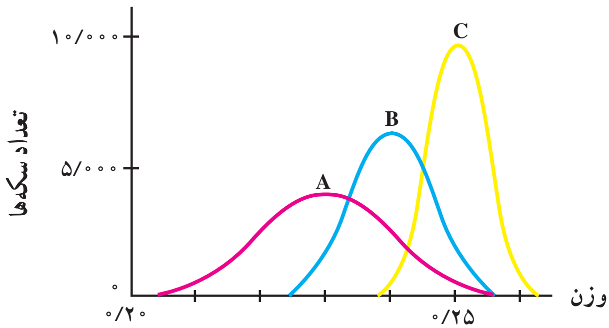
از فرشتگانی که در آسمانهایت جایشان دادی، و از زمینِ خود برترشان بردی. آنان از دیگر آفریدگانت تو را بهتَر شناسند، و از عقاب تو بیشتر می‌هراسند، و به تو نزدیکترَند و نه پراکَندهٔ گردش زمان. آنان با مرتبتی که از آن برخوردارند، و از منزلتی که نزدِ تو دارند، و یکدله تو را دوست دارند و تو را فراوَان طاعت می‌گزارند، و اندک غفلتی در فرمان تو نیارند.

شاخص‌های پراکندگی

ضرابخانهٔ سکه‌های دولتی تصمیم گرفته است که وزن پنج تومانی‌هایی که در جریان بازار قرار می‌دهد، روی نمودار بیاورد. برای این منظور، چندین سکه پنج‌تومانی که تازه ضرب شده‌اند به‌طور تصادفی انتخاب و به دقت وزن می‌شوند. همچنین به همان تعداد سکه پنج تومـانی که پنج سال پیش به بازار عرضه شده بود و به‌همان تعداد سکهٔ پنج تومانی که ۱۰ سال پیش به بازار عـرضه شده بود به دقت وزن می‌شوند. سه نمـودار A و B و C بـه‌دست می‌آید.



ــ فکر می‌کنید کدام نمودار مربوط به سکه‌های نو ارائه شده به بازار و کدام نمودار مربوط به سکه‌های پنج ساله و ده ساله است؟

ــ با گذشت زمان میانگین وزن سکه‌های پنج تومانی چه تغییری می‌کند؟

ــ با گذشت زمان مد وزن سکه‌های پنج تومانی چه تغییری می‌کند؟

ــ چاق و یا لاغر شدن نمودارها را چگونه تعبیر می‌کنید؟

پراکندگی

در فصل قبل با شاخص‌های مرکزی آشنا شدیم. مثلاً شما براساس میانگین زمان‌هایی که طول کشیده است تا در ۲۰ روز گذشته از خانه به مدرسه بروید، از منزل خارج شده‌اید. در بین راه اعدادی را که معرف این مدت زمان بوده‌اند در ذهن خود مرور می‌کنید. میانگین اعداد ۳۵ دقیقه شده است و شما درست ۳۵ دقیقه قبل از شروع کلاس‌ها از خانه خارج شده‌اید، ولی اعدادی مانند ۴۰، ۳۸، ۳۰، ۳۲ و یا حتی ۴۵ در بین اعداد دیده می‌شود. شما مایل نیستید که دیر به مدرسه برسید، به ویژه آن‌که روز امتحان هم باشد. یادآوری اعداد بالا آسودگی خیال را به هم می‌ریزد. از خود می‌پرسید نکند امروز یکی از آن روزهایی باشد که بیش از ۳۵ دقیقه طول بکشد. اگر کمی بیش‌تر طول بکشد، (در حد ۵ دقیقه) اشکال عمده‌ای نخواهد بود ولی تأخیر یک ربع و یا بیش‌تر مشکل‌ساز است. هرچه‌قدر در بین داده‌های شما، اعداد دورتر از میانگین وجود داشته باشد، ناراحتی شما بیش‌تر خواهد بود. اگر قبلاً این اعداد را بهتر وارسی می‌کردید و صرفاً میانگین را مدنظر قرار نمی‌دادید الآن وضع بهتری داشتید، اگر از ابتدا درمی‌یافتید که اعداد تفاوت دارند می‌توانستید در این روز سرنوشت‌ساز تصمیم بهتری بگیرید. معیاری که می‌تواند این تفاوت‌ها و میزان آن‌ها و به‌خصوص دوری آن‌ها از میانگین را برای ما اندازه‌گیری کند شاخص پراکندگی نامیده می‌شود. ما به چند شاخص پراکندگی در زیر اشاره می‌کنیم :

دامنهٔ تغییرات

دامنهٔ تغییرات که در فصل دوم با آن آشنا شدید یکی از شاخص‌های پراکندگی است. هرچه قدر داده‌ها از هم دورتر باشند، دامنهٔ تغییرات آن‌ها بیش‌تر خواهد بود. دیدیم اگر دامنهٔ تغییرات برابر صفر باشد تمام داده‌ها با هم برابرند. یعنی هیچ تفاوتی بین آن‌ها موجود نیست. حال به دو دسته دادهٔ زیر که نمره‌های دانش‌آموزان در درس فیزیک در دو نمونهٔ ۱۰ عضوی را به نمایش می‌گذارند توجه کنید.

الف) ۱۹، ۱۸، ۱۷، ۱۶، ۱۵، ۱۴، ۱۳، ۱۲، ۱۱، ۱۰

ب) ۱۹، ۱۴، ۱۵، ۱۵، ۱۵، ۱۵، ۱۴، ۱۴، ۱۵، ۱۰

دامنهٔ تغییرات هر دو دسته یک اندازه است :

$$R = ۱۹ - ۱۰ = ۹$$

ولی پراکندگی داده‌های دستهٔ دوم کم‌تر است. در داده‌های دسته دوم ما عملاً اعداد ۱۴ و ۱۵ را داریم. به ۱۰ و ۱۹ بیش‌تر باید با شک و تردید نگریست. این دو مقدار ممکن است به خاطر اندازه‌گیری‌های نادرست حاصل شده باشند و یا ممکن است علل دیگری برای وجود آن‌ها در بین داده‌ها باشد. اما در داده‌های دستهٔ اول این دو مقدار کاملاً طبیعی به نظر می‌رسند، زیرا مقادیر نزدیک به آن‌ها نیز داریم. با این توضیحات ما در داده‌های دسته دوم بیش‌تر روی ۱۴ و ۱۵ حساب می‌کنیم و این دو مقدار هم به هم خیلی نزدیک‌اند. این توضیح می‌رساند که دامنهٔ تغییرات در بعضی مواقع به خوبی نمی‌تواند پراکندگی موجود در داده‌ها را نشان دهد. بنابراین به دنبال معیاری هستیم که بهتر از دامنه تغییرات بتواند پراکندگی و تفاوت‌ها را نشان دهد. قبل از آن که به معرفی معیار بعدی برویم ذکر یک نکته ضروری است که وقتی با دید مجرد ریاضی به دو دسته عدد بالا نگاه می‌کنید تفاوت هر دو را در حد ۹ می‌بینید، ولی با دید آماری مسئله کمی فرق می‌کند. در آمار حضور دو مقدار ۱۰ و ۱۹ را در دستهٔ «ب» زیاد جدی نمی‌گیریم چون این‌ها مقادیر تک افتاده‌ای هستند. به‌نظر می‌رسد این دو مقدار به جامعه تعلق نداشته باشند. ما در این درس سعی می‌کنیم علاوه‌بر نقطه نظراتی که شما دربارهٔ مسائل اطراف خود دارید پایه‌های دید آماری را نیز ایجاد کنیم و آن را توسعه بخشیم.

چارک‌های بالا و پایین

گفتیم که دامنهٔ تغییرات ممکن است در بعضی مواقع تعبیرهای نامناسب از جامعه ارائه کند، مثلاً در جامعه‌ای که عملاً داده‌ها به هم نزدیک‌اند، به‌علت وجود دو مقدار خیلی کوچک و خیلی بزرگ در جامعه، دامنهٔ تغییرات عدد بزرگی به‌دست آید، و حال آن‌که جامعه از دید آماری جامعه متمرکزی باشد. تعداد بازدیدکنندگان از کتابخانهٔ مدرسه در ۱۲ روز کاری در زیر آمده است.

۱۱، ۱۵، ۶، ۱۰، ۹، ۵، ۶، ۷، ۸، ۲، ۱، ۰

دامنهٔ تغییرات این داده‌ها برابر ۱۵ است و به نظر زیاد می‌رسد ولی داده‌ها عموماً از ۵ تا ۱۱ پراکنده شده‌اند که پراکندگی آن‌ها زیاد نیست. از این رو برای حذف تأثیر داده‌های بزرگ و کوچک روی تعدادی از اعداد بالای داده‌ها و تعدادی از اعداد پایین داده‌ها را حذف می‌کنند. این‌که چه تعدادی حذف شود بستگی به تعداد داده‌ها و نزدیکی داده‌ها به‌هم دارد. در بعضی مواقع یک دهم داده‌ها از بالا و یک دهم داده‌ها از پایین را حذف می‌کنند و در برخی مواقع یک چهارم داده‌های بالایی و یک‌چهارم داده‌های پایینی را حذف می‌کنند و بعد از آن دامنه تغییرات داده‌های باقی‌مانده را محاسبه می‌کنند. به‌عنوان مثال در داده‌های بالا اگر یک چهارم داده‌ها را از بالا و پایین حذف کنیم به داده‌های زیر خواهیم رسید :

۱۵، ۱۱، ۱۰، **۹، ۸، ۷، ۶، ۶، ۵**، ۲، ۱، ۰

بنابراین دامنهٔ تغییرات این داده‌ها عبارت خواهد بود از ۹ − ۵ = ۴ و حال آن‌که دامنه تغییرات داده‌های اصلی برای ۱۵ − ۰ = ۱۵ است. اگر توجه کرده باشید در برخی از مسابقات امتیازهایی را که داورها می‌دهند، بیش‌ترین و کم‌ترین امتیاز را حذف می‌کنند و بر اساس امتیازهای باقی‌مانده امتیاز ورزشکار را تعیین می‌کنند.

تمرین

۱ــ دو کارخانه تولید کننده مواد غذایی A و B شکلات در بسته‌بندی ۴۸ گرمی می‌فروشند. پنج بسته شکلات به‌صورت تصادفی از یک فروشگاه مواد غذایی از دو محصول انتخاب شد و تمام وزن بسته‌ها به دقت اندازه‌گیری شد.

نتیجه زیر به دست آمد :

A:۴۸/۰۸ ،۴۸/۳۲ ،۴۷/۹۶ ،۴۷/۸۴ ،۴۷/۹۶

B:۴۹/۱۶ ،۴۸/۸۴ ،۴۸/۸۸ ،۴۹/۰۸ ،۴۹

الف ــ کدام کارخانه شکلات بیش‌تری در بسته‌ها می‌فروشد؟ از چه شاخصی استفاده می‌کنید؟

ب ــ کدام کارخانه در توزیع شکلات یکنواخت‌تر عمل کرده است؟

۲ــ اگر دامنهٔ تغییرات برابر صفر باشد دربارهٔ داده‌ها چه نتیجه‌ای می‌گیرید؟

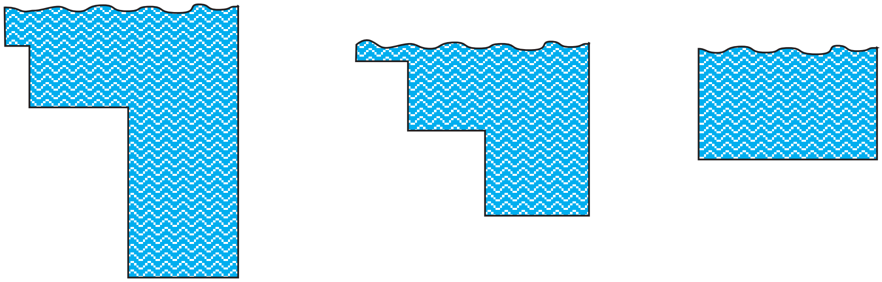
۳ــ تغییر واحد اندازه‌گیری چه تأثیری در دامنه تغییرات می‌گذارد؟ مثلاً اگر شما داده‌ها را برحسب متر اندازه‌گیری کرده باشید و سپس آن‌ها را برحسب سانتی‌متر به‌کار برید دامنه تغییرات چگونه تغییر می‌کند؟

۴ــ اگر به داده‌ها یک مقدار ثابت اضافه کنیم دامنهٔ تغییرات چگونه تغییر می‌کند؟

واریانس

اگر شما خوب شنا نمی‌دانستید و می‌خواستید در استخری شنا کنید که عمق آن متغیر است برای این‌که اطمینان حاصل کنید که در خطر نخواهید بود چه اطلاعاتی را لازم دارید؟

اگر میانگین عمق استخر را به شما داده باشند و این میانگین برابر ۱/۵ متر باشد، آیا برای شنا در این استخر اطمینان دارید؟ به حالات زیر توجه کنید.



در هر سه استخر بالا میانگین عمق برابر ۱/۵ متر است. می‌خواهیم با یک شاخص عددی پراکندگی متغیر عمق را نسبت به عمق میانگین بررسی کنیم. اگر بدانیم که تغییرات عمق در اطراف عمق میانگین، یعنی ۱/۵ متر زیاد نیست با اطمینان بیش‌تری در آن شنا خواهیم کرد. به شرط آن که این پراکندگی با یک شاخص عددی مناسب معرفی شده باشد.

دامنه تغییرات شاخص مناسبی است. اما در تصمیم گیری‌های کلان از ارزش آماری زیادی برخوردار نیست. زیرا ما به شاخص‌هایی نیاز داریم که هم پراکندگی داده‌ها و هم فراوانی آن‌ها را مد نظر قرار دهند.

دیدیم که پراکندگی یعنی این که داده‌ها از مرکز خود چه‌قدر دور هستند، پس یک راه ابتدایی ممکن است این‌طور به نظر برسد که تک‌تک مقادیر را از میانگین کم کنیم. این تفاضل‌ها را *انحراف از میانگین* می‌نامیم و سپس مجموع مقادیر حاصل را به‌دست آوریم، یعنی مجموع مقادیر

$$x_1 - \bar{x} \text{ و } x_2 - \bar{x} \text{ و } \dots \text{ و } x_n - \bar{x}$$

را حساب کنیم. قبلاً دیده‌ایم که مجموع مقادیر بالا برابر صفر است. لذا برای هر نوع داده‌ای اعم از آن که داده‌ها به میانگین نزدیک باشند و یا از آن دور باشند، آن چه که از این روش به‌دست می‌آید برابر صفر است. پس معیار مذکور نمی‌تواند معیار مفیدی باشد. البته علت آن که این مجموع برابر صفر شد به این خاطر است که برخی از داده‌ها از میانگین بزرگ‌تر و برخی دیگر کوچک‌ترند. درنتیجه مقادیر مثبت و منفی حاصل می‌شوند که مجموع آن‌ها یک‌دیگر را خنثی می‌کنند. برای جلوگیری از علامت‌های مثبت و منفی یکی از راه‌ها استفاده از قدرمطلق است. یعنی از مقادیر $|x_1 - \bar{x}|$ و ... و $|x_n - \bar{x}|$

استفاده کنیم. این مقادیر می‌تواند معیاری برای پراکندگی ارائه کند ولی معمولاً کارکردن با قدرمطلق کار آسانی نیست و از این رو از توان دوم مقادیر بالا استفاده می‌کنیم یعنی مجموع مجذورات فاصله از میانگین را در نظر می‌گیریم. پس یک معیار پراکندگی ممکن است به‌صورت زیر باشد :

$$(x_1-\bar{x})^2+\cdots+(x_n-\bar{x})^2$$

ملاحظه می‌شود که اگر داده‌ها برابر باشند مقدار بالا برابر صفر است و بالعکس اگر مقدار بالا برابر صفر باشد داده‌ها برابر بوده و در نتیجه هیچ تفاوتی در بین آن‌ها نیست. اما اشکالی که دستور بالا دارد آن است که اگر تعدادی دادهٔ دیگر اضافه کنیم به مجموع بالا مقادیر مثبت دیگری اضافه خواهد شد و لذا این مجموع بزرگ‌تر خواهد شد. حال آیا درست است که بگوییم با اضافه شدن داده‌ها پراکندگی بیش‌تر می‌شود؟ مسلماً پاسخ منفی است، چون پراکندگی چیزی است که به اصل جامعه مربوط است و این نوع محاسبات نباید در مقدار آن‌ها تأثیر بگذارد. پس برای آن که تأثیر تعداد داده‌ها را از بین ببریم مجموع مذکور را بر تعداد آن‌ها تقسیم می‌کنیم و از این‌جا به معیار زیر می‌رسیم :

واریانس

واریانس برابر میانگین مجذور انحرافات از میانگین است و آن را با σ^2 نشان می‌دهیم. بنابراین :

$$\sigma^2=\frac{(x_1-\bar{x})^2+\cdots+(x_n-\bar{x})^2}{n}$$

(σ از حروف کوچک یونانی است و سیگما خوانده می‌شود. حرف بزرگ آن Σ است)

تعداد ساعاتی که ۴ دانش‌آموز در طول هفته به ورزش اختصاص داده‌اند در زیر آمده است : واریانس این داده‌ها را حساب کنید.

۱ ,۵ ,۷ ,۹

ابتدا میانگین آن‌ها را حساب می‌کنیم.

میانگین : $x = \frac{۱+۵+۷+۹}{۴} = \frac{۲۲}{۴} = ۵/۵$

سپس انحراف از میانگین را حساب می‌کنیم.

انحرافات از میانگین : $-۴/۵,\ -۰/۵,\ ۱/۵,\ ۳/۵$

مجذور انحرافات از میانگین : $۲۰/۲۵,\ ۰/۲۵,\ ۲/۲۵,\ ۱۲/۲۵$

میانگین مجذور انحرافات از میانگین: $\frac{۲۰/۲۵+۰/۲۵+۲/۲۵+۱۲/۲۵}{۴} = \frac{۳۵}{۴} = ۸/۷۵$

پس $\sigma^۲ = ۸/۷۵$

اگر داده‌ها برابر باشند واریانس آن‌ها صفر است و بالعکس.

اگر داده‌ها برابر باشند، میانگین آن‌ها نیز همان مقدار مشترک داده‌ها خواهد بود. پس انحرافات از میانگین برابر صفر است. لذا مجموع مجذورات انحراف صفر خواهد بود. پس واریانسی که از تقسیم این مجموع بر تعداد داده‌ها به‌دست می‌آید برابر صفر است. حال فرض کنید واریانس صفر باشد. پس باید مجموع مجذورات انحرافات صفر شده باشد. مجذور انحرافات عددی نامنفی است و چون مجموع آن‌ها صفر است لازم می‌آید هریک از آن‌ها صفر باشند، چون در غیر این صورت مجموع مجذور انحرافات صفر نخواهد شد. با صفر شدن هریک از انحرافات نتیجه می‌شود که تمام داده‌ها برابر میانگین هستند. پس مساوی‌اند.

نکاتی دربارهٔ واحد میانگین و واریانس : فرض کنید مطالعه‌ای در بارهٔ قد افراد انجام می‌دهید، و داده‌های شما برحسب سانتی‌متر اندازه‌گیری و بیان شدند. از آن‌جایی که میانگین به‌وسیله مجموع داده‌ها به‌دست می‌آید آن هم برحسب سانتی‌متر خواهد بود. یعنی واحد میانگین از نوع همان واحد داده‌هاست. اما در واریانس شرایط فرق می‌کند زیرا در این‌جا شما مقادیر $(x_i - \bar{x})$ را که برحسب سانتی‌متر است به توان ۲ می‌رسانید، پس مجذور انحرافات بر حسب سانتی‌مترمربع است و لذا مجموع آن‌ها و درنتیجه واریانس برحسب سانتی‌مترمربع خواهد بود. بنابراین واحد واریانس از نوع مجذور واحد متغیر است.

دستور دیگری برای محاسبه واریانس: دیدیم که برای محاسبهٔ واریانس باید میانگین مجذور انحرافات از میانگین را حساب کنیم، یعنی

$$\sigma^2 = \frac{1}{n}\left[(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \cdots + (x_n - \bar{x})^2\right]$$

محاسبه واریانس با استفاده از این دستور در بعضی مواقع مشکلاتی دارد.
می‌توان ثابت کرد عبارت بالا را می‌توان به‌صورت زیر نیز نوشت

$$\sigma^2 = \frac{1}{n}\left[x_1^2 + \cdots + x_n^2\right] - \bar{x}$$

اثبات این دستور چندان مشکل نیست، توجه کنید که

$$\sigma^2 = \frac{1}{n}\left[(x_1-\bar{x})^2+\cdots+(x_n-\bar{x})^2\right]$$

$$= \frac{1}{n}\left[(x_1^2-2x_1\bar{x}+\bar{x}^2)+\cdots+(x_n^2-2x_n\bar{x}+\bar{x}^2)\right]$$

$$= \frac{1}{n}\left[(x_1^2+\cdots+x_n^2)-2(x_1+\cdots+x_n)\bar{x}+n\bar{x}^2\right]$$

$$= \frac{x_1^2+\cdots+x_n^2}{n}-\frac{2(x_1+\cdots+x_n)}{n}\bar{x}+\bar{x}^2$$

$$= \frac{x_1^2+\cdots+x_n^2}{n}-2\bar{x}^2+\bar{x}^2$$

$$= \frac{x_1^2+\cdots+x_n^2}{n}-\bar{x}^2$$

توجه کنید در $x_1^2+\cdots+x_n^2$ ابتدا مقادیر x_i ها را به توان ۲ می‌رسانیم و سپس آن‌ها را با هم جمع می‌کنیم و در $\bar{x}^2$، $\bar{x}$ (میانگین داده‌ها) را به توان دو می‌رسانیم.

حال واریانس داده‌های زیر را حساب می‌کنیم.

۱ ،۵ ،۶ ،۷ ،۹

ابتدا میانگین را حساب می‌کنیم :

$$\bar{x} = \frac{۱+۵+۶+۷+۹}{۵} = \frac{۲۸}{۵} = ۵/۶$$

$$x_1^2+\ldots+x_5^2 = ۱+۲۵+۳۶+۴۹+۸۱$$
$$= ۱۹۲$$

پس :

$$\sigma^2 = \frac{۱۹۲}{۵}-(۵/۶)^2$$
$$= ۳۸/۴ - ۳۱/۳۶ = ۷/۰۴$$

با استفادهٔ مستقیم از دستور واریانس نیز به‌همین نتیجه می‌رسیدیم

$$\sigma^2 = \frac{(۱-۵/۶)^2+(۵-۵/۶)^2+(۶-۵/۶)^2+(۷-۵/۶)^2+(۹-۵/۶)^2}{۵}$$

$$= \frac{۲۱/۱۶+۰/۳۶+۰/۱۶+۱/۹۶+۱۱/۵۶}{۵}$$

$$= ۷/۰۴$$

انحراف معیار

این که واحد واریانس از نوع مجذور واحد متغیر است می‌تواند مشکلات و سوء تفاهماتی را در پی داشته باشد. فرض کنید شما قدها را برحسب متر اندازه‌گیری کرده باشید و همکلاسی شما قدها را برحسب سانتی‌متر، در این صورت واریانسی که همکلاسی شما حساب می‌کند ۱۰۰۰۰ برابر واریانسی است که شما حساب کرده‌اید. همکلاسی شما ممکن است این‌طور نتیجه بگیرد که پراکندگی در قد بسیار زیاد و چشم‌گیر است و حال آن که نظر شما ممکن است غیر از این باشد. برای آن که این اختلاف نظرها را از بین ببریم سعی می‌کنیم که تفاوت عمده در واحد واریانس و واحد میانگین را با جذر گرفتن از واریانس از بین ببریم. جذر واریانس را انحراف معیار می‌گوییم.

انحراف معیار که با نماد σ نشان داده می‌شود برابر جذر واریانس است.

واحد انحراف معیار همان واحد متغیر است.

تعداد ساعت‌هایی در هفته که ۴ نفر به‌صورت داوطلبانه دریک بیمارستان خدمت می‌کنند در زیر آمده است. انحراف معیار داده‌ها را حساب کنید.　　۱، ۳، ۴، ۵، ۶

$$\bar{x} = \frac{۱+۳+۴+۵+۶}{۵} = \frac{۱۹}{۵} = ۳/۸$$

$$\sigma^۲ = \frac{(-۲/۸)^۲ + (-۰/۸)^۲ + (۰/۲)^۲ + (۱/۲)^۲ + (۲/۲)^۲}{۵}$$

$$= \frac{۷/۸۴ + ۰/۶۴ + ۰/۰۴ + ۱/۴۴ + ۴/۸۴}{۵}$$

$$= \frac{۱۴/۸۰}{۵} = ۲/۹۶$$

در نتیجه:　$\sigma \quad \sqrt{۳/۹۶} \approx ۱/۷$

تمرین

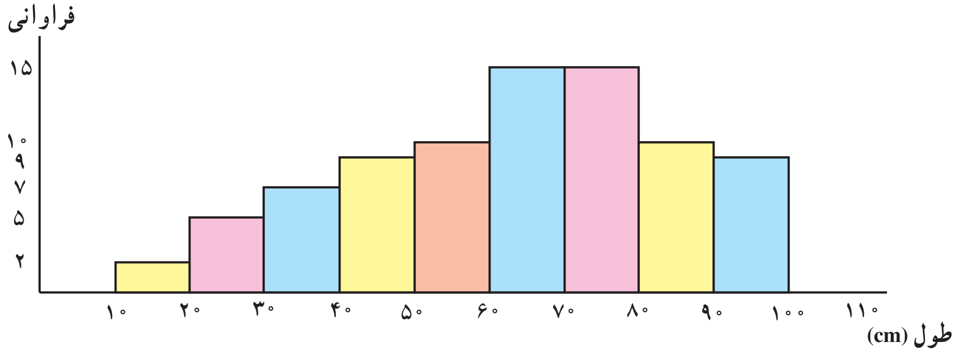
۱ــ نمودار زیر طول ۸۲ گیاه پرورش داده شده در یک آزمایشگاه را نشان می‌دهد.

ــ طول چند گیاه بین ۵۰ تا ۶۰ سانتی‌متر است؟

ــ طول چند گیاه بین ۷۰ تا ۹۰ سانتی‌متر است؟

ــ میانگین و انحراف معیار طول گیاهان را به‌دست آورید.

ــ توضیح دهید چه ویژگی نمودار بیان‌گر آن است که میانه و میانگین برابر نیستند.



۲ــ درجه حرارت بدن ۸ بیمار (به نزدیک‌ترین عدد صحیح گرد شده است) در زیر آمده است :

۳۸ ۳۸ ۳۹ ۳۹ ۴۰ ۴۰ ۴۱ ۴۱

میانگین و انحراف معیار را حساب کنید.

درجه حرارت‌های واقعی به قرار زیر می‌باشند :

۳۸/۰ ۳۸/۴ ۳۹/۰ ۳۹/۴ ۳۹/۸ ۴۰/۲ ۴۰/۹ ۴۱/۲

ــ میانگین و انحراف معیار را حساب کنید.

ــ گرد کردن داده‌ها چه تأثیری در میانگین و انحراف معیار داشت؟

ــ آیا گرد کردن همواره تأثیری بر میانگین و انحراف معیار دارد؟

۳ــ در یک نمونهٔ ۹۰۰ نفری از مردان دربارهٔ مسافتی که روزانه از منزل به محل کار طی می‌کنند سؤال شد. اعداد صفحه بعد نتیجهٔ این بررسی را نشان می‌دهد.

مسافت طی شده (کیلومتر)	درصد
کمتر از ۱	۱۷
۱ و کمتر از ۳	۳۱
۳ و کمتر از ۸	۳۸
۸ و کمتر از ۱۰	۷
۱۰ و کمتر از ۱۵	۵
۱۵ و کمتر از ۵۰	۲
	۱۰۰

ــ نمودار مستطیلی این داده‌ها را رسم کنید.

ــ میانگین و انحراف معیار را برای این داده‌ها به‌دست آورید.

ــ اگر میانگین مسافت طی شده توسط زنان کارمند ۵ کیلومتر باشد، با فرض آن که انحراف معیار هر دو گروه یکسان است، آیا می‌توانید بگویید که مردان فاصلهٔ بیش‌تری طی می‌کنند یا زنان؟

۴ــ دو گروه داده به‌صورت زیر داریم :

گروه اول : ۶ ۵ ۴ ۳ ۲

گروه دوم : ۸ ۷ ۶ ۵ ۴

الف ــ واریانس هر گروه را حساب کنید.

ب ــ این دو گروه داده چه رابطه‌ای با هم دارند؟

ج ــ رابطه موجود بین دو گروه داده، چه اثری روی واریانس‌ها داشته است؟

د ــ با توجه به تعریف واریانس، نتیجهٔ حاصل از قسمت ج را توضیح دهید.

۵ــ دو گروه داده به‌صورت زیر داریم :

گروه اول : ۵ ۴ ۳ ۲ ۱

گروه دوم : ۱۵ ۱۲ ۹ ۶ ۳

الف ــ واریانس هر گروه را حساب کنید.

ب ــ این دو گروه چه رابطه‌ای با هم دارند؟

ج ــ رابطه موجود بین دو گروه داده، چه تأثیری روی واریانس داشته است؟

د ــ نتیجه حاصل از قسمت ج را توضیح دهید.

۶ــ در حالت کلی ثابت کنید که اگر داده‌ها را با عددی جمع کنیم، در واریانس آن‌ها تغییری حاصل نخواهد شد. به زبان نمادها $\sigma_{a+x}^{2} = \sigma_{x}^{2}$

۷ــ در حالت کلی ثابت کنید که اگر داده‌ها را در عددی ضرب کنیم واریانس آن‌ها در مجذور این عدد ضرب خواهد شد. به زبان نمادها $\sigma_{ax}^{2} = a^{2}\sigma_{x}^{2}$

۸ــ با ترکیب دو تمرین بالا نتیجه بگیرید که $\sigma_{ax+b}^{2} = a^{2}\sigma_{x}^{2}$

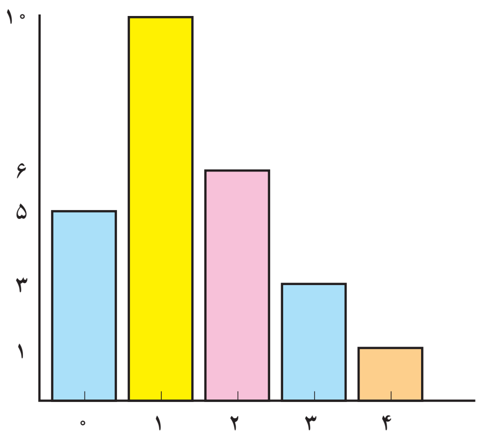
۹ــ کدامیک از برابری‌های زیر دربارهٔ انحراف معیار صحیح است؟

$\sigma_{ax+b} = a\sigma_{x} \quad \sigma_{ax+b} = -a\sigma_{x} \quad \sigma_{ax+b} = |a|\sigma_{x}$

۱۰ــ در یک بررسی تعداد فرزندان ۲۵ نفر از زنان مورد سؤال بود. نتیجه در نمودار زیر نشان داده شده است.

الف ــ نشان دهید متوسط تعداد فرزندان برابر ۱/۴ است.

ب ــ نشان دهید انحراف معیار این داده‌ها تقریباً ۱/۰۶ است.



ــ گروه ۲۵ نفری دیگری از زنان مورد بررسی قرار گرفتند. در این گروه متوسط تعداد فرزندان ۲/۴ و انحراف معیار آن‌ها ۲ است. با توجه به اطلاعات به‌دست آمده در الف و ب، تفاوت بین تعداد فرزندان این دو گروه را بررسی کنید.

ضریب تغییرات

مسلماً قد همکلاسی‌های شما متفاوت است، همچنین وزن آن‌ها نیز چنین است ولی در سن همکلاسی‌های شما اختلاف زیادی نیست. یعنی انتظار می‌رود واریانس سن کم، ولی واریانس قد و وزن زیاد باشد. حال سؤال این است که تغییرات در قد همکلاسی‌ها بیش‌تر است یا در وزن آن‌ها. سن تقریباً ثابت است، زیاد متغیر نیست ولی می‌خواهیم بدانیم از قد و وزن کدام بیش‌تر متغیر است. اگر این دو صفت برحسب یک واحد می‌بودند مقایسه واریانس آن‌ها می‌توانست جواب‌گوی این سؤال باشد ولی چون واحدها یکی نیستند مقایسه میزان پراکندگی این دو متغیر از طریق واریانس‌ها ممکن نیست. برای آن‌که بتوانیم به سؤال بالا جواب دهیم باید شاخصی معرفی کنیم که بدون واحد باشد تا امکان مقایسه فراهم آید. اگر به خاطر داشته باشید در جدول فراوانی برای آن که تأثیر اندازه نمونه را از بین ببریم به محاسبهٔ درصد فراوانی نسبی پرداختیم. در این‌جا نیز برای از بین بردن واحد اندازه‌گیری از معیار **ضریب تغییرات** استفاده می‌کنیم.

ضریب تغییرات که با نماد CV نشان می‌دهیم عبارت است از خارج قسمت انحراف معیار بر میانگین

تذکر: از آن‌جایی که ضریب تغییرات معیاری برای میزان پراکندگی است باید مثبت باشد، چون پراکندگی منفی معنی ندارد. پس لازم است $\bar{x}$ ، که در مخرج کسر بالا آمده است، مثبت باشد. برای اطمینان از مثبت بودن $\bar{x}$ ، ضریب اطمینان را فقط برای داده‌های مثبت تعریف می‌کنیم.

ضریب تغییرات داده‌های زیر را محاسبه می‌کنیم :

۱ ,۳ ,۴ ,۵ ,۶

قبلاً دیدیم که $\bar{x} = ۳/۸$ و $\sigma = ۱/۷$ پس

$$CV = \frac{۱/۷}{۳/۸}$$

تذکر: هرچه‌قدر داده‌ها بزرگ‌تر باشند زمانی پراکندگی در آن‌ها محسوس است که تفاوت داده‌ها بیش‌تر باشد. ما اختلاف بین ۳ و ۳/۵ را به‌خوبی درک می‌کنیم ولی اختلاف بین ۳۰۰۰ و ۳۰۰۰/۵ چندان برای ما معنی‌دار نیست. شما در خرید اجناس ارزان قیمت انتظار تخفیف زیاد ندارید ولی در خرید اجناس گران قیمت از بحث و مجادله بر روی مبالغ کم خودداری می‌کنید. دلیل این امر آن است که در مقادیر بزرگ می‌توان از اختلاف‌های کم صرف‌نظر کرد. مثلاً از نظر آماری داده‌های زیر تقریباً با هم برابرند و تفاوت زیادی با هم ندارند.

۱۱۱ ,۱۱۱/۵ ,۱۱۲ ,۱۱۲/۵

ولی داده‌های زیر با هم تفاوت دارند

۱ ,۱/۵ ,۲ ,۲/۵

وقتی از ۱ به ۱/۵ می‌رویم پنجاه درصد ۱ به آن اضافه می‌شود ولی وقتی که از ۱۱۱ به ۱۱۱/۵ می‌رویم حدود چهارهزارم ۱۱۱ به آن اضافه می‌شود که ناچیز است. بنابراین پراکندگی به طریقی باید به‌نسبت بزرگی داده‌ها تعدیل شود. از این رو اگر ما انحراف معیار را بر $\bar{x}$ تقسیم کنیم میزان پراکندگی را برای یک واحد از میانگین حساب کرده‌ایم و به این ترتیب تأثیر بزرگی داده‌ها را از بین برده‌ایم. با ملاحظه بالا می‌توان توصیف زیر را برای ضریب تغییرات به‌کار برد.

ضریب تغییرات عبارت است از میزان پراکندگی به ازای یک واحد از میانگین

تمرین

۱ــ اگر میانگین برابر ۴ و انحراف معیار برابر ۶ باشد ضریب تغییرات چه‌قدر است؟

۲ــ ضریب تغییرات سن دانش‌آموزان کلاس شما ۱۰ سال دیگر چه تغییری می‌کند؟ کم‌تر می‌شود یا بیش‌تر؟

۳ــ اگر ۲۰٪ نمره هر دانش‌آموز به نمره او اضافه شود چه تأثیری روی ضریب تغییرات نمره‌ها حاصل می‌شود؟

۴ــ دو دسته دادهٔ زیر قیمت کالایی را در دو بازار جداگانه نشان می‌دهد.

۱۰، ۱۱، ۱۰، ۹، ۱۲، ۱۱، ۱۰، ۱۲، ۹، ۱۳، ۸

۱۰، ۹، ۱۱، ۱۰، ۱۰، ۱۱، ۹، ۱۰، ۸، ۱۳، ۱۰

الف ــ در کدام دسته پراکندگی بیش‌تر مشهود است؟

ب ــ دامنهٔ تغییرات را در هر دو دسته حساب کنید.

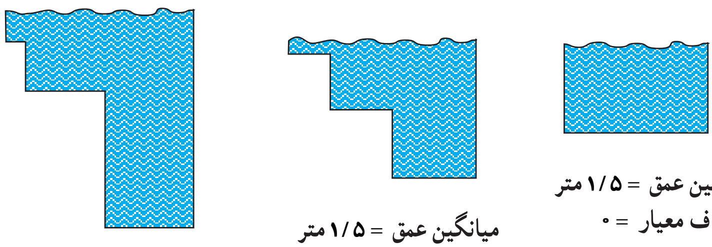
ج ــ آیا مقدار دامنهٔ تغییرات با پراکندگی مشاهده شده در «الف» هم‌خوانی دارد؟

د ــ ترجیح می‌دهید از کدام بازار خرید کنید؟ چرا؟

هـ ــ اگر داده‌های واقعی در اختیار شما نباشند، فقط به صرف داشتن دامنهٔ تغییرات می‌توانستید تصمیم‌گیری کنید؟

و ــ کدام یک از شاخص‌های پراکندگی در تصمیم‌گیری شما می‌تواند مؤثرتر باشد؟ این شاخص را معرفی کرده و محاسبه کنید.

۵ــ در محله‌ای سه استخر شنا با مشخصات زیر موجود است:



میانگین عمق = ۱/۵ متر
انحراف معیار = ۰

میانگین عمق = ۱/۵ متر
ضریب تغییرات = ۰/۵

میانگین عمق = ۱/۵ متر
انحراف معیار = ۱/۰۵

الف ــ برای فردی مبتدی با قد ۱/۴۵ متر شنا کردن در کدام یک از این استخرها امنیت بیش‌تری دارد؟ چرا؟

ب ــ در کدام استخر با خطر بیش‌تری روبه‌رو است؟ چرا؟

۶ـ دو دستهٔ دادهٔ زیر ارزش واردات برحسب کشورهای مبدأ را در سال‌های ۶۵ و ۷۰ نشان می‌دهد. (میلیون ریال)

کشور و منطقه	۱۳۵۵	۱۳۶۰	۱۳۶۵	۱۳۶۶	۱۳۶۷	۱۳۶۸	۱۳۶۹	۱۳۷۰
امارات	۳۲۴۳	۳۲۳۷۱	۴۱۳۴۳	۲۷۴۰۸	۱۹۱۹۹	۶۸۶۷۵	۶۵۴۵۲	۹۵۶۶۱
اندونزی	۴۱	۳۴	۲۱۱	۱۱۰۵	۷۳۴	۵۰۱۵	۴۳۳۵	۶۹۱۱
بنگلادش	۱۰۵۰	۳۰۹۸	۲۲۲۸	۲۶۷۷	۱۱۸۶	۳۱۹۶	۳۳۲۴	۳۰۹۶
پاکستان	۴۴۶۸	۶۷۲۱	۷۰۴۱	۴۳۰۴	۲۷۴۲	۱۰۴۵۵	۳۲۹۶	۱۲۴۰۳
تایلند	۱۳۷۵	۱۹۸۷۵	۴۰۲۵	۹۵۶۹	۴۸۹۳	۱۳۴۳۳	۱۳۲۷۷	۷۲۰۶
تایوان	۹۲۰	۱۰۵۱۰	۷۳۳۰	۴۶۱۹	۴۷۵۴	۵۲۹۱	۲۸۹۸	۱۰۲۹۸
ترکیه	۲۹۳۷	۲۳۳۶۹	۴۵۲۱۶	۳۳۳۳۱	۲۹۵۲۹	۵۰۶۴۵	۴۸۹۳۶	۶۹۶۳۳
چین	۷۰۹۳	۹۶۳۱	۱۹۹۳	۳۷۴۷	۲۴۸۶	۳۴۶۰	۱۱۶۱۳	۱۹۲۶۲
ژاپن	۱۵۵۵۰۰	۱۲۹۶۰۶	۹۷۶۲۷	۷۴۰۶۸	۵۸۰۶۶	۷۰۴۳۹	۱۳۰۲۴۸	۲۱۶۹۲۵
سنگاپور	۲۷۷۳	۸۲۵۴	۵۷۰۰	۳۸۱۶	۵۳۲۸	۵۵۵۴	۷۴۴۶	۹۰۳۰
سوریه	۲۷۳	۱۷۸	۹۱۴	۲۵۸۲	۶۸۸۳	۸۸۷۸	۲۰۰۵	۹۰۷
عربستان	۲۲	۱۶۹۱	۲۰۶۶	۷۳۲	۹۱۰	۶۳۸۲	۴۶۱	۲۰۸۶

ـ تغییرات ارزش واردات در کدام سال بیش‌تر بوده‌است؟

۷ـ با مراجعه به جدول صفحه بعد به سؤالات زیر پاسخ دهید.

۱۰ مرکز استان را به تصادف انتخاب کنید و بر اساس میزان بارندگی سالانه در این مراکز، میزان بارندگی سالانه در ایران در این سال را حدوداً به‌دست آورید.

ـ در این جامعه چند متغیر وجود دارد؟ نام ببرید و نوع آن‌ها را مشخص کنید.

ـ آیا میزان بارندگی سالانه در مراکز استان‌ها، به میزان بارندگی سالانه در ایران نزدیک است یا خیر؟

ـ برای اندازه‌گیری تفاوت‌ها از چه معیارهایی استفاده می‌کنید؟

ـ میزان بارندگی به طور متوسط در هر ماه در رشت چقدر است؟

سطر مربوط به بارندگی در استان گیلان (به مرکز رشت) معرف چه متغیری است؟ آن متغیر را معرفی کنید و نوع آن را مشخص کنید.

مقدار بارندگی سالانه مراکز استانها (میلیمتر)

مرکز استان	۱۳۶۵	۱۳۷۰	۱۳۷۲	۱۳۷۳	۱۳۷۴	۱۳۷۵	۱۳۷۶
اراک	۳۴۸	۳۲۰	۳۲۰	۳۵۶	۳۶۷	۳۰۳	۲۵۶
اردبیل	۰۰۰	۰۰۰	۳۲۸	۳۴۴	۲۳۰	۳۴۷	۳۴۲
ارومیه	۳۸۶	۲۶۹	۵۳۴	۴۸۶	۳۶۷	۲۶۳	۲۶۹
اصفهان	۱۳۸	۱۱۲	۷۷	۷۹	۲۰۷	۶۷	۲۱۱
اهواز	۲۸۶	۲۹۲	۱۶۸	۲۰۶	۲۳۹	۱۵۲	۵۳۹
ایلام	۰۰۰	۶۷۶	۷۴۱	۸۶۱	۵۴۹	۵۷۸	۷۱۷
بندرعباس	۲۳۰	۲۴۴	۳۶	۷۰	۵۰۷	۱۶۷	۴۳۷
بوشهر	۱۱۹	۱۷۱	۶۴	۲۸۷	۵۸۸	۲۱۷	۸۰۷
تبریز	۳۲۳	۲۳۶	۳۳۹	۳۳۶	۱۵۸	۲۵۹	۲۰۲
تهران	۲۹۲	۲۶۰	۲۱۹	۲۵۵	۳۱۹	۱۵۶	۲۴۴
خرم‌آباد	۵۹۴	۵۱۶	۵۸۷	۵۸۹	۵۳۷	۴۲۸	۶۰۹
رشت	۱۱۸۶	۱۲۹۸	۱۶۸۵	۱۳۱۹	۱۲۱۰	۱۰۱۵	۱۲۹۷
زاهدان	۴۲	۴۴	۵۴	۵۶	۱۹۰	۷۳	۱۴۹
زنجان	۳۲۱	۲۸۶	۴۲۵	۳۳۵	۲۶۶	۲۴۹	۳۳۱
سمنان	۲۲۲	۱۹۴	۶۷	۱۲۲	۱۹۲	۱۲۸	۱۶۱
سنندج	۵۴۸	۳۸۲	۵۹۹	۵۶۱	۵۱۳	۴۱۳	۵۷۶
شهرکرد	۴۹۲	۴۹۱	۲۸۲	۳۶۸	۳۷۹	۲۵۸	۴۲۳
شیراز	۴۹۵	۲۷۲	۱۵۵	۴۵۱	۵۷۸	۲۰۲	۴۹۸
قزوین	۰۰۰	۰۰۰	۰۰۰	۰۰۰	۰۰۰	۰۰۰	۳۳۶
قم	۰۰۰	۰۰۰	۰۰۰	۰۰۰	۰۰۰	۷۸	۱۳۸
کرمان	۱۵۹	۱۱۵	۵۵	۱۷۶	۲۵۶	۱۲۳	۱۷۹
کرمانشاه	۵۶۹	۴۵۱	۶۵۲	۵۵۹	۴۳۳	۳۹۶	۴۸۸
گرگان	۰۰۰	۰۰۰	۰۰۰	۰۰۰	۰۰۰	۰۰۰	۵۳۶
مشهد	۲۵۰	۳۲۶	۳۲۷	۱۶۷	۲۲۴	۱۷۰	۳۴۸
همدان	۴۱۷	۲۹۲	۲۸۸	۴۳۸	۳۱۹	۲۱۳	۳۳۶
یاسوج	۱۲۲۸	۱۲۰۲	۴۱۷	۱۰۶۳	۱۰۲۵	۴۷۴	۱۱۹۸
یزد	۶۴	۶۰	۲۱	۳۴	۱۰۴	۲۹	۴۲

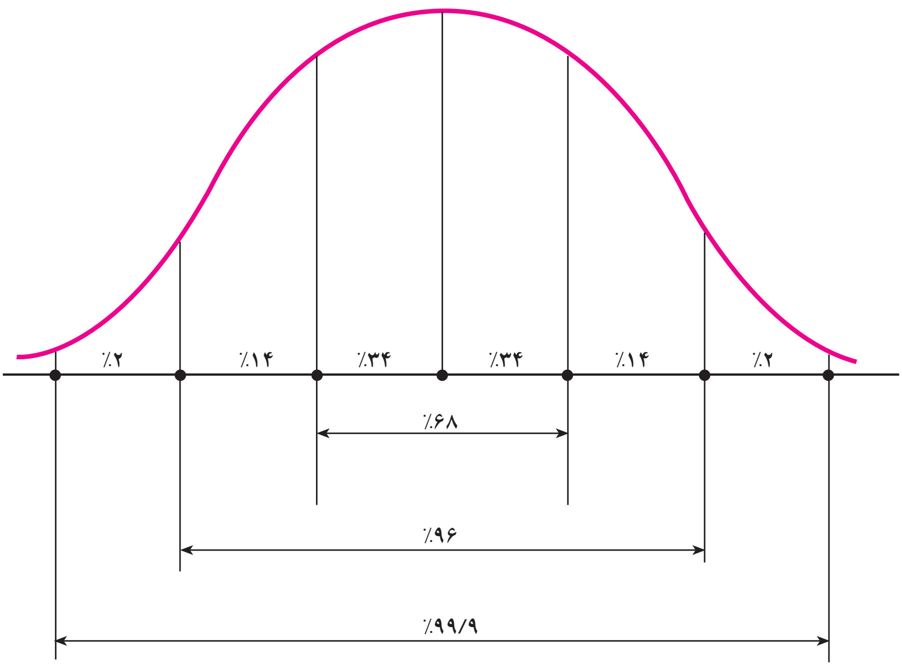
مأخذ: سازمان هواشناسی کشور.

پراکندگی در منحنی نرمال

قبلاً در فصل نمودارها و همچنین در فصل شاخص‌های مرکزی به منحنی نرمال اشاره کردیم. اگر σ انحراف معیار این جامعه باشد پراکندگی افراد در جامعه به‌صورت زیر است. برای درک بهتر موضوع ابتدا فرض می‌کنیم مثلاً میانگین برابر ۵ و انحراف معیار برابر ۲ باشد در این صورت می‌خواهیم بدانیم چند درصد افراد در فاصله‌ای به مرکز میانگین و به فاصله یک انحراف معیار از آن قرار دارند. درواقع می‌خواهیم بدانیم چند درصد افراد در فاصله (۲+۵ ، ۲–۵) یعنی (۳,۷) قرار دارند.

در این منحنی نشان داده می‌شود که این درصد برابر ۶۸٪ است.

همچنین درصد افرادی که در فاصله‌ای به مرکز میانگین و به طول دو برابر انحراف معیار از میانگین قرار دارند برابر ۹۶٪ است و سرانجام تقریباً ۱۰۰٪ افراد در فاصله‌ای به مرکز میانگین و به‌طول ۳ برابر انحراف معیار قرار دارند. این اطلاعات را به‌طور مختصر در شکل زیر ملاحظه می‌کنید (مقادیر روی نمودار تقریبی هستند.) :



محاسبات شاخص‌های عددی در MINITAB

ــ داده‌های زیر ساعاتی را که ۱۰ دانشجو (انتخاب شده به‌طور تصادفی) صرف تحقیق در طول هفته می‌کنند نشان می‌دهد :

۸، ۰، ۴، ۱۰، ۷، ۲، ۰، ۵، ۱۴، ۷

با استفاده از MINITAB میانگین، میانه، دامنه تغییرات و انحراف معیار این داده‌ها را به‌دست آورید.

شما می‌توانید از Calc و Stat در لیست انتخاب استفاده نمایید.

Stat مقادیر چند شاخص عددی را هم‌زمان ارائه می‌دهد ولی Calc مقدار شاخص‌ها را یکی‌یکی به‌دست می‌آورد. با استفاده از MINITAB WINDOWS برای انتخاب Stat مراحل زیر را دنبال کنید :

قدم اول: داده‌ها را در ستون C1 وارد کنید.

قدم دوم: روی Stat در لیست انتخاب، کلیک کنید.

قدم سوم: روی Basic Statistics کلیک کنید.

قدم چهارم: روی Descriptive Statistics کلیک کنید.

قدم پنجم: C1 را در کار زیر Variable تایپ کنید.

قدم ششم: روی OK کلیک کنید. نتایج و خروجی Minitab روی صفحه نمایان می‌شود.

اگر از فرمان‌های زبانی MINITAB استفاده می‌کنید، با استفاده از SET داده‌ها را در ستون C1 وارد کنید. سپس فرمان زیر را تایپ کنید :

```
MTB > DESCRIBE   C1
```

با اجرای این فرمان خروجی Minitab ظاهر می‌شود.

فعالیت

فعالیت مثال بالا را اجرا کنید و مقادیر به‌دست آمده را در جدول زیر بنویسید.

Descriptive Statistics

Varable	N	Mean	Median	Tr Mean	Stder	SE MEAN
hours	?	?	?	?	?	?

Variable	Min	Max	Q1	Q3
hours	?	?	?	?

همان‌طور که گفته شد می‌توان از دستور Calc نیز استفاده کرد، اگر از Minitab Windows استفاده می‌کنید به‌ترتیب زیر عمل کنید :

قدم اول: داده‌ها را در ستون C1 وارد کنید.

قدم دوم: روی Calc در لیست انتخاب، کلیک کنید.

قدم سوم: روی Column Statistics کلیک کنید.

قدم چهارم: در صفحه Column Statistics روی شاخصی که می‌خواهید محاسبه کنید کلیک کنید. به‌طور مثال برای پیدا کردن میانگین روی Mean کلیک کنید.

قدم پنجم: در کنار Input Variable در کادر، C1 را تایپ کنید.

قدم ششم: اگر می‌خواهید پاسخ را به‌عنوان یک عدد ثابت ذخیره کنید در کادر در کنار Store Result in، K1 را تایپ کنید. اگر از K1 قبلاً استفاده کرده‌اید K2 را تایپ کنید. اگر نمی‌خواهید نتیجه را ذخیره کنید، این مرحله را اجرا نکنید.

قدم هفتم: روی OK کلیک کنید. نتیجه روی صفحه نمایان خواهد شد.

اگر از فرمان‌های زبانی Minitab استفاده می‌کنید، با استفاده از SET داده‌ها را در ستون C1 وارد کنید. سپس فرمان زیر را تایپ کنید :

```
MTB > Mean   C1
```

در پاسخ، Minitab خروجی زیر را ارائه می‌کند :

```
Column Mean
Mean of C1 = .........
```

به‌طور متشابه می‌توانید میانه، دامنه تغییرات و انحراف معیار را با کلیک کردن روی شاخص مورد نظر محاسبه کنید و فرمان‌های زبانی Minitab برای شاخص‌های فوق عبارتند از :MEDIAN C1، RANGE C1، STANDARD DEVIATION که باید پس از PROMPT تایپ شوند.

نمودار جعبه‌ای

داده‌های زیر، درآمد هفتگی نمونهٔ ۱۲ تایی از خانواده‌ها را نشان می‌دهد.

۵۲۰۰۰	۲۲۰۰۰	۶۰۰۰۰	۳۲۰۰۰	۱۷۰۰۰	۲۳۰۰۰
۴۶۰۰۰	۲۷۰۰۰	۹۲۰۰۰	۴۲۰۰۰	۳۸۰۰۰	۲۹۰۰۰

با استفاده از Minitab نمودار جعبه‌ای این داده‌ها را رسم کنید. اگر از Minitab Windows

استفاده می‌کنید، مراحل زیر را دنبال کنید :

قدم اول: داده‌ها را در ستون C1 وارد کنید.

قدم دوم: روی GRAPH در لیست انتخاب کلیک کنید.

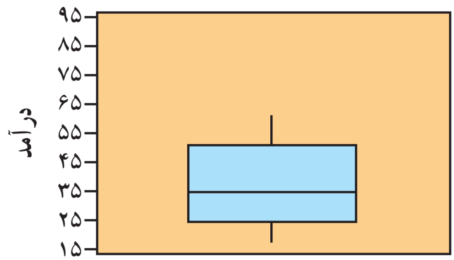
قدم سوم: روی BOXPLOT کلیک کنید.

قدم چهارم: C1 را در کادر زیر Y تایپ کنید.

قدم پنجم: روی OK کلیک کنید. نمودار جعبه‌ای مورد نظر نمایان می‌شود. اگر از فرمان‌های زبانی Minitab استفاده می‌کنید ابتدا با استفاده از فرمان SET داده‌ها را در ستون C1 وارد کنید. سپس فرمان زیر را تایپ کنید :

```
MTB > Boxplot   C1
```

با اجرای دستورات فوق نمودار جعبه‌ای زیر رسم می‌شود :



تمرین

۱ــ نمودار ساقه و برگ زیر موجود است.

ــ میانگین، انحراف معیار این داده‌ها را با استفاده از Minitab به‌دست آورید.

ــ نمودار جعبه‌ای این داده‌ها را با استفاده از Minitab رسم کنید.

ساقه	برگ
۰	۳۶۹
۱	۲۸۵۱۰۵
۲	۵۱۶
۴	۱
۵	
۶	۲