جمع‌بندی نکته‌های مبحث آمار

مفاهیم اولیه

دسته‌بندی داده ها و جدول

جامعه‌ی آماری: به مجموعه‌ای از اشیاء که می‌خواهیم اطلاعاتی درباره‌ی آنها به‌دست آوریم گوییم، به تعداد اعضای جامعه، اندازه‌ی جامعه گوییم، یک جامعه می‌تواند محدود یا نامحدود، شمارا یا غیر قابل شمارش باشد، به عنوان مثال شرکت‌کنندگان در مسابقه‌ی پرش در یک مسابقه، یک جامعه‌ی متناهی شمارا خواهد بود، اما اعداد حقیقی بازه‌ی $[۳ , ۴]$ شمارا نیست.

متغیر تصادفی: مشخصه‌ی ویژه‌ای از افراد جامعه را که می‌خواهیم مورد مطالعه قرار دهیم، متغیر تصادفی می‌نامیم، به عنوان مثال ارتفاع پرش ورزشکاران در مسابقه‌ی پرش ارتفاع ، یک متغیر تصادفی است، متغیرها به دو دسته‌ی کمی و کیفی تقسیم می‌شوند:

الف – متغیر کمی
- ۱- متغیر کمی پیوسته: این نوع متغیرها قابل اندازه‌گیری‌اند، مانند حجم، طول، وزن و ...
- ۲- متغیر کمی گسسته: که قابل شمارش هستند، مانند تعداد افراد دارای مدرک لیسانس در یک اداره

ب – متغیر کیفی
- ۱- متغیر کیفی ترتیبی: در این نوع متغیرها نوعی ترتیب مطرح است، مانند مراحل تحصیلی، مراحل رشد یک انسان
- ۲- متغیر کیفی اسمی: در این نوع متغیرها ترتیب مطرح نیست، مانند گروه خونی افراد، سردی و گرمی هوا

بنابراین به‌طور کلی متغیرهای کمی، قابل اندازه‌گیری یا شمارش هستند ولی متغیرهای کیفی قابل اندازه‌گیری یا شمارش نیستند.

نمونه: اگر به علت وسعت جامعه، نتوانیم تمام آن را مطالعه کنیم، بخشی از آن را به عنوان **نمونه** انتخاب می‌کنیم، اطلاعات حاصل از نمونه را داده می‌نامیم و با x_i نمایش می‌دهیم، تعداد افراد یا اشیایی را که در نمونه انتخاب می‌کنیم، حجم یا اندازه‌ی نمونه می‌نامیم. آمارگیری را به دو روش انجام می‌دهیم، یک به روش **نمونه‌گیری** که بخشی از جامعه را مورد مطالعه قرار می‌دهیم و دیگری به روش **سرشماری** که تمام افراد جامعه مورد مطالعه قرار می‌گیرند.

دسته‌بندی داده‌ها و جدول

برای دسته‌بندی و تحلیل داده‌ها نیاز به تعاریف زیر داریم:

①**دامنه‌ی تغییرات:** فاصله‌ی کمترین داده‌ها (a) و بیشترین داده‌ها (b) را دامنه‌ی تغییرات می‌نامیم.

$$\boxed{R = b - a}$$

در مثال نمونه‌ی ۱: دامنه‌ی تغییرات $R = ۵ - ۱ = ۴$ است.

در مثال نمونه‌ی ۲: دامنه‌ی تغییرات $R = ۳۹ - ۴ = ۳۵$ است.

② **فاصله‌ی طبقات:** در اغلب موارد (همانند مثال نمونه‌ی ۲) نیاز داریم که داده‌ها را دسته‌بندی می‌کنیم، در این صورت اگر تعداد طبقات را K و فاصله‌ی طبقات یا طول دسته‌ها را با C نمایش دهیم، آنگاه:

$$\text{طول طبقات} : C = \frac{R}{K}$$

$$\text{تعداد طبقات} : K = \frac{R}{C}$$

■ **مثال نمونه‌ی ۱:** تعداد افراد ساکن در هر آپارتمان یک مجتمع ۶ واحدی، به‌صورت زیر آمارگیری شده است:

۳ , ۳ , ۱ , ۲ , ۵ , ۴

■ **مثال نمونه‌ی ۲:** در یک روز بهاری، دمای هوای ۳۰ استان کشور برحسب درجه به‌صورت زیر آمارگیری شده است:

۴ , ۵ , ۶ , ۸ , ۱۱ , ۱۱ , ۱۳ , ۱۳ , ۱۳ , ۱۵ , ۱۵ , ۱۵ , ۱۵ , ۱۷
۱۸ , ۱۸ , ۲۰ , ۲۰ , ۲۰ , ۲۱ , ۲۲ , ۲۵ , ۲۷ , ۲۷ , ۳۱ , ۳۱ , ۳۳
۳۷ , ۳۸ , ۳۹

در این صورت برای داده‌های با می‌نیمم a و ماکزیمم b، در طبقه‌ی اول، داده‌ها در بازه‌ی $[a , a + C)$ و در طبقه‌ی دوم، داده‌ها در بازه‌ی $[a + C , a + ۲C)$ و ... قرار می‌گیرند، به $a + C$ در این طبقه، کران پایین طبقه‌ی دوم و به $a + ۲C$ در این طبقه، کران بالای طبقه گوییم.

■ **مثال:** اگر در مثال نمونه‌ی ۲، تعداد طبقات را ۶ فرض کنیم، طبقات را بنویسید.

◄ **حل:** از آنجایی که $K = ۶$ است، پس $C = \frac{۳۵}{۶} \approx ۵/۸۳$ که به ۶ گرد می‌کنیم، لذا طبقات به‌صورت زیر خواهند بود.

$[۴ , ۱۰)$ و $[۱۰ , ۱۶)$ و $[۱۶ , ۲۲)$ و $[۲۲ , ۲۸)$ و $[۲۸ , ۳۴)$ و $[۳۴ , ۴۰)$

جمع‌بندی نکته‌های مبحث آمار

③ **مرکز دسته (نشان دسته):** در هر طبقه میانگین کران بالا و کران پایین طبقه را، مرکز دسته می‌نامیم و با x_i نمایش می‌دهیم، پس:

$$\text{مرکز دسته‌ی } i \text{ام} : x_i = \frac{\text{کران پایین} + \text{کران بالا}}{۲}$$

❖ **نتیجه‌ی ۱:** اگر C طول دسته و x_i، نشان دسته‌ی iام باشد، آنگاه:

$$\text{دسته‌ی } i \text{ام} = \left[x_i - \frac{C}{۲} , x_i + \frac{C}{۲}\right)$$

بنابراین کران بالا $x_i + \frac{C}{۲}$ و کران پایین $x_i - \frac{C}{۲}$ خواهد بود.

❖ **نتیجه‌ی ۲:** در طبقه‌بندی داده‌ها همواره خواهیم داشت:

اختلاف دو کران بالای متوالی = اختلاف دو کران پایین متوالی = اختلاف دو مرکز متوالی = طول دسته

■ **مثال:** در مثال نمونه‌ی ۲، مراکز دسته‌ها عبارتند از ۳۷ , ۳۱ , ۲۵ , ۱۹ , ۱۳ , ۷ .

④ **فراوانی مطلق:** به تعداد دفعاتی که یک داده‌ی آماری تکرار می‌شود، فراوانی مطلق گوییم و با f_i نمایش می‌دهیم.

در مثال نمونه‌ی ۱، فراوانی تعداد ساکنین سه نفره در آپارتمان‌ها، ۲ است.

در مثال نمونه‌ی ۲، فراوانی مطلق دسته‌ی دوم، یعنی بازه‌ی [۱۰ , ۱۶)، ۹ است، یعنی $f_۲ = ۹$.

⑤ **فراوانی نسبی:** نسبت فراوانی مطلق هر دسته به کل داده‌ها را فراوانی نسبی آن دسته می‌نامیم و با F_i نمایش می‌دهیم، لذا:

$$F_i = \frac{f_i}{n} \quad \text{و} \quad n = \sum f_i$$

و درصد فراوانی نسبی برابر $۱۰۰ \times \frac{f_i}{n}$ است.

➢ **تذکر (۱):** در یک جدول توزیع فراوانی، مجموع فراوانی‌های نسبی همواره ۱ است و جمع درصد فراوانی‌های نسبی ۱۰۰ است.

⑥ **فراوانی تجمعی:** برابر است با فراوانی آن دسته بعلاوه‌ی مجموع فراوانی‌های دسته‌های قبل از آن و آن را با F_{c_i} نمایش می‌دهیم، بنابراین:

❖ **نتیجه‌ی ۱:** فراوانی مطلق طبقه‌ی اول برابر فراوانی تجمعی طبقه‌ی اول است.

❖ **نتیجه‌ی ۲:** فراوانی تجمعی طبقه‌ی آخر برابر مجموع فراوانی‌ها یا کل داده‌هاست.

❖ **نتیجه‌ی ۳:** اگر فراوانی تجمعی طبقه‌ی ماقبل را از فراوانی تجمعی آن طبقه کم کنیم، فراوانی مطلق آن طبقه به‌دست می‌آید:

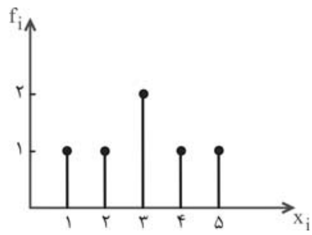
$$f_i = F_{c_i} - F_{c_{i-۱}}$$

نمودارهای آماری

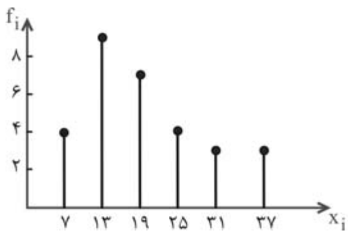
نمودارهای آماری وسیله‌هایی هستند برای مقایسه‌ی سریع‌تر و تحلیل مناسب‌تر اطلاعات، در زیر مهم‌ترین آنها را می‌بینیم.

① **نمودار میله‌ای**

در این نمودار، روی محور طول‌ها نماینده‌ی طبقه (مرکز دسته) یا خود داده و روی محور عرض‌ها، فراوانی مطلق هر طبقه یا داده را مشخص می‌کنیم. این نمودار بیشتر برای متغیرهای گسسته مناسب است اما در متغیرهای پیوسته با استفاده از مرکز دسته‌ها از آن استفاده می‌کنیم. در شکل‌های رو به رو، نمودار میله‌ای مثال‌های نمونه‌ی ۱ و ۲ رسم شده است.



(مثال ۱)

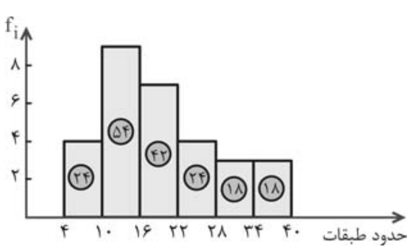


(مثال ۲)

② **نمودار مستطیلی**

در این نمودار، روی محور طول‌ها حدود طبقات و روی محور عرض‌ها فراوانی مطلق هر طبقه را قرار می‌دهیم، این نمودار برای متغیرهای پیوسته مناسب است. در شکل رو به رو، نمودار مستطیلی مربوط به مثال نمونه‌ی ۲ رسم شده است. با توجه به شکل مساحت سطح هاشورخورده برابر است با:

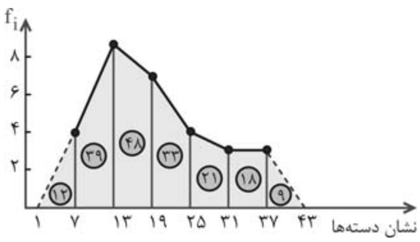
$$S = ۲۴ + ۵۴ + ۴۲ + ۲۴ + ۱۸ + ۱۸ = ۱۸۰$$



جمع‌بندی نکته‌های مبحث آمار

③ نمودار چند بر فراوانی

در این نمودار، روی محور طول‌ها مراکز دسته‌ها (یا خود داده‌ها) و روی محور عمودی فراوانی مطلق نظیر آن را مشخص می‌کنیم و نقاط به دست آمده را به هم وصل می‌کنیم، دو دسته‌ی مجازی با فراوانی ($f_i = 0$) به ابتدا و انتهای مراکز اضافه می‌کنیم تا سطح چندبر فراوانی به‌دست آید.

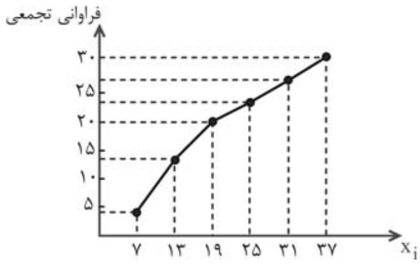


$S = 12 + 39 + 48 + 33 + 21 + 18 + 9 = 180$

❖ **نتیجه‌ی ۱:** با توجه به نمودارها دیده می‌شود که سطح زیر نمودار چند بر فراوانی با مجموع مساحت مستطیل‌های نمودار مستطیلی برابر است.

④ نمودار فراوانی تجمعی

در این نمودار، بر روی محور xها نشان دسته‌ها و بر روی محور yها، فراوانی تجمعی هر دسته را در نظر می‌گیریم و نقاط به‌دست آمده را به هم وصل می‌کنیم، در مثال نمونه‌ی ۲ شکل رو به رو را خواهیم داشت.

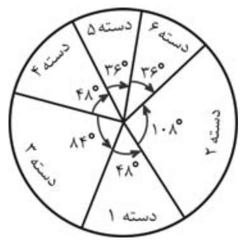


نکته
۱- نمودار فراوانی تجمعی همواره صعودی است.
۲- اگر قسمتی از نمودار فراوانی تجمعی به‌صورت خطی موازی محور x ها باشد، مفهوم آن این است که فراوانی مطلق آن طبقه صفر است.

⑤ نمودار دایره‌ای

اگر تعداد دسته‌ها در یک جامعه‌ی آماری K باشد، در این صورت سطح دایره رابه K قطاع متناسب با فراوانی نسبی هر دسته تقسیم می‌کنیم، زاویه‌ی متناظر به هر دسته در قطاع مربوط برابر است با:

$$n = f_1 + f_2 + \dots + f_n \qquad \boxed{\alpha_i = \underbrace{\frac{f_i}{n}}_{\text{فراوانی نسبی}} \times 360^\circ}$$



در مثال نمونه‌ی (۲)، به عنوان نمونه برای دسته‌ی اول و دوم خواهیم داشت:

$$\alpha_1 = \frac{4}{30} \times 360^\circ = 48^\circ \quad \text{و} \quad \alpha_2 = \frac{9}{30} \times 360^\circ = 108^\circ$$

در شکل قبلی، نمودار دایره‌ای مثال ۲ رسم شده است.

❖ **نتیجه‌ی ۱:** در نمودار دایره‌ای، دسته‌ای که فراوانی بیشتری دارد، قطاع بزرگ‌تری به آن اختصاص می‌یابد.

❖ **نتیجه‌ی ۲:** به‌طور کلی:

فراوانی نسبی دسته‌ی iام × °۳۶۰ = زاویه‌ی مرکزی دسته‌ی iام برحسب درجه

⑥ نمودار ساقه و برگ:

اگر داده‌های آماری عددی باشند، آنها را به‌ترتیب صعودی مرتب می‌کنیم، نمودار ساقه و برگ از دو قسمت ساقه و برگ تشکیل شده است، در قسمت ساقه ارقام مشترک و در قسمت برگ ارقام غیر مشترک را قرار می‌دهیم، در قسمت برگ داده‌ها از کوچک به بزرگ مرتب می‌شوند.

در مثال نمونه‌ی ۲، نمودار ساقه و برگ به‌صورت زیر است:

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

در این نمودار برای عدد دو رقمی ۱۳، ساقه را ۱ و برگ را ۳ در نظر گرفتیم یا برای ۲۱، ساقه را ۲ و برگ را ۱ در نظر می‌گیریم.

❖ **نتیجه‌ی ۱:** در قسمت برگ، تعداد اعداد برابر کل داده‌هاست.

جمع‌بندی نکته‌های مبحث آمار

شاخص‌های آماری

شاخص‌های مرکزی

مهم‌ترین شاخص‌های مرکزی را در زیر می‌بینیم.

(۱) **مُد:** داده‌ای که فراوانی آن از همه‌ی داده‌ها بیشتر است را مُد یا نما گوییم.

در مثال نمونه‌ی (۲)، عدد ۱۵ چهار بار تکرار شده است، پس مُد آن ۱۵ است.

تذکر (۲): ممکن است یک جامعه‌ی چند مُدی باشد یا اصلاً مُد نداشته باشد.

(۲) **میانه:** چنانچه داده‌ها را به‌ترتیب صعودی مرتب کنیم، عنصر وسط (وقتی تعداد داده‌ها فرد است) یا میانگین دو عنصر وسط (وقتی تعداد داده‌ها زوج است) را میانه می‌گوییم.

در مثال نمونه‌ی (۲)، تعداد داده‌ها ۳۰ تاست، پس دو عنصر پانزدهم و شانزدهم در وسط قرار می‌گیرند که ۱۸ و ۱۸ خواهند بود، پس میانه ۱۸ است.

نکته اگر به داده‌های آماری K واحد بیفزاییم، مُد و میانه به اندازه‌ی K واحد افزایش می‌یابند.
اگر داده‌ها را K برابر کنیم، مُد، میانه و K برابر می‌شوند.

(۳) **چارک‌ها:** میانه‌ی نیمه‌ی اول داده‌ها را چارک اول $(Q_۱)$ و میانه‌ی نیمه‌ی دوم داده‌ها را چارک سوم $(Q_۳)$ می‌نامیم. چارک دوم همان میانه است.

■ **مثال:** چارک اول و دوم و سوم داده‌های «۱۳ , ۱۱ , ۹ , ۸ , ۵ , ۴ , ۲ , ۱ , ۱» را بیابید.

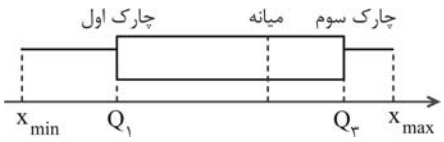
◀ **حل:** از آنجایی که ۹ داده داریم، چارک دوم دقیقاً میانه یعنی ۵ است، چارک اول میانه‌ی چهار عدد اول است و چارک سوم میانه‌ی چهار عدد سمت راست چارک دوم است.

۱ , ۱ , ۲ , ۴ ,   ۵   , ۸ , ۹ , ۱۱ , ۱۳

$$\text{چارک اول} = \frac{۱+۲}{۲} = \frac{۳}{۲} \qquad \text{میانه} \qquad \text{چارک سوم} = \frac{۹+۱۱}{۲} = ۱۰$$

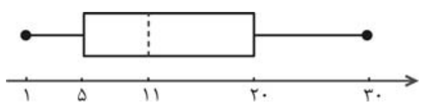
(۴) **نمودار جعبه‌ای:** نموداری است که بر روی آن پارامترها به شکل زیر نمایش داده می‌شود.

در این نمودار باید پنج مقدار ،کوچکترین داده (a)، چارک اول $Q_۱$، میانه، چارک سوم $(Q_۳)$ و بزرگترین داده (b) را در نظر بگیریم.



■ **مثال:** نمودار جعبه‌ای داده‌های ۳۰ , ۲۰ , ۱۷ , ۱۱ , ۱۱ , ۹ , ۹ , ۵ , ۱ را رسم کنید.

◀ **حل:** در این داده‌ها $a = ۱$ و $b = ۳۰$ و $۱۱ =$ میانه و چارک اول $Q_۱ = ۵$ و چارک سوم $Q_۳ = ۲۰$ است، پس:



(۵) **میانگین:** برای داده‌های $x_۱ , x_۲ , x_۳ , \dots , x_n$، میانگین n داده برابر است با:

$$\bar{x} = \frac{x_۱ + x_۲ + x_۳ + \dots + x_n}{n} = \frac{\sum x_i}{n}$$

در مثال نمونه‌ی (۱)، میانگین ۴ , ۵ , ۲ , ۱ , ۳ , ۳ برابر است با:

$$\bar{x} = \frac{۳+۳+۱+۲+۵+۴}{۶} = ۳$$

تذکر (۳): با استفاده از جدول توزیع فراوانی میانگین برابر است با:

$$\boxed{\bar{x} = \frac{f_۱x_۱ + f_۲x_۲ + \dots + f_nx_n}{f_۱ + f_۲ + \dots + f_n} = \frac{\sum f_ix_i}{\sum f_i}}$$

مرکز	$x_۱$	$x_۲$	...	x_n
فراوانی	$f_۱$	$f_۲$	...	f_n

که در آن $x_۱ , x_۲ , x_۳ , \dots , x_n$، مرکز دسته‌ها را نمایش می‌دهد.

در مثال نمونه‌ی (۲)، با توجه به جدول توزیع فراوانی میانگین برابر است با:

مرکز	۷	۱۳	۱۹	۲۵	۳۱	۳۷
فراوانی	۴	۹	۷	۴	۳	۳

$$\bar{x} = \frac{۷\times۴ + ۱۳\times۹ + ۱۹\times۷ + ۲۵\times۴ + ۳۱\times۳ + ۳۷\times۳}{۴+۹+۷+۴+۳+۳}$$

$$\bar{x} = ۱۹/۴$$

نکته ۱- اگر همه‌ی داده‌های آماری با هم برابر باشند، میانگین برابر یکی از آنهاست.
۲- اگر همه‌ی داده‌ها را در a ضرب کنیم، میانگین a برابر می‌شود و اگر به هر یک از داده‌ها عدد b را اضافه کنیم به میانگین b واحد اضافه می‌شود. $(a \neq ۰)$

■ **مثال:** میانگین $x_1 , x_2 , \dots , x_{10}$ برابر ۵ است، میانگین $-2x_1+1 , -2x_2+1 , \dots , -2x_{10}+1$ را بیابید.

◀ **حل:** با توجه به نکته‌ی (۲)، میانگین برابر $-2\times 5+1=-9$ است.

نکته ۳- اگر داده‌های $x_1 , x_2 , \dots , x_n$ تشکیل دنباله‌ی حسابی بدهند، آنگاه میانگین برابر است با: $\bar{x}=\dfrac{x_1+x_n}{2}$

■ **مثال:** میانگین اعداد زوج ۲ تا ۲۰۰ را بیابید.

◀ **حل:** چون اعداد فوق تشکیل دنباله‌ی حسابی می‌دهند بنابراین $\bar{x}=\dfrac{2+200}{2}=101$ خواهد بود.

نکته ۴- اگر داده‌های آماری $1 , 2 , \dots , n$ باشند آنگاه $\bar{x}=\dfrac{n+1}{2}$

نکته ۵- اگر میانگین داده‌های $x_1 , x_2 , \dots , x_n$ برابر $\bar{x}$ و $a_1 , a_2 , \dots , a_n$، دنباله‌ی حسابی تشکیل دهند، آنگاه میانگین $x_1+a_1 , x_2+a_2 , \dots , x_n+a_n$ برابر است با: $\bar{y}=\bar{x}+\dfrac{a_1+a_n}{2}$

شاخص‌های آماری

شاخص‌های پراکندگی

شاخص‌های پراکندگی، میزان اختلاف بین داده‌ها را نمایش می‌دهند، مهم‌ترین شاخص‌های پراکندگی به شرح زیر است:

① **دامنه‌ی تغییرات:** اختلاف بزرگترین و کوچکترین داده است. یعنی $R=b-a$.

نکته ۱- اگر به همه‌ی داده‌های آماری یک عدد بیفزاییم، دامنه‌ی تغییرات تغییر نمی‌کند ولی اگر همه‌ی داده‌های آماری را در یک عدد ثابت ضرب کنیم، دامنه‌ی تغییرات در قدر مطلق آن عدد ضرب می‌شود.
۲- اگر دامنه‌ی تغییرات صفر باشد، همه‌ی داده‌ها با هم برابرند، در این حالت میانگین، میانه و مُد بر هم منطبق‌اند.

② **انحراف از میانگین:** یعنی اختلاف هر داده از میانگین یا $x_i-\bar{x}$.

■ **مثال:** در داده‌های $1 , 2 , 3 , 4 , 5$ انحراف از میانگین داده‌ها را بیابید.

◀ **حل:** در این داده‌ها میانگین $\bar{x}=\dfrac{1+5}{2}=3$ است، پس انحراف از میانگین داده‌ها برابر است با:

$$1-3=\underline{-2} \quad \text{و} \quad 2-3=\underline{-1} \quad \text{و} \quad 3-3=\underline{0} \quad \text{و} \quad 4-3=\underline{1} \quad \text{و} \quad 5-3=\underline{2}$$

به وضوح دیده می‌شود که مجموع انحراف از میانگین‌ها صفر است.

❖ **نتیجه‌ی ۱:** در حالت کلی مجموع انحراف از میانگین‌ها صفر است، یعنی $\sum_{i=1}^{n}(x_i-\bar{x})=0$. $\sum_{i=1}^{n} f_i(x_i-\bar{x})=0$

③ **واریانس:** واریانس داده‌های $x_1 , x_2 , \dots , x_n$ را به‌صورت زیر تعریف می‌کنیم

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

■ **مثال:** واریانس داده‌های $1 , 2 , 3 , 4 , 5$ را بیابید.

◀ **حل:** میانگین داده‌ها $\bar{x}=3$ است، پس:

$$\sigma^2=\frac{(1-3)^2+(2-3)^2+(3-3)^2+(4-3)^2+(5-3)^2}{5}=\frac{4+1+0+1+4}{5}=2$$

➢ **تذکر (۱۴):** اگر جدول توزیع فراوانی موجود باشد، آنگاه واریانس برابر است با:

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

■ **مثال:** واریانس داده‌های جدول روبه‌رو رابیابید.

مرکز دسته	۱	۳	۵	۷	۹
فراوانی	۳	۶	۴	۲	۱

◀ **حل:**

$$\bar{x}=\frac{\sum f_i x_i}{n}=\frac{1\times 3+3\times 6+5\times 4+2\times 7+9\times 1}{16}=4$$

$$\sigma^2=\frac{\sum f_i(x_i-\bar{x})^2}{n}=\frac{3(1-4)^2+6(3-4)^2+4(5-4)^2+2(7-4)^2+1(9-4)^2}{16}=\frac{80}{16}=5$$

۴ **انحراف معیار:** به جذر واریانس، انحراف معیار می‌گوییم، بنابراین:

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

نکته

۱- اگر همه‌ی داده‌های آماری با هم برابر باشند، واریانس و انحراف معیار صفرند.

۲- اگر به همه‌ی داده‌های آماری عددی را بیفزاییم واریانس و انحراف معیار تغییر نمی‌کنند.

۳- اگر همه‌ی داده‌های آماری را در عدد $a \neq ۰$ ضرب کنیم واریانس در $a^۲$ و انحراف معیار در $|a|$ ضرب می‌شود.

۴- اگر $x_۱ , x_۲ , \dots , x_n$ تشکیل دنباله‌ی حسابی با قدر نسبت d بدهند، آنگاه واریانس برابر است با: $\sigma^۲ = \frac{d^۲(n^۲ - ۱)}{۱۲}$

■ **مثال:** واریانس داده‌های ۹ , ۷ , ۵ , ۳ , ۱ را بیابید.

◄ **حل:** از آنجایی که $d = ۲$ و $n = ۵$ است، پس $\sigma^۲ = \frac{۲^۲(۵^۲ - ۱)}{۱۲}$ یا $\sigma^۲ = ۸$.

۵ **ضریب تغییرات:** نسبت انحراف معیار به میانگین را ضریب تغییرات گویند، هرچه ضریب تغییرات به صفر نزدیک باشد داده‌ها، استانداردترند.

$$\boxed{C.V = \frac{\sigma}{\bar{x}}}$$

نکته اگر ضریب تغییرات داده‌ها صفر باشد، داده‌ها با هم برابرند.