درس دوم

آمار توصیفی

مقدمه

آمار توصیفی به خلاصه‌سازی داده‌ها در قالب نمودار، جدول و یا شاخص‌هایی در قالب معیارهای گرایش به مرکز و معیارهای پراکندگی که در ادامه با آنها آشنا خواهید شد، می‌پردازد. آمارتوصیفی اطلاعاتی از چگونگی داده‌های جمع‌آوری شده فراهم می‌آورد که بسیار مفید است.

معیارهای گرایش به مرکز

معمولاً سعی می شود، دانسته‌های نهفته در داده‌ها را به‌صورت یک یا چند عدد شاخص درآورد، تا بتوان هم اندیشهٔ کلی دربارهٔ ویژگی مورد مطالعه به‌دست آورد و هم نتیجه مطالعات را بسادگی گزارش کرد. میانگین و میانه به عنوان معیارهای گرایش به مرکز در این کتاب معرفی می‌شوند.

میانگین

میانگین ساده‌ترین و در عین حال پرکاربردترین معیار گرایش به مرکز است که در پایهٔ هشتم با آن آشنا شده‌اید.

میانگین، متوسط یا مرکز ثقل داده‌هاست که آن را با $\overline{X}$ نشان می‌دهیم و برابر است با :

$$\overline{X}=\frac{x_۱+x_۲+\cdots+x_N}{N}$$

که در آن x_i داده‌ها و N برابر با تعداد کل داده‌ها است.

فعالیت

محمد، جرم ۵ نفر از دوستان خود را پرسید و آنها را در جدول زیر یادداشت کرد. سپس میانگین جرم دوستان خود را حساب کرد :

دوست	رضا	نیما	سام	احمد	علی
جرم (کیلوگرم)	۵۵	۶۱	۵۷	۵۵	۶۲

نحوهٔ محاسبه میانگین

۱ محمد ابتدا مجموع جرم دوستان خود را محاسبه کرد :

۶۲+۵۵+۵۷+۶۱+۵۵=۲۹۰

۲ سپس عدد حاصل را بر عدد ۵ (تعداد دوستان) تقسیم کرد :

$\frac{۲۹۰}{۵}=۵۸$

* $\bar{x}_{ax} = \frac{ax_1 + ax_2 + \cdots + ax_N}{N} = \frac{a(x_1 + x_2 + \cdots + x_N)}{N} = a\bar{x}$

* $\bar{x}_{x+a} = \frac{(x_1+a) + (x_2+a) + \cdots + (x_N+a)}{N} = \frac{x_1 + x_2 + \cdots + x_N + Na}{N}$

$= \frac{x_1 + x_2 + \cdots + x_N}{N} + \frac{Na}{N} = \bar{x} + a$

میانگین جرم دوستان محمد برابر است با ...۵۸

ویژگی‌های میانگین

اگر هر یک از داده‌های آماری با مقدار ثابتی جمع شود، میانگین آنها نیز با همان مقدار ثابت جمع خواهد شد. چرا؟ *

اگر هر یک از داده‌های آماری در مقدار ثابتی ضرب شود، میانگین آنها نیز در همان مقدار ثابت ضرب خواهد شد. چرا؟ *

کار در کلاس

۱ در فعالیت قبل، میانگین جرم دوستان محمد چند گرم است؟ ۵۸۰۰۰ = ۵۸×۱۰۰۰

۲ هوای اهواز در هر ساعت از یک روز بهاری گزارش شد. اگر میانگین دمای هوا ۲۸ درجهٔ سانتی‌گراد باشد، میانگین دمای هوا چند درجهٔ فارنهایت است؟ (راهنمایی $F = \frac{9}{5}C + 32$)

$F = \frac{9}{5}(28) + 32 = \frac{252}{5} + 32 = 50{,}4 + 32 = 82{,}4$

میانه

پس از مرتب کردن داده‌ها، مقداری را که تعداد داده‌های بعد از آن با تعداد داده‌های قبل از آن برابر است میانه می‌نامیم و آن را با Q_2 نمایش می‌دهیم.

مثال : در فعالیت قبل، میانهٔ داده‌ها کدام است؟

محمد برای پاسخ به این سؤال :

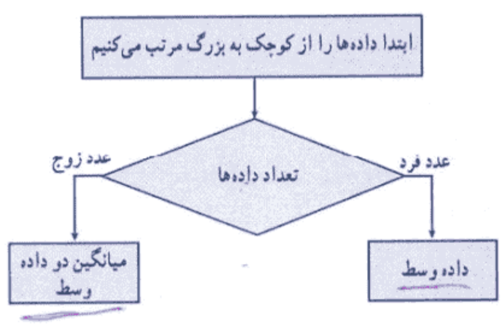
الف) داده‌ها را از کوچک به بزرگ مرتب کرد :

۵۵ ۵۵ ۵۷ ۶۱ ۶۲

ب) جرم رضا و احمد از سام کمتر است. در حالی‌که جرم علی و نیما از سام بیشتر است.

در مثال فوق، عدد ۵۷ را میانهٔ داده‌ها می‌نامند، زیرا پس از مرتب کردن داده‌ها از کوچک به بزرگ، در وسط داده‌ها قرار می‌گیرد.

روش محاسبهٔ میانه :



مثال :اعداد زیر نمره‌های درس ریاضی سمیرا در طول یک سال است.

ـ میانگین و میانهٔ نمره‌های او را حساب کنید.

۱۹ ۱۷ ۱۸ ۱۸ ۲۰ ۵

الف) محاسبهٔ میانگین

$$\bar{X}=\frac{۱۹+۱۷+۱۸+۱۸+۲۰+۵}{۶}\simeq ۱۶/۱۷$$

ب) محاسبهٔ میانه

۵, ۱۷, ۱۸, ۱۸, ۱۹, ۲۰

$$Q_۲=\frac{۱۸+۱۸}{۲}=۱۸$$

پ) به نظر شما کدام معیار توانایی دانش‌آموز در این درس را بهتر ارزیابی می‌کند؟ چرا؟ میانه، چون نمره ۵ باعث شده میانگین خیلی کم شود. درصورتی که این دانش آموز اکثراً نمرات خوبی داشته است.

میانگین داده‌ها تحت تأثیر داده‌های خیلی بزرگ یا خیلی کوچک که در آمار به آنها داده‌های دورافتاده می‌گوییم، قرار می‌گیرد. درصورتی که میانهٔ داده‌ها تحت تأثیر داده‌های دور افتاده قرار نمی‌گیرد. بنابراین درصورت وجود دادهٔ دورافتاده، از میانه استفاده می‌کنیم در غیراین صورت از میانگین استفاده خواهیم کرد.

کار در کلاس

داده‌های زیر مربوط به تعداد ضربان قلب ۱۲ دانش‌آموز پایهٔ یازدهم، قبل از یک مسابقهٔ دو است.

۱۰۰ ۹۱ ۸۲ ۷۵ ۱۰۵ ۹۸ ۹۸ ۱۰۱ ۸۹ ۹۲ ۹۷ ۸۶

ـ میانهٔ داده‌ها را مشخص کنید: ۷۵ ۸۲ ۸۶ ۸۹ ۹۱ ۹۲ ۹۷ ۹۸ ۹۸ ۱۰۰ ۱۰۱ ۱۰۵

$$Q_۲=\frac{۹۲+۹۷}{۲}=۹۴/۵$$

ـ میانگین داده‌ها را مشخص کنید.

$$\bar{x}=\frac{۷۵+۸۲+۸۶+۸۹+۹۱+۹۲+۹۷+۹۸+۹۸+۱۰۰+۱۰۱+۱۰۵}{۱۲}=\frac{۱۱۱۴}{۱۲}\approx ۹۲/۸۳$$

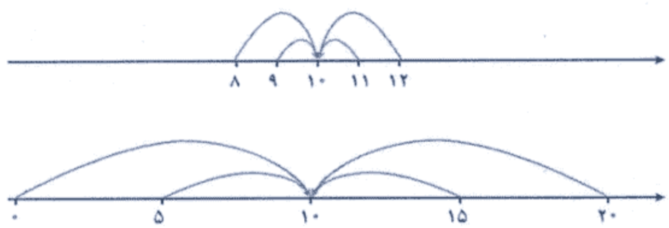
معیارهای پراکندگی

میانه و میانگین اطلاعاتی پیرامون مرکز داده‌ها در اختیار ما قرار می‌دهند. گاه در توصیف داده‌ها لازم است از چگونگی پراکندگی آنها نیز اطلاعی داشته باشیم. در این درس با دامنهٔ تغییرات، واریانس، انحراف معیار، چارک اول و چارک سوم به‌عنوان معیارهای پراکندگی آشنا خواهیم شد.

فعالیت

نمرهٔ درس ریاضی دانش‌آموزان دو کلاس A و B، به تفکیک گزارش شده است :

۱۲	۱۱	۱۰	۹	۸	A
۲۰	۱۵	۱۰	۵	۰	B



الف) میانهٔ نمرهٔ این دو کلاس را محاسبه کنید.

۸ ۹ ۱۰ ۱۱ ۱۲ $\quad Q_2 = ۱۰$

۰ ۵ ۱۰ ۱۵ ۲۰ $\quad Q'_2 = ۱۰$

ب) میانگین نمرهٔ این دو کلاس را محاسبه کنید.

$\bar{x}_A = \frac{۸+۹+۱۰+۱۱+۱۲}{۵} = \frac{۵۰}{۵} = ۱۰$ $\qquad \bar{x}_B = \frac{۰+۵+۱۰+۱۵+۲۰}{۵} = \frac{۵۰}{۵} = ۱۰$

پ) به نظر شما یک معلم ریاضی ترجیح می‌دهد در کدام کلاس تدریس کند؟ چرا؟

همان‌طور که در فعالیت می‌بینید، تنها توجه به معیارهای گرایش به مرکز نمی‌تواند اطلاعات کاملی از داده‌ها در اختیار ما قرار دهد و لازم است به چگونگی پراکندگی داده‌ها نیز توجه شود.

دامنهٔ تغییرات

دامنهٔ تغییرات ساده‌ترین شاخص پراکندگی است که اختلاف بین بزرگ‌ترین و کوچک‌ترین داده‌ها را نشان می‌دهد و آن را با نماد R نمایش می‌دهیم.

مثال : در فعالیت بالا برای محاسبهٔ دامنهٔ تغییرات نمره ریاضی کلاس A و کلاس B به‌صورت زیر عمل کنید :

	کلاس B	کلاس A
کوچک‌ترین داده	۰	۸
بزرگ‌ترین داده	۲۰	۱۲
دامنه تغییرات	۲۰ − ۰=۲۰	۱۲ −۸ =۴

در فعالیت بالا دامنهٔ تغییرات نمرهٔ ریاضی کلاس A، ۴ نمره و دامنهٔ تغییرات نمرهٔ ریاضی کلاس B، ۲۰ نمره است. ملاحظه می‌شود که در کلاس A پراکندگی داده‌ها کمتر از کلاس B است.

تهیه کننده :
گروه ریاضی دوره ی دوم متوسطه و انجمن معلمان ریاضی، استان خوزستان
khuzmath1394@chmail.ir

کار در کلاس

معلم از ۷ نفر از دانش‌آموزان خواست تا تعداد کتاب‌های غیردرسی را که در طول تابستان گذشته مطالعه کرده‌اند، گزارش کنند.

الف) دامنه تغییرات آنها را محاسبه کنید.

۱۵ ۸ ۹ ۱۲ ۱۴ ۴ ۱ $R = 15 - 1 = 14$

ب) دو دانش‌آموز دیگر به جمع آنها اضافه شدند و آنها نیز تعدادکتاب‌های غیردرسی را که در طول تابستان گذشته مطالعه کرده بودند، به ترتیب ۵ و ۱۱ اعلام کردند. مجدداً دامنهٔ تغییرات این ۹ داده را محاسبه کنید. $R = 15 - 1 = 14$

پ) از مقایسهٔ پاسخ الف و ب چه نتیجه‌ای می‌گیرید؟ هردو با هم برابرند. چون دامنه تغییرات فقط به بزرگترین داده و کوچکترین داده بستگی دارد.

همان طور که می‌بینید، دامنهٔ تغییرات تنها به بزرگ‌ترین و کوچک‌ترین داده‌ها وابسته است و با تغییر تعداد و مقدار داده‌های میانی، مقدار آن تغییر نخواهد کرد. پس این معیار نمی‌تواند بیانگر خوبی برای پراکندگی داده‌ها باشد.

واریانس

می‌خواهیم شاخص بهتری برای بیان پراکندگی داده‌ها پیدا کنیم. از آنجا که میانگین، معیاری برای مرکز داده‌ها است، شاخصی که بیانگر اختلاف داده‌ها از میانگین باشد و معایب وارد بر دامنه تغییرات را برطرف سازد، می‌تواند شاخص بهتری برای بیان پراکندگی داده‌ها باشد.

فعالیت

الف) در ادامهٔ فعالیت قبل اختلاف از میانگین را برای نمره‌های ریاضی کلاس A و B به کمک جدول‌های زیر محاسبه کنید.

تهیه کننده:
گروه ریاضی دوره ی دوّم متوسطه و انجمن معلمان ریاضی، استان خوزستان
khuzmath1394@chmail.ir

کلاس A	
x_i	$(x_i - \overline{X})$
۸	۸−۱۰=−۲
۹	۹−۱۰=−۱
۱۰	۱۰−۱۰=۰
۱۱	۱۱−۱۰=۱
۱۲	۱۲−۱۰=۲

کلاس B	
y_i	$(y_i - \overline{Y})$
۰	۰−۱۰=−۱۰
۵	۵−۱۰=−۵
۱۰	۱۰−۱۰=۰
۱۵	۱۵−۱۰=۵
۲۰	۲۰−۱۰=۱۰

ب) مجموع اختلاف داده‌ها از میانگین را برای هر کلاس محاسبه کنید.

اختلاف داده‌ها از میانگین کلاس A $= (-2) + (-1) + 0 + 1 + 2 = 0$

مجموع اختلاف داده‌ها از میانگین کلاس B $= (-10) + (-5) + 0 + 5 + 10 = 0$

در هر دو کلاس، مجموع اختلاف داده‌ها از میانگین داده‌ها صفر شد. با مراجعه به تعریف میانگین، بدیهی است این نتیجه اتفاقی نبوده است.

> همواره برای هرمجموعه‌ای از داده‌ها، مجموع اختلاف داده‌ها از میانگین صفر خواهد شد.

بنابراین برای ساختن شاخصی که پراکندگی حول میانگین را نشان دهد، باید از قدرمطلق اختلاف داده‌ها از میانگین یا از مجذور اختلاف داده‌ها از میانگین استفاده کرد. استفاده از مجذور اختلاف داده‌ها از میانگین متداول‌تر است.

الف) مجذور اختلاف از میانگین برای نمره‌های ریاضی کلاس A و B را به کمک جداول زیر محاسبه کنید.

کلاس B

y_i	$(y_i - \overline{Y})$	$(y_i - \overline{Y})^۲$
۰	-۱۰	۱۰۰
۵	-۵	۲۵
۱۰	۰	۰
۱۵	۵	۲۵
۲۰	۱۰	۱۰۰

کلاس A

x_i	$(x_i - \overline{X})$	$(x_i - \overline{X})^۲$
۸	-۲	۴
۹	-۱	۱
۱۰	۰	۰
۱۱	۱	۱
۱۲	۲	۴

ب) مجموع مجذور اختلاف داده‌ها از میانگین را برای هر کلاس محاسبه کنید.

مجموع مجذور اختلاف داده‌ها از میانگین برای کلاس A	$(۸-۱۰)^۲ + ... + (۱۲-۱۰)^۲ = ۱۰$
مجموع مجذور اختلاف داده‌ها از میانگین برای کلاس B	$(۰-۱۰)^۲ + ... + (۲۰-۱۰)^۲ = ۲۵۰$

پ) میانگین مجذور اختلاف داده‌ها از میانگین را برای هر کلاس محاسبه و مقایسه کنید.

میانگین مجذور اختلاف داده‌ها از میانگین برای کلاس A	$\frac{(۸-۱۰)^۲+\cdots+(۱۲-۱۰)^۲}{۵} = \frac{۱۰}{۵} = ۲$
میانگین مجذور اختلاف داده‌ها از میانگین برای کلاس B	$\frac{(۰-۱۰)^۲+\cdots+(۲۰-۱۰)^۲}{۵} = \frac{۲۵۰}{۵} = ۵۰$

> میانگین مجذور اختلاف داده‌ها از میانگین آنها را واریانس می نامند و از نماد $\sigma^۲$ برای نمایش آن استفاده می‌شود :
> $$\sigma^۲ = \frac{(x_۱ - \overline{X})^۲ + \cdots + (x_N - \overline{X})^۲}{N}$$

تذکر : واحد واریانس برابر با توان دوم واحد دادهٔ مورد نظر است.

تهیه کننده:
گروه ریاضی دوره ی دوم متوسطه و انجمن معلمان ریاضی، استان خوزستان
khuzmath1394@chmail.ir

$\bar{x} = \frac{15+8+8+12+14+4+1}{7} = \frac{62}{7} \approx 8.85 \to 9$

$\sigma^2 = \frac{(6)^2+(1)^2+(1)^2+(3)^2+(5)^2+(5)^2+(8)^2}{7} = \frac{161}{7} = 23$

همان‌طور که در فعالیت قبل دیده می‌شود، واریانس بزرگ (کلاس B) نشان دهندهٔ دور بودن داده‌ها از میانگین آنها و واریانس کوچک (کلاس A) نشان دهندهٔ نزدیکی داده‌ها به میانگین آنهاست. چنانچه همه داده‌ها با هم برابر باشند، واریانس آنها صفر خواهد بود. بنابراین واریانس معیار خوبی برای سنجش پراکندگی و تغییرپذیری داده‌ها نسبت به میانگین است.

$\bar{x} = \frac{15+8+8+12+14+4+1+5+11}{9} = \frac{78}{9} \approx 8.67 \to 9$

کار در کلاس

واریانس تعداد کتاب‌های غیردرسی مطالعه شده در «کاردرکلاس» قبل، توسط ۷ و ۹ دانش‌آموز را محاسبه کنید.

$\sigma^2 = \frac{(6)^2+(1)^2+(1)^2+(3)^2+(5)^2+(5)^2+(8)^2+(4)^2+(2)^2}{9} = \frac{181}{9} \approx 20.11$

واریانس	دامنه تغییرات	تعداد کتاب‌های مطالعه شده توسط هر دانش‌آموز
۲۳	۱۴	۱۵ ۸ ۸ ۱۲ ۱۴ ۴ ۱
۲۰/۱۱	۱۴	۱۵ ۸ ۸ ۱۲ ۱۴ ۴ ۱ ۵ ۱۱

همان‌طور که در این «کاردرکلاس»، دیده می‌شود، واریانس برخلاف دامنهٔ تغییرات با تغییر تعداد و مقادیر داده‌ها تغییر می‌کند.

پس اختلاف از میانگین تغییری نمی‌کند

ویژگی‌های واریانس

* اگر هر یک از داده‌های آماری با مقدار ثابتی جمع شود، واریانس آنها تغییر نخواهد کرد. چرا؟ چون میانگین هم به همان مقدار اضافه می‌شود
* اگر هر یک از داده‌های آماری در مقدار ثابتی ضرب شود، واریانس آنها در مجذور همان مقدار ثابت ضرب خواهد شد. چرا؟

چون میانگین هم در همان عدد ضرب می‌شود پس اختلاف از میانگین‌ها هم در همان عدد ضرب می‌شود و چون که اختلاف از میانگین‌ها به توان ۲ می‌رسد این عدد هم به توان ۲ می‌رسد.

کار در کلاس

۱) در اولین فعالیت، واریانس جرم دوستان محمد چند گرم به توان دو است؟

$\sigma^2 = \frac{(62-58)^2+(55-58)^2+(57-58)^2+(61-58)^2+(55-58)^2}{5} = \frac{16+9+1+9+9}{5} = \frac{44}{5} = 8.8$

۲) هوای اهواز در هر ساعت از یک روز بهاری گزارش شد. اگر واریانس دمای هوا ۶ درجه سانتی‌گراد به توان دو باشد، واریانس دمای هوا چند درجهٔ فارنهایت به توان دو است؟ (راهنمایی $F = \frac{9}{5}C + 32$) جمع با ۳۲ تأثیری ندارد. اما باید در توان دو $\frac{9}{5}$ ضرب شود.

$\sigma^2 = 6 \times \left(\frac{9}{5}\right)^2 = \frac{486}{25} = 19.44$

تهیه کننده:
گروه ریاضی دوره‌ی دوم متوسطه و انجمن معلمان ریاضی، استان خوزستان
khuzmath1394@chmail.ir

انحراف معیار

معیارهای گرایش به مرکز و پراکندگی فعالیت قبل در جدول زیر آمده است.

اشتباه تایپی

	میانگین	میانه	دامنه تغییرات	واریانس	جذر واریانس
کلاس A	۱۰	۱۰	۴	~~۲/۵~~ ۲	~~۱/۶~~ $\sqrt{2} \approx 1.41$
کلاس B	۱۰	۱۰	۲۰	~~۶۲/۵~~ ۵۰	~~۷/۹~~ $\sqrt{50} \approx 7.07$

همان‌طور که در جدول و نمودار بالا دیده می‌شود، واریانس پراکندگی حول میانگین را بیشتر از حدانتظار نشان می‌دهد؛ زیرا در محاسبهٔ واریانس از میانگین مجذور اختلاف از میانگین داده‌ها استفاده می‌شود. درحالی‌که جذر واریانس شاخص بهتری برای پراکندگی حول میانگین داده‌ها است.

$$* \ \sigma^2_{a+x} = \frac{((a+x_1)-(a+\bar{x}))^2+\cdots+((a+x_N)-(a+\bar{x}))^2}{N} = \frac{(a+x_1-a-\bar{x})^2+\cdots+(a+x_N-a-\bar{x})^2}{N}$$

$$= \frac{(x_1-\bar{x})^2+\cdots+(x_N-\bar{x})^2}{N} = \sigma^2_x$$

$$* \ \sigma^2_{ax} = \frac{(ax_1-a\bar{x})^2+\cdots+(ax_N-a\bar{x})^2}{N} = \frac{(a(x_1-\bar{x}))^2+\cdots+(a(x_N-\bar{x}))^2}{N}$$

$$= \frac{a^2(x_1-\bar{x})^2+\cdots+a^2(x_N-\bar{x})^2}{N} = a^2\left[\frac{(x_1-\bar{x})^2+\cdots+(x_N-\bar{x})^2}{N}\right] = a^2\sigma^2_x$$

جذر واریانس را انحراف معیار می‌نامند و آن‌را با نماد σ نمایش می‌دهند :

$$\sigma = \sqrt{\frac{(x_1 - \overline{X})^2 + \dots + (x_N - \overline{X})^2}{N}}$$

برای گزارش پراکندگی کدام شاخص را ترجیح می‌دهید؟ چرا؟ انحراف معیار ، چون واحد آن با واحد داده‌ها برابر است .

مجدداً این سؤال را مطرح می‌کنیم که در این فعالیت، به نظر شما یک معلم ریاضی ترجیح می‌دهد در کدام کلاس تدریس کند؟ چرا؟

ضریب تغییرات

ضریب تغییرات که با cv نمایش داده می‌شود، نسبت انحراف معیار به میانگین ($cv = \frac{\sigma}{\overline{X}}$) است و معمولاً به صورت درصد بیان می‌شود. مزیت این ضریب آن است که به واحد اندازه‌گیری بستگی ندارد. بنابراین اگر داده‌های مربوط به یک کمیت در دو جامعه با واحدهای متفاوت بیان شده باشد و یا با واحدهایی که نمی‌شناسیم بیان شده باشند می‌توان برای مقایسهٔ پراکندگی داده‌ها در دو جامعه از این ضریب استفاده کرد.

مثال : ضریب تغییرات نمرهٔ ریاضی کلاس A و کلاس B در فعالیت قبل محاسبه شد.

	میانگین	انحراف معیار	ضریب تغییرات
کلاس A	۱۰	~~۱/۶~~ ۱٫۴۱	$\frac{\cancel{۱/۶}\ ۱٫۴۱}{۱۰}$ = ~~۰/۱۶~~ ۰/۱۴ یا ~~۱۶٪~~ ۱۴٪
کلاس B		۷٫۰۷	$\frac{\cancel{۷/۹}\ ۷٫۰۷}{۱۰}$ = ~~۰/۷۹~~ ۰/۷۰ یا ~~۷۹٪~~ ۷۰٪

اشتباه تایپی

معلم ریاضی ترجیح می‌دهد در کلاس A، که ضریب تغییرات کمتری دارد، تدریس کند.

کار در کلاس

دمای هوای یک هفتهٔ اسفند مشهد وکیش، به ترتیب به فارنهایت و سانتی‌گراد گزارش شده است. دمای هوای این هفته در کدام شهر از ثبات بیشتری برخوردار است (ضریب تغییرات کمتری دارد)؟

	شنبه	یکشنبه	دوشنبه	سه‌شنبه	چهارشنبه	پنجشنبه	جمعه
مشهد (فارنهایت)	۵۰	۵۳	۴۹	۴۱	۳۹	۳۷	۳۷
کیش (سانتی‌گراد)	۲۷	۲۶	۲۴	۲۳	۲۲	۲۲	۲۱

$\overline{x}_{\text{مشهد}} = \frac{۵۰+۵۳+۴۹+۴۱+۳۹+۳۷+۳۷}{۷} = \frac{۳۰۶}{۷} \approx ۴۳٫۷۱ \to ۴۴$

$\sigma^2_{\text{مشهد}} = \frac{(۶)^2+(۹)^2+(۵)^2+(۳)^2+(۵)^2+(۷)^2+(۷)^2}{۷} = \frac{۲۷۴}{۷} \approx ۳۹٫۱۴$ $\sigma = \sqrt{۳۹٫۱۴} \approx ۶٫۲۵$

$\overline{x}_{\text{کیش}} = \frac{۲۷+۲۶+۲۴+۲۳+۲۲+۲۲+۲۱}{۷} = \frac{۱۶۵}{۷} \approx ۲۳٫۵۷ \to ۲۴$

$\sigma^2_{\text{کیش}} = \frac{(۳)^2+(۲)^2+(۰)^2+(۱)^2+(۲)^2+(۲)^2+(۳)^2}{۷} = \frac{۳۱}{۷} \approx ۴٫۴۳$ $\sigma = \sqrt{۴٫۴۳} \approx ۲٫۱۰$

چارک‌ها

چارک‌ها (چارک اول، چارک دوم و چارک سوم) مقادیری هستند که داده‌های مرتب شده را به چهار قسمت مساوی تقسیم می‌کنند. بدیهی است چارک دوم همان میانه است. چارک اول را با Q_1 و چارک سوم را با Q_3 نمایش می‌دهند.

۲۲ ۲۴ [۴۸] ۵۱ ۶۰ [۷۰] ۷۵ ۸۰ [۸۷] ۹۳ ۹۵

چارک اول — میانه — چارک سوم

می‌بینید که ۲۵ درصد داده‌ها از ۴۸ (چارک اول)، ۵۰ درصد داده‌ها از ۷۰ (میانه) و ۷۵ درصد داده‌ها از ۸۷ (چارک سوم) کمتر است.

محاسبهٔ چارک‌ها :

ابتدا میانه داده‌ها را بدست آورید

↓

برای داده‌های مرتب شده قبل از میانه، یک میانه بدست آورید و آن را چارک اول بنامید.

↓

برای داده‌های مرتب شده بعد از میانه، یک میانه بدست آورید و آن را چارک سوم بنامید.

مثال : تعداد تصادف‌های اتومبیل‌ها در ۱۵ روز اول تابستان در شهری به‌صورت زیر گزارش شده است.

۱۲ ۱۰ ۱۵ ۲۳ ۱۴ ۲۷ ۱۶ ۳۴ ۴۳ ۴۱ ۳۲ ۱۸ ۲۵ ۳۱ ۱۹

چارک‌ها را مشخص کنید :

۱۰ ۱۲ ۱۴ [۱۵] ۱۶ ۱۸ ۱۹ [۲۳] ۲۵ ۲۷ ۳۱ [۳۲] ۳۴ ۴۱ ۴۳

چارک اول — میانه — چارک سوم

توجه به این نکته نیز ضروری است که با توجه به تعداد داده‌ها، ممکن است چارک‌ها دقیقاً خود داده‌ها نباشند و در فاصلهٔ بین دو دادهٔ متوالی قرار گیرند.

۱۵ - ۱۵ - ۲۰ - ۲۰ - ۲۵ - ۳۰ - ۳۰ - ۳۰ - ۴۵ - ۴۵ - ۴۵ - ۴۵ - ۶۰ - ۶۰ - ۶۰ - ۶۰ - ۶۰ - ۹۰
Q_1 Q_2
۷۵ - ۷۵ - ۸۰ - ۹۰ - ۹۰ - ۹۰ ۹۰ - ۱۰۰ - $\boxed{۱۰۰}$ - ۱۱۵ - ۱۲۰ - ۱۲۰ - ۱۲۰ - ۱۸۰ - ۱۸۰ - ۲۰۰ - ۲۰۰
Q_3

کار در کلاس

معلم یک کلاس می‌خواهد متوسط مدت زمان استفاده دانش‌آموزان از اینترنت را برآورد کند. وی از ۳۵ دانش‌آموز کلاس خود پرسید، در یک شبانه‌روز چند دقیقه از اینترنت استفاده می‌کنند؟ در زیر پاسخ آنها گزارش شده است.

۴۵	۹۰	۶۰	۲۰۰	۱۵	۱۸۰	۴۵	۸۰	۳۰	۱۲۰
۷۵	۹۰	۹۰	۶۰	۲۰	۱۲۰	۱۱۵	۶۰	۳۰	۲۰
۴۵	۴۵	۶۰	۶۰	۶۰	۱۰۰	۹۰	۷۵	۲۰۰	۲۵
					۱۵	۳۰	۱۸۰	۱۰۰	۱۲۰

چارک اول، میانه و چارک سوم مدت زمان استفاده از اینترنت دانش‌آموزان این کلاس را مشخص کنید.

$Q_2 = ۶۰$ $Q_1 = ۴۵$ $Q_3 = ۱۰۰$

تمرین

۱ درستی یا نادرستی جمله‌های زیر را مشخص کنید. با فرض مثبت بودن c

ـ اگر مقدار ثابت c از داده‌ها کم شود، انحراف معیار به اندازهٔ $\sqrt{c}$ کاهش می‌یابد. نادرست

ـ اگر مقدار ثابت c به داده‌ها اضافه شود، ضریب تغییر بزرگ‌تر می‌شود. نادرست

ـ اگر مقدار ثابت $\frac{1}{c}$ در داده‌ها ضرب شود، انحراف معیار $\frac{1}{c}$ برابر می‌شود. درست

ـ اگر مقدار ثابت c در داده‌ها ضرب شود، ضریب تغییر ثابت می‌ماند. درست

۲ کارخانه‌ای دو نوع لاستیک تولید می‌کند. میانگین طول عمر برای نوع A و B به ترتیب ۱۱۰۰۰ کیلومتر و ۱۰۰۰۰ کیلومتر و انحراف معیار برای نوع A و B به ترتیب ۲۰۰۰ کیلومتر و ۱۰۰۰ کیلومتر است. کدام نوع لاستیک بهتر است؟

۳ جدول زیر پول توجیبی (ده هزار ریال) هفتگی پنج دوست نزدیک مینا و مریم را نشان می‌دهد. الف) میانگین و میانهٔ پول توجیبی را برای دوستان مریم و مینا محاسبه کنید. ب) انحراف معیار پول توجیبی را برای دوستان مریم و مینا محاسبه کنید. پ) برنامه‌ریزی برای یک سفر یک روزه با دوستان برای مینا ساده‌تر است یا مریم؟

مینا	۲۳	۲۴	۲۵	۲۶	۲۷
مریم	۱۵	۲۰	۲۵	۳۰	۳۵

خواندنی

علاوه بر چارک، از دهک و صدک نیز استفاده می‌شود. دهک‌ها (دهک اول، دهک دوم .... و دهک نهم) مقادیر نه داده هستند که داده‌های مرتب شده را به ده قسمت مساوی تقسیم می‌کنند. دهک پنجم همان میانه است.

به نقل از روزنامه همشهری ۹۲/۱۱/۹، شناسایی سه دهک پایین درآمدی باید در دستور کار دولت قرار گیرد. با توجه به این جمله چه افرادی باید شناسایی شوند؟

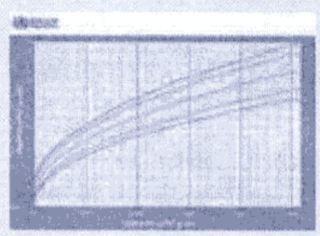
اگر دولت بخواهد یارانهٔ افرادی را که درآمد آنها بیشتر از دهک هشتم است، حذف کند و به یارانهٔ افرادی که درآمد آنها از دهک سوم کمتر است، ۵۰درصد اضافه کند، آیا برای دولت مقرون به‌صرفه است؟

صدک‌ها (صدک اول، صدک دوم ....و صدک نود و نهم) مقادیر نود و نه داده‌اند که داده‌های مرتب شده را به صد قسمت مساوی تقسیم می‌کنند. صدک دهم همان دهک اول و صدک بیست‌وپنجم همان چارک اول است.

کودکانی که زیر صدک سوم قد رشد می‌کنند، فارغ از اندازهٔ قد باید ارزیابی شوند.

ـ منظور از صدک سوم قد چیست؟

ـ بر اساس جملهٔ فوق چه کودکانی باید مورد ارزیابی قرار گیرند؟



۴ میانگین، میانه و انحراف معیار نرخ تورم (مراجعه به خواندنی) سال‌های ۹۴ـ۸۴ را بر اساس جدول زیر محاسبه کنید.

سال	۱۳۸۴	۱۳۸۵	۱۳۸۶	۱۳۸۷	۱۳۸۸	۱۳۸۹	۱۳۹۰	۱۳۹۱	۱۳۹۲	۱۳۹۳	۱۳۹۴
نرخ تورم	۱۰/۴	۱۱/۹	۱۸/۴	۲۵/۴	۱۰/۸	۱۲/۴	۲۱/۵	۳۰/۵	۳۴/۷	۱۵/۶	۱۱/۹

۵ در جدول زیر ارتفاع از سطح دریا برای بعضی از شهرهای استان مرکزی و کهگیلویه و بویراحمد دیده می‌شود.

(راهنمایی: $1m = 3/281\,ft$، فوت: ft ، متر: m)

	مرکزی				کهگیلویه و بویراحمد		
شهر	اراک	محلات	خمین	شازند	یاسوج	دهدشت	دنا
فاصله از سطح دریا	۱۷۰۸ (m)	۱۷۷۵ (m)	۱۸۳۰ (m)	۱۹۲۰ (m)	۶۱۳۵/۴۷ (ft)	۳۲۴۸/۱۹ (ft)	۷۲۱۸/۲۰ (ft)

الف) میانگین ارتفاع از سطح دریا در شهرهای استان مرکزی چقدر است؟

ب) انحراف معیار ارتفاع از سطح دریا در شهرهای استان مرکزی چقدر است؟

پ) ارتفاع از سطح دریا برای شهرهای کدام استان بیشتر است؟

خواندنی

شاخص تورم (شاخص بهای کالاها و خدمات مصرفی) معیار سنجش تغییرات قیمت کالاها و خدماتی است که توسط خانوارها در یک جامعه به مصرف می‌رسد. این شاخص به عنوان وسیله‌ای برای اندازه‌گیری سطح عمومی قیمت کالاها و خدمات مورد مصرف خانوارها، یکی از بهترین معیارهای سنجش تغییر قدرت خرید پول داخل کشور به‌شمار می‌رود. برای محاسبهٔ شاخص تورم، سال ۱۳۹۰ به عنوان سال پایه، ۲۹۴ قلم کالا و ۹۱ قلم خدمت با توجه به اهمیت آنها به طریق علمی انتخاب شده است. برای محاسبهٔ شاخص تورم از رابطهٔ زیر استفاده می‌شود:

$$I_t = \frac{P_t^1 Q_t^1 + \cdots + P_t^{385} Q_t^{385}}{P_\circ^1 Q_\circ^1 + \cdots + P_\circ^{385} Q_\circ^{385}} \times 100$$

که در آن

I_t : شاخص تورم در زمان t

P_t^i : قیمت کالا یا خدمت iام در زمان t

$P_\circ^i$: قیمت کالا یا خدمت iام در زمان پایه

Q_t^i : مقدار مصرف کالا یا خدمت iام در زمان t

$Q_\circ^i$: مقدار مصرف کالا یا خدمت iام در زمان پایه

سال	شاخص تورم	نرخ تورم
۱۳۸۴	۳۹/۸۰	۱۰/۴
۱۳۸۵	۴۴/۵۳	۱۱/۹
۱۳۸۶	۵۲/۷۴	۱۸/۴
۱۳۸۷	۶۶/۱۲	۲۵/۴
۱۳۸۸	۷۳/۲۳	۱۰/۸
۱۳۸۹	۸۲/۳۱	۱۲/۴
۱۳۹۰	۱۰۰/۰۰	۲۱/۵
۱۳۹۱	۱۳۰/۵۴	۳۰/۵
۱۳۹۲	۱۷۵/۸۸	۳۴/۷
۱۳۹۳	۲۰۳/۲۴	۱۵/۶
۱۳۹۴	۲۲۷/۴۶	۱۱/۹

برای محاسبهٔ نرخ تورم (Inf_t) از رابطهٔ زیر استفاده می‌شود:

$$Inf_t = \frac{I_t - I_{t-1}}{I_{t-1}} \times 100$$

I_t شاخص تورم در سال موردنظر و I_{t-1} شاخص تورم در سال قبل از آن است.

تمرین فصل ۷ – صفحه ۱۶۲

تمرین ۲ : $\bar{x}_B = ۱۰۰۰۰ , \sigma_B = ۱۰۰۰ \quad CV_B = \frac{\sigma_B}{\bar{x}_B} = \frac{۱۰۰۰}{۱۰۰۰۰} = ۰/۱$ $\quad$ $\bar{x}_A = ۱۱۰۰۰ , \sigma_A = ۲۰۰۰ \quad CV_A = \frac{\sigma_A}{\bar{x}_A} = \frac{۲۰۰۰}{۱۱۰۰۰} \approx ۰/۱۸$

لاستیک نوع B بهتر است چون ضریب تغییراتش کمتر است.

تمرین ۳ : الف)

$$\bar{x}_{maryam} = \frac{۱۵+۲۰+۲۵+۳۰+۳۵}{۵} = \frac{۱۲۵}{۵} = ۲۵ \qquad \bar{x}_{mina} = \frac{۲۳+۲۴+۲۵+۲۶+۲۷}{۵} = \frac{۱۲۵}{۵} = ۲۵$$

ب)

$$\sigma^۲_{mina} = \frac{(۲۳-۲۵)^۲+(۲۴-۲۵)^۲+(۲۵-۲۵)^۲+(۲۶-۲۵)^۲+(۲۷-۲۵)^۲}{۵} = \frac{۴+۱+۰+۱+۴}{۵} = \frac{۱۰}{۵} = ۲$$

$$\sigma^۲_{maryam} = \frac{(۱۵-۲۵)^۲+(۲۰-۲۵)^۲+(۲۵-۲۵)^۲+(۳۰-۲۵)^۲+(۳۵-۲۵)^۲}{۵} = \frac{۱۰۰+۲۵+۰+۲۵+۱۰۰}{۵} = \frac{۲۵۰}{۵} = ۵۰$$

پ) برای مینا چون پراکندگی پول توجیبی آن کمتر است.

تمرین ۴ :

$Q_۲ = ۱۵/۶$ $\qquad$ $۱۰/۴ , ۱۰/۸ , ۱۱/۹ , ۱۱/۹ , ۱۲/۴ , \underset{Q_۲}{۱۵/۶} , ۱۸/۴ , ۲۱/۵ , ۲۵/۴ , ۳۰/۵ , ۳۴/۷$

$$\bar{x} = \frac{۱۰/۴+۱۰/۸+۱۱/۹+۱۱/۹+۱۲/۴+۱۵/۶+۱۸/۴+۲۱/۵+۲۵/۴+۳۰/۵+۳۴/۷}{۱۱} = \frac{۲۰۳/۵}{۱۱} = ۱۸/۵$$

$$\sigma^۲ = \frac{(۱۰/۴-۱۸/۵)^۲+(۱۰/۸-۱۸/۵)^۲+۲(۱۱/۹-۱۸/۵)^۲+(۱۲/۴-۱۸/۵)^۲+(۱۵/۶-۱۸/۵)^۲}{۱۱}$$

$$\underline{+(۱۸/۴-۱۸/۵)^۲+(۲۱/۵-۱۸/۵)^۲+(۲۵/۴-۱۸/۵)^۲+(۳۰/۵-۱۸/۵)^۲+(۳۴/۷-۱۸/۵)^۲}$$

$$= \frac{۶۵/۶۱+۴۳/۵۶+۰/۰۱+۴۷/۶۱+۵۹/۲۹+۳۷/۲۱+۹+۱۴۴+۲۶۲/۴۴+۸/۴۱}{۱۱} = \frac{۶۷۷/۱۴}{۱۱} \approx ۶۱/۵۶$$

$$\sigma = \sqrt{۶۱/۵۶} \approx ۷/۸۵$$

تهیه کننده:
گروه ریاضی دوره ی دوم متوسطه و انجمن معلمان ریاضی، استان خوزستان
khuzmath1394@chmail.ir

تمرین ۵ :

	مرکزی				کهگیلویه و بویر احمد		
شهر	اراک	محلات	خمین	شازند	یاسوج	دهدشت	دنا
فاصله از سطح دریا	$۱۷۰۸(m)$	$۱۷۷۵(m)$	$۱۸۳۰(m)$	$۱۹۲۰(m)$	$۶۱۳۵/۴۷(ft)$	$۳۲۴۸/۱۹(ft)$	$۷۲۱۸/۲۰(ft)$
	$\downarrow \times ۳/۲۸۱$ $۵۶۰۳/۹۴۸(ft)$	$\downarrow \times ۳/۲۸۱$ $۵۸۲۳/۷۷۵(ft)$	$\downarrow \times ۳/۲۸۱$ $۶۰۰۴/۲۳(ft)$	$\downarrow \times ۳/۲۸۱$ $۶۲۹۹/۵۲(ft)$	$\downarrow \div ۳/۲۸۱$ $۱۸۷۰(m)$	$\downarrow \div ۳/۲۸۱$ $۹۹۰(m)$	$\downarrow \div ۳/۲۸۱$ $۲۲۰۰(m)$

الف)

$$\bar{x} = \frac{۱۷۰۸+۱۷۷۵+۱۸۳۰+۱۹۲۰}{۴} = \frac{۷۲۳۳}{۴} = ۱۸۰۸/۲۵ \quad \rightarrow \quad ۱۸۰۸$$

ب)

$$\sigma^۲ = \frac{(۱۷۰۸-۱۸۰۸)^۲+(۱۷۷۵-۱۸۰۸)^۲+(۱۸۳۰-۱۸۰۸)^۲+(۱۹۲۰-۱۸۰۸)^۲}{۴} = \frac{(-۱۰۰)^۲+(-۳۳)^۲+(۲۲)^۲+(۱۱۲)^۲}{۴}$$

$$= \frac{۱۰۰۰۰+۱۰۸۹+۴۸۴+۱۲۵۴۴}{۴} = \frac{۲۴۱۱۷}{۴} \approx ۶۰۲۹/۲۵ \qquad \sigma = \sqrt{۶۰۲۹/۲۵} \approx ۷۷/۶۴$$

پ) شهرهای استان مرکزی ارتفاع بیشتری از سطح دریا دارند.

$$\bar{x} = \frac{۱۸۷۰+۹۹۰+۲۲۰۰}{۳} = \frac{۵۰۶۰}{۳} \approx ۱۶۸۶/۷$$