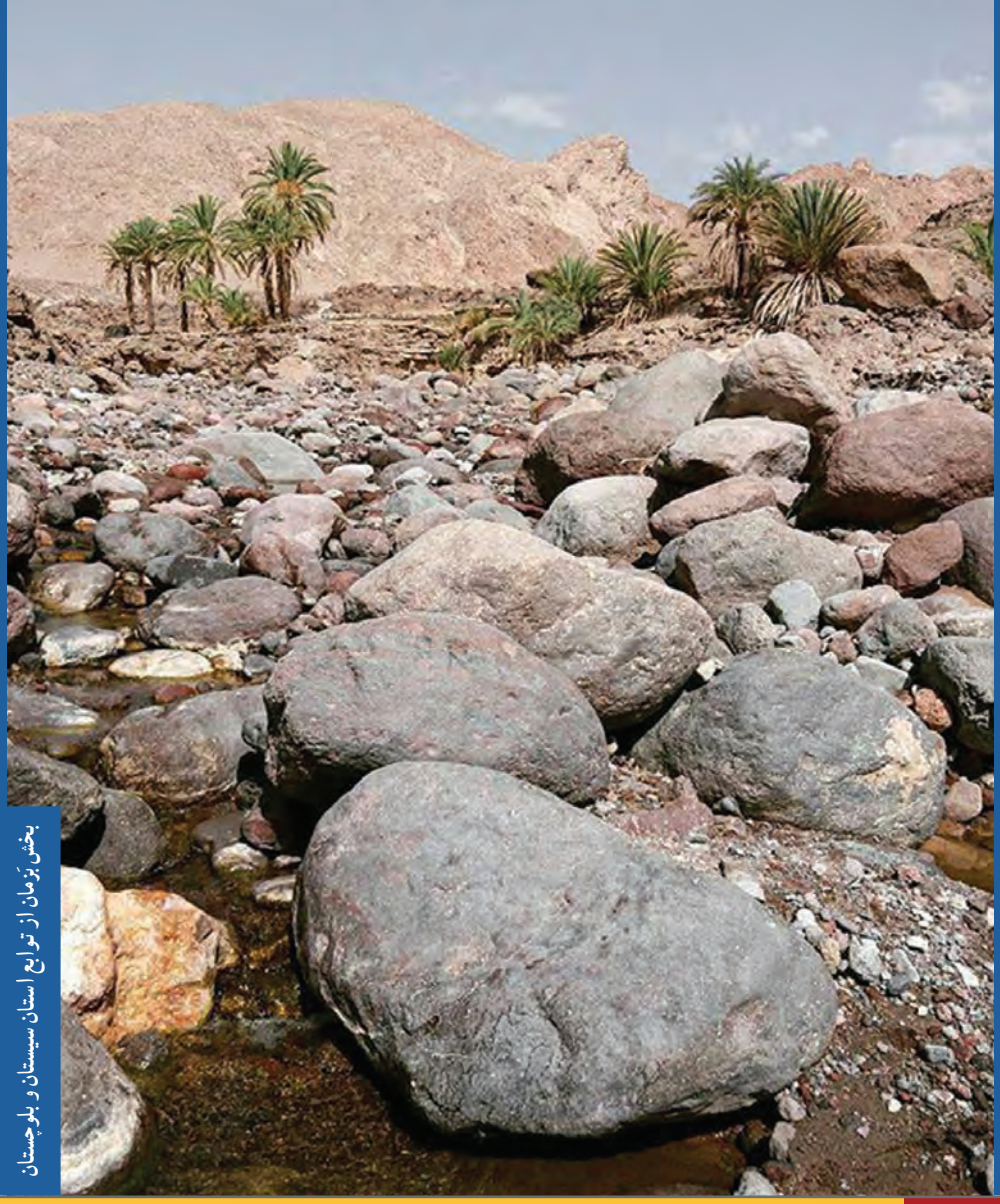
ماسوله ـ استان گیلان



بخش بَرمان از توابع استان سیستان و بلوچستان

اوج علم آمار استفاده از علم احتمال برای ابداع آمار استنباطی بوده است. آمار استنباطی استخراج نتایجی براساس بخش کوچکی از داده‌ها در یک مسئله است که می‌توان آن را به همه داده‌های آن مسئله تعمیم داد. با وجود آنکه پایه‌های احتمال بر نظریه‌های ریاضی استوار شده است، ولی آمار را نمی‌توان به عنوان شاخه‌ای از ریاضی در نظر گرفت.

# ۴ آمار استنباطی

۱ گردآوری داده‌ها

۲ برآورد

# درس ۱ گردآوری داده‌ها

## فعالیت

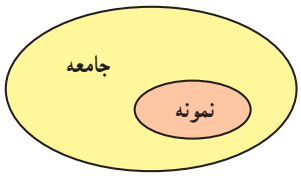
می‌خواهیم برخی از ویژگی‌های مگس‌های سفید مزاحم در شهر تهران را بررسی کنیم. آیا برای انجام این کار می‌توانیم ویژگی‌های همهٔ مگس‌های سفید را اندازه‌گیری کنیم؟ آیا همهٔ آنها در دسترس‌اند؟ آیا زمان و هزینه لازم برای این کار در اختیار داریم؟

**واحد آماری**[۱] به هر یک از افراد یا اشیا می‌گویند که داده‌های مربوط به آنها در یک بررسی آماری گردآوری می‌شود.

مجموعهٔ کلّ واحدهای آماری را **جامعهٔ آماری**[۲] می‌نامند.

هر زیر مجموعه از جامعهٔ آماری را که با روش مشخصی انتخاب شده باشد، یک **نمونه** می‌نامند.

**نمونه‌گیری**[۳]، فرایند انتخاب نمونه‌ای از یک جامعه، به منظور تعمیم اطلاعات آن به جامعه است.




بیشتر مطالعات آماری بر روی بخشی از جامعه است. رابطهٔ بین جامعه و بخشی از آن‌که نمونه نامیده می‌شود، در شکل نشان داده شده است.

## کار در کلاس

در فعالیت قبل هر مگس سفید ..... است. همهٔ مگس‌های سفید، که کل واحدهای آماری هستند، ..... را تشکیل می‌دهند. اگر سنّ همهٔ مگس‌های سفید را در اختیار داشته باشیم، داده‌های ..... را داریم. ۱۰۰ مگس سفید معرف یک ..... است.

---

۱ــ Statistical Unit

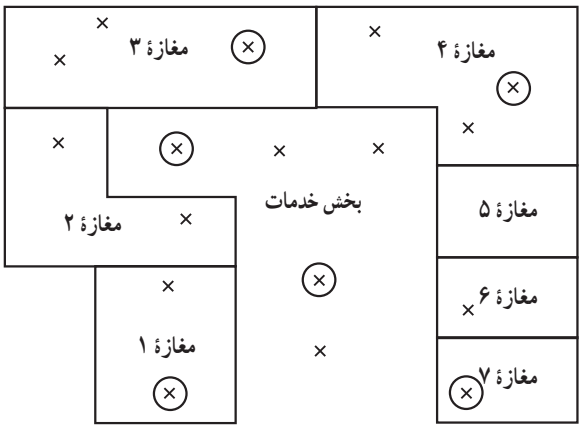
۲ــ Statistical Population

۳ــ Sampling

**نمونه‌گیری تصادفی ساده**[1] نوعی روش نمونه‌گیری که در آن همهٔ واحدهای آماری برای انتخاب شدن در نمونه، احتمال یکسان دارند.

## فعالیت

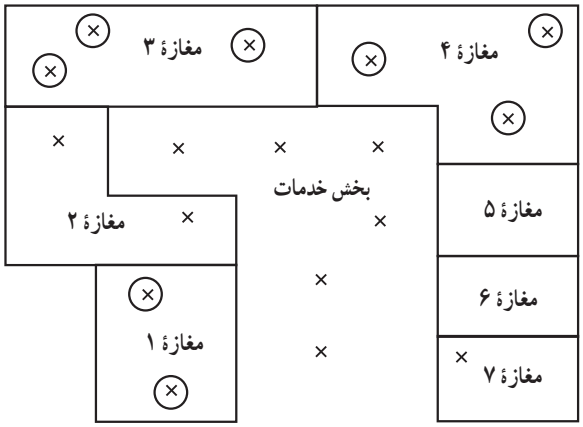
**۱** می‌خواهیم متوسط درآمد کارکنان یک مجتمع تجاری را محاسبه کنیم. اگر این مجتمع از ۷ مغازه و یک بخش خدمات تشکیل شده باشد، که روی هم ۱۷ کارکن دارند، چگونه از بین ۱۷ نفر، ۶ نفر را به تصادف انتخاب می‌کنید؟ یک راه ساده برای انجام این کار نوشتن اسامی کارکنان یا شمارهٔ کارمندی آنها روی ۱۷ برگهٔ کوچک و انتخاب تصادفی ۶ تا از آنهاست. آیا این روش نمونه‌گیری، نمونه‌گیری تصادفی ساده است؟ آیا همهٔ واحدهای جامعه احتمال برابری برای انتخاب دارند؟




در شکل روبه‌رو نقشه‌ای از مجتمع تجاری ترسیم شده که کارکنان با × و دور انتخاب‌شدگان یک دایره رسم شده است.

انجام نمونه‌گیری تصادفی ساده در عمل با دشواری‌هایی همراه است. اگر اندازه جامعهٔ بزرگ باشد، یعنی تعداد واحدهای آماری زیاد باشند، دسترسی به فهرستی از اعضای جامعه و دسترسی به اعضای انتخابی، دشوار و ممکن است هزینه‌بر باشد.

**۲** هر یک از ۷ مغازه و بخش خدمات را به‌صورت یک گروه فرض می‌کنیم. حال از بین ۸ گروه در نظر گرفته شده، سه تا از آنها را به تصادف انتخاب می‌کنیم و در هر یک سرشماری انجام می‌دهیم. آیا این روش نمونه‌گیری **سریع‌تر** است؟




**نمونه‌گیری خوشه‌ای**[2] : نمونه‌گیری که در آن، واحدهای نمونه‌گیری اوّلیه در جامعه، گروه‌ها یا خوشه‌ها باشند. سپس همهٔ واحدهای آماری خوشه‌های انتخاب شده را به عنوان نمونه درنظر می‌گیرم.

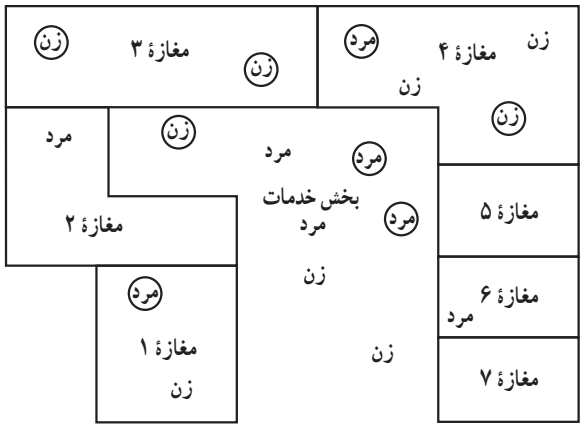
۱ــ Simple Random Sampling

۲ــ Cluster Sampling

**سؤال :** می‌خواهیم میانگین نمرات ریاضی دانش آموزان شهر تهران را محاسبه کنیم. اگر فهرست همهٔ دانش‌آموزان را نداشته باشیم، اما فهرست مدارس موجود باشد، نمونه‌گیری خوشه‌ای، راه مناسبی برای گردآوری داده‌هاست. اگر بودجهٔ کافی یا زمان لازم برای نمونه‌گیری تصادفی ساده نداشته باشیم آیا این روش مقرون به صرفه است؟

**۳** اگر بخواهیم یک نمونهٔ ۸ تایی شامل دقیقاً ۴ مرد و ۴ زن از مجتمع تجاری بگیریم، چگونه این کار را انجام می‌دهیم؟ زمانی که جامعه به دو یا چند بخش تقسیم می‌شود که عضو مشترکی ندارند، می‌توان از هر بخش جداگانه نمونه‌گیری کرد. این کار با افزایش هزینه یا زمان همراه است، ولی انتظار داریم که ..... را نیز افزایش دهد.

در واقع داده‌ها به دو طبقهٔ مجزّا تقسیم شده‌اند که طبقات از نظر ویژگی مورد بررسی همگن هستند. در صورتی که در نمونه‌گیری خوشه‌ای، درون خوشه‌ها هرچه ویژگی مورد بررسی تفاوت بیشتری داشته باشند بهتر است.




**نمونه‌گیری طبقه‌ای**[۱] : روش نمونه‌گیری که در آن با طبقه‌بندی جامعه به زیرجامعه‌های مجزّا یک نمونهٔ تصادفی ساده از هر طبقه انتخاب می‌شود.

علاقه‌مند به نمونه‌گیری از نمرات درس ریاضی دانش آموزان استان تهران هستیم. اگر فهرست همهٔ دانش آموزان را در اختیار داشته باشیم، می‌توانیم از نمونه‌گیری تصادفی ساده استفاده کنیم؛ ولی این روش نمونه‌گیری هیچ تضمینی ندارد که دانش‌آموزان از تمامی شهرهای استان در نمونه حضور داشته باشند. درصورتی‌که اگر از هر شهر متناسب با تعداد دانش‌آموزان آن شهر نمونه‌گیری تصادفی ساده انجام دهیم، مشکل قبلی رفع می‌شود. به عبارت دیگر از نمونه‌گیری طبقه‌ای استفاده می‌کنیم. حال فرض کنید فقط فهرست مدارس را داشته باشیم. چه روش نمونه‌گیری را پیشنهاد می‌کنید؟ (راهنمایی : شما می‌توانید از دو روش نمونه‌گیری پشت سر هم استفاده کنید.)

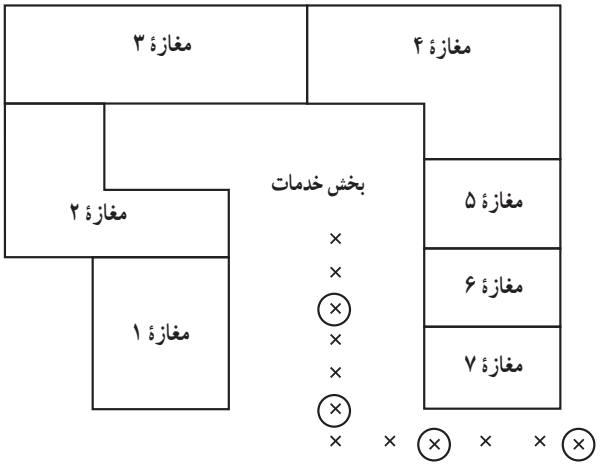
**۴** فرض کنید در مجتمع، ۱۲ نفر حضور دارند. صبر می‌کنیم که مجتمع تجاری تعطیل شود و هنگام خروج کارکنان می‌خواهیم نمونهٔ ۴ نفری انتخاب کنیم. برای این منظور، همانند شکل صفحهٔ بعد عمل کرده‌ایم. ابتدا از ۳ نفر یکی را به تصادف انتخاب می‌کنیم. در این شکل، نفر اوّل انتخاب شده است. حال با همین رویه برای سه نفر بعد هم، نفر اوّل را انتخاب می‌کنیم و ادامه می‌دهیم. این روش نمونه‌گیری شباهت بیشتری به کدام یک از روش‌های نمونه‌گیری قبلی دارد؟ خوشه‌ای یا طبقه‌ای؟ این کار باعث چه نوع صرفه‌جویی می‌شود؟

---

۱ــ Stratified Sampling

به نظر شما این نوع نمونه‌گیری در کدام‌یک از مثال‌های زیر امکان دارد :

- گردآوری اطلاعات از مبدأ و مقصد مسافران در خروجی ــ ورودی یک شهر
- کنترل کیفیت یک خط تولید
- انتخاب نمونه از ماهی‌های یک حوضچه
- زمانی که فهرستی از واحدهای جامعه وجود نداشته باشد.
- فهرست واحدهای آماری ترتیب تصادفی داشته باشند.




**نمونه‌گیری یا سامانمند**[1] **(سیستماتیک)**، نوعی نمونه‌گیری طبقه‌ای است که در آن اندازه طبقات باهم برابر است. فقط از طبقهٔ اوّل، واحد آماری به تصادف انتخاب می‌شود و با همان رویه از طبقات دیگر، این کار انجام می‌گیرد.

آیا اعضای جامعه برای انتخاب شدن در نمونه‌گیری سامانمند شانسی برابر دارند؟ چرا؟

## کار در کلاس

جدول زیر را کامل کنید.

| روش نمونه‌گیری | مزیت | محدودیت |
|---|---|---|
| تصادفی ساده | | |
| خوشه‌ای | | |
| طبقه‌ای | | |
| سامانمند | | |

## فعالیت

از مگس‌های سفید با چه روشی می‌توان نمونه‌گیری کرد؟ فهرستی از آنها نداریم، تعداد آنها را هم نمی‌دانیم. می‌توان چند منطقه از تهران را به تصادف انتخاب کرد و در هر منطقه نمونهٔ در دسترس را انتخاب و بررسی کنیم. آیا این روش نمونه‌گیری به تمامی واحدهای جامعه شانس انتخاب می‌دهد؟

**نمونه‌گیری احتمالی** : نمونه‌گیری است که همهٔ واحدهای آماری احتمالی معلوم برای انتخاب در نمونه داشته باشد. و از روشی تصادفی برای انتخاب واحدهای نمونه استفاده شود.

---

۱ــ Systematic

نمونه‌گیری‌های چهار فعالیت قبل، همگی احتمالی‌اند. در کدام‌یک همهٔ واحدهای آماری احتمال برابری برای انتخاب دارند.

## کار در کلاس

راه حلی ارائه کنید که نمونه‌گیری‌های غیر احتمالی زیر را احتمالی می‌کند، هرچند که به صورت غیر واقعی باشد.

| مثال | نمونه‌گیری غیر احتمالی | نمونه‌گیری احتمالی |
|---|---|---|
| نمونه‌گیری از یک قفس بزرگ خرگوش‌های یک آزمایشگاه | بدون برنامه‌ریزی خرگوش‌هایی را برمی‌دارد که دستش به آنها می‌خورد. | |
| در مطالعاتی که در آنها فرایند سنجش برای شخصی که سنجیده می‌شود ناخوشایند یا دردسرآفرین است. | داوطلبانی که حاضر به پاسخ به سؤالات شما در یک نظرسنجی می‌شوند. | |
| نمونه‌گیری از زغال‌سنگ‌های یک واگن | نمونه در دسترس انتخاب می‌شود. | |

## فعالیت

شاید در نگاه اوّل این‌طور به نظر برسد که انجام نمونه‌گیری تصادفی ساده کاری آسان است، در حالی که در دنیای واقعی، گاهی چنین نیست. روش‌های نمونه‌گیری که به ذهن می‌رسند، گاهی مشکلات و ایرادهایی دارند که در نگاه اوّل دیده نمی‌شوند و استفاده از آنها ما را به نتایجی بسیار دور از واقعیت می‌رساند. این موضوع را با چند مثال بهتر متوجه خواهید شد.

فرض کنید آمارگیری می‌خواهد بداند در یک شهر خانواده‌ها چند نفره‌اند. او برای این کار صد نفر را به‌تصادف انتخاب می‌کند و از آنها می‌پرسد : «خانوادهٔ شما چند نفر است؟»

آیا این روش برای نمونه‌گیری درست است؟ جواب منفی است! دلیل آن هم این است که واحدهای آماری مورد نظر در این مسئلهٔ خانواده‌ها هستند نه افراد. آیا خانواده‌های مختلف احتمال حضور برابر در این نمونه‌گیری را دارند؟ واضح است که احتمال حضور هر خانواده متناسب با تعداد اعضای آن است و مثلاً احتمال حضور یک خانوادهٔ شش نفره دو برابر احتمال حضور یک خانوادهٔ سه نفره است و این، یعنی شرایط نمونه‌گیری ساده برقرار نیست.

نتیجهٔ چنین ایرادی در نمونه‌گیری این است که هر چه تعداد نمونه‌ها را افزایش دهیم، نتایج به مقداری اشتباه نزدیک‌تر می‌شود. مثلاً فرض کنید آمار واقعی تعداد افراد خانواده‌ها چنین باشد :

| تعداد افراد | ۱ | ۲ | ۳ | ۴ | ۵ و بیشتر |
|---|---|---|---|---|---|
| درصد | ۸/۵ | ۲۰/۷ | ۲۸/۵ | ۲۷/۶ | ۱۴/۷ |

در این جامعه تعداد خانواده‌های دو نفره تقریباً $\frac{\cdots\cdots}{\cdots\cdots} = ۱/۴$ برابر تعداد خانواده‌های پنج نفره (و بیشتر) است، ولی با آمارگیری نادرستی که توضیح داده شد به نتیجهٔ دیگری خواهیم رسید؛ احتمال حضور یک خانوادهٔ پنج نفره (و بیشتر) در

نمونه‌ها بیشتر از ..... برابر احتمال حضور یک خانوادهٔ دو نفره است و لذا عددی که در روش نادرست آماری به‌دست می‌آید کمتر از $\frac{۱/۴}{۲/۵}=۰/۵۶$ است. نتیجه اینکه هر چند واقعیت این است که نسبت خانواده‌های دو نفره بسیار بیشتر از خانواده‌های ۵ نفره (و بیشتر) است، ولی ما با نمونه‌گیری اشتباه به نتیجه‌ای بسیار متفاوت می‌رسیم.

برای برطرف کردن این مشکل راه‌های مختلفی دارد. مثلاً اینکه فقط از سرپرست خانواده‌ها در مورد تعداد اعضای خانواده‌ها بپرسیم. (جدول صفحهٔ قبل، برگرفته از آمار واقعی کشور در سرشماری[۱] سال ۱۳۹۵ است.)

## کار در کلاس

فرض کنید در شهری جمعیت کلاس‌های پایهٔ ششم دبستان به شکل زیر باشد :

| تعداد دانش‌آموز | ۲۵ | ۲۶ | ۲۷ | ۲۸ | ۲۹ | ۳۰ | ۳۱ | ۳۲ | ۳۳ | ۳۴ | ۳۵ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| تعداد کلاس | ۷ | ۹ | ۱۰ | ۱۱ | ۷ | ۱۱ | ۸ | ۸ | ۷ | ۵ | ۴ |

الف) چه تعداد کلاس پایهٔ ششم در این شهر وجود دارد؟ تعداد دانش‌آموزان پایهٔ ششم چند تاست؟

ب) چه درصدی از کلاس‌های پایهٔ ششم بیشتر از ۳۰ دانش‌آموز دارند؟

پ) اگر به‌تصادف یک دانش‌آموز ششم دبستانی را انتخاب کنیم، احتمال اینکه کلاسی که در آن درس می‌خواند بیشتر از ۳۰ دانش‌آموز داشته باشد، چقدر است؟

ت) فرض کنید فردی برای اینکه بفهمد کلاس‌های ششم دبستان چند نفری‌اند، تعداد زیادی دانش‌آموز ششم دبستانی را به‌تصادف انتخاب کند و از آنها بپرسد «کلاس شما چند نفره است؟» این کار چه ایرادی دارد؟

ث) اگر با روش قبل درصد کلاس‌های با بیش از ۳۰ دانش‌آموز را محاسبه کنیم، نتیجه از جواب واقعی چند درصد فاصله خواهد داشت؟

**مثال** : فرض کنید می‌خواهیم میزان آلایندگی خودروهای در حال تردّد در شهری را بررسی کنیم. برای این کار چگونه باید نمونه‌گیری کنیم؟

اگر نمونه‌گیری را در تعمیرگاه‌ها انجام دهیم، هرچند هر خودرویی ممکن است گاهی سر از تعمیرگاه درآورد، ولی این نمونه‌گیری، تصادفی ساده نیست؛ زیرا احتمال اینکه یک خودرو در نمونهٔ ما باشد، متناسب با ساعاتی است که در تعمیرگاه بوده است و لذا درصد خودروی‌های آلاینده بسیار بیشتر از واقعیت نشان داده خواهد شد.

اگر نمونه‌گیری را در خیابان انجام دهیم، مشکل آن کمتر است، ولی بازهم نمونه‌گیری ما مشکل دارد؛ زیرا خودروهایی که از آنها بیشتر استفاده می‌شود، احتمال بیشتری دارد که به‌عنوان نمونه انتخاب شوند.

## کار در کلاس

فرض کنید بخواهیم میزان مطالعهٔ غیردرسی دانش‌آموزان یک مدرسه را بررسی کنیم. برای این کار این سؤال را طراحی کرده‌ایم و می‌خواهیم از یک نمونهٔ ۳۰ تایی آن را بپرسیم :

---

۱ــ سر شماری (Census) فرایند انتخاب همهٔ واحدهای آماری جامعه است.

«در یک سال گذشته چند کتاب غیردرسی خوانده‌اید؟»

روش‌های زیر را نقد کنید :

الف) پرسیدن سؤال از تعدادی از دانش‌آموزانی که در کتابخانه هستند.

ب) گذاشتن تعدادی پرسش‌نامه در محل رفت و آمد دانش‌آموزان.

پ) پرسیدن از دانش‌آموزانی که صبح وارد مدرسه می‌شوند و مایل‌اند به سؤال مذکور جواب دهند.

شما چه روشی را پیشنهاد می‌کنید که به نمونه‌گیری تصادفی ساده نزدیک‌تر باشد؟

## کار در کلاس

از جمله مسائلی که مردم در مورد آنها به نظرسنجی‌ها علاقه زیادی نشان می‌دهند، انتخابات است. با این وجود، گاهی مردم و گاهی گروه‌های سیاسی از روش‌هایی برای کشف نظر مردم استفاده می‌کنند که آنها را گمراه می‌کند.

در این مورد روش‌های زیر را نقد کنید :

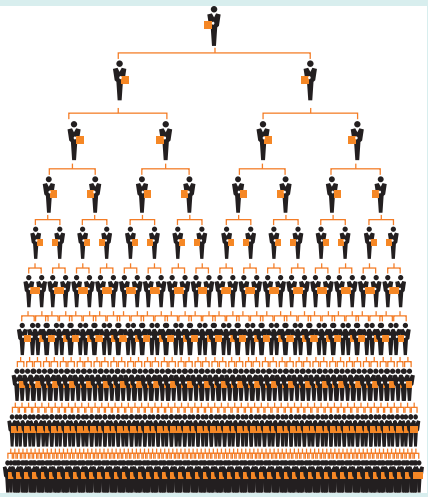
الف) پرسیدن نظر دوستان و اطرافیان.

ب) طراحی یک نظرسنجی در وبگاهی پرطرف‌دار و لحاظ کردن ساز و کاری که از یک آدرس بیش از یک بار رأی گرفته نشود.

در تمام مثال‌هایی که ذکر شد، روش نمونه‌گیری به شکلی بود که داده‌هایی به سمتی انحراف پیدا می‌کردند و لذا افزایش تعداد نمونه‌ها نیز به کاهش این انحراف کمک نمی‌کرد. در علم آمار اصطلاحی خاص برای این مشکل وجود دارد :

اگر یک روش نمونه‌گیری از نمونه‌گیری ایده‌آل فاصله بگیرد و به سمتی خاص انحراف پیدا کند می‌گویند آن روش نمونه‌گیری **اریب** است. لذا آمارشناسان تلاش می‌کنند تا با شناسایی منابع تولید اریبی، نمونه‌گیری‌ها را تا جایی که می‌توانند **نااریب** کنند.

## خواندنی



شرکت‌های هرمی که سال‌ها در دنیا کلاهبرداری کرده‌اند از اواخر دهۀ ۷۰ شمسی (حدود سال ۱۳۷۸) وارد ایران شدند. یک شرکت هرمی از مشتریان خود می‌خواهد که افراد جدیدی را به عضویت شرکت درآورند و به آنها وعده می‌دهد که اگر زیرشاخه تولیدشده به‌اندازۀ کافی بزرگ شود، جوایزی به او تعلق خواهد گرفت. معمولاً اعضای این شرکت‌ها برای اینکه بفهمند شانس رسیدن به جایزه‌ها چقدر است به اطرافیان خود و به‌خصوص بالاسری‌های خود نگاه می‌کنند. مثال سادۀ صفحۀ بعد نشان می‌دهد که این نوع نمونه‌گیری تا چه حد گمراه کننده است :

شرکتی هرمی از مشتریان خود می‌خواهد که بعد از خرید یک میلیون تومان کالا، دو نفر را به‌عنوان زیرشاخه خود معرفی کنند تا آنها هم همین کار را انجام دهند. شرکت به هر کس که زیرمجموعه‌هایش تا سه ردیف رشد کند جایزه می‌دهد. عموم مشتریان تصور می‌کنند تا قبل از یک ماه به جایزه خواهند رسید.

درخت دودویی صفحهٔ قبل نشان می‌دهد که مجموعهٔ مشتریان این شرکت چگونه رشد می‌کند؛ اگر این درخت $n$ ردیف داشته باشد، تعداد اعضای آن برابر است با :

$$۱+۲+۴+\ldots+۲^{n-۱}=۲^{n}-۱$$

همهٔ این اعضا، مگر آنهایی که در سه ردیف آخر قرار دارند، از شرکت جایزه گرفته‌اند. تعداد آنها برابر است با :

$$۱+۲+۴+\ldots+۲^{n-۴}=۲^{n-۳}-۱$$

پس نسبت کسانی که جایزه گرفته‌اند، به کل اعضای برابر کسر زیر است:

$$\frac{۲^{n-۳}-۱}{۲^{n}-۱}\cong\frac{۲^{n-۳}}{۲^{n}}=\frac{۱}{۸}$$

یعنی از هر ۸ نفر فقط یکی موفق به گرفتن جایزه شده است. اکنون از زاویهٔ دید کسی به ماجرا نگاه کنید که در ردیف آخر این مجموعه است و به بالاسری‌های خود نگاه می‌کند. خودش و افراد بالای سرش $n$ نفر هستند و غیر از سه نفر پایینی همه جایزه گرفته‌اند. لذا این عضو، نسبت جایزه گرفته‌ها را به‌کل اعضا به این شکل برآورد می‌کند:

$$\frac{n-۳}{n}=۱-\frac{۳}{n}$$

و این عبارت وقتی $n$ بزرگ باشد، تقریباً برابر یک است. مثلاً به ازای $n=۱۲$ که تعداد اعضای شرکت $۲^{۱۲}-۱=۴۰۹۵$ نفر است. یک عضو تازه وارد تصور می‌کند که نسبت کسانی که جایزه گرفته‌اند برابر با:

$$۱-\frac{۳}{۱۲}=۰/۷۵$$

است؛ یعنی ۷۵ درصد. در حالی که مقدار واقعی کمتر از ۱۲/۵ درصد است!

اگر تعداد ردیف‌ها به $n=۱۸$ برسد، تعداد اعضای شرکت $۲^{۱۸}-۱=۲۶۲۱۴۳$ نفر می‌شود. در این صورت، عضو تازه واردی که فقط به خودش و بالاسری‌های خودش نگاه می‌کند، تصور می‌کند که نسبت کسانی که جایزه گرفته‌اند برابر $۱-\frac{۳}{۱۸}=۰/۸۳$ یعنی ۸۳ درصد است؛ در حالی که در واقع این نسبت تقریباً همان ۱۲/۵ درصد است!

این شرکت‌ها، اگر از طریق قانون، متوقف نشوند، آن‌قدر رشد می‌کنند که دیگر ادامهٔ کار برای اعضا به‌صرفه نیست و در آن مقطع مشخص می‌شود که مثلاً از حدود یک میلیون عضو، ۸۷۵ هزار نفر ضرر کرده‌اند! از آن بدتر وضعیت کسانی است که جایزه‌ای از شرکت گرفته‌اند، ولی زیرشاخه‌های آنها که عموماً دوست و فامیلشان هستند ضرر کرده‌اند؛ این افراد آبرو و اعتبار خود را از دست داده‌اند.

**فعالیت**

می‌خواهیم طول قد دانش‌آموزان یک مدرسه را گردآوری کنیم. برای این منظور چه راهی پیشنهاد می‌کنید؟

**آمارگیری** : گردآوری داده‌ها به یکی از روش‌های ممکن
**آمارگیر** : کسی که آمارگیری را انجام می‌دهد.

اگر قرار شد آمارگیر باشیم، می‌توانیم جدولی به صورت زیر تکمیل کنیم.
مثالی از جدول طراحی شده برای ثبت داده‌ها

| اندازهٔ طول قد | خط‌نشان برای شمارش | تعداد دانش‌آموزان |
|---|---|---|
| ۱۴۰> طول قد | | |
| ۱۴۹/۵> طول قد ≤ ۱۳۹/۵ | | |
| ۱۵۹/۵> طول قد ≤ ۱۴۹/۵ | | |
| ۱۶۹/۵> طول قد ≤ ۱۵۹/۵ | | |
| ۱۶۹/۵≤ طول قد | | |

آمارگیری زحمت زیادی برای آمارگیر دارد. آیا راه‌حل ساده‌تری برای انجام آن دارید؟ یکی از مرسوم‌ترین روش‌های آمارگیری، استفاده از پرسش‌نامه است. پرسش‌نامه شبیه همان جدولی است که هنگام ثبت‌نام در مدرسه، آن را تکمیل کرده‌اید. واحدهای جامعه یا نمونه می‌توانند پرسش‌نامه تکمیل کنند.

**مثالی از پرسش‌نامهٔ طراحی شده**

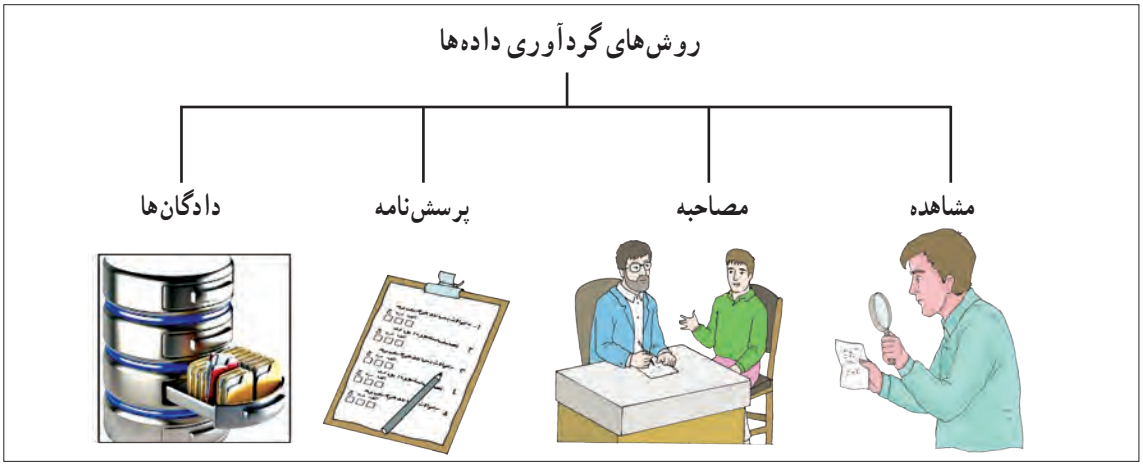
سلام.
می‌خواهیم طول قد دانش‌آموزان مدرسه را آمارگیری کنیم.
لطفاً یکی از گزینه‌ها را انتخاب کنید.
طول قد شما چقدر است؟

| | | |
|---|---|---|
| ۱۴۰> طول قد | [کمتر از ۱۳۹/۵] | ☐ |
| ۱۴۹/۵> طول قد ≤ ۱۳۹/۵ | [۱۳۹/۵ـ۱۴۹/۵] | ☐ |
| ۱۵۹/۵> طول قد ≤ ۱۴۹/۵ | [۱۴۹/۵ـ۱۵۹/۵] | ☐ |
| ۱۶۹/۵> طول قد ≤ ۱۵۹/۵ | [۱۵۹/۵ـ۱۶۹/۵] | ☐ |
| ۱۶۹/۵≤ طول قد | [۱۶۹/۵ و بیشتر] | ☐ |

## کار در کلاس

۱ چه راه دیگری برای آمارگیری طول قد دانش‌آموزان یک مدرسه پیشنهاد می‌کنید؟

۲ فرض کنید زمان لازم را برای گردآوری همهٔ داده‌های دانش‌آموزان در اختیار نداشته باشید. اگر بخواهیم نمونه‌ای را انتخاب و آمارگیری کنیم، چه راهی پیشنهاد می‌کنید که نمونه به صورت تصادفی انتخاب شود؟




۱ **مشاهده**[۱] : گردآوری داده‌ها بدون نیاز به فرد پاسخگو، مانند شمارش تعداد وسایل نقلیهٔ عبوری از یک تقاطع در هر ساعت یا اندازه‌گیری وزن محصولات یک باغ میوه.

۲ **پرسش‌نامه**[۲] : مجموعهٔ سؤالات از پیش تعیین‌شده که توسط تعدادی پاسخگو تکمیل می‌شود. این روش مرسوم‌ترین ابزار گرفتن اطلاعات از مردم است. مرکز آمار ایران هر ۱۰ سال یک‌بار با استفاده از پرسش‌نامه اطلاعات همهٔ خانوارهای ساکن در ایران را گردآوری می‌کند. به این فرایند، سرشماری نفوس و مسکن می‌گوییم.

۳ **مصاحبه**[۳] : معمولاً بین دو نفر صورت می‌گیرد : یکی مصاحبه‌گر (همان آمارگیر) و دیگری مصاحبه‌شونده، یا پاسخگو است. مثلاً اگر بخواهیم دربارهٔ مسائل فرهنگی کاهش شدآمد(ترافیک) پژوهش کنیم، مصاحبه از صاحب‌نظران راه‌حلّ مناسبی برای گردآوری داده‌هاست. از این روش، بیشتر زمانی استفاده می‌شود که آمارگیر از همهٔ پاسخ‌های ممکن اطلاع کافی ندارد.

۴ **دادگان**[۴] : شامل مجموعه‌ای از اطلاعات ذخیره شده است. در بسیاری از موارد، داده‌ها را می‌توان از اطلاعاتی که قبلاً ذخیره شده است، به دست آورد. اگر قرار است تحقیقی در مورد نمره‌های دروس ریاضی استان‌ها انجام شود، اطلاعات ثبتی اداره کل آموزش و پرورش راه‌گشا خواهد بود. از سوی دیگر به دلیل تولید داده‌ها به‌صورت خودکار، در بسیاری از مؤسسات و سامانه‌ها، استفاده از این روش برای گردآوری داده‌ها به سرعت رواج یافته است.

---

۱ــ Observation

۲ــ Questionnaire

۳ــ Interview

۴ــ Database

## کار در کلاس

**الف)** کدام روش برای گردآوری هر یک از داده‌ها مناسب است؟

۱ تعداد قلم‌های هر دانش‌آموز در یک کلاس.

۲ ساعات خواب دانش آموزان کلاس درس شما در شب گذشته.

۳ طول قدّ دانش آموزان یک کلاس.

**ب)** می‌خواهیم طول قدّ دانش‌آموزان یک کلاس یا مدرسه را به یکی از سه روش زیر **آمارگیری** کنیم. هریک از این روش‌ها محدودیت‌هایی دارند. چگونه می‌توان این محدودیت‌ها را از بین برد؟

**پرسش‌نامه** : اگر تعداد واحدهای نمونه زیاد باشد، این روش زمان‌بر است.

**مشاهده** : اگر به‌دقت زیادی نیاز داشته باشیم، مناسب نیست.

**دادگان‌ها** : همیشه اطلاعات ثبتی در اختیار نیست.

## خواندنی

آمارگیری را می‌توان به روش‌هایی بسیار سریع‌تر یا کم‌هزینه‌تر مانند آمارگیری پستی، تلفنی، اینترنتی یا پیامکی انجام داد. همچنین می‌توان با ابزاری نظیر «**گوگل فُرم**» یک پرسش‌نامه طراحی کرد و آن را به نشانی نمونهٔ انتخابی ارسال کرد و نتایج را از «گوگل فُرم» بازیابی کرد.

## فعالیت

قرار است دربارهٔ افرادی که از کوه دنا بالا رفته‌اند، پژوهشی آماری انجام دهیم. واحدهای آماری این پژوهش، همهٔ افرادی هستند که توانسته‌اند به قلّه برسند. هدف از این پژوهش می‌تواند فرهنگی، یا علمی باشد. بسته به نوع پژوهش، یک یا چند ویژگی این افراد (مانند طول قد یا جنسیت) مورد نیاز است. به هر یک از این ویژگی‌ها که مورد پژوهش قرار می‌گیرد، متغیّر می‌گویند. سایر متغیّرها می‌توانند مواردی مانند : سن، وزن، ملیّت، میزان تحصیلات و درآمد باشند. متغیّرهای مورد بررسی در یک پژوهش ممکن است کمّی یا کیفی باشند.

### یادآوری

**متغیّر** : هر ویژگی از اشخاص یا اشیا که قرار است بررسی شود.

**متغیّر کمّی** : متغیّری است که مقادیر عددی می‌گیرد و برای آن عملیات ریاضی از قبیل جمع، تفریق و معدل‌گیری قابل انجام است.

**متغیّر کیفی** : متغیّری است که صرفاً برای دسته‌بندی افراد یا اشیا در گروه‌ها به‌کار می‌رود و لزوماً مقدار عددی نمی‌گیرد.

در مثال کوهنوردان دنا، سن، وزن، قد و درآمد یک کوهنورد، متغیرهای کمّی‌اند. متغیّرهای کیفی معمولاً از نوع مشاهدات غیر عددی‌اند و در مثال کوه‌نوردان دنا ، جنسیت و ملیّت را در بر می‌گیرند. به‌عنوان مثال، جنسیت برای دسته‌بندی افراد به مرد و زن استفاده می‌شود.

**پارامتر یا پارامتر جامعه** : یک مشخصهٔ عددی است که توصیف‌کنندهٔ جنبه‌ای خاص از جامعه است و در صورتی‌که داده‌های کل جامعه در اختیار باشند، قابل محاسبه است.

مثلاً اگر داده‌های مربوط به تک‌تک کوهنوردان را داشته باشیم؛ یعنی به داده‌های جامعه دسترسی داریم. نسبت مردان در کل جامعهٔ کوهنوردان، معرّف یک **پارامتر** است.

اگر داده‌های بعضی از کوهنوردان را داشته باشیم؛ یعنی داده‌های نمونه را در اختیار داریم. نسبت مردانِ کوهنورد به این داده‌های نمونه‌ای را، آماره (مقدار آماره) گویند. آماره ها از نمونه‌ای به نمونهٔ دیگر تغییر می‌کنند؛ این در حالی است که پارامترهای جامعه همیشه ثابت‌اند. چرا؟

در بسیاری از موارد، آمارگیری از کلّ جامعه امکان پذیر نیست. بنابراین، به‌رغم اینکه پارامتر مقدار ثابتی دارد، این مقدار مجهول است و به همین دلیل از آماره‌ها برای تخمین پارامترها استفاده می‌کنند.

**آماره**[1] **یا آمارهٔ نمونه** : مشخصه‌ای عددی که توصیف‌کنندهٔ جنبه‌ای خاص از نمونه است و از داده‌های نمونه به دست می‌آید.

**مثال** : ادارهٔ کشاورزی استان خوزستان در حال ارزیابی هندوانه‌های آمادهٔ برداشت است. در این بررسی، هندوانه‌ها همان واحدهای آماری‌اند. اگر پژوهشگران وزن هندوانه‌ها را مورد بررسی قرار دهند، متغیّر، «وزن» آنهاست. وزن یک متغیّر کمّی است؛ زیرا با مقادیر عددی ارائه می‌شود. اگر وزن تک‌تک هندوانه‌های این زمین بررسی شود، سرشماری از جامعه انجام داده‌ایم (که امکان‌پذیر نیست). متوسط وزنِ همهٔ هندوانه‌های قابل برداشت در این زمین، «پارامتر» است.

حال فرض کنیم پژوهشگران تصمیم دارند بر اساس معیار «مزه» هندوانه‌ها را مورد بررسی قرار دهند. در این حالت، مزهٔ هندوانه‌ها را می‌توان به سه دسته تقسیم کرد : بد، قابل قبول و خوب. حال که می‌خواهیم مزهٔ هندوانه‌ها را امتحان کنیم، مطالعه به بخشی از کلّ هندوانه‌ها محدود می‌شود. در اینجا «مزه» متغیری کیفی است؛ زیرا نمی‌توانیم همهٔ هندوانه‌ها را مزه مزه کنیم، فقط بخشی از هندوانه‌ها مورد مطالعه قرار می‌گیرند؛ پس باید «نمونه» بگیریم. نسبت تعداد هندوانه‌های دارای مزهٔ «خوب» در نمونه، یک «آماره» است.

فرایند نتیجه‌گیری دربارهٔ پارامترهای جامعه بر اساس نمونه، **آمار استنباطی**[2] است.

---

۱ــ Statistic

۲ــ Inferential Statistics

## خواندنی

### روش‌های تولید ارقام (اعداد) تصادفی

سال‌ها پیش، اعداد تصادفی به‌صورت دستی و پس از آن به کمک الگوریتم‌های ریاضی ارقام شبه تصادفی تولید شده است و در کتاب‌هایی با همین عنوان، در اختیار محققان قرار می‌گرفت. امروزه ماشین‌حساب‌های علمی و نرم‌افزارهای رایانه‌ای این اعداد را در کسری از ثانیه تولید می‌کنند. از رقم‌های عدد پی نیز می‌توان به‌عنوان ارقام و اعداد تصادفی طبیعی استفاده کرد. خروجی تمامی روش‌های یادشده به‌صورت زیر است :

۴۶۹۱۴۳۹۳۲ ۸۲۶۵۸۰۲۵۴ ۶۴۸۱۶۲۹۹۷ ۶۱۱۵۷۳۱۳۹ ۹۴۱۷۳۶۷۴۳ ۲۳۶۲۰۳۶۴۷ ۳۴۳۴۲۶۹۴۹

## تمرین

۱ در نمونه‌گیری تصادفی ساده، احتمال اینکه فرد به‌خصوصی در اوّلین انتخاب عضو نمونه باشد، چقدر است؟ اگر مسئله با جای‌گذاری باشد، احتمال اینکه او در دومین انتخاب عضو نمونه باشد، چقدراست؟ اگر مسئله بدون جای‌گذاری باشد، و از نتیجه انتخاب اوّل اطلاع نداشته باشیم، احتمال اینکه او در دومین انتخاب عضو نمونه باشد، چقدر است؟

۲ آیا در نمونه‌گیری خوشه‌ای احتمال انتخاب واحدهای آماری برابر است؟ چرا؟ احتمال انتخاب خوشه‌ها چطور؟ آیا این روش نمونه‌گیری احتمالی است؟

۳ روش‌های نمونه‌گیری احتمالی چه مزیتی بر نمونه‌گیری‌های غیر احتمالی دارند؟

۴ برای هریک از روش‌های نمونه‌گیری احتمالی دو مثال واقعی بیاورید.

۵ اگر اندازهٔ جامعه بزرگ باشد، نمونه‌گیری با جای‌گذاری و بدون جای‌گذاری تقریباً مثل هم هستند. در این صورت، آیا می‌توانید راه حل کلی برای انتخاب تصادفی $n$ نمونه از یک فهرست $N$ تایی ارائه کنید؟

۶ آیا احتمال انتخاب واحدهای آماری در نمونه‌گیری طبقه‌ای برابر است؟ در هر طبقه چطور؟

۷ فرق بین داده و متغیّر چیست؟

۸ فرق بین آماره با پارامتر چیست؟

۹ در یک جامعهٔ آماری، آیا ممکن است که یک پارامتر تغییر کند؟ اگر سه نمونه با اندازهٔ یکسان از یک جامعه داشته باشیم، می‌توان سه مقدار متفاوت از یک آماره به دست آورد؟

۱۰ در یک مطالعه از ۱۲۶۱ مشتری غذاهای گیاهی، سؤال شده است که برای کدام وعده غذایی (ناهار یا شام) سفارش داده‌اند؟

الف) متغیّر را مشخص کنید. این متغیّر کمّی است یا کیفی؟

ب) کدام روش گردآوری داده‌ها برای مطالعه مناسب است؟

پ) جامعهٔ آماری در اینجا چیست؟ در این مطالعه پارامتر و آماره چه چیزی می‌توانند باشند؟

**۱۱** کدام روش گردآوری داده‌ها برای موارد زیر مناسب است؟ یک دلیل برای انتخاب خود ذکر کنید.

- میزان رضایت مشتریان بانک از نحوهٔ برخورد و رسیدگی به درخواست‌های آنها.
- سنّ همهٔ دانش‌آموزان مدرسه بر حسب ماه در پایهٔ دهم.
- تعداد سرنشینان خودروهای سواری در یکی از محورهای خروجی شهر.

**۱۲** فرض کنید جامعه‌ای از $N=۱۰۰$ عضو تشکیل‌شده و می‌خواهیم نمونه‌ای به‌اندازهٔ $n=۲۰$ از آن انتخاب کنیم. در هر یک از حالت‌های زیر احتمال انتخاب هر عضو جامعه به‌عنوان نمونه چقدر است؟ نام هر روش نمونه‌گیری را بگویید.

الف) اگر جامعه به دو قسمت ۵۰ تایی تقسیم شود و بخواهیم از هر قسمت نمونهٔ تصادفی ۱۰ تایی انتخاب کنیم.

ب) اگر جامعه به‌تصادف به ۱۰ قسمت مساوی تقسیم شود و دو قسمت را به‌عنوان نمونه انتخاب کنیم.

پ) اگر جامعه به‌تصادف به ۲۰ قسمت مساوی تقسیم شود، و از قسمت اوّل یک عضو به‌تصادف انتخاب شود. فرض کنید عضو انتخابی دومین عضو باشد و از قسمت‌های بعدی نیز دومین عضو انتخاب شود.

**۱۳** دلایل اریبی در نمونه‌گیری‌های زیر را ذکر کنید. کدام روش گردآوری داده‌ها برای آنها مناسب‌تر است؟

الف) نمونه‌گیری راحت : افراد در دسترس را به‌عنوان نمونه انتخاب می‌کنیم.

ب) نمونهٔ غیر تصادفی : عامل شانس در انتخاب نمونه نقشی ندارد.

پ) نمونه‌گیری ایمیلی (رایانامه‌ای) : پرسش‌نامه‌ای به ایمیل‌های انتخاب شده ارسال می‌شود.

ت) نمونه‌گیری تلفنی : از دفترچهٔ راهنمای تلفن تعدادی شماره به‌تصادف انتخاب می‌شود.

ث) برخی از اعضای انتخاب شده در نمونه حاضر به پاسخگویی نمی‌باشند.

**۱۴** نوع روش نمونه‌گیری مناسب‌تر را انتخاب کنید.

الف) شرکت واردکنندهٔ خودروهای سنگین برای بررسی عملکرد سامانه ترمز آنها می‌خواهد ده درصد از خودروهایی را که به‌مرور زمان وارد کشور می‌شوند بازرسی کند.

ب) مدیر مدرسهٔ ۶۰۰ نفری می‌خواهد نظر دانش‌آموزان را برای تغییر ساعت تعطیلی مدرسه براساس یک نمونهٔ ۱۲ تایی بداند.

پ) در قسمت قبل اگر مدرسه، شش پایه داشته باشد و ما حدس بزنیم که نظر ۶ پایه باهم تفاوت دارد (با فرض برابر بودن تعداد دانش آموزان در پایه‌های مختلف).

**۱۵** چگونه از ارقام ۰ تا ۹ عدد تصادفی انتخاب می‌کنید؟ آیا با روش پیشنهادی شما می‌توان عدد تصادفی بین اعداد ۰ تا ۹۹ انتخاب کرد؟ آیا امکان توسعهٔ روش پیشنهادی شما به انتخاب تصادفی از فهرستی ۱۰۰۰ تایی امکان‌پذیر است؟

# درس ۲ برآورد

## فعالیت

یک شرکت تولید لیوان شیشه‌ای می‌خواهد تعداد لیوان‌هایی را که در یک بسته قرار می‌دهد مشخص کند. تعداد لیوان‌ها در هر بسته به میانگین تعداد اعضای خانوارهای کشور بستگی دارد که بُعد خانوار نام دارد. مثلاً در ۷ سال پیش بُعد خانوار (میانگین تعداد اعضای خانواده‌ها) ۴ بوده است. لذا بسته‌بندی لیوان‌ها از ۶ به ۴ کاهش داده شد. از آنجا که فروش شرکت کم شده، به نظر کارشناسان، دلیل آن تغییر بُعد خانوار در کشور است. بُعد خانوار هر کشور از اطلاعات سرشماری قابل دسترسی است که ۷ سال پیش انجام شده است. سرشماری یکی از مهم‌ترین طرح‌های آمارگیری در هر کشوری است، که در ایران هر ۱۰ سال یک بار انجام می‌شود، لذا داده‌های جدید آن تا ۳ سال آینده در دسترس نیست. از آنجا که سرشماری روش مقرون به‌صرفه‌ای برای گردآوری داده‌ها به‌منظور پاسخگویی به این سؤال نیست، شرکت تصمیم می‌گیرد که بُعد خانوار خریدارهای محصول این شرکت را به‌وسیلهٔ نمونه‌گیری انجام دهد.

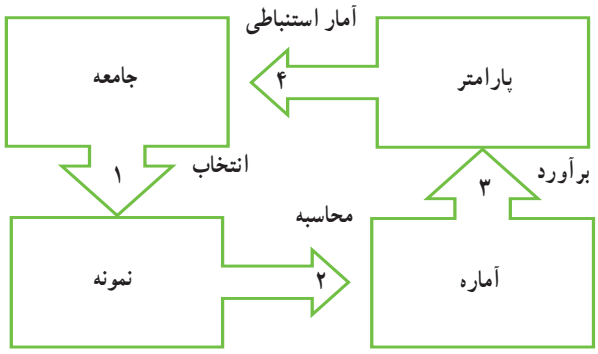
در اینجا صورت ساده‌تر آن را در نظر می‌گیریم. فرض کنید، بُعد خانوار ۹ خریدار محصول به‌صورت زیر باشد. میانگین بُعد این نمونه چقدر است؟

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| ۳ | ۲ | ۷ | ۲ | ۵ | ۳ | ۳ | ۱ | ۴ |

**برآورد نقطه‌ای[۱] پارامتر جامعه** برابر است با مقدار عددی حاصل از جای‌گذاری اعداد نمونه تصادفی در آمارهٔ نظیر آن پارامتر. به بیان دیگر مقدار عددی آماره را برآورد یا برآورد نقطه‌ای می‌نامند.

در این فعالیت میانگین تعداد اعضای خانوار پارامتر است. آماره، .....؛ و برآورد نقطه‌ای پارامتر ..... است.

۱ــ Point Estimation




## کار در کلاس

فرض کنید، جامعه از ۶ نفر تشکیل شده باشد با درآمد ماهیانه برحسب میلیون تومان به‌صورت زیر :

| | | | | | |
|---|---|---|---|---|---|
| ۲ | ۵ | ۳ | ۰ | ۱ | ۴ |

می‌خواهیم بر اساس نمونه‌ای به‌اندازهٔ ۱، **میانگین** این جامعهٔ ۶ عضوی را برآورد کنیم. در واقع باید از بین ۶ نفر، یکی را به‌تصادف انتخاب کنیم. اگر شخصی انتخاب شود که درآمدش ۵ باشد، این عدد برآورد میانگین درآمد همهٔ افراد است. ممکن است فرد انتخابی درآمدی نداشته باشد. آن‌گاه صفر به‌عنوان نمونهٔ انتخاب شده و برآورد میانگین درآمد این افراد برابر ۰ می‌شود. نمونه‌های مختلف منجر به برآوردهای متفاوتی می‌شوند.

- در این مثال، پارامتر جامعه چیست و مقدار آن چقدر است؟
- آیا بر اساس هر یک از نمونه‌ها برآورد به مقدار پارامتر نزدیک است؟
- چه راه‌حلی پیشنهاد می‌کنید که برآورد به پارامتر نزدیک‌تر شود؟

درست حدس زده‌اید! اگر اندازهٔ نمونه را بیشتر کنیم امکان نزدیک شدن برآورد به پارامتر بیشتر می‌شود. اندازهٔ نمونه را به ۲ افزایش می‌دهیم. به‌عنوان مثال، اگر نمونه‌گیری تصادفی انجام شده شامل درآمدهای ۰ و ۴ باشد، آن‌گاه برآورد میانگین جامعه عدد ۲ است؛ یعنی پارامتر جامعه که مقدار آن ۲/۵ بوده است را ۲ برآورد کرده‌ایم.

- آیا نمونه‌ای تصادفی به‌اندازهٔ ۲ وجود دارد که مقدار پارامتر را دقیقاً ۲/۵ برآورد کند؟
- آیا امکان دارد با نمونه‌های مختلف برآوردهای برابر به دست آوریم؟
- بدون شمارش بگویید امکان مشاهده چند نمونهٔ دوتایی داریم؟

در جدول زیر، احتمال مشاهده هریک از مقادیر برآورد میانگین برای نمونه‌های دوتایی آمده است.

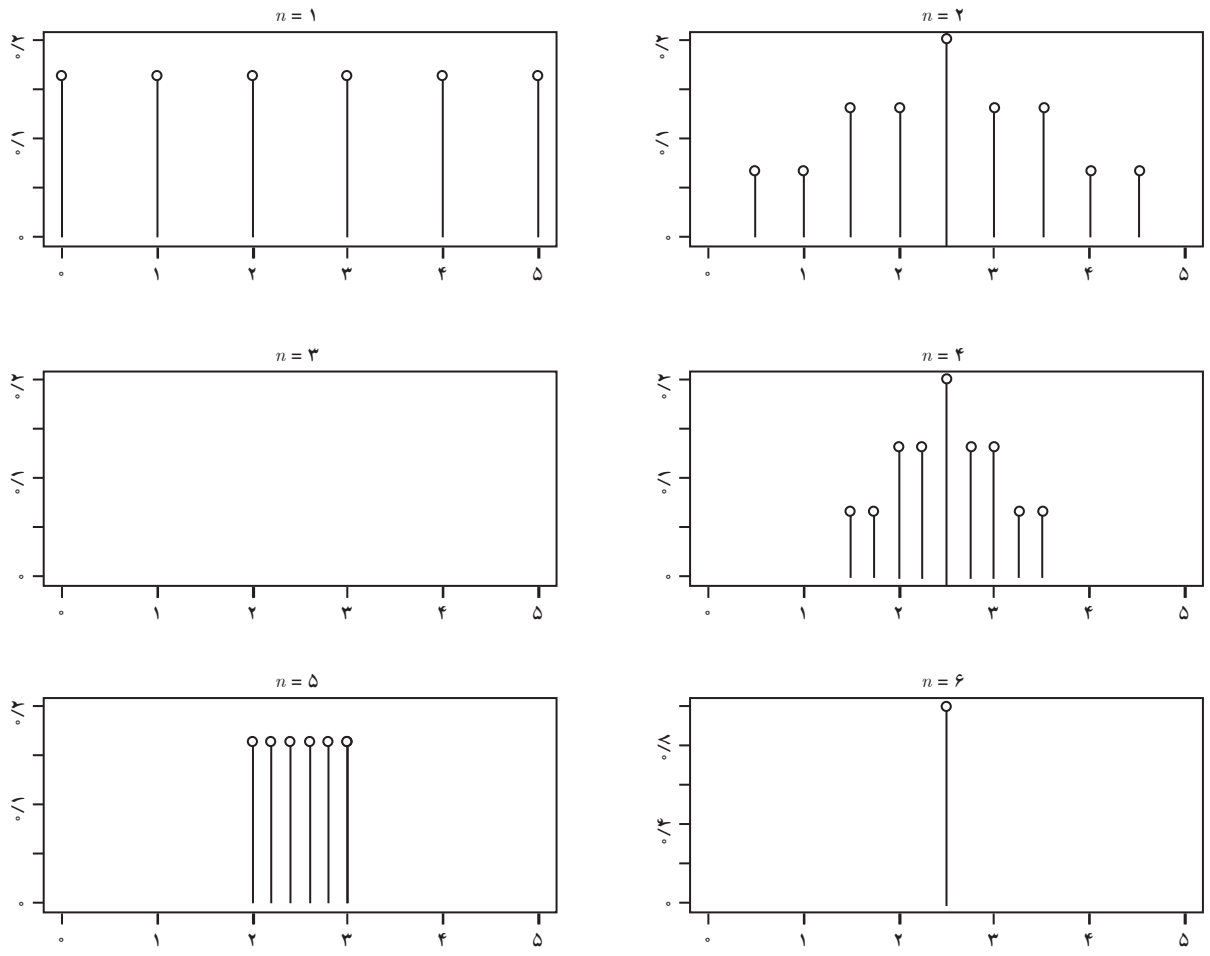
| نمونه | {۰,۱} | {۰,۲} | {۰,۳}{۱,۲} | {۰,۴}{۱,۳} | {۰,۵}{۱,۴}{۲,۳} | {۱,۵}{۲,۴} | {۲,۵}{۳,۴} | {۳,۵} | {۴,۵} |
|---|---|---|---|---|---|---|---|---|---|
| $\overline{x}$ | ۰/۵ | ۱ | ۱/۵ | ۲ | ۲/۵ | ۳ | ۳/۵ | ۴ | ۴/۵ |
| احتمال | $\frac{۱}{۱۵}$ | $\frac{۱}{۱۵}$ | $\frac{۲}{۱۵}$ | $\frac{۲}{۱۵}$ | $\frac{۳}{۱۵}$ | $\frac{۲}{۱۵}$ | $\frac{۲}{۱۵}$ | $\frac{۱}{۱۵}$ | $\frac{۱}{۱۵}$ |

اگر نمونه‌گیری تصادفی ساده به‌اندازهٔ $n = ۳$ از این ۶ عضو جامعه انجام دهیم، همانند جدول قبل مقادیر $\overline{x}$ و احتمال مشاهده هر مقدار را محاسبه و در جدول بنویسید.

| نمونه | $\{۰,۱,۲\}$ | $\{۰,۱,۳\}$ | $\{۰,۱,۴\}$ | $\{۰,۱,۵\}$<br>$\{۰,۲,۴\}$<br>$\{۱,۲,۳\}$ | $\{۰,۲,۵\}$<br>$\{۰,۳,۴\}$<br>$\{۱,۲,۴\}$ | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| $\overline{x}$ | $\frac{۳}{۳}$ | $\frac{۴}{۳}$ | $\frac{۵}{۳}$ | $\frac{۶}{۳}$ | | | | | | |
| احتمال | | | | | | | | | | |

## فعالیت

جدول به‌دست آمده از کار در کلاس قبل را برای $n = ۳$ رسم کنید. برای این منظور، بر روی محور طول‌ها مقادیر برآورد میانگین جامعه، یعنی $\overline{x}$ را مشخص کنید. حال احتمال مشاهده هریک از مقادیر را در نمودار علامت بزنید. این کار برای اندازهٔ نمونه‌های مختلف انجام شده است. هر نمودار مربوط به‌اندازه نمونهٔ به‌خصوص، $n = ۱,۲,۳,۴,۵,۶$ است.



اگر برآورد را بر اساس نمونه‌ای به‌اندازهٔ ۳ محاسبه کنیم، احتمال اینکه برآورد به پارامتر نزدیک‌تر باشد، نسبت به $n = ۱,۲$ بیشتر است. آیا اگر اندازهٔ نمونه از ۳بیشتر شود، احتمال اینکه برآورد به پارامتر نزدیک‌تر شود، باز هم بیشتر می‌شود؟ زمانی که اندازهٔ نمونه به ۶ می‌رسد، برآورد برابر ..... می‌شود.

همان‌طور که در نمودارها دیده‌اید با افزایش اندازهٔ نمونه برآوردها به میانگین جامعه، که پارامتر است، نزدیک‌تر می‌شوند. به بیان دیگر در هر نمودار با زیاد شدن اندازهٔ نمونهٔ انحراف معیار برآوردهای پارامتر کمتر می‌شود. پس هرقدر انحراف معیار برآورد کمتر باشد، آن برآورد بهتر است.

سؤال اساسی آن است که انحراف معیارِ برآورد میانگین جامعه چقدر است؟ خوشبختانه آمارشناسان پاسخ این سؤال را با رابطهٔ زیر داده‌اند. البته برای سادگی محاسبات، فرض شده که جامعه نامتناهی است.

$$\sigma_{\bar{x}} = \sigma / \sqrt{n}$$

انحراف معیار جامعه تقسیم بر جذر اندازهٔ نمونه = انحراف معیار میانگین

هرچند که انحراف معیار جامعه معمولاً معلوم نیست، ولی این رابطه حدس ما را اثبات کرده است. با افزایش اندازهٔ نمونهٔ انحراف معیار برآورد کاهش می‌یابد. به عبارتی دیگر برآورد دقیق‌تر یا خطای کمتری برای برآورد میانگین جامعه داریم.

## کار در کلاس

به فعالیت ابتدای درس باز می‌گردیم. اگر از مطالعات سال‌های گذشته بدانیم که انحراف معیار درآمد هر فرد در کشور ۲میلیون تومان است انحراف معیار برآورد میانگین درآمد افراد جامعه را برای اندازهٔ نمونه‌های ذکر شده محاسبه کنید.

| $n$ | ۲۵ | ۱۰۰ | ۱۰۰۰۰ |
|---|---|---|---|
| $\sigma_{\bar{x}}$ | | | |

- انحراف معیار برآورد میانگین درآمد افراد جامعه با نمونهٔ ۱۰۰ نفری چند برابر انحراف معیار با نمونهٔ ۱۰۰۰۰ نفری است؟
- اگر اندازهٔ نمونه ۱۰ برابر شود، انحراف معیار برآورد میانگین چند برابر می‌شود؟

## برآورد بازه‌ای

اگر بعد از یک آزمون ساده از شما سؤال شود نمرهٔ شما چند می‌شود، حتماً بدون تردید نمره‌ای که انتظار آن را دارید می‌گویید. این یک برآورد ذهنی است (که مرتبط به درس فعلی ما نمی‌شود، ولی برای درک ادامهٔ درس مفید است). حال فرض کنید آزمون ساده نبوده و به صحیح بودن برخی پاسخ‌های خود شک دارید. بازهم به‌صورت ذهنی پاسخ می‌دهید، ولی پاسخ شما با شک و تردید همراه است. معمولاً ترجیح می‌دهید به‌جای ذکر یک نمره، بازه‌ای که برای نمرهٔ خود به‌صورت ذهنی ترسیم کرده‌اید بیان کنید. به خاطر اینکه اطمینان خود را نیز از بازه ذکر شده بیان کنید، به ذکر یک درصد اطمینان اکتفا می‌کنید. مثلاً می‌گویید نمرهٔ من بین ۱۶ تا ۱۹ است، با اطمینان ۹۰ درصد. هرچه فاصلهٔ دو عدد بازه کمتر باشد و درصد اطمینان ذکر شده بیشتر، برآورد دقیق‌تر است.

**برآورد بازه‌ای یا بازه اطمینان**[1] **پارامتر جامعه** عبارت است از بازه‌ای عددی برای پارامتر به همراه یک درصد اطمینان که به **ضریب اطمینان**[2] شهرت دارد.

## فعالیت

در فعالیت قبل میانگین داده‌ها ۲/۵ محاسبه می‌شود؛ یعنی برآورد میانگین جامعه ۲/۵ به دست آمده است. چقدر به این برآورد اطمینان داریم؟ برای یافتن پاسخ سؤال به یاد آورید که دقت برآورد میانگین جامعه به ..... و ..... بستگی داشت. اگر ..... زیاد می‌شد، یا ..... کم بود، دقت برآورد میانگین بیشتر می‌گردید. بر اساس این دو کمیت پاسخ این سؤال را با رابطهٔ زیر داده‌اند.

**برآورد بازه‌ای برای میانگین جامعه** : اگر نمونه‌ای تصادفی به‌اندازهٔ $n$ در اختیار داشته باشیم، با اطمینان بیش از ۹۵٪ می‌توانیم بگوییم :

$$\bar{x} - ۲\sigma / \sqrt{n} \le \mu \le \bar{x} + ۲\sigma / \sqrt{n}$$

که $\mu$ میانگین جامعه و $\sigma$ انحراف معیار جامعه است.

اگر یک نمونه به اندازهٔ چهار داشته باشیم یک فاصله اطمینان برای میانگین جامعه محاسبه کنید.

مشاهدات ۵,۲,۱,۰

میانگین نمونه $\bar{x} = ....$

انحراف معیار نمونه $\sigma = ۱/۸۷$

$..... < \mu < .....$

## کار در کلاس

خط فقر حداقل درآمدی است که برای زندگی در یک ماه به ازای هر نفر موردنیاز است. خط فقر برابر است با نصف میانگین درآمد افراد جامعه. بر اساس داده‌های فعالیت اوّل خط فقر را برآورد کنید. انحراف معیار جامعه را برآورد کنید. اگر فرض کنیم که انحراف معیار به دست آمده انحراف معیار جامعه است، یک برآورد فاصله‌ای برای خط فقر محاسبه کنید.

## خواندنی

از جمله ساده‌ترین مسائل در آمار استنباطی، برآورد کردن یک نسبت در جامعه است :

- چند درصد از مردم تصمیم دارند در تعطیلات نوروز سال آینده به مسافرت بروند؟
- چند درصد از کشاورزان در سال گذشته از آبیاری قطره‌ای در مزارع خود استفاده کردند؟
- چند درصد از درختان یک باغ آلوده به شته شده‌اند؟
- چند درصد مشتریان یک فروشگاه از برخورد کارمندان راضی هستند؟

---

۱ــ Confidence Interval

۲ــ Confidence Coefficient

فرض کنید برای به‌دست آوردن نسبتی از اعضای جامعه که ویژگی خاصی را دارند، نمونه‌ای $n$ تایی را به تصادف از جامعه انتخاب می‌کنیم. اگر $m$ تا از نمونه‌ها ویژگی مورد مطالعه را داشته باشند، آن‌گاه می‌توانیم بگوییم:

«تقریباً به نسبت $\frac{m}{n}$ از افراد جامعه آن ویژگی را دارند.»

ولی این گزاره چه معنایی می‌دهد؟

در این تقریب، مقدار خطا چقدر است؟ آمارشناسان در پاسخ به سؤال فرمول زیر را ارائه کرده‌اند:

**برآورد بازه‌ای نسبت، با اطمینان ۹۵ درصد:**

اگر از جامعه‌ای نسبتاً بزرگ، $n$ نمونهٔ تصادفی ساده انتخاب کنیم و $m$ تا از آنها ویژگی مورد مطالعه ما را داشته باشند، آن‌گاه نسبت واقعی افرادی از جامعه (پارامتر نسبت) که آن ویژگی را دارند، با اطمینان ۹۵ درصد، در بازهٔ زیر است:

$$p-۲\sqrt{\frac{p(۱-p)}{n}} \le \text{نسبت واقعی} \le p+۲\sqrt{\frac{p(۱-p)}{n}}$$

که در آن $p=\frac{m}{n}$ برآورد پارامتر نسبت ویژگی در جامعه است.

توجه داشته باشید که اندازهٔ بازه اطمینان وابسته به تعداد نمونه، یعنی عدد $n$، است، ولی ارتباطی با تعداد اعضای جامعه ندارد. اگر بخواهیم بدون تغییر تعداد نمونه با اطمینانی بیش از ۹۵ درصد نسبت را برآورد کنیم، لازم است که به‌جای عدد ۲ ضریبی بزرگ‌تر گذاشته شود.

**مثال:** فرض کنید از ۴۸ دانش‌آموز مدرسه پرسیده‌ایم که «آیا برای آمدن به مدرسه از وسایل نقلیهٔ عمومی استفاده می‌کنید؟» ۳۶ نفر به سؤال ما جواب مثبت داده‌اند. در این صورت، چند درصد از دانش‌آموزان مدرسه جوابشان به این سؤال مثبت خواهد بود؟ پاسخی با اطمینان ۹۵ درصد مدّ نظر است.

**حل:** طبق فرض‌های مسئله تعداد نمونه ۴۸ است و ۳۶ نفر ویژگی مورد نظر را داشته‌اند. پس قرار می‌دهیم:

$n=۴۸\ ,\ m=۳۶$

پس $p=\frac{۳۶}{۴۸}=۰/۷۵$. طبق فرمول اخیر باید عبارت زیر را هم محاسبه کنیم:

$$۲\sqrt{\frac{p(۱-p)}{n}}=۲\sqrt{\frac{۰/۷۵(۱-۰/۷۵)}{۴۸}}=۲\sqrt{\frac{۱}{۴۸}\times\frac{۳}{۴}\times\frac{۱}{۴}}=۲\times\frac{۱}{۱۶}=۰/۱۲۵$$

پس با اطمینان ۹۵ درصد می‌توان گفت، نسبت مورد نظر بین ۶۳ درصد و ۸۸ درصد است. طول این بازه ۲۵ درصد است. اگر بخواهیم به‌دقت بهتری برسیم باید تعداد داده‌ها را زیاد کنیم.

با توجه به اینکه در فرمول بازه اطمینان جمله $\sqrt{n}$ در مخرج آمده است، به حقیقت زیر می‌رسیم:

اگر در محاسبهٔ بازه اطمینان نسبت تعداد نمونه‌ها را $k$ برابر کنیم، طول بازهٔ اطمینان تقسیم بر $\sqrt{k}$ می‌شود. پس اگر تعداد نمونه‌ها را ۴ برابر کنیم طول بازهٔ اطمینان نصف می‌شود و اگر تعداد نمونه‌ها را صد برابر کنیم، دقت محاسبه نسبت، یک رقم اعشار بهتر خواهد شد.

یک مؤسسه نظرسنجی معتبر، یک روز قبل از برگزاری انتخابات ریاست‌جمهوری، از یک نمونهٔ ۱۰۰۰ نفری از واجدان شرایط، که به‌طور تصادفی از کلّ کشور انتخاب شده‌اند، پرسیده است که «آیا در انتخابات شرکت خواهید کرد؟»

اگر جواب ۷۰۰ نفر مثبت بوده باشد، یک بازهٔ اطمینان ۹۵ درصدی برای درصد شرکت‌کنندگان در انتخابات به دست آورید.

**حل :** در این مسئله $n=۱۰۰۰$ و $p=\frac{۷۰۰}{۱۰۰۰}=\ldots$ . باید عبارت زیر را نیز محاسبه کنیم :

$$۲\sqrt{\frac{p(۱-p)}{n}}=۲\sqrt{\frac{\cdots(۱-\cdots)}{\cdots}}=۰/۰۲۹$$

پس با اطمینان ۹۵ درصد مشارکت در انتخابات بین ..... درصد و ..... درصد خواهد بود.

**مثال :** همان‌طور که می‌دانید در انتخابات ریاست‌جمهوری اگر نامزدی موفق به کسب بیش از نیمی از آرای رأی‌دهندگان شود، رئیس‌جمهور خواهد شد، ولی اگر هیچ نامزدی این میزان رأی نیاورد دو نامزدی که بیشترین تعداد رأی را دارند به دور دوم خواهند رفت. در برخی سال‌ها رقابت بین نامزدها به شکلی شدید است که از پیش به‌سختی می‌توان حدس زد که نامزد پیشرو آیا در دور اوّل انتخاب خواهد شد یا خیر. در این صورت، نظرسنجی‌ها باید بسیار دقیق‌تر باشد. فرض کنید از پیش بدانیم که آرای یکی از نامزدها نزدیک به ۵۰ درصد است. اگر بخواهیم طول بازهٔ اطمینان ۹۵ درصدی، کمتر از یک درصد باشد نمونهٔ ما باید شامل چند نفر باشد؟

**حل :** طول بازهٔ اطمینان برابر است با :

$$۲\times۲\sqrt{\frac{p(۱-p)}{n}}$$

و با توجه به اینکه می‌دانیم $p\cong ۰/۵$ پس طول بازهٔ اطمینان تقریباً برابر است با :

$$۲\times۲\sqrt{\frac{۰/۵(۱-۰/۵)}{n}}=\frac{۲}{\sqrt{n}}$$

پس اگر بخواهیم طول بازه اطمینان ۹۵ درصدی، برابر یک درصد باشد باید $n$ را آن‌قدر بزرگ بگیریم که $\frac{۲}{\sqrt{n}}$ از ۰/۰۱ کمتر شود؛ یعنی $\frac{۲}{\sqrt{n}}\cong ۰/۰۱$ و این یعنی $n=۲۰۰^۲=۴۰۰۰۰$ .

در حالت کلی، قبل از آمارگیری برآوردی از نسبت مورد مطالعه نداریم. پس اگر بخواهیم به‌دقت خاصی برسیم، چگونه باید از پیش بفهمیم که به چند نمونه احتیاج داریم؟ برای پاسخ به این سؤال به این نکته توجه کنید که عبارت $p(۱-p)$ یک چندجمله‌ای درجه دو است و بیشترین مقدار خود را در $p=۰/۵$ می‌گیرد؛ یعنی $p(۱-p)\le ۰/۵\,(۱-۰/۵)=۰/۵^۲$

لذا به این نتیجه می‌رسیم :

در برآورد بازه‌ای نسبت، بازهٔ زیر شامل بازهٔ اطمینان ۹۵ درصدی است :

$$p-\frac{۱}{\sqrt{n}}\le \text{پارامتر نسبت}\le p+\frac{۱}{\sqrt{n}}$$

که در آن $p=\frac{m}{n}$ . به بیان دیگر «با اطمینان بیشتر از ۹۵ درصد»، نسبت مورد نظر در این بازه است.

## تمرین

۱ در اوّلین کار در کلاس، جداول را برای نمونه‌گیری تصادفی ساده به‌اندازهٔ ۴ و ۵ تشکیل داده و مقادیر $\bar{x}$ را در مقابل احتمال مشاهدهٔ هر مقدار، محاسبه و در جدولی بنویسید.

۲ از اعداد ۰ تا $N$، ۱۰ عدد به‌تصادف انتخاب شده است. اگر اعداد انتخابی به‌صورت زیر باشند با دو روش مختلف $N$ را برآورد کنید.

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| ۹ | ۲ | ۵ | ۷ | ۳ | ۱۲ | ۱۱ | ۹ | ۸ | ۵ |

۳ رئیس یک دانشگاه علاقه‌مند است متوسط سنّ دانشجویانی که در سال جاری ثبت نام کرده‌اند را بداند. برای این منظور، او یک نمونهٔ تصادفی از سنّ ۲۵ دانشجو را انتخاب می‌کند. میانگین سنّ آنها برابر ۲۲ سال برآورد شده است. اگر در بررسی‌های گذشته انحراف معیار طول قد دانشجویان این دانشگاه برابر ۱/۹ سال باشد، بازهٔ اطمینان ۹۵ درصد برای میانگین سنّ جامعه را محاسبه کنید.

۴ طول فاصلهٔ اطمینان، برابر تفاضل حدّ بالا و پایین بازهٔ اطمینان است.

الف) اگر در فرمول بازه اطمینان اندازهٔ نمونه افزایش یابد، طول فاصلهٔ اطمینان ..... می‌یابد. چرا؟

ب) اگر در فرمول بازه اطمینان انحراف معیار جامعه افزایش یابد، طول فاصله اطمینان ..... می‌یابد. چرا؟

۵ داده‌های زیر نمرات ۲۴ دانش‌آموز از ۱۰۰ است.

۷۵ ۷۴ ۷۳ ۷۱ ۷۱ ۷۰ ۶۷ ۷۵

۷۹ ۷۸ ۷۸ ۷۸ ۷۸ ۷۷ ۷۵ ۸۰

۸۷ ۸۶ ۸۶ ۸۳ ۸۲ ۸۲ ۸۱ ۹۱

الف) میانگین و انحراف معیار نمرات را محاسبه کنید.

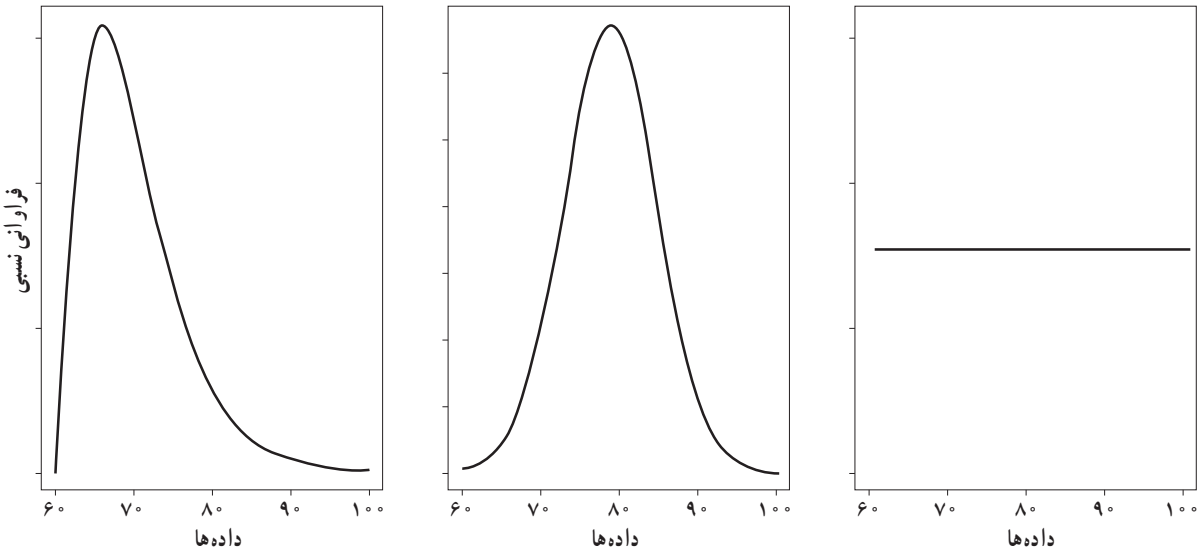
ب) اگر انحراف معیار جامعه ۶ باشد بازه اطمینان ۹۵ درصد برای میانگین نمرات جامعه محاسبه کنید.

پ) چند درصد داده‌ها داخل این بازه قرار می‌گیرند؟

ت) بافت نگاشت فراوانی داده‌ها را رسم کنید. (در فواصل [۶۷,۷۱] و [۷۱,۷۵] و ...)

ث) چندبر فراوانی بافت نگاشت را رسم کنید (وسط مستطیل‌ها را با پاره‌خط به هم متصل کرده و به محور طول‌ها وصل کنید).

ج) اگر داده‌ها زیاد شوند، به نظر شما شکل چندبر فراوانی بافت نگاشت به کدام‌یک از نمودارهای صفحهٔ بعد شباهت بیشتری خواهد داشت؟ (نام نمودارها به ترتیب : یکنواخت، نرمال، نامتقارن یا چوله است)



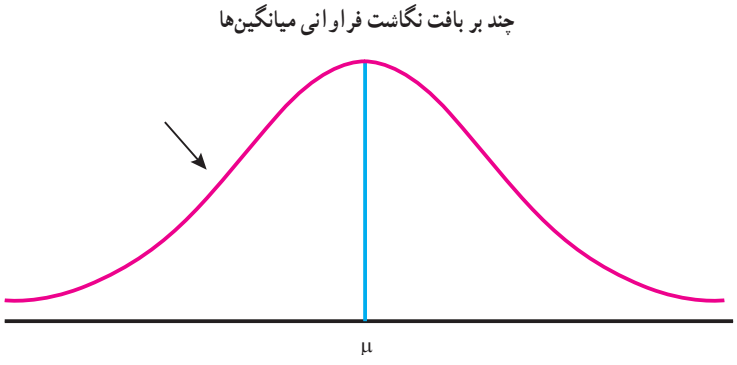

**۶** اگر در سؤال قبل ۱۰۰ بار نمونه‌گیری را تکرار کنیم؛ یعنی ۱۰۰ دفعه نمونه‌ای به‌اندازهٔ ۲۴ بگیریم و چند بر فراوانی بافت نگاشت ۱۰۰ میانگین را رسم کنیم می‌توان نشان داد که تقریباً به صورت یک منحنی به شکل زیر است (توجه کنید منظور از ۱۰۰ عددی بزرگ است، ۱۰۰ یک مثال است). در این شکل $\mu$ نشان دهندهٔ میانگین جامعه است، که در اینجا میانگین نمرات همهٔ دانش‌آموزان است، که مجهول است. حال فرض کنید برای ۱۰۰ نمونهٔ ۲۴ تایی، ۱۰۰ بازهٔ اطمینان ۹۵ درصدی محاسبه کرده‌ایم. در زیر نمودار نرمال ۲۰ تا از آنها رسم شده است. نقاط قرمز رنگ نشان‌دهندهٔ میانگین نمونه و پاره‌خط‌های افقی آبی معرف فاصلهٔ اطمینان مربوطه‌اند. خط سیاه عمودی میانگین جامعه را مشخص کرده است.

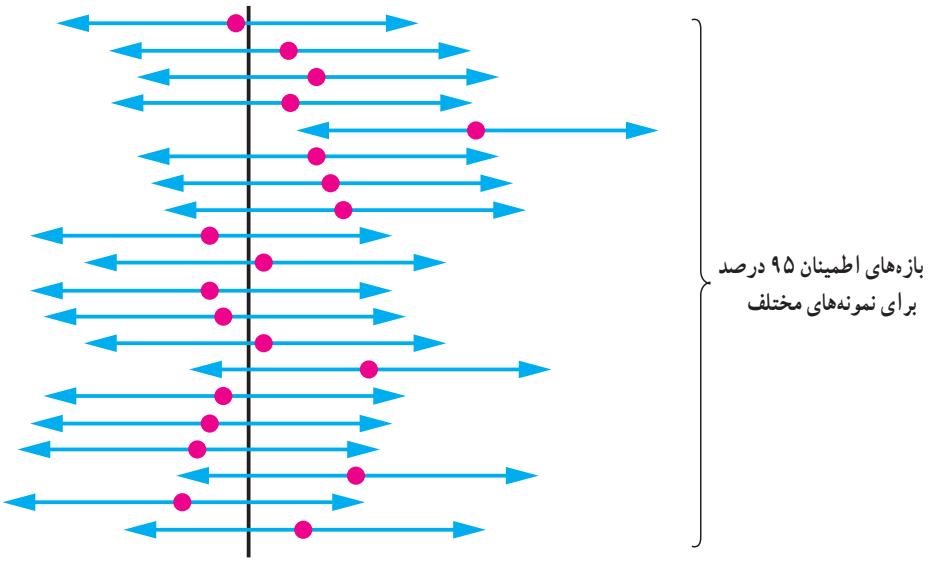
الف) اگر پاره‌خط آبی، خط سیاه را قطع نکند، چه نتیجه‌ای باید گرفت؟

ب) چند درصد از ۲۰ پاره‌خط آبی، خط سیاه را قطع کرده‌اند؟

پ) اگر ۱۰۰ پاره‌خط آبی را رسم می‌کردیم، انتظار داشتید چند تا از آنها خط سیاه را قطع نکنند؟

ت) نتیجه این تمرین تعبیر یک بازهٔ اطمینان ۹۵ درصد است. اگر ۱۰۰ بار نمونه‌گیری را تکرار کنیم و ۱۰۰ بازهٔ اطمینان محاسبه کنیم انتظار داریم ..... از آنها پارامتر میانگین جامعه را در بر گیرند.






**۷** شاخص پوسیدگی دندان ($DMFT$) در ایران برای سال ۱۳۶۰ برابر ۳ بوده است؛ یعنی به‌طور متوسط هر ایرانی دارای یک دندان کشیده شده، یک دندان پوسیده و یک دندان پرشده است. براساس نمونه‌ای به اندازهٔ ۴۰۰، این شاخص در سال ۱۳۹۵ برابر ۶ شده است ($\overline{x} = ۲$). اگر انحراف معیار دندان‌های کشیده شده، پوسیده و پرشده به ترتیب برابر ۱ ، ۲ و ۱/۶ باشد، بازهٔ اطمینان ۹۵ درصدی برای میانگین دندان‌های کشیده شده، پوسیده و پرشده محاسبه کنید.

**۸** پارامتر میانگین جامعه را با چه آماره‌هایی می‌توان برآورد کرد؟ (۵ آماره نام ببرید)

**۹** پارامتر واریانس و انحراف معیار جامعه را با چه آمارهایی می‌توان برآورد کرد؟

**۱۰** در فصل قبل با چه آماره‌هایی آشنا شده‌اید؟ آنها چه پارامترهایی را برآورد می‌کردند؟

**۱۱** مثال‌های این درس را با اندازهٔ نمونه‌های مختلف حل کنید. چه نتیجه‌هایی می‌توان گرفت؟ (مقدار برآورد تغییر نمی‌کند.)

# منابع


۱ گزارش سالانه کیفیت هوای تهران در سال ۱۳۹۳/ سولماز احدی و همکاران/ مرکز نشر شهر
۲ آمار و احتمال مهندسی/ نادر نعمت‌الهی/ قفنوس
۳ آمار و احتمال مقدماتی/ جواد بهبودیان/ آستان قدس
۴ آمار و احتمال مقدماتی/ مجتبی گنجعلی، احسان بهرامی سامانی/ نگارندهٔ دانش
۵ مبانی احتمال/ سعید قهرمانی/ ترجمهٔ غلامحسین شاهکار، ابوالقاسم بزرگ‌نیا / دانشگاه صنعتی شریف
۶ مبانی احتمال/ شلدون راس/ ترجمهٔ احمد پارسیان، علی زینل همدانی/ شیخ‌بهایی
۷ مبانی ریاضیات گسسته/ حمیدرضا امیری و یدالله ایلخانی‌پور/ انتشارات مدرسه
۸ دانشنامهٔ ریاضی/ پرویز شهریاری و همکاران/ کانون فرهنگی آموزش
۹ ورودی به نظریهٔ احتمال/ عین‌الله پاشا/ انتشارات مدرسه
۱۰ آنالیز ریاضی ــ جلد اوّل/ غلامحسین مصاحب/ امیرکبیر
۱۱ منطق ریاضی/ محمد اردشیر/ هرمس
۱۲ منطق و استدلال ریاضی/ عبدالحسین مصحفی/ فاطمی
۱۳ منطق، برهان و مجموعه‌ها/ مارو وین آل. بیتینگر/ ترجمهٔ محمود محمدی و میثم خطیبی/ مبتکران
۱۴ Elementary Statistics / Allan G Bluman / Higher Education
۱۵ Detecting activation in / MRI dat/ KJ Worsley
۱۶ H.Baxter et al. (2006) GCSE Mathematics for OCR , Hodder Murray


سازمان پژوهش و برنامه‌ریزی آموزشی جهت ایفای نقش خطیر خود در اجرای سند تحول بنیادین در آموزش و پرورش و برنامه درسی ملی جمهوری اسلامی ایران، مشارکت معلمان را به‌عنوان یک سیاست اجرایی مهم دنبال می‌کند. برای تحقق این امر در اقدامی نوآورانه سامانه تعاملی بر خط اعتبارسنجی کتاب‌های درسی راه‌اندازی شد تا با دریافت نظرات معلمان دربارهٔ کتاب‌های درسی نونگاشت، کتاب‌های درسی را در اولین سال چاپ، با کمترین اشکال به دانش‌آموزان و معلمان ارجمند تقدیم نماید. در انجام مطلوب این فرایند، همکاران گروه تحلیل محتوای آموزشی و پرورشی استان‌ها، گروه‌های آموزشی و دبیرخانهٔ راهبری دروس نقش سازنده‌ای را بر عهده داشتند. ضمن ارج نهادن به تلاش تمامی این همکاران، اسامی دبیران و هنرآموزانی که تلاش مضاعفی را در این زمینه داشته و با ارائهٔ نظرات خود سازمان را در بهبود محتوای این کتاب یاری کرده‌اند به شرح زیر اعلام می‌شود.

## کتاب آمار و احتمال ــ کد ۱۱۱۲۱۵

| ردیف | نام و نام خانوادگی | استان محل خدمت | ردیف | نام و نام خانوادگی | استان محل خدمت |
|---|---|---|---|---|---|
| ۱ | فرزانه کدخدایی | لرستان | ۲۲ | حمیده محمدزاده | خراسان‌شمالی |
| ۲ | علیرضا رافعی‌بروجنی | چهارمحال وبختیاری | ۲۳ | حمیده شریفی | هرمزگان |
| ۳ | نادر بلال‌زاده | آذربایجان‌شرقی | ۲۴ | عزیز اسدی | کردستان |
| ۴ | حجت‌الله ضیایی‌بروجنی | چهارمحال وبختیاری | ۲۵ | حمید حسینخانی | کرمان |
| ۵ | پژمان شعبانی‌نژاد ناورود | گیلان | ۲۶ | محمد عادلی | کرمان |
| ۶ | سمیه بهرامی | شهرتهران | ۲۷ | جواد ابراهیم‌نژاد | سیستان وبلوچستان |
| ۷ | سکینه علەزاده | ایلام | ۲۸ | رضا قلاسی | خراسان‌جنوبی |
| ۸ | ایرج نوری | ایلام | ۲۹ | مهدی قسوره | خراسان‌جنوبی |
| ۹ | مجتبی رفیعی | شهرستان‌های تهران | ۳۰ | شجاع‌علی گرجیان‌مهلبانی | مازندران |
| ۱۰ | عصمت چاهکی | مرکزی | ۳۱ | علی‌رضا بیات | همدان |
| ۱۱ | جبار منبری | کردستان | ۳۲ | علی خالقی | فارس |
| ۱۲ | ژاله اخوان | شهرستان‌های تهران | ۳۳ | ابوذر نخعی‌مطلق | سیستان و بلوچستان |
| ۱۳ | حجت رنگین | قزوین | ۳۴ | جلال سرحدی | مازندران |
| ۱۴ | مریم امیدیان | سمنان | ۳۵ | جمیله یوسف‌زاده | گیلان |
| ۱۵ | رقیه عاصفی | شهرتهران | ۳۶ | جمال نوین | یزد |
| ۱۶ | ام البنین ربیعی | سمنان | ۳۷ | مسعودرضا عرب‌یارمحمدی | گلستان |
| ۱۷ | افشین خاصه‌خان | آذربایجان‌شرقی | ۳۸ | سیما مهرجویا | آذربایجان‌غربی |
| ۱۸ | مهناز رضایی | خراسان‌شمالی | ۳۹ | حسین بهنامی | خراسان‌رضوی |
| ۱۹ | امین زارعی | هرمزگان | ۴۰ | مینا عظیمی | همدان |
| ۲۰ | مریم مددی | اصفهان | ۴۱ | فرخ حسن‌زاده | کهگیلویه و بویراحمد |
| ۲۱ | غلامرضا روئین‌تن | فارس | | | |