

مشخصه‌های پراکندگی

هدفهای رفتاری: در پایان این فصل، فراگیر باید بتواند:

- ۱- دلایل به کار بردن مشخصه‌های پراکندگی را بیان کند.
- ۲- مشخصه‌های مهم پراکندگی را معرفی کند.
- ۳- نحوه محاسبه واریانس را برای توزیعهای گسسته و پیوسته، به کار برد.
- ۴- خواص مشخصه‌های پراکندگی و کاربرد آنها را بیان کند.
- ۵- برای توزیعهای با دامنه نامعین، پراکندگی را با استفاده از چارکها محاسبه کند.
- ۶- مشخصه‌های نسبی پراکندگی را معرفی کند.
- ۷- مشخصه‌های چولگی را برای توزیعهای با دامنه معلوم و نامعین محاسبه کند.
- ۸- با مشخصه‌های پراکندگی بتواند جامعه‌ها را مقایسه کند.

در مطالعه تغییرپذیری صفت متغیر برای اعضای جامعه، تنها نمی‌توان به متوسط اندازه صفت، اکتفا نمود، زیرا قبلاً دیدیم که یک میانگین همواره به یک جامعه تعلق ندارد، بلکه دو یا چند توزیع متفاوت، ممکن است میانگینهای یکسان داشته باشند، ولی پراکندگی آنها متفاوت باشد. برای مشخص نمودن پراکندگی صفت متغیر، می‌توان از انحرافات مقادیر صفت از همدیگر یا انحرافات مقادیر صفت از مشخصه‌های دیگر مانند مشخصه‌های مرکزی استفاده کرد.

مشخصه‌های پراکندگی

طول دامنه تغییرات

در آمار برای اندازه‌گیری پراکندگی، مشخصه‌های مختلفی مورد استفاده قرار می‌گیرند. یکی از

ساده‌ترین آنها، که به آسانی با توزیع صفت متغیر، محاسبه می‌شود «طول دامنه تغییرات» صفت می‌باشد. طول دامنه تغییرات از تفاضل بزرگترین اندازه صفت و کوچکترین مقدار آن به دست می‌آید.

$$R = X_{\max} - X_{\min} \quad (1)$$

فرض کنیم مقادیر مشاهده شده صفت متغیر در جامعه به قرار زیر به دست آمده باشد:

$$X: 3, 5, 7, 9, 7, 4, 5, 8, 11, 10$$

دامنه تغییرات صفت، مطابق فرمول ۱ به صورت زیر محاسبه می‌شود:

چون بزرگترین اندازه صفت، ۱۱ و کوچکترین اندازه آن، ۳ می‌باشد، بنابراین:

$$R = 11 - 3 = 8$$

خواهد بود. ولی دامنه تغییرات، به عنوان مشخصه پراکندگی، تصویر صحیحی درباره چگونگی تراکم مقادیر، حول مقدار متوسط صفت، ارائه نمی‌دهد و فقط از تغییر کوچکترین و بزرگترین مقدار صفت، متأثر می‌شود. از این رو، برای ارزیابی نوسانات مقدار صفت نسبت به مقدار متوسط، مشخصه‌های پراکندگی دیگری مورد استفاده قرار می‌گیرند که از نظر کلی به دامنه تغییرات ارجحیت دارند.

متوسط قدر مطلق انحرافات (انحراف متوسط^۱)

اگر بخواهیم نوسانات مقادیر صفت را حول میانگین، با میانگین انحرافات ارزیابی کنیم، یعنی مشخصه پراکندگی را به صورت:

$$\frac{\sum F_i (X_i - \bar{X})}{N}$$

بیان کنیم، چون صورت کسر، مطابق خاصیت ۳ میانگین، همواره مساوی صفر است، درباره پراکندگی واقعی صفت، چیزی به دست نخواهد آمد.

برای برطرف کردن این مشکل، از قدر مطلق انحرافات و یا از مجذور انحرافات استفاده می‌شود. محاسبه میانگین قدر مطلق انحرافات که آن را با نماد AD نشان می‌دهیم، به قرار زیر خواهد بود:

$$AD = \frac{\sum F_i |X_i - \bar{X}|}{N} \quad (2)$$

و اگر فراوانیها برای مقادیر صفت، برابر ۱ باشد، رابطه بالا به صورت صفحه بعد در خواهد آمد:

۱- Average Deviation = AD

$$AD = \frac{\sum |X_i - \bar{X}|}{N} \quad (3)$$

روش محاسبه آن را با ذکر یک مثال بیان می‌کنیم.

مثال ۱: برای توزیع صفت متغیر که با جدول زیر بیان شده است، متوسط قدرمطلق انحرافات را

محاسبه می‌کنیم:

X	۲	۳	۴	۵	۶	-
F _i	۱۰	۲۰	۳۰	۱۶	۴	۸۰

ابتدا میانگین حسابی را برای این توزیع محاسبه می‌کنیم، مقدار آن برابر $\bar{X} = 3/8$ است.

بنابراین، طبق رابطه ۲ خواهیم داشت:

$$\begin{aligned} AD &= \frac{\sum F_i |X_i - \bar{X}|}{N} \\ &= \frac{10|2 - 3/8| + 20|3 - 3/8| + 30|4 - 3/8| + 16|5 - 3/8| + 4|6 - 3/8|}{80} \\ &= \frac{18 + 16 + 6 + 19/2 + 8/8}{80} = \frac{68}{80} = 0/85 \end{aligned}$$

تمام محاسبات می‌تواند در جدول زیر انجام پذیرد (جدول ۱):

جدول ۱

X	F _i	F _i X _i	X _i - \bar{X}	X - \bar{X}	F _i X _i - \bar{X}
۲	۱۰	۲۰	۲ - ۳/۸	۱/۸	۱۸
۳	۲۰	۶۰	۳ - ۳/۸	۰/۸	۱۶
۴	۳۰	۱۲۰	۴ - ۳/۸	۰/۲	۶
۵	۱۶	۸۰	۵ - ۳/۸	۱/۲	۱۹/۲
۶	۴	۲۴	۶ - ۳/۸	۲/۲	۸/۸
-	۸۰	۳۰۴	-	-	۶۸

$$\bar{X} = \frac{\sum F_i X_i}{N} = \frac{304}{80} = 3/8$$

$$AD = \frac{\sum F_i |X_i - \bar{X}|}{N} = \frac{68}{80} = 0/85$$

یعنی به طور متوسط، هریک از مقادیر صفت، به اندازه $0/85$ واحد از متوسط خود انحراف دارند.

اگر صفت متغیر با توزیع فاصله‌ای، بیان شده باشد، همانطور که در محاسبه متوسط حسابی، گفته شد، ابتدا وسط فاصله‌ها را تعیین کرده، سپس مانند توزیع گسسته، عمل می‌کنیم.
مثال ۲: متوسط قدرمطلق انحرافات را برای مؤسّسات تولیدکننده مواد شیمیایی در مثال ۹ از فصل سوم، محاسبه می‌کنیم (جدول ۲):

جدول ۲

X	F _i	X _i	F _i X _i	X _i - \bar{X}	F _i X _i - \bar{X}
۱۱-۱۳	۷	۱۲	۸۴	۳/۸۵	۲۶/۹۵
۱۳-۱۵	۹	۱۴	۱۲۶	۱/۸۵	۱۶/۶۵
۱۵-۱۷	۱۰	۱۶	۱۶۰	۰/۱۵	۱/۵۰
۱۷-۱۹	۸	۱۸	۱۴۴	۲/۱۵	۱۷/۲۰
۱۹-۲۱	۶	۲۰	۱۲۰	۴/۱۵	۲۴/۹۰
-	۴۰	-	۶۳۴	-	۸۷/۲۰

$$\bar{X} = \frac{634}{40} = 15/85$$

$$AD = \frac{\sum F_i |X_i - \bar{X}|}{N} = \frac{87/20}{40} = 2/18$$

متوسط قدرمطلق انحرافات، یک اندازه طبیعی برای پراکندگی توزیع می‌باشد، لیکن به علت فاقد بودن بعضی ویژگیهای ریاضی، از این مشخصه، کمتر استفاده می‌شود.

متوسط مجذور انحرافات یا واریانس^۱ (پراش)

در عمل به جای استفاده از علامت قدرمطلق، از مجذور انحرافات استفاده می‌کنند. بنابراین یکی دیگر از معیارهای پراکندگی، متوسط مجذور انحرافات است که آن را واریانس می‌نامند و با نماد $V(X)$ نشان داده و با فرمول زیر تعریف می‌شود:

۱- Variance

$$V(X) = \frac{F_1(X_1 - \bar{X})^2 + F_2(X_2 - \bar{X})^2 + \dots + F_k(X_k - \bar{X})^2}{F_1 + F_2 + \dots + F_k}$$

$$= \frac{\sum F_i(X_i - \bar{X})^2}{N} \quad (4)$$

و اگر مقادیر صفت بدون فراوانی باشند (فراوانی هر کدام یک باشد):

$$V(X) = \frac{\sum (X_i - \bar{X})^2}{N} \quad (5)$$

محاسبه واریانس را با مثال زیر نشان می‌دهیم:

مثال ۳: برای توزیع صفت متغیر که با جدول زیر بیان شده است، واریانس را محاسبه می‌کنیم:

X	۲	۳	۴	۵	-
F _i	۱۰	۴۰	۳۰	۲۰	۱۰۰

متوسط حسابی برای این توزیع $\bar{X} = 3/6$ است، بنابراین با استفاده از فرمول ۴ خواهیم داشت:

$$V(X) = \frac{\sum F_i(X_i - \bar{X})^2}{N}$$

$$= \frac{10(2 - 3/6)^2 + 40(3 - 3/6)^2 + 30(4 - 3/6)^2 + 20(5 - 3/6)^2}{10 + 40 + 30 + 20}$$

$$= 0/84$$

تمامی محاسبات می‌تواند در جدول زیر انجام پذیرد (جدول ۳):

جدول ۳

X _i	F _i	F _i X _i	X _i - \bar{X}	(X _i - \bar{X}) ²	F _i (X _i - \bar{X}) ²
۲	۱۰	۲۰	-۱/۶	۲/۵۶	۲۵/۶
۳	۴۰	۱۲۰	-۰/۶	۰/۳۶	۱۴/۴
۴	۳۰	۱۲۰	۰/۴	۰/۱۶	۴/۸
۵	۲۰	۱۰۰	۱/۴	۱/۹۶	۳۹/۲
-	۱۰۰	۳۶۰	-	-	۸۴

$$\bar{X} = \frac{360}{100} = 3/6$$

$$V(X) = \frac{\sum F_i(X_i - \bar{X})^2}{N} = \frac{84}{100} = 0.84$$

در محاسبه واریانس، اغلب به علت اینکه متوسط حسابی، با رقمهای اعشاری باید از مقادیر صفت کم شود، به توان رساندن این اعداد اعشاری، محاسبه را دشوارتر می‌سازد. از این رو، فرمول واریانس را به صورت زیر تبدیل می‌کنند:

$$V(X) = \frac{\sum F_i(X_i - \bar{X})^2}{N} = \frac{\sum F_i X_i^2}{N} - \bar{X}^2 \quad (6)$$

$$V(X) = \frac{\sum F_i(X_i - \bar{X})^2}{N} = \frac{\sum F_i X_i^2 - \frac{(\sum F_i X_i)^2}{N}}{N} \quad (7)$$

و اگر فراوانی هر کدام از اندازه‌ها مساوی یک باشد، فرمولهای ۶ و ۷ به صورت زیر نوشته خواهد شد:

$$V(X) = \frac{\sum (X_i - \bar{X})^2}{N} = \frac{\sum X_i^2}{N} - \bar{X}^2 \quad (8)$$

$$V(X) = \frac{\sum (X_i - \bar{X})^2}{N} = \frac{\sum X_i^2 - \frac{(\sum X_i)^2}{N}}{N} \quad (9)$$

فرمولهای ۶ تا ۹ را فرمولهای کاربردی واریانس نامند.
مثال ۴: واریانس را برای توزیع صفت متغیر که در مثال ۳ بیان شده است، با استفاده از فرمول ۶ محاسبه می‌کنیم (جدول ۴):

جدول ۴

X_i	F_i	$F_i X_i$	$F_i X_i^2$
۲	۱۰	۲۰	۴۰
۳	۴۰	۱۲۰	۳۶۰
۴	۳۰	۱۲۰	۴۸۰
۵	۲۰	۱۰۰	۵۰۰
—	۱۰۰	۳۶۰	۱۳۸۰

$$\bar{X} = \frac{360}{100} = 3.6$$

$$V(X) = \frac{\sum F_i X_i^2}{N} - \bar{X}^2$$

$$= \frac{1380}{40} - (3/6)^2 = 13/8 - 12/96 = 0/84$$

اگر صفت متغیر، به صورت توزیع فاصله‌ای بیان شده باشد، مانند محاسبه متوسط حسابی، ابتدا، وسط فاصله‌ها را تعیین کرده، سپس مانند توزیع گسسته، عمل می‌کنند.

مثال ۵: واریانس را برای توزیع صفت متغیر که با جدول ۲ در مثال ۲ بیان شده است، محاسبه می‌کنیم. تمامی محاسبات در جدول ۵ آورده شده است:

جدول ۵

X	F _i	X _i	F _i X _i	F _i X _i ^۲
۱۱-۱۳	۷	۱۲	۸۴	۱۰۰۸
۱۳-۱۵	۹	۱۴	۱۲۶	۱۷۶۴
۱۵-۱۷	۱۰	۱۶	۱۶۰	۲۵۶۰
۱۷-۱۹	۸	۱۸	۱۴۴	۲۵۹۲
۱۹-۲۱	۶	۲۰	۱۲۰	۲۴۰۰
-	۴۰	-	۶۳۴	۱۰۳۲۴

$$V(X) = \frac{\sum F_i X_i^2 - \frac{(\sum F_i X_i)^2}{N}}{N}$$

$$= \frac{10324 - \frac{(634)^2}{40}}{40} = \frac{275/1}{40} = 6/8775$$

انحراف معیار^۱

واریانس، دارای ویژگی‌های مهم ریاضی است که در اینجا درباره آنها بحث نمی‌کنیم لیکن به علت اینکه، واریانس، پراکندگی را به صورت مجذور انحرافات بیان می‌کند، برای اینکه متوسط انحراف برای یک عضو مشخص گردد، از واریانس جذر گرفته و آن را انحراف معیار صفت متغیر می‌نامند و با حرف یونانی σ نشان می‌دهند.

^۱ - Standard Deviation

$$\begin{aligned}
 &= .+ \sqrt{V(X)} \\
 &= +\sqrt{\frac{\sum F_i (X_i - \bar{X})^2}{N}} \quad (۱۰)
 \end{aligned}$$

با وجود اینکه در آمار نظری، به عنوان اندازه پراکندگی، به طور اساسی، واریانس مورد استفاده قرار می‌گیرد، ولی انحراف معیار، مناسب‌ترین و متداولترین مشخصه پراکندگی است که در عمل از آن استفاده می‌شود. از این رو، در بحثهای بعدی، پراکندگی را با انحراف معیار، ارزیابی خواهیم کرد.

مثال ۶: انحراف معیار را برای مثال ۵ محاسبه می‌کنیم:

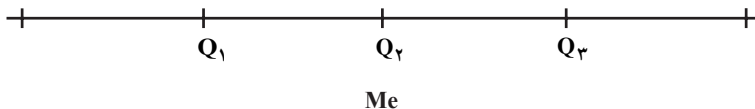
$$= .+ \sqrt{V(X)} = \sqrt{6/8775} = 2/622$$

انحراف معیار، نیز مانند متوسط قدر مطلق انحرافات، متوسط انحرافات اندازه‌های صفت را از میانگین صفت نشان می‌دهد.

انحراف چارکی^۱

به همان دلایلی که میانه به عنوان مشخصه مرکزی تعریف شده است، برای تغییرپذیری جامعه‌هایی که دامنه توزیع صفت در آنها نامشخص باشد، از مشخصه‌های آماری که از دامنه توزیع متأثر نمی‌باشند، استفاده می‌کنند. چنین مشخصه‌های آماری، چارکها^۲ هستند که میانه یکی از آنها می‌باشد. برای تعیین پراکندگی توزیع صفت در چنین حالتی از انحراف چارکی استفاده می‌شود که به صورت زیر تعریف می‌گردد:

اگر طول دامنه تغییرات صفت را به چهار قسمت مساوی از نظر حجم تقسیم کنیم، سه عددی که این دامنه را به چهار قسمت افراز می‌کند به ترتیب چارک اول، چارک دوم و چارک سوم نامند و با Q_1 ، Q_2 و Q_3 نشان داده می‌شوند.



متوسط تفاضل چارک دوم و چارک اول با تفاضل چارک سوم و چارک دوم را انحراف چارکی نامند و آن را با نماد QD نشان می‌دهند.

^۱ - Quartile Deviation

^۲ - Quartiles

فرمول آن به صورت زیر بیان می‌شود :

$$QD = \frac{(Q_2 - Q_1) + (Q_3 - Q_2)}{2} = \frac{Q_3 - Q_1}{2} \quad (11)$$

چارکها، مقدار صفت متغیر در مجموعه مرتب شده مقادیر صفت هستند که ۲۵ درصد اعضای جامعه، کوچکتر از Q_1 و ۲۵ درصد اعضای جامعه، بین Q_1 و Q_2 و ۲۵ درصد اعضای جامعه بین Q_2 و Q_3 و ۲۵ درصد بقیه اعضا، بالاتر از Q_3 قرار می‌گیرند.

چارکهای اول و سوم، یعنی Q_1 و Q_3 ، مشابه میانه، طبق فرمولهای زیر محاسبه می‌شوند :

$$Q_1 = X_i + \frac{\frac{N}{4} - FC_i}{F_i} \times I \quad (12)$$

که در آن

X_i - نقطه شروع فاصله گروهی از صفت متغیر است که چارک اول در آن قرار دارد.

I - طول فاصله‌ای است که چارک اول در آن قرار دارد.

FC_i - فراوانی انباشته گروه قبل از گروه چارک اول می‌باشد.

F_i - فراوانی مطلق گروه چارک اول است.

$$Q_3 = X_i + \frac{\frac{3N}{4} - FC_i}{F_i} \times I \quad (13)$$

که در آن

X_i - نقطه شروع فاصله گروهی از صفت متغیر است که چارک سوم در آن قرار دارد.

I - طول فاصله‌ای است که چارک سوم در آن قرار دارد.

FC_i - فراوانی انباشته گروه قبل از گروه چارک سوم می‌باشد.

F_i - فراوانی مطلق گروه چارک سوم است.

بدیهی است، برای محاسبات چارک اول و چارک سوم، باید ستون فراوانیهای تجمعی را در

جدول توزیع صفت، محاسبه کنیم.

مثال ۷: فرض کنید توزیع فراوانی کارگران برحسب مزد و حقوق ماهانه آنها در یک مؤسسه با

جدول زیر بیان شده باشد. می‌خواهیم تغییرپذیری (پراکندگی) مزد و حقوق ماهانه را برای این جامعه

محاسبه کنیم :

جدول ۶

مزد و حقوق ماهانه کارگران X	تعداد کارگران F _i	فراوانی انباشته FC _i
تا ۴۰۰۰۰	۱۲	۱۲
۴۰۰۰۰-۴۵۰۰۰	۳۸	۵۰
۴۵۰۰۰-۵۰۰۰۰	۴۴	۹۴
۵۰۰۰۰-۵۵۰۰۰	۵۰	۱۴۴
۵۵۰۰۰-۶۰۰۰۰	۶۲	۲۰۶
۶۰۰۰۰-۷۰۰۰۰	۴۲	۲۴۸
۷۰۰۰۰-۸۰۰۰۰	۲۵	۲۷۳
۸۰۰۰۰-۱۰۰,۰۰۰	۱۴	۲۸۷
۱۰۰,۰۰۰ تومان به بالا	۱۳	۳۰۰
-	۳۰۰	-

ملاحظه می‌شود که مشخصه‌هایی مانند «طول دامنه تغییرات»، «متوسط قدر مطلق انحرافات»، واریانس و انحراف معیار به‌عنوان مشخصه پراکندگی برای این توزیع نمی‌تواند محاسبه گردد. زیرا حدود بالا و پایین جامعه نامشخص است. از این رو «انحراف چارکی» را برای این توزیع فراوانیها، محاسبه می‌کنیم. ابتدا ستون فراوانیهای انباشته را محاسبه می‌کنیم که نتیجه آن در ستون سوم جدول ۶ آمده است.

محاسبه چارک اول (Q_1): برای این منظور $\frac{N}{4}$ را تعیین می‌کنیم:

$$\frac{N}{4} = \frac{300}{4} = 75$$

با مقایسه مقدار چارک اول (۷۵) با ستون فراوانیهای انباشته جدول، مشاهده می‌شود که در فراوانی انباشته (۹۴) متعلق به فاصله ۴۵۰۰۰-۵۰۰۰۰ قرار دارد، بنابراین چارک اول در فاصله ۴۵۰۰۰-۵۰۰۰۰ می‌باشد. با استفاده از فرمول ۱۲ مقدار Q_1 را محاسبه می‌کنیم.

$$X_i = 45000 \quad \text{در اینجا:}$$

$$I = 50000 - 45000 = 5000$$

$$FC_i = 50$$

$$F_i = 44$$

$$\begin{aligned}
 Q_1 &= X_i + \frac{\frac{N}{4} - FC_i}{F_i} \times I \\
 &= 450000 + \frac{75 - 50}{44} \times 50000 \\
 &= 450000 + \frac{1250000}{44} = 450000 + 2841 = 47841
 \end{aligned}$$

مشابه روش بالا، چارک سوم (Q_3) را محاسبه می‌کنیم. برای این منظور $\frac{3N}{4}$ را تعیین می‌کنیم:

$$\frac{3N}{4} = \frac{3 \times 300}{4} = 225$$

با $\frac{3N}{4}$ ، فاصله‌ای را که شامل چارک سوم است، جستجو می‌کنیم. چون ۲۲۵ در فراوانی انباشته (۲۴۸) که متعلق به فاصله ۷۰۰۰۰-۶۰۰۰۰ است، قرار دارد، از این رو، چارک سوم در فاصله ۷۰۰۰۰-۶۰۰۰۰ می‌باشد. با استفاده از فرمول ۱۳، چارک سوم را محاسبه می‌کنیم.

$$X_i = 60000 \quad \text{در اینجا:}$$

$$I = 70000 - 60000 = 10000$$

$$FC_i = 206$$

$$F_i = 42$$

$$\begin{aligned}
 Q_3 &= X_i + \frac{\frac{3N}{4} - FC_i}{F_i} \times I \\
 &= 60000 + \frac{225 - 206}{42} \times 10000 \\
 &= 60000 + \frac{190000}{42} = 60000 + 4524 = 64524
 \end{aligned}$$

حال با استفاده از فرمول ۱۱، انحراف چارکی (QD) را محاسبه می‌کنیم:

$$\begin{aligned}
 QD &= \frac{Q_3 - Q_1}{2} \\
 &= \frac{64524 - 47841}{2} = 8341/5
 \end{aligned}$$

چون تمامی مقادیر صفت متغیر، بر روی انحراف چارکی، تأثیر نمی‌کنند (حداکثر ۵۰ درصد از مقادیر صفت در آن تأثیر دارند)، از این مشخصه در مواردی که محاسبه انحراف معیار، دشوار یا غیرممکن است، استفاده می‌شود.

مشخصه‌های نسبی پراکندگی

برای مقایسه تغییرات صفت در دو یا چند جامعه با میانگینهای مختلف، از مشخصه‌های نسبی پراکندگی، استفاده می‌شود. این مشخصه‌ها از نسبت مشخصه‌های پراکندگی به مشخصه‌های مرکزی صفت، محاسبه می‌شوند.

یکی از مهمترین مشخصه‌های نسبی پراکندگی، ضریب تغییرات است که به صورت زیر تعریف می‌شود:

ضریب تغییرات^۱

نسبت انحراف معیار صفت متغیر بر متوسط حسابی صفت که برحسب درصد بیان شده باشد، ضریب تغییرات، نامیده می‌شود و آن را با C.V نشان می‌دهند.

$$C.V = \frac{s}{\bar{X}} \times 100 \quad (14)$$

که در آن:

– انحراف معیار صفت متغیر است.

\bar{X} – متوسط حسابی صفت متغیر، می‌باشد.

مثال ۸: ضریب تغییرات را برای صفت متغیر که در مثال ۵ آورده شده است، و برای آن

$$\bar{X} = 15/85 \text{ و } s = 2/622 \text{ می‌باشند، محاسبه می‌کنیم:}$$

$$C.V = \frac{s}{\bar{X}} \times 100 = \frac{2/622}{15/85} \times 100 = 16/54\%$$

یعنی تغییرات صفت، ۱۶/۵ درصد اندازه متوسط آن را تشکیل می‌دهد.

ضریب تغییرات نه تنها برای بیان تغییرپذیری نسبی به کار می‌رود، بلکه به عنوان «مشخصه همگنی» جامعه نیز مورد استفاده قرار می‌گیرد. هرگاه ضریب تغییرات صفت در جامعه از ۳۳ درصد تجاوز نکند، آن جامعه را از نظر صفت متغیر، «همگن» در نظر می‌گیرند.

مثلاً جامعه‌ای که در مثال بالا آورده شد، نسبت به صفت متغیر، همگن می‌باشد.

وقتی به عنوان اندازه تغییرپذیری، متوسط قدر مطلق انحرافات (A.D)، مورد استفاده قرار

می‌گیرد، مشخصه نسبی پراکندگی به صورت انحراف خطی نسبی، محاسبه می‌شود:

^۱ – Coefficient of variation

انحراف خطی نسبی

نسبت متوسط قدر مطلق انحرافات بر متوسط حسابی صفت متغیر که بر حسب درصد بیان شده باشد، «انحراف خطی نسبی» نامیده می‌شود. اگر آن را با $C.V_{AD}$ نشان دهیم، به صورت زیر نوشته می‌شود:

$$C.V_{AD} = \frac{AD}{\bar{X}} \times 100 \quad (15)$$

که در آن:

AD - متوسط قدر مطلق انحرافات صفت متغیر است.

\bar{X} - متوسط حسابی صفت متغیر، می‌باشد.

مثال ۹: انحراف خطی نسبی را برای صفت متغیر در مثال ۲، محاسبه می‌کنیم که برای آن

$\bar{X} = 15/85$ و $AD = 2/18$ قبلاً محاسبه شده است.

$$\begin{aligned} C.V_{AD} &= \frac{AD}{\bar{X}} \times 100 \\ &= \frac{2/18}{15/85} \times 100 = 13/75- \end{aligned}$$

درحالاتی که استفاده از مشخصه‌های تغییر پذیری (پراکندگی) مذکور، امکان پذیر نباشد، مناسب

است از مشخصه نسبی پراکندگی دیگری به نام «انحراف چارکی نسبی» استفاده گردد.

انحراف چارکی نسبی

نسبت انحراف چارکی بر میانه صفت متغیر «انحراف چارکی نسبی» نامیده می‌شود که آن را با q نشان می‌دهیم:

$$q = \frac{QD}{Me} \times 100 \quad (16)$$

مثال ۱۰: انحراف چارکی نسبی را برای صفت متغیر در مثال ۷ محاسبه می‌کنیم

که برای آن: $QD = 8341/5$ و $Me = 55484$ می‌باشد.

$$q = \frac{QD}{Me} \times 100 = \frac{8341/5}{55484} \times 100 = 15-$$

مشخصه چولگی (انحراف از قرینگی)

برای پی بردن به خصوصیات عمومی توزیع، اغلب، مشخصه‌های انحراف از قرینگی یا چولگی محاسبه می‌شود.

قبل از اینکه مشخصه‌های چولگی را معرفی کنیم، باید بدانیم توزیع قرینه چیست، تا براساس آن، غیرقرینه بودن، مشخص گردد.

هرگاه فراوانیهای متعلق به دو مقدار صفت که از مرکز توزیع، در فاصله‌های مساوی قرار دارند، یکسان باشند، آن توزیع را «قرینه» می‌خوانند. در غیر این صورت، توزیع را «غیرقرینه» یا «چوله» می‌نامند.

مثلاً در توزیع فراوانی‌هایی که با جدول زیر بیان شده است، فراوانیها، نسبت به وسط توزیع، دوه‌دو باهم برابرند:

X	۵	۱۰	۱۵	۲۰	۲۵
F _i	۱۰	۴۰	۸۰	۴۰	۱۰

در صورتی که در توزیع فراوانی‌هایی که در جدول زیر بیان شده است، فراوانیها برای مقادیر صفت که نسبت به وسط توزیع در فاصله‌های مساوی قرار گرفته‌اند، یکسان نیستند.

X _i	۲	۴	۶	۸	۱۰	۱۲
F _i	۱۰	۲۰	۴۰	۴۵	۲۵	۱۰

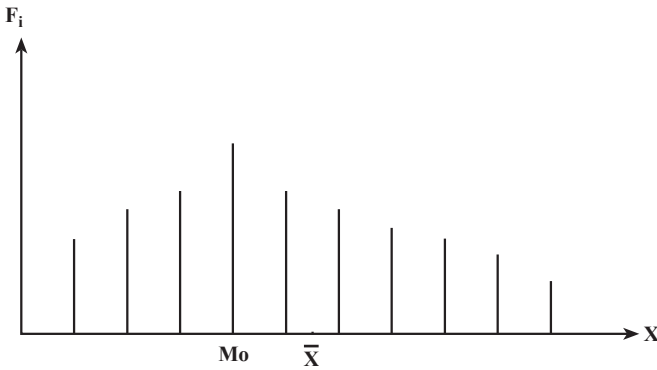
برای مقایسه درجه‌دوری توزیعها از حالت قرینه بودن، از مشخصه‌نسبی که «ضریب چولگی» نام دارد، استفاده می‌شود.

— ضریب چولگی پیرسن: در فصل گذشته، ملاحظه شد که برای توزیعهای قرینه، مقدار میانگین، میانه و نما باهم برابرند. در ارتباط با آن، ساده‌ترین مشخصه انحراف از قرینگی یا چولگی که ضریب چولگی پیرسن نام دارد، بر مناسبات بین مشخصه‌های مرکزی استوار گردیده است.

هرچه تفاوت بین مشخصه‌های مرکزی $(\bar{X} - Mo)$ ، بیشتر باشد، چولگی نیز بیشتر است. ضریب چولگی پیرسن به صورت زیر تعریف می‌شود:

$$P_{AS} = \frac{\bar{X} - Mo}{.} \quad (17)$$

کمیت P_{AS} ، می تواند مثبت، منفی یا صفر باشد. مقدار مثبت این ضریب، وجود چولگی راست را نشان می دهد یعنی دامنه راست توزیع نسبت به نما، بیشتر کشیده است تا دامنه چپ آن (مانند نمودار شکل ۱):



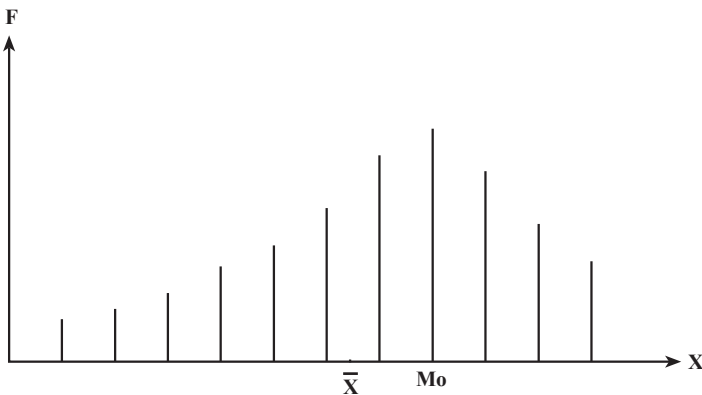
شکل ۱- وجود چولگی راست

در چولگی راست، بین مشخصه های مرکزی رابطه زیر برقرار است:

$$Mo < Me < \bar{X} \quad (18)$$

علامت منفی در این ضریب، گواه وجود چولگی چپ است که در شکل ۲ نشان داده شده

است:



شکل ۲- وجود چولگی چپ

در این حالت، بین مشخصه‌های مرکزی، رابطه زیر برقرار است :

$Mo \# Me \# \bar{X}$	(۱۹)
-----------------------	------

معمولاً وقتی ضریب چولگی کوچکتر از ۱/۵ باشد، توزیع را تقریباً قرینه و وقتی که ضریب چولگی بزرگتر از ۵/۵ باشد آن را خیلی چوله در نظر می‌گیرند. طبیعی است که در توزیعهای قرینه، ضریب چولگی مساوی با صفر خواهد بود.

مثال ۱۱: ضریب چولگی پیرسن (P_{AS}) را برای توزیع زیر محاسبه می‌کنیم :

X	۱۰-۱۵	۱۵-۲۰	۲۰-۲۵	۲۵-۳۰	۳۰-۳۵	-
F_i	۴	۱۰	۲۰	۵	۱	۴۰

برای محاسبه ضریب چولگی پیرسن، میانگین حسابی، نما و انحراف معیار توزیع را باید به دست آورد. تمامی محاسبات در جدول ۷ آورده شده است.

جدول ۷

X	F_i	X_i	$F_i X_i$	$F_i X_i^2$
۱۰-۱۵	۴	۱۲/۵	۵۰	۶۲۵
۱۵-۲۰	۱۰	۱۷/۵	۱۷۵	۳۰۶۲/۵
۲۰-۲۵	۲۰	۲۲/۵	۴۵۰	۱۰۱۲۵
۲۵-۳۰	۵	۲۷/۵	۱۳۷/۵	۳۷۸۱/۲۵
۳۰-۳۵	۱	۳۲/۵	۳۲/۵	۱۰۵۶/۲۵
-	۴۰	-	۸۴۵	۱۸۶۵۰

$$\bar{X} = \frac{\sum F_i X_i}{N} = \frac{845}{40} = 21.125$$

$$V(X) = \frac{\sum F_i X_i^2 - \frac{(\sum F_i X_i)^2}{N}}{N}$$

$$= \frac{18650 - \frac{(845)^2}{40}}{40} = \frac{799.375}{40} = 19.98$$

$$= .+ \sqrt{V(X)} = \sqrt{19/98} = 4/47$$

$$\begin{aligned} Mo &= X_i + \frac{F_i - F_{i-1}}{(F_i - F_{i-1}) + (F_i - F_{i+1})} \times I \\ &= 20 + \frac{20 - 10}{(20 - 10) + (20 - 5)} \times 5 \\ &= 20 + \frac{50}{25} = 22 \end{aligned}$$

در نتیجه، ضریب چولگی پیرسن طبق فرمول ۱۷، به صورت زیر محاسبه می شود:

$$P_{AS} = \frac{\bar{X} - Mo}{.} = \frac{21/125 - 22}{4/47} = -.0/19$$

ولی باید متذکر شد که ضریب چولگی پیرسن همیشه نمی تواند مورد استفاده قرار گیرد. مثلاً اگر توزیع دو نما داشته باشد، این ضریب نمی تواند محاسبه گردد.

— ضریب چولگی بر پایه میانگین توان سوم انحرافات:

یکی از متداول ترین ضرایب چولگی که نقص ضریب چولگی پیرسن را ندارد و از تمامی عناصر توزیع (یعنی از مقادیر صفت و فراوانیها) در محاسبه آن استفاده می شود ضریب چولگی بر پایه میانگین توان سوم انحرافات می باشد که با M_{AS} نشان داده می شود، به صورت زیر بیان می شود:

$$M_{AS} = \frac{\sum F_i (X_i - \bar{X})^3}{N^3} = \frac{\sum F_i (X_i - \bar{X})^3}{N^3} \quad (20)$$

مثال ۱۲: ضریب چولگی بر پایه میانگین توان سوم انحرافات را برای توزیع فراوانیهای صفت متغیر که با جدول زیر بیان شده است، محاسبه می کنیم:

X	۳	۴	۵	۶	۷	۸	—
F_i	۱	۶	۱۳	۱۱	۱۰	۹	۵۰

تمامی محاسبات در جدول ۸، آورده شده است:

جدول ۸

X	F _i	F _i X _i	F _i X _i ^۲	X _i - \bar{X}	(X _i - \bar{X}) ^۲	F _i (X _i - \bar{X}) ^۲
۳	۱	۳	۹	-۳	-۲۷	-۲۷
۴	۶	۲۴	۹۶	-۲	-۸	-۴۸
۵	۱۳	۶۵	۳۲۵	-۱	-۱	-۱۳
۶	۱۱	۶۶	۳۹۶	۰	۰	۰
۷	۱۰	۷۰	۴۹۰	۱	۱	۱۰
۸	۹	۷۲	۵۷۶	۲	۸	۷۲
-	۵۰	۳۰۰	۱۸۹۲	-	-	-۶

$$\bar{X} = \frac{300}{50} = 6$$

$$V(X) = \frac{\sum F_i X_i^2}{N} - \bar{X}^2$$

$$= \frac{1892}{50} - 6^2 = 37.84 - 36 = 1.84$$

$$= \sqrt{V(X)} = \sqrt{1.84} = 1.36$$

$$\sigma^2 = 1.36 \times 1.84 = 2.50$$

با استفاده از کمیتهای محاسبه شده، طبق فرمول ۲۰، ضریب چولگی را محاسبه می کنیم:

$$M_{AS} = \frac{\sum F_i (X_i - \bar{X})^3}{\sum F_i} = \frac{-6}{2.5} = -0.048$$

بنابراین، چون ضریب چولگی، کوچکتر از ۱/۰ به دست آمده است، می توان گفت، توزیع تقریباً قرینه است.

سؤالها و تمرینها ؟

- ۱- چرا در مطالعه صفت متغیر، برای اعضای جامعه، تنها به کمیتهای متوسط نمی توان اکتفا کرد؟
- ۲- اندازه پراکندگی چیست؟ ساختن اندازه پراکندگی انحرافات مقادیر صفت از متوسط صفت، چه اهمیتی دارد؟
- ۳- طول دامنه تغییرات چیست؟
- ۴- طول دامنه تغییرات را برای داده های زیر به دست آورید :
۳, ۴, ۷, ۸, ۱۰, ۹, ۵, ۷, ۱۱, ۶, ۱۳, ۱۰, ۱۵
- ۵- طول دامنه تغییرات را برای طول قد دانش آموزان که در جدول زیر داده شده است، محاسبه کنید.

X طول قد	۱۴۴-۱۴۸	۱۴۸-۱۵۲	۱۵۲-۱۵۶	۱۵۶-۱۶۰	۱۶۰-۱۶۴	
F _i فراوانی	۲	۶	۱۰	۱۵	۷	۴۰

- ۶- نقص طول دامنه تغییرات به عنوان یک مشخصه پراکندگی در چیست؟
- ۷- چرا از انحرافات مقادیر صفت، نسبت به متوسط حسابی، نمی توانیم متوسط انحرافات را محاسبه کنیم؟
- ۸- متوسط قدر مطلق انحرافات را تعریف کنید.
- ۹- متوسط قدر مطلق انحرافات را برای توزیع قد دانش آموزان در تمرین ۵ محاسبه کنید.
- ۱۰- چرا در عمل، از متوسط قدر مطلق انحرافات، کمتر استفاده می شود؟
- ۱۱- متوسط مجذور انحرافات یا واریانس، چه مشخصه ای است؟
- ۱۲- واریانس به عنوان مشخصه پراکندگی، برای برطرف کردن چه مشکلی در آمار، وارد گردیده است؟
- ۱۳- فرمولهای کاربردی یا فرمولهای محاسباتی واریانس، کدامها هستند؟ آنها را بنویسید.
- ۱۴- واریانس را برای طول قد دانش آموزان در تمرین ۵ محاسبه کنید.

۱۵- واریانس را برای اندازه‌های صفت که در زیر داده شده‌اند، محاسبه کنید.

$X: 3, 5, 3, 6, 7, 2, 8, 10$

۱۶- چرا به جای واریانس، برای نشان دادن پراکندگی صفت، از انحراف معیار، استفاده

می‌شود؟

۱۷- انحراف معیار چیست؟ تعریف کنید.

۱۸- انحراف معیار را برای تمرین ۵ (طول قد دانش‌آموزان)، محاسبه کنید.

۱۹- برای توزیع فراوانیهای صفت که در زیر داده شده، انحراف معیار و متوسط قدر مطلق

انحرافات را محاسبه کرده، آنها را باهم مقایسه کنید. به نظر شما کدام دقیق‌تر است؟

X	۲	۴	۶	۸	۱۰	۱۲	-
F_i	۴	۷	۱۰	۶	۲	۱	۳۰

۲۰- انحراف چارکی چیست؟ تعریف کنید.

۲۱- لزوم وارد کردن انحراف چارکی به عنوان مشخصه پراکندگی، در چیست؟

۲۲- چارکها، چه مفهومی را بیان می‌کنند؟

۲۳- چارک اول چیست؟

۲۴- چارک دوم چیست؟

۲۵- چارک سوم چیست؟

۲۶- انحراف چارکی را برای توزیع فراوانیهای زیر محاسبه کنید :

X	۴-۸	۸-۱۲	۱۲-۱۶	۱۶-۲۰	۲۰-۲۴	۲۴ به بالا
F_i	۳	۵	۸	۱۲	۶	۶

۲۷- مشخصه‌های نسبی پراکندگی، چه مفهومی دارند؟

۲۸- چرا برای بیان پراکندگی صفت، از مشخصه‌های نسبی پراکندگی استفاده می‌شود؟

۲۹- ضریب تغییرات چیست؟

۳۰- ضریب تغییرات را برای تمرین ۱۵ محاسبه کنید.

۳۱- ضریب تغییرات را برای تمرین ۱۹ محاسبه کنید.

- ۳۲- ضریب «انحراف خطی نسبی»، چیست؟ در چه مواردی از آن استفاده می‌کنند؟
- ۳۳- انحراف خطی نسبی را برای تمرین ۵ محاسبه کنید.
- ۳۴- انحراف خطی نسبی را برای تمرین ۱۹ محاسبه کنید.
- ۳۵- انحراف چارکی نسبی چیست؟ در چه مواردی از آن استفاده می‌کنند؟
- ۳۶- انحراف چارکی نسبی را برای تمرین ۲۶ محاسبه کنید.
- ۳۷- برای پی‌بردن به خصوصیات عمومی توزیعها، از چه نوع مشخصه‌هایی استفاده می‌شود؟
- ۳۸- چه نوع توزیعی را، توزیع قرینه گویند؟ تعریف کنید.
- ۳۹- توزیع غیرقرینه یا توزیع چوله، چه نوع توزیعی است؟
- ۴۰- مشخصه چولگی یا ضریب چولگی چیست؟ چه خصوصیات از توزیع صفت را بیان می‌کند؟

- ۴۱- ضریب چولگی پیرسن، بر چه اساسی ساخته می‌شود؟
- ۴۲- ضریب چولگی پیرسن را برای تمرین ۱۵ محاسبه کنید.
- ۴۳- ضریب چولگی پیرسن را برای تمرین ۱۹ محاسبه کنید.
- ۴۴- ضریب چولگی منفی، چگونه تعبیر می‌شود؟
- ۴۵- ضریب چولگی مثبت، چگونه تعبیر می‌شود؟
- ۴۶- ضریب چولگی براساس متوسط توان سوم انحرافات چیست؟ فرمول آن را بنویسید.
- ۴۷- ضریب چولگی براساس متوسط توان سوم انحرافات را برای تمرین ۱۵ محاسبه کنید.
- ۴۸- ضریب چولگی براساس متوسط توان سوم انحرافات را برای تمرین ۱۹ محاسبه کنید.

خودآزمونهای چهارگزینه‌ای

- ۱- کدام یک از مشخصه‌های پراکندگی زیر، از نظر دقت، اعتبار بیشتری دارد؟
- الف - دامنه تغییرات ب - واریانس
- ج - متوسط قدرمطلق انحرافات د - انحراف معیار
- ۲- از روی یک نمونه $n=100$ در جامعه، کمیت‌های زیر محاسبه شده است:
- $$F_i X_i = 1200 \quad F_i (X_i - \bar{X})^2 = 36 \quad n=100$$
- واریانس جامعه کدام است؟

الف - ۰/۳۶ ب - ۰/۱۶ ج - ۰/۶ د - ۰/۱۲

۳- اگر مجموع توان دوم، 10 اندازه از صفت متغیر $X^2 = 1700$ و میانگین آنها $\bar{X} = 12$ باشد، واریانس X کدام است؟

الف - ۶ ب - ۱۶ ج - ۲۶ د - صفر

۴- با معلوم بودن $\bar{X} = 10$ و $X^2 = 1200$ ، $V(X) = 20$ ، حجم جامعه چقدر بوده است؟

الف - ۸ ب - ۲۰ ج - ۱۲ د - ۱۰

۵- انحراف چارکی برای توصیف کدام مشخصه صفت، به کار می‌رود؟

الف - مشخصه مرکزی ب - مشخصه پراکندگی

ج - مشخصه چولگی د - هیچ کدام

۶- اگر بین مشخصه‌های مرکزی رابطه $Mo \# Me \# \bar{X}$ برقرار باشد، کدام گزینه برای صفت متغیر درست است؟

الف - چولگی به چپ است. ب - چولگی ندارد.

ج - چولگی به راست است. د - توزیع قرینه است.

۷- کدام یک از روابط زیر، فرمول ضریب چولگی بر پایه میانگین توان سوم انحرافات است؟

الف - $\frac{\sum F_i (X_i - \bar{X})^3}{N}$ ب - $\frac{\sum F_i (X_i - \bar{X})^2}{N}$

ج - $\frac{\sum F_i (X_i - \bar{X})^3}{N}$ د - $\frac{\sum F_i (X_i - \bar{X})^3}{N}$

۸- کدام گزینه تعریف ضریب تغییرات صفت می‌باشد؟

الف - نسبت میانگین بر انحراف معیار صفت

ب - نسبت میانه بر انحراف معیار صفت

ج - نسبت انحراف معیار بر میانه صفت

د - نسبت انحراف معیار بر میانگین صفت

۹- انحراف خطی نسبی از کدام گزینه به دست می آید؟

الف - $\frac{AD}{\bar{X}}$ ب - $\frac{QD}{\bar{X}}$ ج - $\frac{\dot{\quad}}{Me}$ د - $\frac{V(X)}{\bar{X}}$

۱۰- کدام یک از گزینه‌های زیر ضریب انحراف چارکی را بیان می کند؟

الف - $\frac{Q_3 - Q_2}{2}$ ب - $\frac{QD}{Me}$

ج - $\frac{Q_3 - Q_1}{2}$ د - $\frac{Q_3 - Q_1}{Me}$