درس ۲ برآورد

ملاسعیدی- زیر نظر استاد انارکی(اصفهان)
آیدی تلگرام و ایبتا: @sinxcosx
وبلاگ: sinxcosx.blogfa.com
سایت: sinxcosx.ir
تلفن: 09168324500

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؟

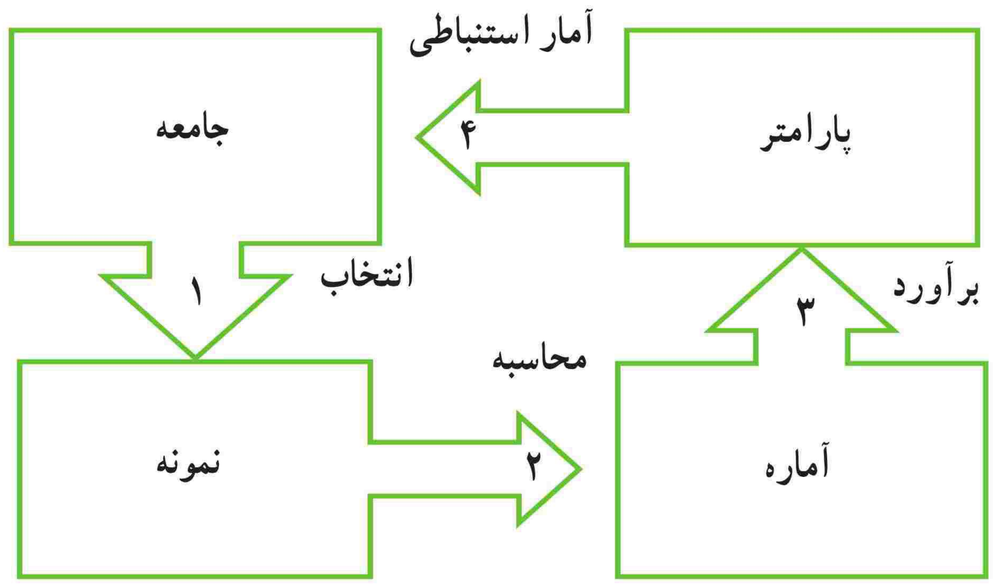
۳	۲	۷	۲	۵	۳	۳	۱	۴

برآورد نقطه‌ای[1] 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 **جامعه** ، پارامتر است. آماره، **میانگین تعداد اعضای خانوار نمونه انتخابی** ؛ و برآورد نقطه‌ای پارامتر $\frac{۳۰}{۹}$ است.

۱ـ Point Estimation

ملاسعیدی- زیر نظر استاد انارکی(اصفهان)
آیدی تلگرام و ایبتا: @sinxcosx
وبلاگ: sinxcosx.blogfa.com
سایت: sinxcosx.ir
تلفن: 09168324500



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؟ **پارامتر ، میانگین درآمد ماهیانه هر ۶ نفر جامعه است** $\frac{۴+۱+۰+۳+۵+۲}{۶}=۲/۵$

■ آیا بر اساس هر یک از نمونه‌ها برآورد به مقدار پارامتر نزدیک است؟ **خیر**

■ چه راه‌حلی پیشنهاد می‌کنید که برآورد به پارامتر نزدیک‌تر شود؟ **اندازه نمونه افزایش یابد**

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؟ $\{۰,۵\},\{۱,۴\},\{۲,۳\}$

■ آیا امکان دارد با نمونه‌های مختلف برآوردهای برابر به دست آوریم؟ **بله ، پاسخ سوال قبل گویای این مطلب است**

■ بدون شمارش بگویید امکان مشاهده چند نمونهٔ دوتایی داریم؟ $\binom{۶}{۲}=۱۵$

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	$\frac{۱}{۱۵}$	$\frac{۱}{۱۵}$	$\frac{۲}{۱۵}$	$\frac{۲}{۱۵}$	$\frac{۳}{۱۵}$	$\frac{۲}{۱۵}$	$\frac{۲}{۱۵}$	$\frac{۱}{۱۵}$	$\frac{۱}{۱۵}$

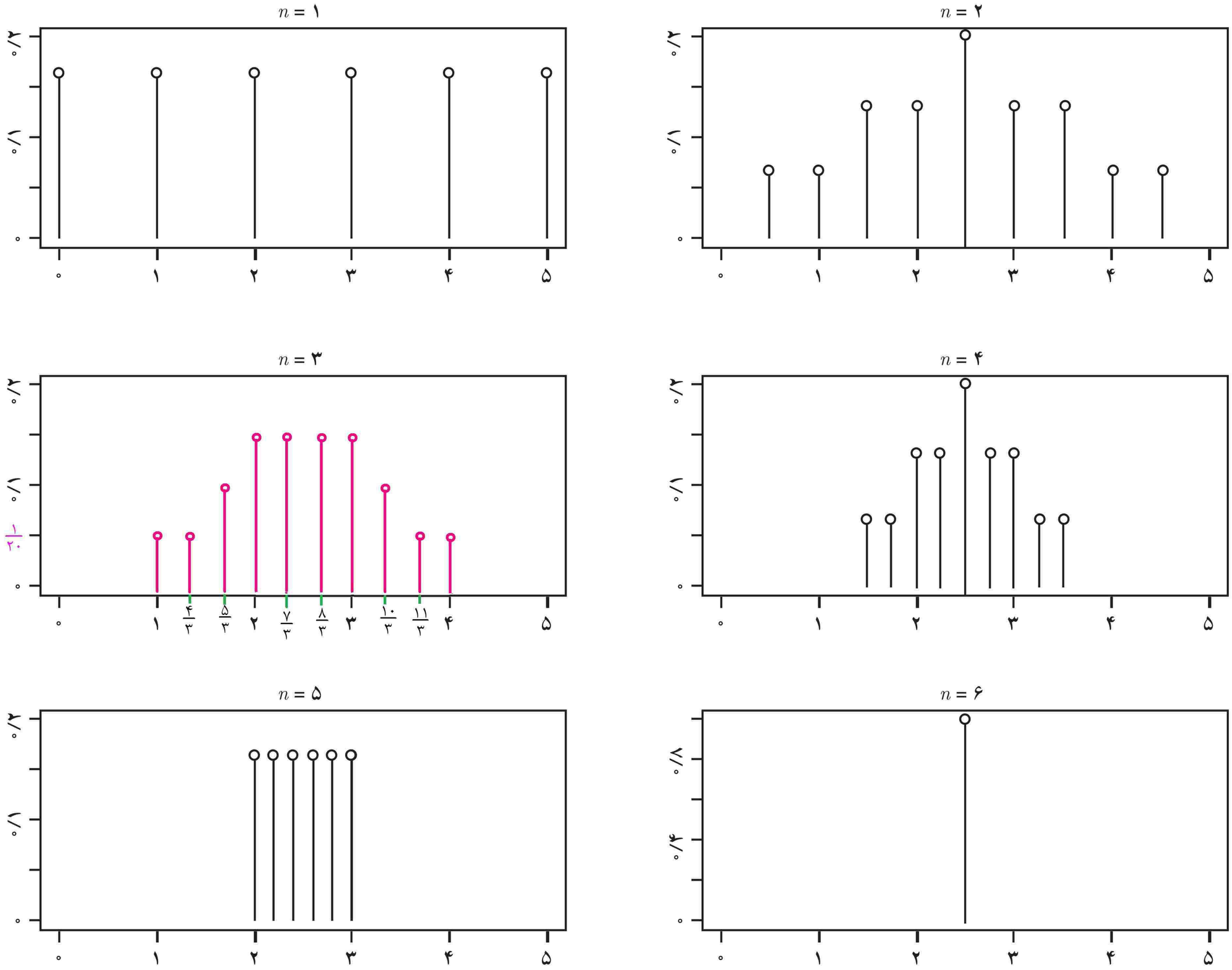
ملاسعیدی- زیر نظر استاد انارکی(اصفهان)
آیدی تلگرام و ایتا: @sinxcosx
وبلاگ: sinxcosx.blogfa.com
سایت: sinxcosx.ir
تلفن: 09168324500

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} {۰,۲,۳}	{۰,۱,۵} {۰,۲,۴} {۱,۲,۳}	{۰,۲,۵} {۰,۳,۴} {۱,۲,۴}	{۰,۳,۵} {۱,۲,۵} {۱,۳,۴}	{۰,۴,۵} {۱,۳,۵} {۲,۳,۴}	{۱,۴,۵} {۲,۳,۵}	{۲,۴,۵}	{۴,۳,۵}
$\overline{x}$	$\frac{۳}{۳}$	$\frac{۴}{۳}$	$\frac{۵}{۳}$	$\frac{۶}{۳}$	$\frac{۷}{۳}$	$\frac{۸}{۳}$	$\frac{۹}{۳}$	$\frac{۱۰}{۳}$	$\frac{۱۱}{۳}$	$\frac{۱۲}{۳}$
احتمال	$\frac{۱}{۲۰}$	$\frac{۱}{۲۰}$	$\frac{۲}{۲۰}$	$\frac{۳}{۲۰}$	$\frac{۳}{۲۰}$	$\frac{۳}{۲۰}$	$\frac{۳}{۲۰}$	$\frac{۲}{۲۰}$	$\frac{۱}{۲۰}$	$\frac{۱}{۲۰}$

فعالیت

نمودار 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.



ملاسعیدی- زیر نظر استاد انارکی(اصفهان)
آیدی تلگرام و ایتا: @sinxcosx
وبلاگ: sinxcosx.blogfa.com
سایت: sinxcosx.ir
تلفن: 09168324500

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؟ **بله** زمانی‌که اندازهٔ نمونه به ۶ می‌رسد، برآورد برابر**پارامتر جامعه یعنی۲/۵**می‌شود.

همان‌طور که در نمودارها دیده‌اید با افزایش اندازهٔ نمونه ، برآوردها به میانگین جامعه**(پارامتر)** نزدیک‌تر می‌شوند.

به بیان دیگر در هر نمودار با زیاد شدن اندازهٔ نمونه ، 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 **بسیار بزرگ** است.

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

انحراف معیار جامعه تقسیم بر جذر اندازهٔ نمونه = انحراف معیار **برآورد** 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**میانگین**کاهش می‌یابد. به عبارتی دیگر برآورد دقیق‌تر یا خطای کمتری برای برآورد میانگین جامعه داریم.

کار در کلاس

به **کار در کلاس قبل** باز می‌گردیم. اگر از مطالعات سال‌های گذشته بدانیم که انحراف معیار درآمد **افراد** در کشور ۲میلیون تومان است انحراف معیار برآورد میانگین درآمد افراد جامعه را برای **نمونه ها با اندازه های** ذکر شده محاسبه کنید.

n	۲۵	۱۰۰	۱۰۰۰۰
$\sigma_{\bar{x}}$	$\frac{۲}{\sqrt{۲۵}} = ۰/۴$	$\frac{۲}{\sqrt{۱۰۰}} = ۰/۲$	$\frac{۲}{\sqrt{۱۰۰۰۰}} = ۰/۰۲$

- انحراف معیار برآورد میانگین درآمد افراد جامعه با نمونهٔ ۱۰۰ نفری چند برابر انحراف معیار با نمونهٔ ۱۰۰۰۰ نفری است؟ **۱۰ برابر**
- اگر اندازهٔ نمونه ۱۰ برابر شود، انحراف معیار برآورد میانگین چند برابر می‌شود؟ $\frac{۱}{\sqrt{۱۰}}$ **برابر می شود زیرا :**

$n \to ۱۰n$

$\sqrt{n} \to \sqrt{۱۰} \times \sqrt{n}$

$\frac{\sigma}{\sqrt{n}} \to \frac{۱}{\sqrt{۱۰}} \times \frac{\sigma}{\sqrt{n}}$

$\sigma_{\bar{x}} \to \frac{۱}{\sqrt{۱۰}} \times \sigma_{\bar{x}}$

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 (بازه اطمینان[1] پارامتر جامعه) عبارت است از بازه‌ای عددی برای پارامتر به همراه درصدی اطمینان که به **ضریب اطمینان**[2] شهرت دارد.

فعالیت

اکنون جامعه را کل افراد کشور و ۶ داده ی کار درکلاس ابتدای درس را یک نمونه در نظر بگیرید . در این کار در کلاس 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 ... n ... و ... σ ... بستگی داشت. اگر ... n ... زیاد می‌شد، یا ... σ ... 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 μ میانگین جامعه و σ انحراف معیار جامعه است.

(معمولاً σ را نداریم و انحراف معیار نمونه ی انتخابی را که برآوردی از σ است در محاسبه به جای آن قرار می دهیم .)

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

مشاهدات ۰,۵,۲,۱

میانگین نمونه $\bar{x} = ۲$

انحراف معیار نمونه $\sigma = ۱/۸۷$

$$۰/۱۳ \le \mu \le ۳/۷۸$$

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 **کار در کلاس** اوّل خط فقر را برآورد کنید. $\frac{\bar{x}}{۲} = \frac{۲/۵}{۲} = ۱/۲۵$

انحراف معیار جامعه را برآورد کنید.

$$\sigma^۲ = \frac{(۴-۲/۵)^۲+(۱-۲/۵)^۲+(۰-۲/۵)^۲+(۳-۲/۵)^۲+(۵-۲/۵)^۲+(۲-۲/۵)^۲}{۶} = ۲/۹۱۶ \Rightarrow \sigma = \sqrt{۲/۹۱۶} = ۱/۷۰۸$$

اگر فرض کنیم که انحراف معیار به دست آمده انحراف معیار جامعه است، یک برآورد فاصله‌ای برای خط فقر محاسبه کنید.

$$\frac{۵}{۲} - \frac{۲\times ۱/۷۰۸}{\sqrt{۶}} \le \mu \le \frac{۵}{۲} + \frac{۲\times ۱/۷۰۸}{\sqrt{۶}} \xrightarrow{\div ۲} ۰/۵۵ \le \frac{\mu}{۲} \le ۱/۹۵$$

فعالیت

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؟

ملاسعیدی- زیر نظر استاد انارکی(اصفهان)
آیدی تلگرام و ایتا: @sinxcosx
وبلاگ: sinxcosx.blogfa.com
سایت: sinxcosx.ir
تلفن: 09168324500

۱ـ Confidence Interval

۲ـ Cofidence 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 **اعضای نمونه** 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 بیش از ۹۵ درصد:

اگر از جامعه‌ای نسبتاً بزرگ، **نمونه ای** تصادفی ساده **با اندازه ی n** انتخاب کنیم و m تا از آنها ویژگی مورد مطالعه ما را داشته باشند، آن‌گاه نسبت واقعی افرادی از جامعه که آن ویژگی را دارند، با اطمینان **بیش از** ۹۵ درصد، در بازهٔ زیر است:

$$\left(p-۲\sqrt{\frac{p(۱-p)}{n}}\ ,\ p+۲\sqrt{\frac{p(۱-p)}{n}}\right)$$

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

توجه داشته باشید که **طول** بازه اطمینان وابسته به **اندازه** نمونه، یعنی عدد n، است، ولی ارتباطی با تعداد اعضای جامعه ندارد. **با افزایش ضریب رادیکال (مقدار عدد ۲) و بدون تغییر اندازه ی نمونه ، می توان با اطمینان بیشتری نسبت را برآورد کرد.**

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 **بیش از** ۹۵ درصد مدّ نظر است.

حل. طبق فرض‌های مسئله **اندازه** 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 **بیش از** ۹۵ درصد می‌توان گفت، نسبت مورد نظر بین **۶۲/۵** درصد و **۸۷/۵** درصد است. طول این بازه ۲۵ درصد است.

اگر بخواهیم به‌دقت بهتری برسیم باید **اندازه نمونه را بزرگتر 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 **اندازه** نمونه را k برابر کنیم، طول بازهٔ اطمینان تقسیم بر $\sqrt{k}$ می‌شود. پس اگر **اندازه** نمونه را ۴ برابر کنیم طول بازهٔ اطمینان نصف می‌شود و اگر **اندازه** نمونه را صد برابر کنیم، دقت محاسبه نسبت، یک رقم اعشار بهتر خواهد شد.

کار در کلاس

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 **بیش از** ۹۵ درصدی برای درصد شرکت‌کنندگان در انتخابات به دست آورید.

حل : در این مسئله $n=۱۰۰۰$ و $p=\frac{۷۰۰}{۱۰۰۰}=۰/۷$. باید عبارت زیر را نیز محاسبه کنیم :

$$۲\sqrt{\frac{p(۱-p)}{n}}=۲\sqrt{\frac{۰/۷(۱-۰/۷)}{۱۰۰۰}}=۰/۰۲۹$$

پس با اطمینان **بیش از** ۹۵ درصد مشارکت در انتخابات بین $(۰/۷-۰/۰۲۹)\times۱۰۰=۶۷/۱$ درصد و $(۰/۷+۰/۰۲۹)\times۱۰۰=۷۲/۹$ 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 **بیش از** 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 $p\cong۰/۵$ 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 بازهٔ اطمینان **بیش از** ۹۵ درصدی، **کمتر از** یک درصد باشد باید n را آن‌قدر بزرگ بگیریم که $\frac{۲}{\sqrt{n}}$ از ۰/۰۱ کمتر شود؛ یعنی $\frac{۲}{\sqrt{n}}\leq۰/۰۱$ و این یعنی $n\geq۲۰۰^۲=۴۰۰۰۰$.

در حالت کلی، قبل از آمارگیری برآوردی از نسبت مورد مطالعه نداریم. پس اگر بخواهیم به‌دقت خاصی برسیم، چگونه باید از پیش بفهمیم که به‌نمونه‌ی **چند عضوی** 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 $p(۱-p)\leq۰/۵(۱-۰/۵)=۰/۵^۲$

لذا به این نتیجه می‌رسیم :

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

$$\left(p-\frac{۱}{\sqrt{n}}\,,\,p+\frac{۱}{\sqrt{n}}\right)$$

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.

۵	۸	۹	۱۱	۱۲	۳	۷	۵	۲	۹

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 **سن** دانشجویان این دانشگاه برابر ۱/۹ سال باشد، بازهٔ اطمینان **بیش از** 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 **میانگین جامعه** اندازهٔ نمونه افزایش یابد، طول فاصلهٔ اطمینان ..... می‌یابد. چرا؟

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

۵ داده‌های زیر نمرات ۲۴ دانش‌آموز از ۱۰۰ **امتیاز** 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.

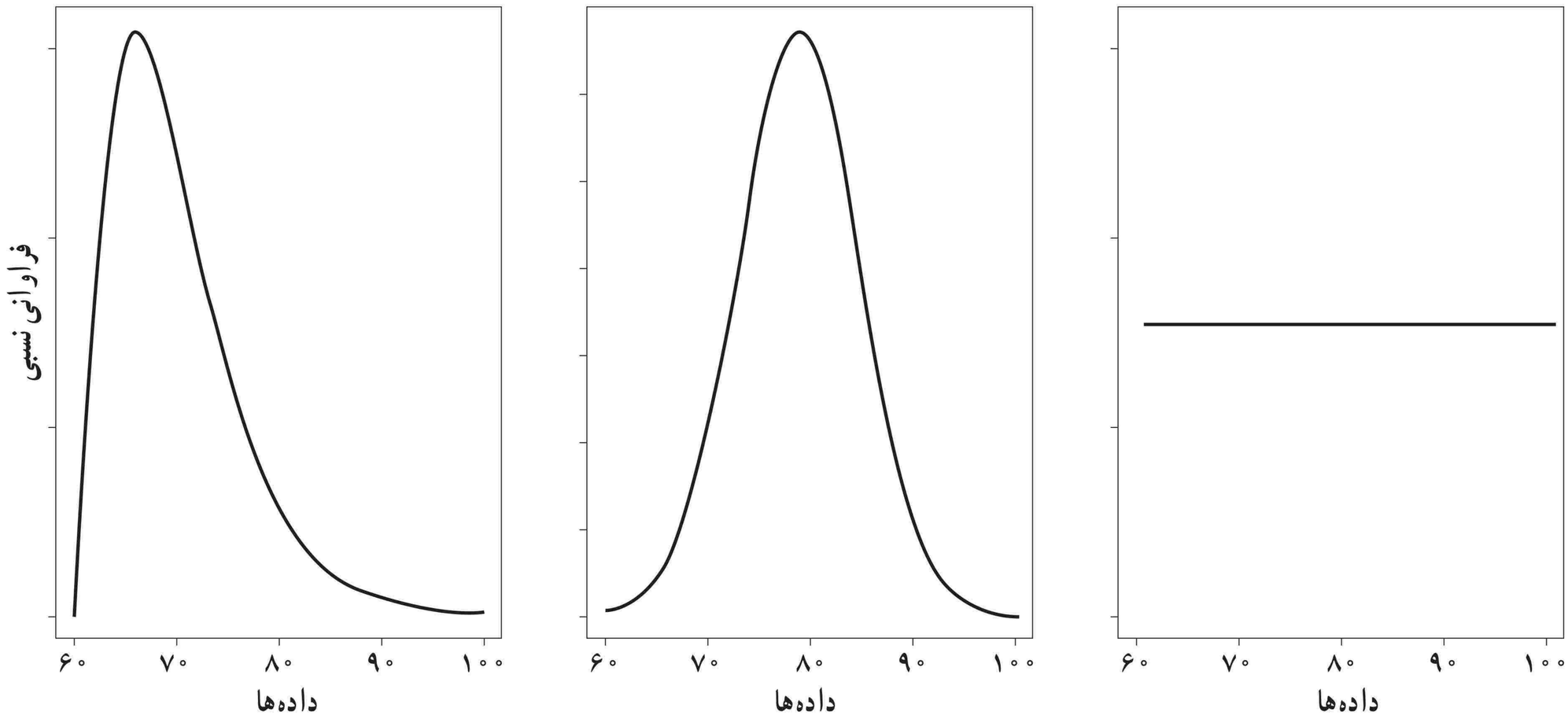
ب) اگر انحراف معیار جامعه ۶ باشد بازه اطمینان ۹۵ درصد برای میانگین نمرات جامعه محاسبه کنید.

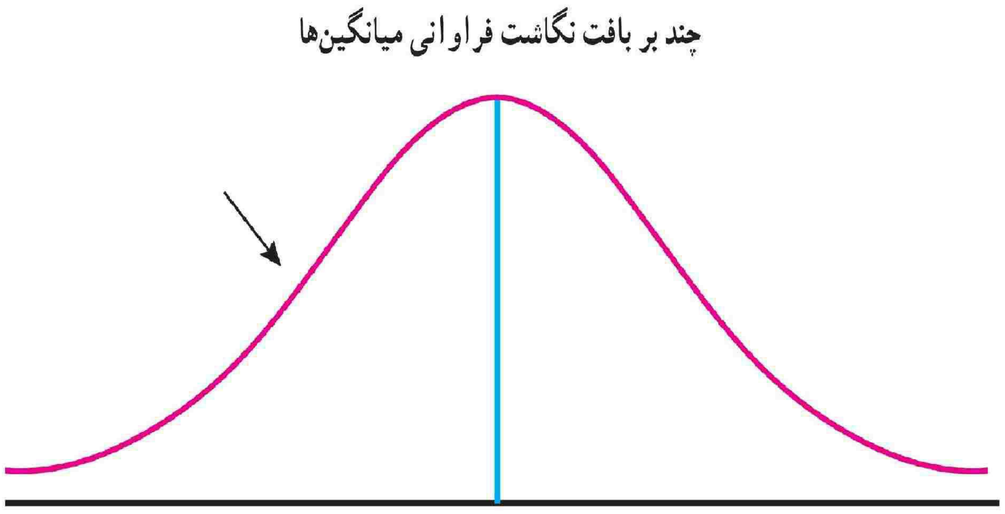
پ) چند درصد داده‌ها داخل این بازه قرار می‌گیرند؟

ت) بافت نگاشت فراوانی داده‌ها را رسم کنید. (در فواصل (۶۷,۷۱] و (۷۱,۷۵] و ... و[۸۷و۹۱])

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 **منحنی** فراوانی بافت نگاشت ۱۰۰ میانگین را رسم کنیم می‌توان نشان داد که تقریباً به صورت یک منحنی به شکل **روبرو** است (توجه کنید منظور از ۱۰۰ عددی بزرگ است، ۱۰۰ یک مثال است). در این شکل μ نشان دهندهٔ میانگین جامعه است، که در اینجا میانگین نمرات همهٔ دانش آموزان است، که مجهول است. حال فرض کنید برای ۱۰۰ نمونهٔ ۲۴ تایی، ۱۰۰ بازهٔ اطمینان ۹۵ درصدی محاسبه کرده‌ایم. در **صفحه بعد** 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؟ **میانگین در بازه ی آن نمونه وجود ندارد**

ب) چند درصد از ۲۰ پاره‌خط آبی، خط سیاه را قطع کرده‌اند؟ $\frac{۱۹}{۲۰} \times ۱۰۰ = ٪۹۵$

پ) اگر ۱۰۰ پاره‌خط آبی را رسم می‌کردیم، انتظار داشتید چند تا از آنها خط سیاه را قطع نکنند؟ **۱۰۰-۹۵ = ۵**

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 ...۹۵... از آنها پارامتر میانگین جامعه را در بر گیرند.