



۸) حداقل مجذور خطا مربوط به میانگین است. $(\sum (x_i - \bar{x})^2 = \min)$

۹) حداقل مجذور قدرمطلق انحرافات مربوط به میانه است. $(\sum |x_i - Md| = \min)$

۱۰) حداقل تعداد تخمین‌های اشتباه مربوط به مد است.

۴) چندک‌ها:

۱- چارک‌ها: اگر داده‌ها را به ترتیب مرتب و آن‌ها را به چهار قسمت مساوی تقسیم کنیم چارک‌ها به دست می‌آیند

چارک اول Q_1 ، چارک دوم Q_2 و چارک سوم Q_3 می‌باشد.

۲- دهک‌ها: اگر داده‌ها را به ترتیب مرتب و آن‌ها را به ده قسمت مساوی تقسیم کنیم دهک‌ها به دست می‌آیند.

دهک اول را با D_1 و ... و دهک نهم را با D_9 نشان می‌دهیم.

۳- صدک‌ها: اگر داده‌ها را به ترتیب مرتب و آن‌ها را به صد قسمت مساوی تقسیم کنیم صدک‌ها به دست می‌آیند.

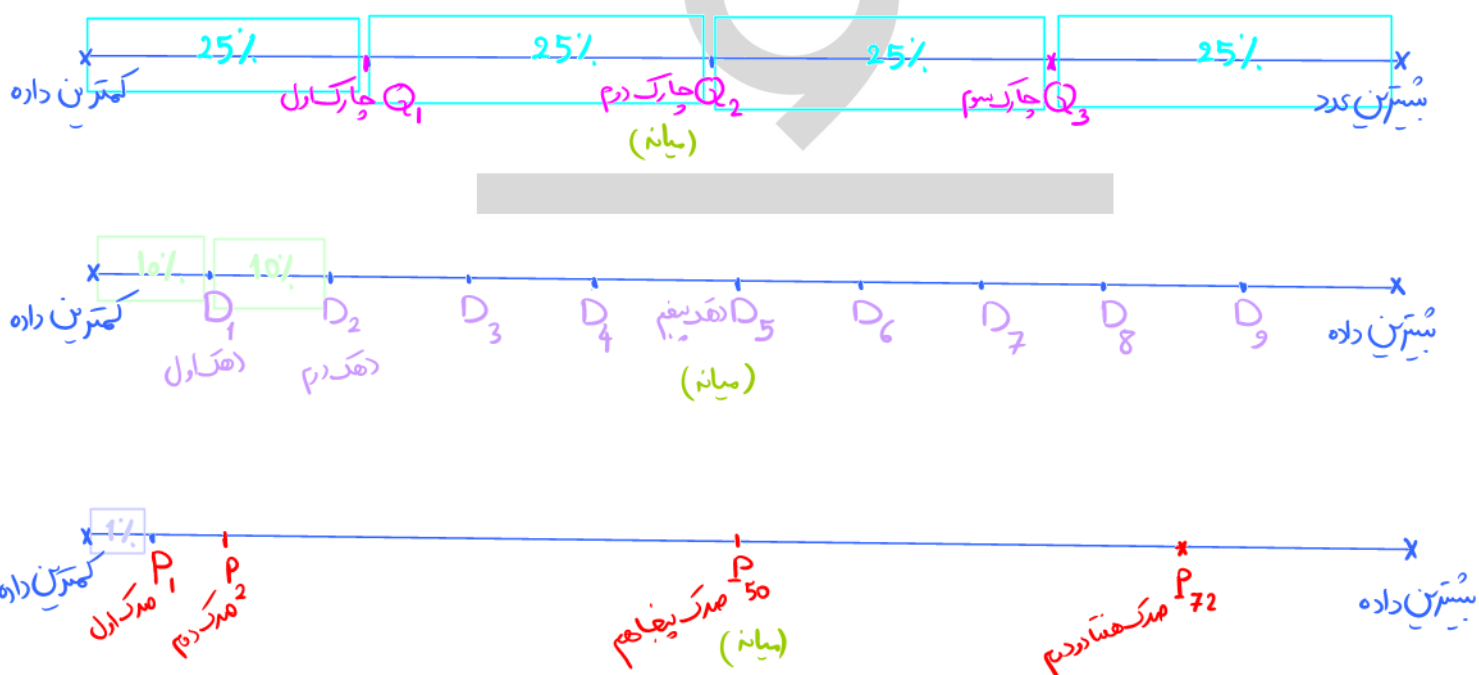
صدک اول را با P_1 و ... و صدک نودونهم را با P_{99} نشان می‌دهیم.

چارک اول، داده‌ای است که تقریباً ۲۵٪ داده‌ها کوچکتر از آن هستند.

دهک دوم عددی است که تقریباً ۲۰٪ اعداد کوچکتر از آن هستند.

صدک چهاردهم داده‌ای است که تقریباً ۱۴٪ اعداد کوچکتر از آن هستند.

میانه = صدک پنجاهم = دهک پنجم = چارک دوم



تمرین: (دکتری - ۹۹) در یک مطالعه آماری در خصوص نمرات درس زبان انگلیسی زبان آموخته‌گان در یک موسسه خصوصی کمیت‌های زیر حاصل شده‌اند، به طوری که $\bar{x}$ ، Q_1 ، Q_3 ، s^2 ، Mo ، Md ، میانۀ، نما، واریانس، چارک اول و چارک سوم توزیع می‌باشد. کدام مورد درست است؟

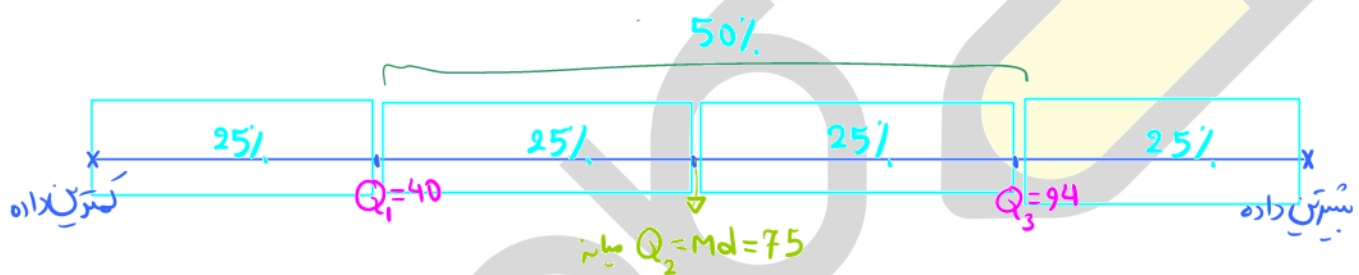
$$\bar{x} = 70, \quad Md = 75, \quad Mo = 83, \quad s^2 = 400, \quad Q_1 = 40, \quad Q_3 = 94$$

(۲) نمرات نیمی از آنان کمتر از 75 است.

(۱) نمرات $\frac{1}{4}$ از آنان بین 40 تا 94 است.

(۴) نمرات بیشتر آنان 94 و بیشتر است.

(۳) نمرات $\frac{3}{4}$ از آنان بیش از 94 است.



نکته: هس ۵۰٪ داده‌ها بین چارک اول و سوم هستند.



برای محاسبه چندک‌ها به ترتیب زیر عمل می‌کنیم:

مرحله اول: داده‌ها را از کوچک به بزرگ مرتب می‌کنیم.

مرحله دوم: اگر اندیس Q ، D یا P را با a نشان دهیم، در چارک‌ها $\frac{aN}{4} + \frac{1}{4}$ ، در دهک‌ها $\frac{aN}{10} + \frac{1}{10}$ و در صدک‌ها $\frac{aN}{100} + \frac{1}{100}$ را محاسبه می‌کنیم.

مرحله سوم: دو حالت رخ می‌دهد:

حالت اول: اگر حاصل عدد صحیح n شد، n آمین عدد همان چندک خواسته شده است.

حالت دوم: اگر حاصل عدد اعشاری شد، بخش صحیح آن را با n و بخش اعشار را با r نشان می‌دهیم. در این صورت:

$$\text{چندک خواسته شده} = x_n + r(x_{n+1} - x_n)$$

تمرین: برای داده‌های زیر چارک سوم و دهک چهارم را بیابید.

هشتمین عدد Q_3 ← چارک سوم در
۰, ۱, ۲, ۲, ۲, ۴, ۴, ۵, ۵, ۷

حل Q_3 چارک سوم $\Rightarrow n=3 \Rightarrow \frac{3 \times 10}{4} + \frac{1}{4} = 8$ ← چهارم
هشتمین عدد $Q_3 = 5$

نقد داده‌ها × اندیس $+ \frac{1}{2}$
۱۰۰ ← صدک
۱۰ ← دهک
۴ ← چارک

D_4 دهک چهارم $\Rightarrow n=4 \Rightarrow \frac{4 \times 10}{10} + \frac{1}{10} = 4.5$
(چارک سوم - بیستمین عدد) + ۰.۵ (چارک سوم عدد)

$$D_4 = 2 + 0.5(2 - 2) = 2$$





محاسبه چندک‌ها برای داده‌های طبقه‌بندی شده:

اگر داده‌ها و جدول‌های فراوانی خلاصه شده باشند برای محاسبه چندک خواسته شده به ترتیب عمل می‌کنیم:

مرحله اول: فراوانی تجمعی هر طبقه را محاسبه می‌کنیم.

مرحله دوم: برای چارک‌ها $\frac{aN}{4}$ ، برای دهک‌ها $\frac{aN}{10}$ و برای صدک‌ها $\frac{aN}{100}$ را محاسبه می‌کنیم.

مرحله سوم: اولین طبقه‌ای که فراوانی تجمعی آن از عدد حاصل بزرگتر یا با آن مساوی است را به عنوان طبقه چندک در نظر می‌گیریم.

$$Q_a = \text{حد پایین طبقه} + \frac{\text{فراوانی تجمعی رده قبل} - \frac{aN}{4}}{\text{فراوانی رده}} \times \text{طول رده}$$

$$D_a = \text{حد پایین طبقه} + \frac{\text{فراوانی تجمعی رده قبل} - \frac{aN}{10}}{\text{فراوانی رده}} \times \text{طول رده}$$

$$P_a = \text{حد پایین طبقه} + \frac{\text{فراوانی تجمعی رده قبل} - \frac{aN}{100}}{\text{فراوانی رده}} \times \text{طول رده}$$





شاخص‌های پراکندگی

پارامترهای پراکندگی در واقع معیارهایی برای تعیین میزان پراکندگی داده‌ها از یکدیگر و یا میزان پراکندگی داده‌ها نسبت به میانگین است.

(۱) دامنه تغییرات: تفاوت بین بیشترین و کمترین داده را دامنه تغییرات می‌نامند.

$$R = x_{\max} - x_{\min}$$

سؤال $2, \overset{\text{min}}{\textcircled{1}}, 5, 7, 5, \overset{\text{max}}{\textcircled{15}}, 9 \Rightarrow R = 15 - 1 = 14$

سؤال

حداکثر صفا	$\overset{\text{min}}{\textcircled{2}} - 5$	$5 - 8$	$8 - \overset{\text{max}}{\textcircled{11}}$
فراوانی	7	3	9

 $\Rightarrow R = 11 - 2 = 9$

سؤال

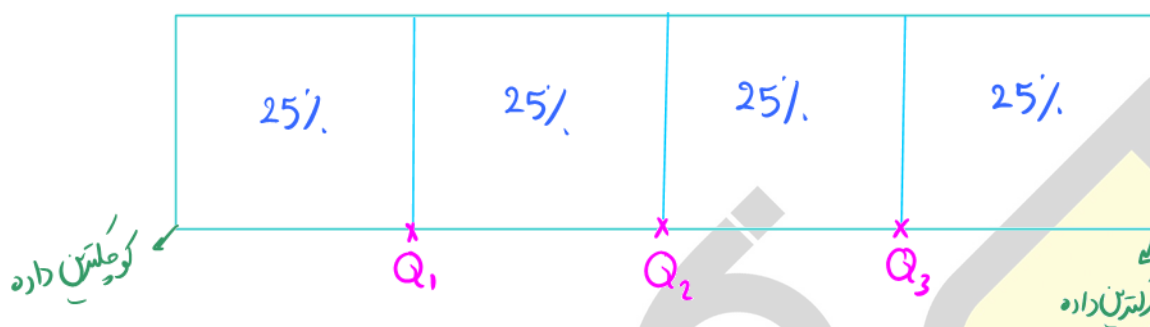
مرکز صفا	$\overset{\text{min}}{\textcircled{-2}}$	0	1	$\overset{\text{max}}{\textcircled{4}}$
فراوانی	1	5	2	1

 $\Rightarrow R = 4 - (-2) = 6$



(۲) دامنه چارک‌ها: تفاوت چارک سوم و چارک اول را دامنه چارک‌ها می‌نامند که شامل ۵۰٪ اندازه‌های متوسط است.

$$IQR = Q_3 - Q_1$$



چارک اول - چارک سوم = $IQR = Q_3 - Q_1$ (دامنه چارک‌ها)

$\div 2$

(انحراف چارکی) $SIQR = \frac{Q_3 - Q_1}{2}$

(۳) انحراف چارکی: اگر دامنه چارک‌ها را نصف کنیم انحراف چارکی به دست می‌آید.

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

معمولاً در داده‌هایی که شامل داده‌های پرت می‌باشند یا حدود باز دارند از انحراف چارکی برای بررسی پراکندگی داده‌ها استفاده می‌شود زیرا تحت تأثیر این اعداد پرت قرار نمی‌گیرد.



تمرین: (ارشد - ۹۸) نمرات مسئولیت پذیری تعدادی از کارکنان یک شرکت خدماتی به صورت زیر تنظیم شده است. انحراف چارگی این توزیع کدام است؟

فاصله طبقات	< ۲۰	۲۰ - ۲۵	۲۵ - ۳۰	۳۰ - ۳۵	≥ ۳۵
فراوانی	۸	۱۵	۱۹	۱۸	۱۶
	۵/۲۵ (۲✓)		۴/۷۵ (۳)		۴/۳۳ (۲)

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

ردیف	حدود صغائر	فراوانی	فراوانی تجمعی
۱	< ۲۰	۸	۸
۲	۲۰ - ۲۵	۱۵	۲۳
۳	۲۵ - ۳۰	۱۹	۴۲
۴	۳۰ - ۳۵	۱۸	۶۰
۵	≥ ۳۵	۱۶	۷۶
جمع		N = ۷۶	

$$Q_1 \Rightarrow \text{اندیس} = 1$$

$$N = 76 \text{ مقدار داده ها}$$

$$\frac{1 \times 76}{4} = 19$$

به سون فراوانی تجمعی نگاه می کنیم. اولین ده ای که فراوانی تجمعی آن یا ۱۹ باشد یا از ۱۹ بزرگتر باشد.

$$\text{طول رده} \times \frac{\text{فراوانی تجمعی رده قبل} - \text{مقدار درست آماره}}{\text{فراوانی رده}} + \text{حد پایین رده} = \text{میدک}$$

$$Q_1 = 20 + \frac{19 - 8}{15} \times 5 = 23.7$$

$$Q_3 \Rightarrow \text{اندیس} = 3 \Rightarrow \frac{3 \times 76}{4} = 57 \Rightarrow \text{به سون فراوانی تجمعی نگاه می کنیم اولین رده ای که فراوانی تجمعی آن ۵۷ یا بزرگتر است}$$

$$Q_3 = 30 + \frac{57 - 42}{18} \times 5 = 34.2$$

$$\text{انحراف چارگی} = \frac{34.2 - 23.7}{2} = 5.25$$





تمرین: (ارشد - ۹۷) در داده‌های دسته‌بندی شده با جدول فراوانی زیر، انحراف چارگی کدام مورد است؟

x	< 9	9 - 13	13 - 17	17 - 21	21 - 25	≥ 25
f	3	7	8	12	15	7

۴/۴۵ (۴) ✓

۴/۲۵ (۳)

۳/۹۵ (۲)

۳/۷۵ (۱)

حدود میانی	فراوانی	فراوانی نسبی
< 9	3	3
9 - 13	7	10
13 - 17	8	18
17 - 21	12	30
21 - 25	15	45
≥ 25	7	52
جمع	N=52	

$$Q_1 \Rightarrow \text{اندیس} = 1 \Rightarrow \frac{1 \times 52}{4} = 13$$

چارک اول

$$\text{طول رده} \times \frac{\text{فراوانی نسبی رده قبل} - \text{مقدار رده قبل}}{\text{فراوانی رده}} = \text{مقدار رده}$$

$$Q_1 = 13 + \frac{13 - 10}{8} \times 4 = 14.5$$

$$Q_3 \Rightarrow \text{اندیس} = 3 \Rightarrow \frac{3 \times 52}{4} = 39$$

چارک سوم

$$Q_3 = 21 + \frac{39 - 30}{15} \times 4 = 23.4$$

$$SIQR = \frac{23.4 - 14.5}{2} = 4.45$$

(انحراف چارگی)





تمرین: (ارشد - ۹۶) نمرات مسئولیت پذیری کارکنان یک شرکت در جدول زیر دسته بندی شده است، انحراف چارکی این نمرات کدام است؟

فاصله طبقاتی	< ۱۱	۱۱ - ۱۳	۱۳ - ۱۵	۱۵ - ۱۷	≥ ۱۷
فراوانی	۵	۷	۱۵	۱۹	۱۰
	۱/۷۵ (۴)		۱/۶۶ (۳)		۱/۶۱ (۲)

طول رد، $\times$ $\frac{\text{تجایی قبل - عدد رده قبلی}}{\text{فراوانی رده قبلی}} + \text{حد پایین رده} = \text{حد کوچک}$

فراوانی مطلق	فراوانی نسبی	فراوانی نسبی
< ۱۱	۵	۵
۱۱ - ۱۳	۷	۱۲
۱۳ - ۱۵	۱۵	۲۷
۱۵ - ۱۷	۱۹	۴۶
≥ ۱۷	۱۰	۵۶

رده مربوط به چارک اول

رده مربوط به چارک سوم

$Q_1 = \text{چارک اول}$

$$a=1 \left. \begin{array}{l} \\ N=56 \end{array} \right\} \Rightarrow \frac{aN}{4} = \frac{56}{4} = 14 \Rightarrow Q_1 = 13 + \frac{14 - 12}{15} \times 2 = 13.267$$

$Q_2 = \text{چارک دوم (میان)}$

$$a=2 \left. \begin{array}{l} \\ N=56 \end{array} \right\} \Rightarrow \frac{aN}{4} = 28 \Rightarrow Q_2 = 15 + \frac{28 - 27}{19} \times 2 = 15 + \frac{2}{19}$$

$Q_3 = \text{چارک سوم}$

$$a=3 \left. \begin{array}{l} \\ N=56 \end{array} \right\} \Rightarrow \frac{aN}{4} = 42 \Rightarrow Q_3 = 15 + \frac{42 - 27}{19} \times 2 = 15 + \frac{15}{19} = 16.579$$



$$\text{انحراف} = \frac{16.579 - 13.267}{2} = 1.656$$

۴) واریانس: یکی از مهم‌ترین شاخص‌های پراکندگی واریانس می‌باشد. واریانس را معمولاً با یکی از نمادهای زیر نشان می‌دهند:

$$\sigma^2, V(x), \text{Var}(x), D(x)$$

حالت اول: اگر فراوانی داده‌ها، همگی یک باشد:

$$\sigma^2 = \frac{\sum (x_i - \mu)^2}{N} \quad \text{یا} \quad \sigma^2 = \frac{\sum x_i^2}{N} - \left(\frac{\sum x_i}{N} \right)^2$$

حالت دوم: اگر فراوانی مطلق داده‌ها، داده شده باشد:

$$\sigma^2 = \frac{\sum F_i (x_i - \mu)^2}{N} \quad \text{یا} \quad \sigma^2 = \frac{\sum F_i x_i^2}{N} - \left(\frac{\sum F_i x_i}{N} \right)^2$$

حالت سوم: اگر فراوانی نسبی داده‌ها، داده شده باشد:

$$\sigma^2 = \sum f_i (x_i - \mu)^2 \quad \text{یا} \quad \sigma^2 = \sum f_i x_i^2 - \left(\sum f_i x_i \right)^2$$

۵) انحراف معیار: جذر مثبت واریانس را انحراف معیار نامیده و آن را با σ نشان می‌دهند.

تمرین: برای داده‌های زیر واریانس را بیابید.

۱, ۲, ۳, ۴