 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:

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

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

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

۱)بیشتر می شود.　　۲)کمتر می شود　　✓ ۳)ثابت می ماند　　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 می شوند بنابراین ضریب تغییرات ثابت می ماند.

مثال۳۷: در ۶۰ داده‌ی آماری، میانگین ۳ و انحراف معیار ۱/۲ است اگر به تمام داده‌ها ۹ واحد اضافه شود ضریب تغییرات داده-های 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 داده های 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:

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

۷.۶مقایسه ی دقت یا پراکندگی دو گروه

برای مقایسه‌ی دو گروه ابتدا میانگین داده‌ها را بدست می آوریم که به یکی از دو حالت زیر برخورد می کنیم:

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

حالت۲: اگر میانگین دو گروه با هم برابر نباشد ضریب تغییرات را بدست می آوریم. 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 می کنیم.

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

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

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 می گیریم.　　داده های B = ۶۳+ داده های A

$$\begin{cases} A: ۱,۲,۳,۴,۵ & \begin{cases} \bar{x}_A = ۳ \\ \sigma_A = \sqrt{۲} \end{cases} \Rightarrow CV_A = \frac{\sqrt{۲}}{۳} \\ B: ۶۱,۶۲,۶۳,۶۴,۶۵ & \begin{cases} \bar{x}_B = ۶۳ \\ \sigma_B = \sqrt{۲} \end{cases} \Rightarrow CV_B = \frac{\sqrt{۲}}{۶۳} \end{cases} \Rightarrow CV_A > CV_B$$