

شاخص‌های پراکندگی یا تغییرپذیری variability

✓ چگونگی پراکندگی داده‌ها را حول شاخص‌های متوسط داده‌ها توصیف می‌کنند یا میزان پراکندگی یا تغییراتی را که در بین نمره‌های یک توزیع وجود دارد نشان می‌دهد.

✓ این شاخص‌ها شامل: دامنه تغییرات R ، انحراف چارکی Q ، واریانس S^2 ، انحراف استاندارد S می‌باشند.

✓ فرض کنید دو گروه با میانگین ۳۵ (هر کدام) بزرگترین و کوچک‌ترین نمره گروه اول ۸ و ۹۵، بزرگترین و کوچک‌ترین نمره گروه دوم ۳۰ و ۴۱ می‌باشد. با وجود برابری میانگین دو گروه ولی پراکندگی گروه اول بیشتر و گروه دوم همگون‌ترند.

دامنه تغییرات (Range)

✓ دامنه تغییرات تفاوت میان بزرگترین و کوچکترین داده حاصل در یک سری عددی است.
✓ از ساده‌ترین و بی‌اعتبارترین شاخص‌هاست. زیرا تحت تاثیر دو عدد اول و آخر توزیع است.

$$✓ R = x_{max} - x_{min} = X_H - X_L$$

✓ نکته: در داده‌هایی که کمیت‌های پیوسته هستند و اعداد حد دارند، در صورتیکه حد واقعی منظور نشده باشد به این تفاضل یک واحد اضافه می‌کنیم (به دلیل در نظر گرفتن حد پایین واقعی کوچکترین عدد و حد بالای واقعی بزرگترین عدد، اضافه شده است).

✓ این شاخص دارای ۲ ایراد است:

✓ برای نمونه‌های بزرگ شاخص بی‌ثباتی است.

✓ به جز تحت شرایط خاص، مستقل از حجم نمونه است.

به چند دلیل بندرت از دامنه تغییر به عنوان تنها اندازه تغییر پذیری استفاده می‌شود.

✓اول: مقدار این شاخص فقط به مقدار دو نمره که در منتهی الیه بالا و منتهی الیه پایین توزیع قرار دارند بستگی دارد. این مسئله باعث بی ثباتی دامنه تغییر می‌گردد.

✓ مثلاً دامنه تغییر در دو گروه ممکن است فقط به دلیل آنکه نمره یک فرد در یکی از گروه‌ها فوق‌العاده زیاد یا فوق‌العاده کم است، با هم خیلی تفاوت داشته باشند.

✓دوم : دامنه تغییر، الگوی پراکندگی نمره‌ها در داخل توزیع را به طور کامل منعکس نمی‌کند.

✓مثلاً اگر نمره واقع شده در منتهی الیه بالای توزیع خیلی بزرگ باشد، دامنه تغییر نمره خیلی بزرگ خواهد بود. ولی در همین حالت ممکن است سایر نمره‌ها به هم خیلی نزدیک باشند.

دامنه تغییرات

✓مثال: فرض کنید دستمزد ساعتی 7 نفر از کارکنان در سازمان که به طور تصادفی انتخاب شده اند به شرح زیر باشند:

✓ $260-280-240-220-330-200-220 \rightarrow R=(330-200)+1=131$ (الف)

✓ $190-240-175-405-265-290-185 \rightarrow R=(405-175)+1=231$ (ب)

✓نکته: هر چه قدر دامنه تغییر به صفر نزدیکتر باشد بیانگر همگنی دو گروه و هر چه قدر بزرگتر از آن شود بیانگر ناهمگنی دو گروه است.

✓دامنه تغییر چندان معتبر نیست به دو دلیل:

✓1- چون فقط از مقادیر انتهایی توزیع تاثیر می پذیرد.

✓2- تعداد افراد مورد مطالعه در نمونه در آن اثر می گذارد. هر قدر افراد مورد مطالعه کمتر باشد احتمال کوچک بودن دامنه تغییر بیشتر است و بالعکس

دامنه تغییرات

✓ استفاده از دامنه تغییرات مستلزم داشتن مقیاس فاصله‌ای است.

✓ دامنه تغییر مانند نما در گرایش‌های مرکزی، در بین شاخص‌های پراکندگی کم اهمیت‌ترین شاخص است.

✓ مزیت دامنه تغییر: به آسانی محاسبه می‌شود.

✓ R بی‌ثبات‌ترین و بی‌دقت‌ترین شاخص پراکندگی است.

✓ تمامی شاخص‌های پراکندگی در سطح مقیاس فاصله‌ای ونسبی هستند. هر چه جلوتر می‌رویم آماره‌ها از قدرت تبیین بیشتری برخوردارند.

انحراف چارکی Quartile deviation

✓ به آن نصف دامنه چارک‌ها یا چارک متوسط Average Quarter هم می‌گویند
✓ پراکندگی داده‌ها را در اطراف میانه توزیع نشان می‌دهد و از دامنه تغییر با ثبات
تر است.

✓ چارک‌ها نقاطی بر روی مقیاس اندازه‌گیری هستند که کلیه مشاهدات یا نمره‌ها
را به چهار قسمت مساوی تقسیم می‌کنند.

✓ عبارت است از تفاوت نقطه 75 درصدی منهای نقطه 25 درصدی تقسیم بر 2. $Q = \frac{Q_3 - Q_1}{2}$

✓ Q1: چارک اول یا X نقطه ۲۵ درصدی: درست از $\frac{1}{4}$ اعداد بزرگتر هستند.

✓ Q2: چارک متوسط یا میانه یا X نقطه ۵۰ درصدی: نقطه ای است که توزیع را به دو قسمت
مساوی تقسیم می‌کند

✓ Q3: چارک سوم یا X نقطه ۷۵ درصدی: درست از $\frac{3}{4}$ اعداد بزرگترند.

ویژگی های Q

✓ نکته: فاصله ی بین چارک اول و سوم دامنه ی تغییر بین چارک ها نامیده می شود:

$$Q3 - Q1 = \text{دامنه تغییر بین چارک ها}$$

✓ نکته: دامنه تغییر میان دهک اول و نهم را در اصطلاح انحراف دهک ها می نامند: $D = D9 - D1$

✓ اساس محاسبات میانه و انحراف چارکی، نقاط چارکی هستند.

✓ در داده های پرت و داده های دارای توزیع کجی، بهترین شاخص پراکندگی، انحراف چارکی خواهد بود.

✓ وقتی میانگین را داشته باشیم انحراف معیار شاخص مناسبی برای پراکندگی است ولی اگر نمونه دارای کجی خیلی زیاد باشد و شاخص توزیع، میانه باشد نه میانگین؛ باید از انحراف معیار با احتیاط استفاده کرد و در این صورت انحراف چارکی شاخص مناسبی برای نشان دادن پراکندگی است.

✓ از انحراف چارکی زمانی استفاده می شود که مقیاس اندازه گیری حداقل فاصله ای باشد.

الف) محاسبه انحراف چارکی برای اعداد طبقه بندی نشده

✓ مرتب کردن داده‌ها از کوچک به بزرگ

✓ محاسبه میانه یا چارک متوسط

✓ محاسبه اعداد میانه سمت چپ میانه یا چارک اول

✓ محاسبه میانه اعداد سمت راست میانه یا چارک سوم

✓ مثال: در توزیع اندازه‌های زیر چارک اول و سوم را محاسبه کنید.

11-9-7-14-15-19-16

6-8-9-11-12-13-14-17

ب) محاسبه Q برای اعداد طبقه بندی شده

$$Q_1 = L + \frac{\frac{N}{4} - cfb}{fi}(i)$$

$$Q_3 = L + \frac{\frac{3N}{4} - cfb}{fi}(i)$$

$$Q = \frac{Q_3 - Q_1}{2}$$

بعد از ستون cf ستون $cf\%$ را باید تشکیل دهیم و شروع به محاسبه کنیم.

مثال: برای داده‌های زیر Q_1 ، Q_2 ، Q_3 و نقطه ۷۰ درصد را محاسبه کنید

X	Fi	cfi	Cf%
80-84	8	50	۱۰۰
75-79	7	42	۸۴
70-74	5	35	۷۰
65-69	6	30	۶۰
60-64	10	24	۴۸
55-59	8	14	۲۸
50-54	6	6	۱۲

$$Q_1 = 54/5 + \frac{\frac{50}{4} - 6}{8} \times 5 = 58/55$$

تفسیر: ۲۵ درصد افراد نمره زیر

58/55 گرفته‌اند.

$$Q_3 = L + \frac{\frac{3N}{4} - cf}{f} (i)$$

$$= 74/5 + \frac{\frac{150}{4} - 35}{7} \times 5$$

$$= 76/3$$

$$Q = \frac{Q_3 - Q_1}{2} = \frac{76/3 - 58/55}{2} = 17/75$$

تمرین منزل: دامنه تغییرات و انحراف چارکی توزیع
فراوانی زیر را محاسبه کنید؟

X	F	
8	3	
7	4	
6	3	
5	3	
4	2	
3	1	

نقاط درصدی

✓ نقاط درصدی نقاطی در نمره‌های خام هستند که نشان می‌دهند چند درصد نمره‌ها یا اندازه‌ها در زیر آن نمره قرار دارد. نقاط درصدی زمانی به کار برده می‌شوند که پژوهشگر بخواهد از یک جامعه قسمتی یا تعدادی را انتخاب کند.

✓ عکس رتبه درصدی می‌باشد. یعنی رتبه یا درصد نمره مشخص و باید نمره یا داده مورد نظر را پیدا کرد.

✓ نقاط درصدی را با P_x نشان داده و به جای x درصد مورد نظر نوشته می‌شود. (مثلاً نودمین درصد P_{90}).

✓ P_x همانند میانه اندازه‌های ترتیبی هستند.

✓ در صورتیکه ۸۵ درصد از دانشجویان یک کلاس در یک آزمون هوشی کمتر از ۷۵ گرفته باشند، ۷۵، هشتاد و پنجمین نقطه درصدی است.

✓ نکته: نقاط درصدی را می‌توان بدون مراجعه به داده‌ها، از روی نمودار درصد فراوانی تراکمی نیز به دست آورد.

مثال دوم: نقطه ۲۵، ۵۰، ۷۵ درصدی و نقطه ۶۰ درصد را برای داده‌های زیر به دست آورید.

c- l	f	cf
۱-۲	۱	۱
۳-۴	۲	۳
۵-۶	۵	۸
۷-۸	۴	۱۲
۹-۱۰	۸	۲۰
۱۱-۱۲	۸	۲۸
۱۳-۱۴	۱۲	۴۰
۱۵-۱۶	۵	۴۵
۱۷-۱۸	۳	۴۸
۱۹-۲۰	۲	۵۰

$$P_X = L + \frac{P.n - CFb}{Fi} \times I$$

$$P_N = \frac{60 \times 50}{100} = 30$$

$$P_{60} = 12/5 + \frac{30 - 28}{12} \times 2 = 12/83$$

یعنی 60 درصد این داده ها کمتر از 83/12 کسب کرده‌اند.

تمرین منزل: چارک‌ها و نقطه 70 درصد را برای داده‌های زیر به دست آورید.

x	f_i	cf
۲۵-۲۹	۸	۸
۳۰-۳۴	۶	۱۴
۳۵-۳۹	۱۰	۲۴
۴۰-۴۴	۱۳	۳۷
۴۵-۴۹	۵	۴۲
۵۰-۵۴	۸	۵۰

انحراف متوسط یا مجموع قدر مطلق انحرافات از میانگین Average Deviation

Md یا AD متوسط میانگین نمره‌ها را حول میانگین نشان می‌دهد.

انحراف متوسط چون تحت تاثیر تک تک داده‌ها قرار می‌گیرد، از اعتبار بیشتری در مقایسه با دامنه تغییرات و انحراف چارک‌ی برخوردار است.

فاصله بین هر عدد از یکی از شاخص‌های مرکزی (میانه-مد) انحراف نامیده می‌شود.

میانگین انحراف نمره‌ها از میانگین را انحراف متوسط گویند.

$$AD = \frac{\sum |Xi - \bar{X}|}{N}$$

|| قدر مطلق یعنی مقادیر X را بدون در نظر گرفتن علامت جبری (+-) آنها با هم جمع می‌کنیم. ad=ad=

(اگر علامت را منظور کنیم جمع جبری Xها همیشه صفر است)

مثال: انحراف متوسط داده‌های زیر را به دست آورید؟

2-8-14-30-26

نمره‌ها X_i	انحراف نمره‌ها از میانگین $X_i - \bar{X}$	قدر مطلق انحراف نمره‌ها از میانگین $\sum X_i - \bar{X} $
۲	2-16	۱۴
۸	8-16	۸
۱۴	14-16	۲
۳۰	30-16	۱۴
۲۶	26-16	۱۰
	$\sum X - \bar{X} = 0$	$\sum X_i - \bar{X} = 48$

$$\begin{aligned}\bar{X} &= \frac{\sum X}{N} \\ &= \frac{2 + 8 + 14 + 30 + 26}{5} \\ &= 16\end{aligned}$$

$$\begin{aligned}AD &= \frac{\sum |X_i - \bar{X}|}{N} \\ &= \frac{48}{5} = 9.6\end{aligned}$$

محاسبه انحراف متوسط برای داده‌های طبقه بندی شده

cl	Fi	X'	XF i'	$ Xi - \bar{X} $	$ Xi - \bar{X} Fi$
3-5	4	4	16	4-10.7=6.72	26.88
6-8	2	7	14	3.73	7.44
9-11	9	10	90	.72	6.48
12-14	4	13	52	2.28	9.12
15-17	6	16	96	5.28	31.68
C=3	N=25		268		81/6

$$\bar{X} = \frac{\sum FiXi}{N} = \frac{268}{25} = 10.72$$

$$AD = \frac{\sum Fi|Xi - \bar{X}|}{N} = \frac{81/6}{25} = 3/264$$

ویژگی های انحراف متوسط

- ✓ هرچه قدر انحراف از میانگین به صفر نزدیکتر شود نشانه پراکندگی کمتر در صفت یا ویژگی مورد مطالعه در بین آزمودنی‌هاست و برعکس
- ✓ انحراف از میانگین اعداد مساوی صفر می‌شود.
- ✓ انحراف از میانگین زمانی دارای معناست که مقیاس اندازه گیری فاصله‌ای باشد.
- ✓ عیب آن: مقدار آن تحت تاثیر نمره‌های خیلی بزرگ و خیلی کوچک قرار می‌گیرد.