

۴

آمار استنباطی

درس ۱

گردآوری داده‌ها

خرم مهر

فعالیت

صفحه ۱۰۴



می‌خواهیم برخی از ویژگی‌های مگس‌های سفید مزاحم در شهر تهران را بررسی کنیم. آیا برای انجام این کار می‌توانیم ویژگی‌های همه مگس‌های سفید را اندازه‌گیری کنیم؟ آیا همه آنها در دسترس‌اند؟ آیا زمان و هزینه لازم برای این کار در اختیار داریم؟ **جواب هر سه سوال خیر**

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

مجموعه کل واحدهای آماری را جامعه آماری^۲ می‌نامند.

هر زیر مجموعه از جامعه 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. رابطه بین جامعه و بخشی از آن که نمونه نامیده می‌شود، در شکل نشان داده شده است.

کار در کلاس

صفحه ۱۰۴

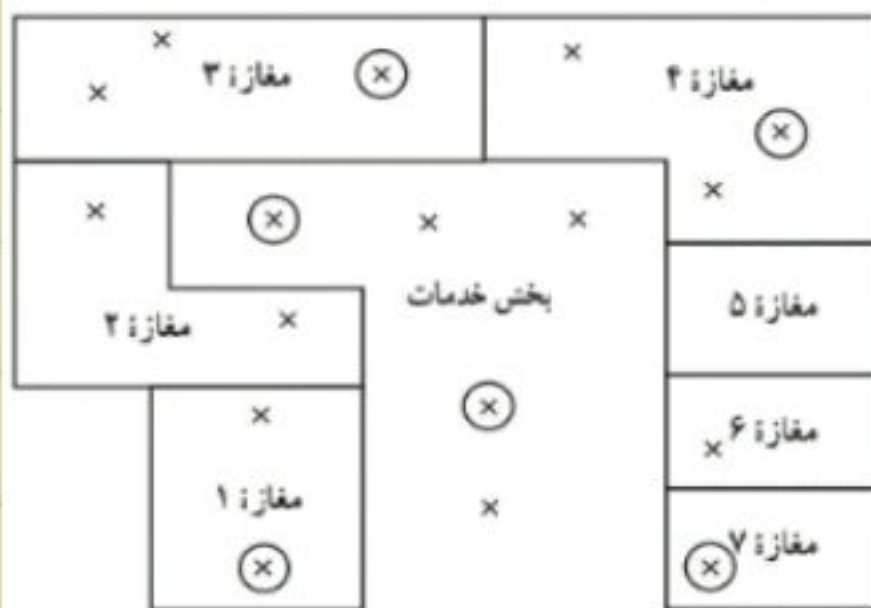
در فعالیت قبل هر مگس سفید **واحد آماری** است. همه مگس‌های سفید، که کل واحدهای آماری هستند، **جامعه آماری** را تشکیل می‌دهند. اگر سن همه مگس‌های سفید را در اختیار داشته باشیم، داده‌های **نمونه**... را داریم. ۱۰۰ مگس سفید معرف یک **نمونه‌گیری** است.

نمونه گیری تصادفی ساده^۱ نوعی روش نمونه گیری که در آن همه واحدهای آماری برای انتخاب شدن در نمونه، احتمال یکسان دارند.

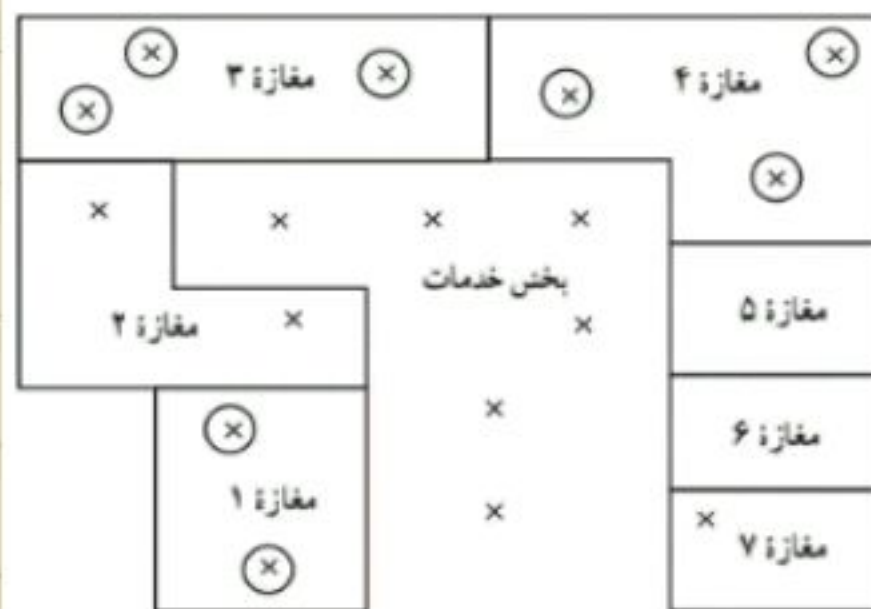
صفحہ ۱۰۵

فعالیت

۱ می خواهیم 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 می کنید؟ یک راه ساده برای انجام این کار نوشتن اسامی کارکنان یا شماره کارمندی آنها روی ۱۷ برگه کوچک و انتخاب تصادفی ۶ تا از آنهاست. آیا این روش نمونه گیری، نمونه گیری تصادفی ← بله



ساده است؟ آیا همه واحدهای جامعه احتمال برابری برای انتخاب دارند؟ ← بله
در شکل روبه رو نقشه ای از مجتمع تجاری ترسیم شده که کارکنان با x و دور انتخاب شدگان یک دایره رسم شده است. انجام نمونه گیری تصادفی ساده در عمل با دشواری هایی همراه است. اگر اندازه جامعه بزرگ باشد، یعنی تعداد واحدهای آماری زیاد باشند، دسترسی به فهرستی از اعضای جامعه و دسترسی به اعضای انتخابی، دشوار و ممکن است هزینه بر باشد.



۲ هر یک از ۷ مغازه و بخش خدمات را به صورت یک گروه فرض می کنیم. حال از بین ۸ گروه در نظر گرفته شده، سه تا از آنها را به تصادف انتخاب می کنیم و در هر یک سرشماری انجام می دهیم. آیا این روش نمونه گیری سریع تر است؟

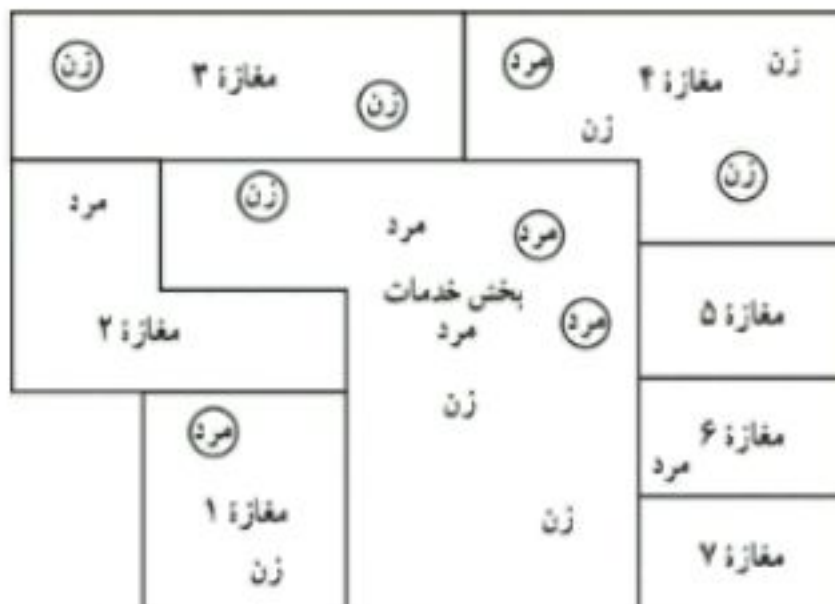
↓
بله

نمونه گیری خوشه ای^۲: نمونه گیری که در آن، واحدهای نمونه گیری اولیه در جامعه، گروه ها یا خوشه ها باشند. سپس همه واحدهای آماری خوشه های انتخاب شده را به عنوان نمونه در نظر می گیریم.

صفحه ۱۰۶

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

۲ اگر بخواهیم یک نمونه 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 با زمان همراه است، ولی انتظار داریم که **دقت** را نیز افزایش دهد.



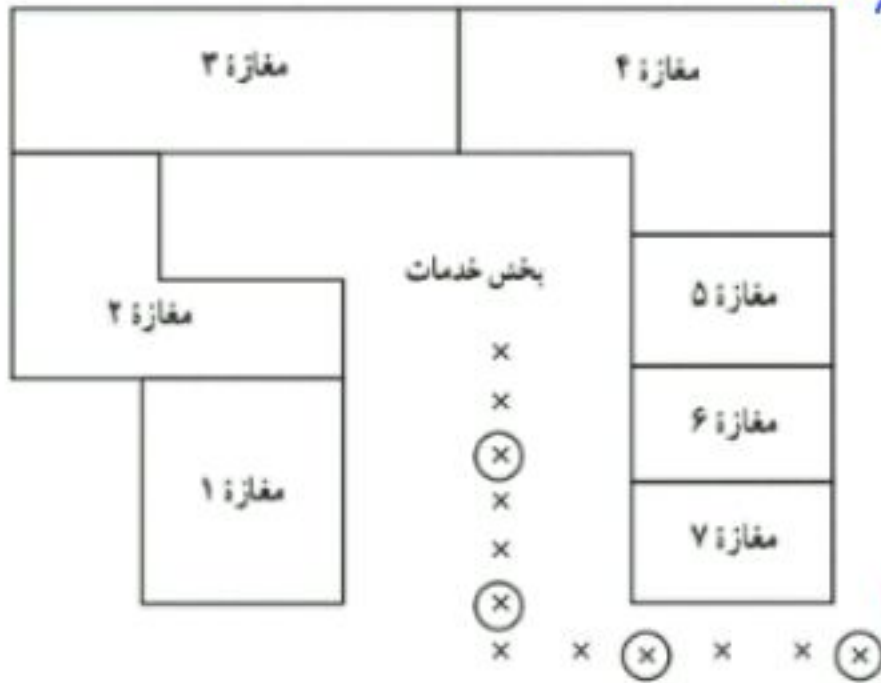
در واقع داده‌ها به دو طبقه مجزا تقسیم شده‌اند که طبقات از نظر ویژگی مورد بررسی همگن هستند. در صورتی که در نمونه‌گیری خوشه‌ای، درون خوشه‌ها هرچه ویژگی مورد بررسی تفاوت بیشتری داشته باشند بهتر است.

نمونه‌گیری طبقه‌ای: روش نمونه‌گیری که در آن با طبقه‌بندی جامعه به زیرجامعه‌های مجزا یک نمونه تصادفی ساده از هر طبقه انتخاب می‌شود.

علاقه‌مند به نمونه‌گیری از نمرات درس ریاضی دانش‌آموزان استان تهران هستیم. اگر فهرست همه دانش‌آموزان 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. در صورتی که 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.) **تصادفی ساده - طبقه‌ای**

۴ فرض کنید در مجتمع، ۱۲ نفر حضور دارند. صبر می‌کنیم که مجتمع تجاری تعطیل شود و هنگام خروج کارکنان می‌خواهیم نمونه ۴ نفری انتخاب کنیم. برای این منظور، همانند شکل صفحه بعد عمل کرده‌ایم. ابتدا از ۳ نفر یکی را به تصادف انتخاب می‌کنیم. در این شکل، نفر اول انتخاب شده است. حال با همین رویه برای سه نفر بعد هم، نفر اول 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؟ **صرفه‌جویی در زمان و هزینه**

صفحه ۱۵۷



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

- گردآوری اطلاعات از مبدأ و مقصد مسافران در خروجی - 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^۱ (سیستماتیک)، نوعی نمونه‌گیری طبقه‌ای است که در آن اندازه طبقات با هم برابر است. فقط از طبقه اول، واحد آماری به تصادف انتخاب می‌شود و با همان روبه از طبقات دیگر، این کار انجام می‌گیرد.

بله

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

چون واحد آماری بصورت تصادفی انتخاب می‌شود

کار در کلاس

صفحه ۱۵۷

جدول زیر را کامل کنید.

محدودیت	مزیت	روش نمونه‌گیری
هزینه بالا - وقت‌گیر بودن	شانس یکسان همه واحدها برای انتخاب	تصادفی ساده
"	دسترسی ساده و هم‌شانس بودن	خوشه‌ای
کلیان نبودن ویژگی مورد آزمون در طبقات	تناسب تعداد نمونه با اندازه جامعه	طبقه‌ای
زمان‌بر بودن	دسترسی ساده - هم‌شانس بودن	سامانمند

فعالیت

صفحه ۱۰۷

از مگس های سفید با چه روشی می توان نمونه گیری کرد؟ فهرستی از آنها نداریم، تعداد آنها را هم نمی دانیم. می توان چند منطقه از تهران را به تصادف انتخاب کرد و در هر منطقه نمونه در دسترس را انتخاب و بررسی کنیم. آیا این روش نمونه گیری به تمامی واحدهای جامعه شانس انتخاب می دهد؟

نوشته ای

خیر

نمونه گیری احتمالی: نمونه گیری است که همه واحدهای آماری احتمالی معلوم برای انتخاب در نمونه داشته باشند. و از روشی تصادفی برای انتخاب واحدهای نمونه استفاده شود.

کار در کلاس

صفحه ۱۰۸

راه حلی ارائه کنید که نمونه گیری های غیر احتمالی زیر را احتمالی می کند، هر چند که به صورت غیر واقعی باشد.

نمونه گیری احتمالی	نمونه گیری غیر احتمالی	مثال
۱	بدون برنامه ریزی خرگوش هایی را برمی دارد که دستش به آنها می خورد.	نمونه گیری از یک قفس بزرگ خرگوش های یک آزمایشگاه
۲	داوطلبانی که حاضر به پاسخ به سؤالات شما در یک نظر سنجی می شوند.	در مطالعاتی که در آنها فرایند سنجش برای شخصی که سنجیده می شود ناخوشایند یا دردسرافرین است.
۳	نمونه در دسترس انتخاب می شود.	نمونه گیری از زغال سنگ های یک واگن

۱ خرگوش ها را به چند دسته تقسیم کنیم سپس از هر دسته خرگوشی را انتخاب کنیم.

۲ از بین کل افراد ابتدا آنها را به چند دسته تقسیم کنیم سپس متناسب با تعداد نمونه گیری طبقه ای انجام دهیم.

۳ از چند قسمت واگن مخصوصا قسمت های میانی و زیرین چند نمونه انتخاب می کنیم

شاید در نگاه اول این طور 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 اول 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 به تصادف انتخاب می‌کند و از آنها می‌پرسد: «خانواده 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 مسئله 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 خانواده شش نفره دو برابر احتمال حضور یک خانواده سه نفره است و این، یعنی شرایط نمونه‌گیری ساده برقرار نیست. نتیجه 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 $\frac{1}{4} = \frac{20.7}{100.7}$ برابر تعداد خانواده‌های پنج نفره (و بیشتر) است، ولی با آمارگیری نادرستی که توضیح داده شد به نتیجه دیگری خواهیم رسید؛ احتمال حضور یک خانواده پنج نفره (و بیشتر) در

نمونه‌ها بیشتر از $2.5 \dots$ برابر احتمال حضور یک خانواده دو نفره است و لذا عددی که در روش نادرست آماری به دست می‌آید کمتر از $\frac{1}{4} = 0.25$ 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 صفحه قبل، برگرفته از آمار واقعی کشور در سرشماری سال ۱۳۹۵ است.)

محمد مهدی

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

۲۵	۲۴	۲۳	۲۲	۲۱	۲۰	۲۹	۲۸	۲۷	۲۶	۲۵	تعداد دانش‌آموز
۴	۵	۷	۸	۸	۱۱	۷	۱۱	۱۰	۹	۷	تعداد کلاس

الف) چه تعداد کلاس پایه ششم در این شهر وجود دارد؟ تعداد دانش‌آموزان پایه ششم چند ناست؟

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

پ) اگر به تصادف 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 به تصادف 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 آلاینده‌گی خودروهای در حال تردد 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 نمونه 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 باز هم نمونه‌گیری ما مشکل دارد؛ زیرا خودروهایی که از آنها بیشتر استفاده می‌شود، احتمال بیشتری دارد که به‌عنوان نمونه انتخاب شوند.

$$\text{الف) تعداد کلاس‌ها} \quad 4 + 5 + 7 + 8 + 8 + 11 + 7 + 11 + 10 + 9 + 7 = 87$$

$$4(25) + 5(24) + 7(23) + 8(22) + 8(21) + 11(20) + 7(29) + 11(28) + 10(27) + 9(26) + 7(25) = 2575$$

تعداد دانش‌آموزان

$$\frac{4 + 5 + 7 + 8 + 8}{87} \times 100 = \frac{32}{87} \times 100 = 36,78\%$$

کلاس‌های بیش از ۳۰ نفره (ب)

$$\frac{\text{تعداد دانش آموزان کلاسهای بسته‌های بسته‌ها: ۳ نفر}}{\text{تعداد کل دانش آموزان}} = \frac{4(35) + 5(34) + 7(33) + 8(32) + 8(31)}{2565}$$

$$= \frac{1045}{2565} = 40.7\%$$

پ)

ت) این روش زمان زیادی نیاز دارد

ث)

کار هر کلاس

صفحه ۱۰۹

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

نیتیه غیردرسی خواهد بود چون اغلب آن مطالعه در

«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. نیتیه غیردرسی چون بسته‌های بسته‌ها می‌تند و آریب

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؟

صفحه بعد

نیتیه غیردرسی و آریب - آنهایی که مطالعه بی‌سختی دارند معمولاً اول مطالعه هستند

از هر کدام از پایه های و از تمام رشته ها، دانش آموزانی انتخاب کنیم (تقریباً) و سوال مورد نظر را از آنها بپرسیم

کار در کلاس

صفحه ۱۱۰

از جمله مسائلی که مردم در مورد آنها به نظرسنجی ها علاقه زیادی نشان می دهند، انتخابات است. با این وجود، گاهی مردم و گاهی گروه های سیاسی از روش هایی برای کشف نظر مردم استفاده می کنند که آنها را همراه می کند.

در این مورد روش های زیر را نقد کنید :

الف) پرسیدن نظر دوستان و اطرافیان.

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

در تمام مثال هایی که ذکر شد، روش نمونه گیری به شکلی بود که داده هایی به سمتی انحراف پیدا می کردند و لذا افزایش تعداد نمونه ها نیز به کاهش این انحراف کمک نمی کرد. در علم آمار اصطلاحی خاص برای این مشکل وجود دارد :

الف) نتایج نظرسنجی از واقعیت دور خواهد بود چون دوستان و اطرافیان معمولاً هم نظرند

ب) «ریک و بگانه قطعاً جستجوی فکری خاصی وجود دارد و نتایج دعوی نخواهد بود»

اگر یک روش نمونه گیری از نمونه گیری ایده آل فاصله بگیرد و به سمتی خاص انحراف پیدا کند می گویند آن روش نمونه گیری اریب است. لذا آمارشناسان تلاش می کنند تا با شناسایی منابع تولید اریبی، نمونه گیری ها را تا جایی که می توانند نا اریب کنند.

محمّد مهدی

فعالیت

ص ۱۱۲

می خواهیم طول قد دانش آموزان یک مدرسه را گردآوری کنیم. برای این منظور چه راهی پیشنهاد می کنید؟

آمارگیری: گردآوری داده ها به یکی از روش های ممکن
آمارگیر: کسی که آمارگیری را انجام می دهد.

اگر قرار شد آمارگیر باشیم، می توانیم جدولی به صورت زیر تکمیل کنیم.

مثالی از جدول طراحی شده برای ثبت داده ها

اندازه طول قد	خط نشان برای شمارش	تعداد دانش آموزان
$140 < \text{طول قد}$		
$139/5 \leq \text{طول قد} < 149/5$		
$149/5 \leq \text{طول قد} < 159/5$		
$159/5 \leq \text{طول قد} < 169/5$		
$169/5 \leq \text{طول قد}$		

به ن
دانش آموزان

بله رفتن اطلاعات از خود

↑ طراحی پرسش نامه

آمارگیری زحمت زیادی برای آمارگیر دارد. آیا راه حل ساده تری برای انجام آن دارید؟ یکی از مرسوم ترین روش های آمارگیری، استفاده از پرسش نامه است. پرسش نامه شبیه همان جدولی است که هنگام ثبت نام در مدرسه، آن را تکمیل کرده اید. واحدهای جامعه یا نمونه می توانند پرسش نامه تکمیل کنند.

کار در کلاس

ص ۱۱۳

۱ چه راه دیگری برای آمارگیری طول قد دانش آموزان یک مدرسه پیشنهاد می کنید؟

۲ فرض کنید زمان لازم را برای گردآوری همه داده های دانش آموزان در اختیار نداشته باشید. اگر بخواهیم نمونه ای را

انتخاب و آمارگیری کنیم، چه راهی پیشنهاد می کنید که نمونه به صورت تصادفی انتخاب شود؟

۱- پرسش نامه و یا مراجعه به فرم های ثبت نام دانش آموزان در بدو ورود به مدرسه

۲- از هر کلاس و از روی لیست اسامی دانش آموزان ۳ نفر را بطور تصادفی انتخاب کنیم



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

- ۱ تعداد قلم‌های هر دانش‌آموز در یک کلاس. ← مصاحبه
- ۲ ساعات خواب دانش‌آموزان کلاس درس شما در شب گذشته. ← پرسش‌نامه
- ۳ طول قد دانش‌آموزان یک کلاس. ← مشاهده - پرسش‌نامه

ب) می‌خواهیم طول قد دانش‌آموزان یک کلاس یا مدرسه را به یکی از سه روش زیر آمارگیری کنیم. هر یک 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 به دقت زیادی نیاز داشته باشیم، مناسب نیست.
 دادگان‌ها: همیشه اطلاعات ثبتی در اختیار نیست.

ارسال پرسش‌نامه بصورت اینترنتی

وهدت چندروزه برای تکمیل و

ارسال

استفاده از رسانه‌های باریت زیاد
 در اندازه‌گیری مخصوصاً
 وسیله دیجیتالی

درخواست از دانش‌آموزان
 برای مراجعه به پایگاه
 اطلاعاتی و تکمیل اطلاعات خود

قرار است درباره افرادی که از کوه دنا بالا رفته‌اند، پژوهشی آماری انجام دهیم. واحدهای آماری این پژوهش، همه افرادی هستند که توانسته‌اند به قله 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، متغیر می‌گویند. سایر متغیرها می‌توانند مواردی مانند: سن، وزن، ملیت، میزان تحصیلات و درآمد باشند. متغیرهای مورد بررسی در یک پژوهش ممکن است کمی یا کیفی باشند.

یادآوری

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

متغیر کمی: متغیری است که مقادیر عددی می‌گیرد و برای آن عملیات ریاضی از قبیل جمع، تفریق و معدل‌گیری قابل انجام است.

متغیر کیفی: متغیری است که صرفاً برای دسته‌بندی افراد یا اشیا در گروه‌ها به کار می‌رود و لزوماً مقدار عددی نمی‌گیرد.

متغیر
 ۶

پارامتر یا پارامتر جامعه: یک مشخصه عددی است که توصیف کننده جنبه‌ای خاص از جامعه است و در صورتی که داده‌های کل جامعه در اختیار باشند، قابل محاسبه است.

آماره^۱ یا آماره نمونه: مشخصه‌ای عددی که توصیف کننده جنبه‌ای خاص از نمونه است و از داده‌های نمونه به دست می‌آید.

فرایند نتیجه‌گیری درباره پارامترهای جامعه بر اساس نمونه، آمار استنباطی^۲ است.

تمرین

ص ۱۱۶

۱ در نمونه‌گیری تصادفی ساده، احتمال اینکه فرد به خصوصی در اولین انتخاب عضو نمونه باشد، چقدر است؟ اگر مسئله با جای‌گذاری باشد، احتمال اینکه او در دومین انتخاب عضو نمونه باشد، چقدر است؟ اگر مسئله بدون جای‌گذاری باشد، و از نتیجه انتخاب اول اطلاع نداشته باشیم، احتمال اینکه او در دومین انتخاب عضو نمونه باشد، چقدر است؟

اگر تعداد اعضای جامعه برابر N باشد، احتمال اینکه فرد خصوصی از بین انتخاب باشد $\frac{1}{N}$ است

چون با جای‌گذاری است ما هم تعداد اعضای جامعه N خواهد بود پس احتمال

انتخاب آن فرد $\frac{1}{N}$ است

پایه قیمت آخر:

شرط لازم برای آنکه فرد خصوصی در انتخاب دوم باشد آنست که در انتخاب اول نباشد

اگر جامعه N عضوی باشد \leftarrow احتمال اینکه شخص مورد نظر انتخاب اول نباشد $\frac{N-1}{N}$

برای دومین انتخاب جامعه $N-1$ عضو دارد \leftarrow احتمال انتخاب شخص مورد نظر $\frac{1}{N-1}$

بنابراین $P = \frac{N-1}{N} \times \frac{1}{N-1} = \frac{1}{N}$ احتمال اینکه شخص مورد نظر دومین انتخاب باشد

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

در مرحله‌ی اول از نمونه‌گیری خوشه‌ای، خوشه‌ها واحد آماری محسوب میشوند که احتمال انتخاب آنها برابر است.
 در مرحله بعد مطالعه فقط روی اعضای خوشه‌ی منتخب انجام میشود و اعضای سایر خوشه‌ها کنار گذاشته میشود.
 بنابراین احتمال انتخاب واحدهای آماری برابر نیست.
 این نمونه‌گیری احتمالی است چون در مرحله‌ی اولی احتمال انتخاب هیچکدام از خوشه‌ها صفر نیست.

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

در نمونه‌گیری احتمالی، شانس انتخاب همچندم از واحدهای آماری ضرر نیست و گاهی هم بلیان است اما در نمونه‌گیری غیر احتمالی بعضی از واحدهای آماری در انتخاب نمونه شانس ندارند.

۴ برای هر یک از روش‌های نمونه‌گیری احتمالی دو مثال واقعی بیاورید.

انتخاب ۳ نفر از دانش‌آموزان کلاس از روی شماره آنها در دفتر کلاسی
 برای رفتن به بلیه اردو

انتخاب بلیه از تماشاگران مسابقه فوتبال از روی شماره صدری برای

دارن جایزه

۱- تصادفی ساده

برای بررسی دلایل انت تحصیلی دانش آموزان متوسطه دوم شهر تهران،
هر منطقه را به عنوان چند خوشه در نظر بگیریم و تعداد مناسبی خوشه انتخاب کنیم

۲- خوشه‌ای

برای بررسی نمرات ریاضی یک مدرسه با ۱۲ کلاس مختلف، ۳ کلاس

را بصورت تصادفی انتخاب کنیم و نمرات ریاضی همه دانش آموزان
۳ کلاس را مورد بررسی قرار دهیم

الگوی مهم دانش آموزان یک شهر را مورد بررسی قرار دهیم، آنها را به

۳- طبقه‌ای

گروه‌های ابتدایی، متوسطه اول و متوسطه دوم طبقه‌بندی کنیم و مناسب

با تعداد افراد هر گروه به صورت تصادفی، افرادی را برای حضور در نمونه انتخاب کنیم

از بین دانش آموزان یک مدرسه در رشته‌های ریاضی و تجربی دانش‌آموزانی

مناسب با تعداد آنها به طور تصادفی، افرادی را انتخاب کنیم و نمرات

مطالعه آنها در یک هفته را مورد بررسی قرار دهیم

تقسیم لیست کلاس ۳۰ نفره کلاس به ۵ دسته مساوی، انتخاب

یک عدد تصادفی از ۱ تا ۴ و ادامه روند تا انتخاب ۵ نفر

۴- سیستماتیک

انتخاب برندگان برنامه ۹۰ از بین شرکت کنندگان پیامکی

۵ اگر اندازه 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؟

با استفاده از اعداد تصادفی تولید شده بوسیله یک ماشین حساب

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

۷ فرق بین داده و متغیر چیست؟

۸ فرق بین آماره با پارامتر چیست؟

۶- اگر طبقات هم اندازه باشند، واحدهای آماری همگی شانس برابر برای

انتخاب شدن در نمونه دارند. در غیر این صورت احتمال انتخاب میان واحدها

دلی در هر حالت، شانس انتخاب واحدهای آماری در هر طبقه یکسان است.

۷- داده واقعی است (درباره یک موضوع) و اگر از هر فرد به فرد دیگر تغییر کند به آن متغیر می‌گویند

نمونه در آماره احتمال دانش آموزان کلاس داده است و چون این نمره

از فردی به فرد دیگر تغییر می‌کند یک متغیر است

۸- پارامتر از جامعه بدست می‌آید و عددی ثابت است

آماره از نمونه بدست می‌آید، عدد ثابتی نیست و تغییر می‌کند

۱۰ در یک جامعه آماری، آیا ممکن است که یک پارامتر تغییر کند؟ اگر سه نمونه با اندازه یکسان از یک جامعه داشته باشیم، می توان سه مقدار متفاوت از یک آماره به دست آورد؟

خیر، پارامتر همیشه ثابت است

بله، آماره از یک نمونه به نمونه دیگر تغییر می کنند

۱۱ در یک مطالعه از ۱۲۶۱ مشتری غذاهای گاهی، سؤال شده است که برای کدام وعده غذایی (ناهار یا شام) سفارش داده اند؟

الف) متغیر را مشخص کنید. این متغیر کمی است یا کیفی؟

ب) کدام روش گردآوری داده ها برای مطالعه مناسب است؟

پ) جامعه آماری در اینجا چیست؟ در این مطالعه پارامتر و آماره چه چیزی می توانند باشند؟

الف) متغیر = وعده غذایی (ناهار یا شام) ← کیفی

ب) پرسش شفاهی

پ) تمام مشتریان غذاهای گاهی جامعه آماری است

مرد جامعه = پارامتر و مدمنونه = آماره

۱۲ کدام روش گردآوری داده ها برای موارد زیر مناسب است؟ یک دلیل برای انتخاب خود ذکر کنید.

۱) میزان رضایت مشتریان بانک از نحوه برخورد و رسیدگی به درخواست های آنها.

۲) سن همه دانش آموزان مدرسه بر حسب ماه در پایه دهم.

۳) تعداد سرنشینان خودروهای سواری در یکی از محورهای خروجی شهر.

۱- پرسش نامی : در دسترس بودن مشتریان در بانک

۲- دادگان : مراجعه به پرونده های دانش آموزان در مدرسه

۳- مشاهده : توجه به خودروهای سواری بدون نیاز به توقف آنها

۱۱ فرض کنید جامعه‌ای از $N=100$ عضو تشکیل شده و می‌خواهیم نمونه‌ای به اندازه $n=20$ 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 نمونه تصادفی ۱۰ تایی انتخاب کنیم.

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

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

الف) طبقه‌ای

$$\text{احتمال انتخاب هر عضو جامعه} = \frac{10}{50} = \frac{1}{5}$$

ب) خوشه‌ای

$$\text{احتمال انتخاب هر عضو جامعه} = \frac{2}{10}$$

پ) سیستماتیک

$$\text{احتمال انتخاب هر اول} = \frac{1}{5}$$

۱۲ دلایل آریبی 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) نمونه غیر تصادفی: عامل شانس در انتخاب نمونه نقشی ندارد.

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

(ت) نمونه‌گیری تلفنی: از دفترچه راهنمای تلفن تعدادی شماره به تصادف انتخاب می‌شود.

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

الف) تمام اسرادر جامعه شانس انتخاب ندارند

بترین روش انتخاب با قرعه‌کشی است (روش تصادفی ساده)

ب) اعمال نظر شخص نمونه‌گیر در احتمال تغییر نتایج بترین روش تصادفی ساده است

پ) حذف کسانی که اعیل ندارند یا بررسی به استرینت ندارند یا اعیل خودشان را چک نمی کنند

تبرین روش انتخاب نمونه از بین مخاطبان شبکه های اجتماعی است

ت) ممکن است بعضی از شماره های در دفتر ثبت نشده باشد و یا تعدادی از شماره های مسدود باشد

تبرین روش انتخاب شماره بصورت اتفاقی است

ث) ممکن است حجم نمونه به اندازه کافی بزرگ نباشد

باید افسردگی را جایگزین آنهایی که پاسخگو نیستند، نسیم

۱۱ نوع روش نمونه گیری مناسبتر را انتخاب کنید.

الف) شرکت وارد کننده خودروهایی سنگین برای بررسی عملکرد سامانه ترمز آنها می خواهد ده درصد از خودروهایی را که به مرور زمان وارد کشور می شوند بازرسی کند.

ب) مدیر مدرسه ۶۰۰ نفری می خواهد نظر دانش آموزان را برای تغییر ساعت تعطیلی مدرسه بر اساس یک نمونه 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 پایه های مختلف).

الف) سیستماتیک: انتخاب خودروها با فاصله های معین زمانی

ب) تصادفی ساده: انتخاب ۱۲ نفر از محیط مدرسه

پ) طبقه ای: ۶ پایه به عنوان ۶ طبقه هستند و از هر طبقه ۱ نفر انتخاب شود

۱۵ چگونه از ارقام ۰ تا ۹ عدد تصادفی انتخاب می کنید؟ آیا با روش پیشنهادی شما می توان عدد تصادفی بین اعداد ۰ تا ۹۹ انتخاب کرد؟ آیا امکان توسعه روش پیشنهادی شما به انتخاب تصادفی از فهرستی ۱۰۰۰ تایی امکان پذیر است؟

لرقام ۰ تا ۹ را ردی ۱۰ کارت بنویسیم و قرعه کشی کنیم و یا اینکه به کمک جدول اعداد تصادفی عدد را انتخاب کنیم

وقتی تعداد اعضای جامعه زیاد باشد، ساده ترین روش

انتخاب عدد با کمک عددی تصادفی و ماشین حساب است.



با از روی هر روز و لحظات

مهره مهره
تهران - پردیس
۱۴۰۲/۰۱/۱۵

