

نمونه گیری و توزیع های نمونه گیری

نمونه گیری و توزیع های نمونه گیری

■ نمونه گیری تصادفی ساده

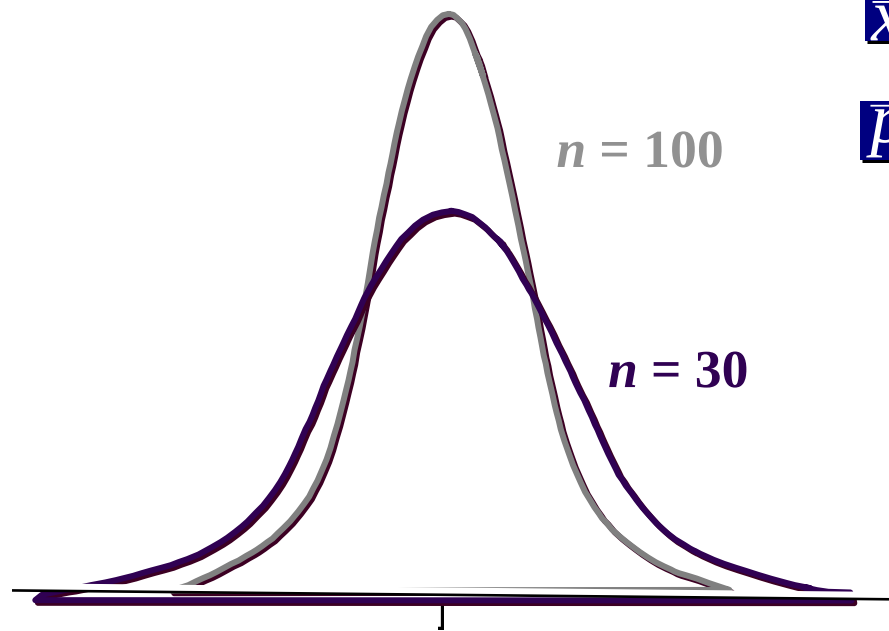
■ برآورد نقطه ای

■ مقدمه ای بر توزیع نمونه گیری

■ توزیع نمونه گیری برای $\bar{x}$

■ توزیع نمونه گیری برای $\bar{p}$

■ روشهای نمونه گیری



استنباط آماری

- مقصود از استنباط آماری بدست آوردن اطلاعات در باره جامعه از اطلاعات نمونه است.
- یک جامعه، مجموعه ای از تمام واحدهای مورد علاقه است.
- یک نمونه یک زیر مجموعه از جامعه است.
- نتایج نمونه فقط برآوردهایی از مقادیر صفات ویژه جامعه را تهیه می کند.
- یک پارامتر، یک صفت ویژه عددی، یک جامعه است.
- با پیشرفت روشهای نمونه گیری، نتایج نمونه، برآوردهای خوبی از صفات ویژه جامعه را فراهم کردند.

نمونه گیری تصادفی ساده

■ جامعه متناهی

- یک نمونه تصادفی ساده از یک جامعه متناهی با اندازه N ، نمونه ای است که انتخاب هر نمونه ممکن با اندازه n دارای احتمال انتخاب یکسان باشد.
- جایگذاری کردن هر واحد نمونه قبل از انتخاب واحدهای بعدی، نمونه گیری با جایگذاری نامیده می شود.

نمونه گیری تصادفی ساده

■ جامعه متناهی

- نمونه گیری بدون جایگذاری روشی است که در اغلب موارد استفاده می شود.
- در طرح های بزرگ نمونه گیری، تولید اعداد تصادفی کامپیوتری اغلب برای عمل انتخاب نمونه استفاده می شود.

نمونه گیری تصادفی ساده

■ جامعه نا متناهی

- یک نمونه تصادفی ساده از یک جامعه نا متناهی، نمونه ای است که انتخاب آن در شرایط زیر صدق کند .
- هر واحد از همان جامعه انتخاب می شود .
- هر واحد به طور مستقل انتخاب می شود.

نمونه گیری تصادفی ساده

■ جامعه نا متناهی

- اگر تهیه لیستی از واحدها یا شمارش واحدها غیر ممکن باشد معمولاً جامعه نا متناهی فرض می شود.
- روش انتخاب عدد تصادفی نمی تواند برای جامعه های نا متناهی استفاده شود.

برآورد نقطه ای

- در برآورد نقطه ای ما از داده های نمونه برای محاسبه مقدار آماره نمونه استفاده می کنیم، که به عنوان برآورد یک پارامتر جامعه به کار می رود (روش گشتاوری).
- میانگین نمونه را به عنوان برآورد نقطه ای میانگین جامعه می نامیم.
- انحراف معیار نمونه برآورد نقطه ای، انحراف معیار جامعه است.
- نسبت نمونه برآورد نقطه ای نسبت جامعه است.

برآورد نقطه ای

- زمانی که مقدار مورد انتظار یک برآورد نقطه ای با پارامتر جامعه یکسان باشد، برآورد نقطه ای نا اریب گفته می شود.

خطای نمونه گیری

- قدر مطلق تفاوت بین یک برآورد نقطه ای نا اریب و پارامتر جامعه مطابق آن را خطای نمونه گیری نامیده می شود.
- خطای نمونه گیری، نتیجه استفاده از یک زیر مجموعه از جامعه (نمونه) برای ساختن برآوردها است نه تمام جامعه.
- خطای نمونه گیری:

برای میانگین نمونه $|\bar{x} - \mu|$

برای انحراف معیار نمونه $|s - \sigma|$

برای نسبت نمونه $|\bar{p} - p|$

مثال: کالج

کالژی سالیانه ۹۰۰ درخواست از دانش آموزان متقاضی دریافت می کند. فرم های درخواستی شامل اطلاعات متنوعی شامل نمرات امتحان ورودی دانش آموزان و اینکه آیا فرد خواهان خوابگاه دانشجویی است یا نه، می باشد.

مثال: کالج

مدیر پذیرش مایل است اطلاعات زیر را بداند:

- متوسط نمرات امتحان ورودی
- نسبت متقاضیان خواهان خوابگاه دانشجویی

مثال: کالج

سه روش زیر را برای بدست آوردن اطلاعات مورد نیاز شرح می دهیم.

- انجام سرشماری از تمامی 900 متقاضی
- انتخاب یک نمونه 30 تائی از متقاضیان با استفاده از جدول اعداد تصادفی
- انتخاب یک نمونه 30 تائی از متقاضیان با استفاده از اعداد تصادفی تولیدی کامپیوتر

مثال: کالج

■ سر شماری از 900 متقاضی

- امتیازات نمرات امتحان ورودی دانش آموزان
- میانگین جامعه

$$\mu = \frac{\sum x_i}{900} = 990$$

- انحراف معیار جامعه

$$\sigma = \sqrt{\frac{\sum (x_i - \mu)^2}{900}} = 80$$

مثال: کالج

■ سر شماری از 900 متقاضی

• متقاضیان خوابگاه دانشجویی

• نسبت جامعه

$$p = \frac{648}{900} = .72$$

مثال: کالج

■ انتخاب نمونه 30 تائی با استفاده از جدول اعداد تصادفی

از جامعه متناهی شامل 900 واحد، ما نیازمند اعداد تصادفی 3-رقمی برای انتخاب تصادفی درخواست هایی که از 1 تا 900 شماره گذاری شده است، هستیم.

مثال: کالج

■ انتخاب نمونه 30 تائی با استفاده از جدول اعداد تصادفی

اعدادی که انتخاب می شوند، شماره متقاضیانی می باشند که می خواهیم از آنها نمونه بگیریم، مگر اینکه:

• اعداد تصادفی بزرگتر از 900 باشد، یا

• اعداد تصادفی باشند که قبلاً انتخاب شده اند

ما به انتخاب اعداد تصادفی ادامه می دهیم تا زمانی که، نمونه 30 تائی تکمیل شود.

مثال: کالج

■ استفاده از اعداد تصادفی برای نمونه گیری

اعداد تصادفی	شماره متقاضیانی که در نمونه قرار می گیرند
شماره	
744	No. 744
436	No. 436
865	No. 865
790	No. 790
835	No. 835
<u>902</u>	متجاوز از 900
190	No. 190
<u>436</u>	شماره ای که قبلا استفاده شده
etc.	etc.

مثال: کالج

■ داده های نمونه

<u>No.</u>	اعداد تصادفی	نام درخواست کننده	نمره ورودی SAT	خانه دانشجویی
1	744	Connie Reyman	1025	Yes
2	436	William Fox	950	Yes
3	865	Fabian Avante	1090	No
4	790	Eric Paxton	1120	Yes
5	835	Winona Wheeler	1015	No
.	.	.	.	.
30	685	Kevin Cossack	965	No

مثال: کالج

- انتخاب نمونه 30 تائی با استفاده از اعداد تصادفی تولیدی کامپیوتر
- برنامه excel یک تابع برای تولید اعداد تصادفی در کار برگ مهیا می کند.
- برای هر متقاضی در جامعه از 1 الی 900 عدد تصادفی تولید می کنیم.
- سپس 30 متقاضی را مطابق با 30 عدد تصادفی تولید شده به عنوان نمونه انتخاب می کنیم.
- هر 900 متقاضی دارای احتمال یکسان برای انتخاب دارند.

با استفاده از Excel یک نمونه تصادفی ساده انتخاب می کنیم

■ فرمول کار برگ

	A	B	C	D
1	Applicant Number	SAT Score	On-Campus Housing	Random Number
2	1	1008	Yes	=RAND()
3	2	1025	No	=RAND()
4	3	952	Yes	=RAND()
5	4	1090	Yes	=RAND()
6	5	1127	Yes	=RAND()
7	6	1015	No	=RAND()
8	7	965	Yes	=RAND()
9	8	1161	No	=RAND()

توجه: سطرهای ۱۰ الی ۹۰۱ نشان داده نشده است.

با استفاده از Excel یک نمونه تصادفی ساده انتخاب می کنیم.

■ مقادیر کار برگ

	A	B	C	D
1	Applicant Number	SAT Score	On-Campus Housing	Random Number
2	1	1008	Yes	0.17167
3	2	1025	No	0.16869
4	3	952	Yes	0.77659
5	4	1090	Yes	0.17894
6	5	1127	Yes	0.86877
7	6	1015	No	0.52630
8	7	965	Yes	0.37638
9	8	1161	No	0.32776

توجه: سطرهای ۱۰ الی ۹۰۱ نشان داده نشده است.

با استفاده از Excel یک نمونه تصادفی

ساده انتخاب می کنیم.

■ مقادیر کار برگ (Sorted)

	A	B	C	D
1	Applicant Number	SAT Score	On-Campus Housing	Random Number
2	12	1107	No	0.00027
3	773	1043	Yes	0.00192
4	408	991	Yes	0.00303
5	58	1008	No	0.00481
6	116	1127	Yes	0.00538
7	185	982	Yes	0.00583
8	510	1163	Yes	0.00649
9	394	1008	No	0.00667

توجه: سطرهای ۱۰ الی ۹۰۱ نشان داده نشده است.

مثال: کالج

■ برآورد های نقطه ای

- میانگین نمونه برآورد نقطه ای میانگین جامعه: $\bar{x}$

$$\bar{x} = \frac{\sum x_i}{30} = \frac{29,910}{30} = 997$$

- انحراف معیار نمونه برآورد نقطه ای انحراف معیار جامعه:

$$s = \sqrt{\frac{\sum (x_i - \bar{x})^2}{29}} = \sqrt{\frac{163,996}{29}} = 75.2$$

- نسبت نمونه برآورد نقطه ای نسبت جامعه: $\bar{p}$

$$\bar{p} = 20/30 = .68$$

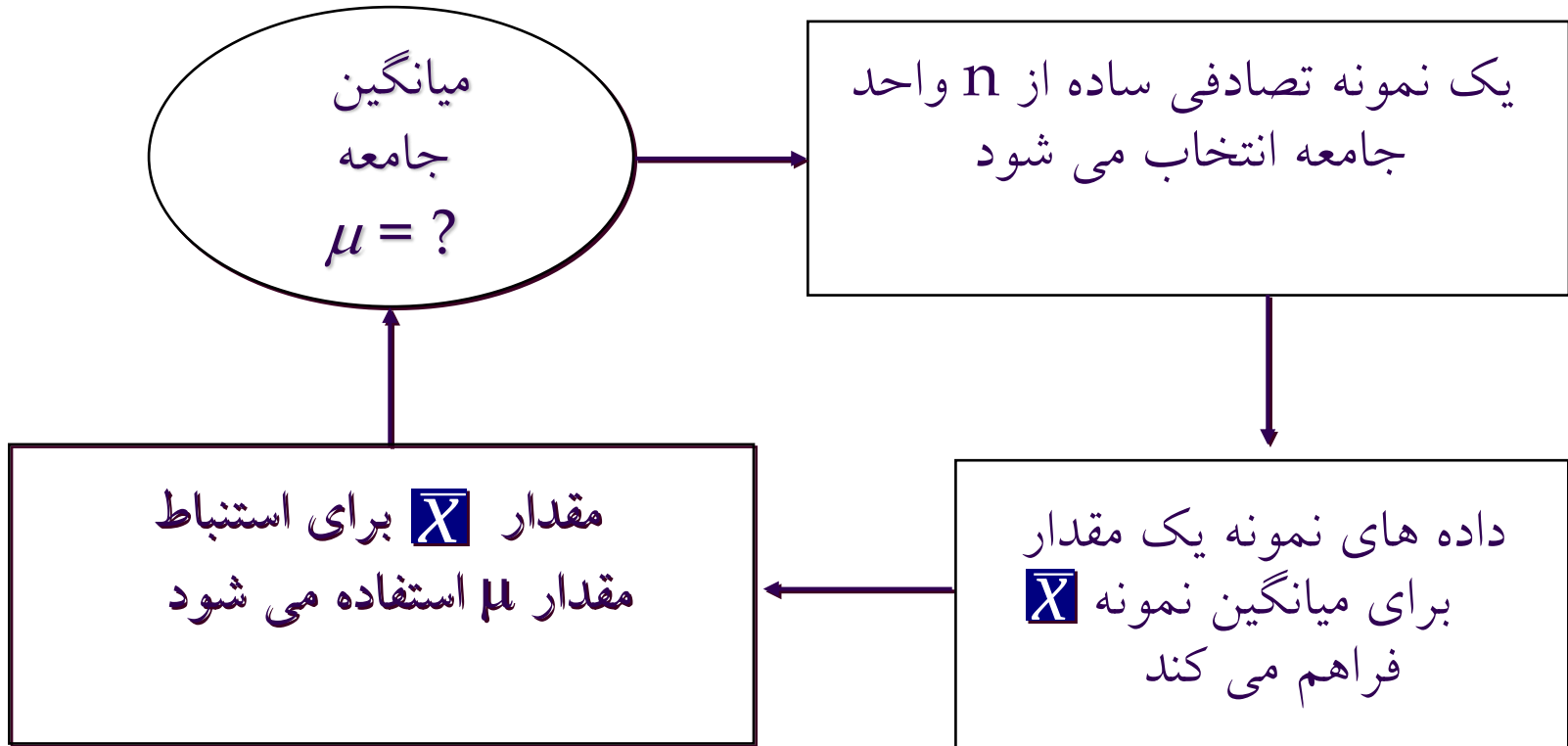
مثال: کالج

■ برآورد های نقطه ای

توجه: اعداد تصادفی متفاوت، نمونه های تصادفی متفاوت را که برآورد های نقطه ای متفاوتی را نتیجه می دهند تشکیل خواهند داد.

توزیع نمونه گیری برای $\bar{x}$

■ مرحله استنباط آماری



توزیع نمونه گیری برای $\bar{X}$

■ توزیع نمونه گیری توزیع احتمال تمام مقادیر ممکن میانگین $\bar{X}$ نمونه است.

■ مقدار مورد انتظار برای $\bar{X}$

$$E(\bar{X}) = \mu$$

که:

μ = میانگین جامعه

توزیع نمونه گیری برای $\bar{X}$

■ انحراف معیار برای $\bar{X}$

جامعه نا متناهی

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}$$

جامعه متناهی

$$\sigma_{\bar{X}} = \left(\frac{\sigma}{\sqrt{n}}\right) \sqrt{\frac{N-n}{N-1}}$$

- یک جمعیت متناهی به نا متناهی تبدیل می شود اگر: $n/N \leq 0.05$
- $\sqrt{(N-n)/(N-1)}$ واحد تصحیح متناهی است.
- از $\sigma_{\bar{X}}$ به عنوان خطای استاندارد برای میانگین استفاده می شود.

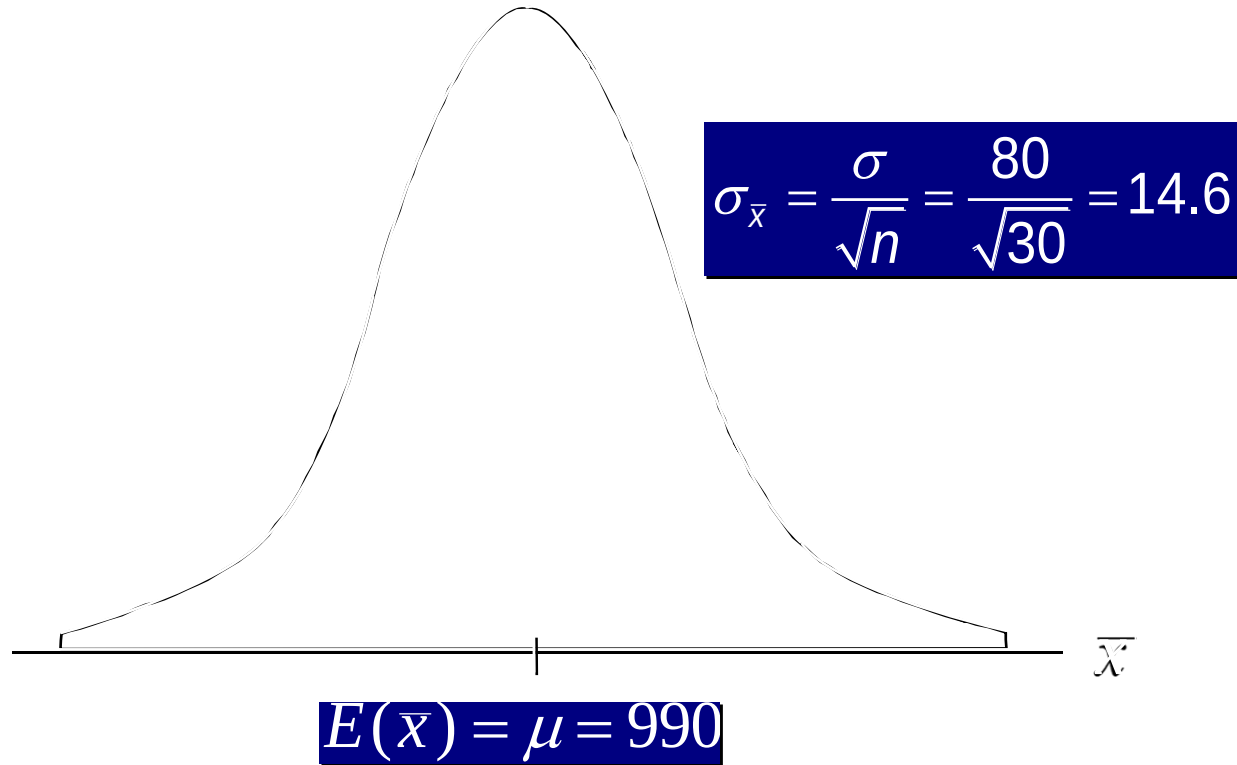
توزیع نمونه گیری برای $\bar{X}$

■ اگر از یک نمونه تصادفی ساده بزرگ ($n \geq 30$) استفاده کنیم. بنا بر قضیه حد مرکزی قادر به این نتیجه گیری هستیم که توزیع نمونه گیری برای $\bar{X}$ را می توان به احتمال نرمال تقریب کرد.

■ زمانی که نمونه تصادفی کوچک باشد ($n < 30$) توزیع نمونه گیری برای $\bar{X}$ را می توان نرمال فرض کرد، تنها اگر فرض کنیم که جامعه دارای یک توزیع احتمال نرمال است .

مثال: کالج

■ توزیع نمونه گیری $\bar{X}$ برای نمرات امتحانی ورودی دانش آموزان .



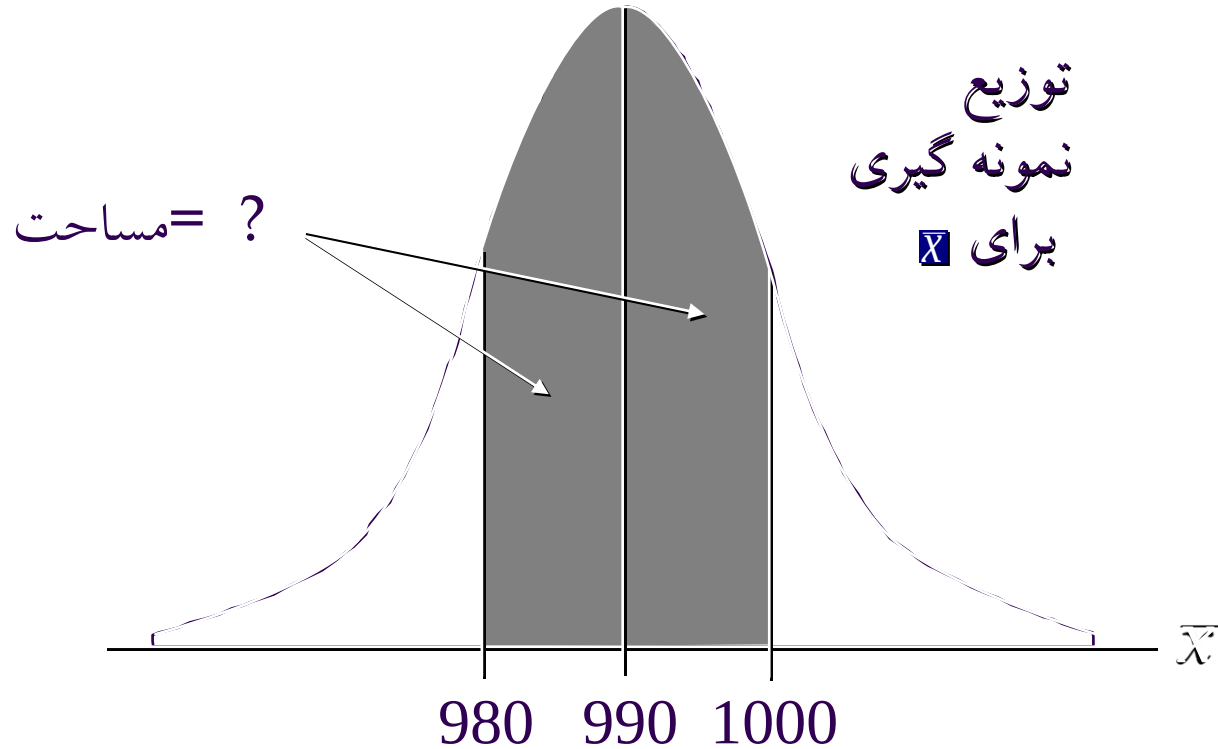
مثال: کالج

■ توزیع نمونه گیری $\bar{X}$ برای نمرات امتحانی ورودی دانش آموزان احتمال اینکه یک نمونه تصادفی ساده 30 تایی، یک برآورد برای میانگین جامعه (نمرات امتحانی ورودی دانش آموزان) تهیه کند که از مقدار واقعی میانگین جامعه، 10 تا بیشتر یا کمتر باشد؟

به عبارت دیگر، چقدر احتمال دارد که $\bar{X}$ بین 980 تا 1000 باشد ؟

مثال: کالج

■ توزیع نمونه گیری $\bar{X}$ برای نمرات امتحانی ورودی دانش آموزان



مثال: کالج

■ توزیع نمونه گیری $\bar{X}$ برای نمرات امتحانی ورودی دانش آموزان

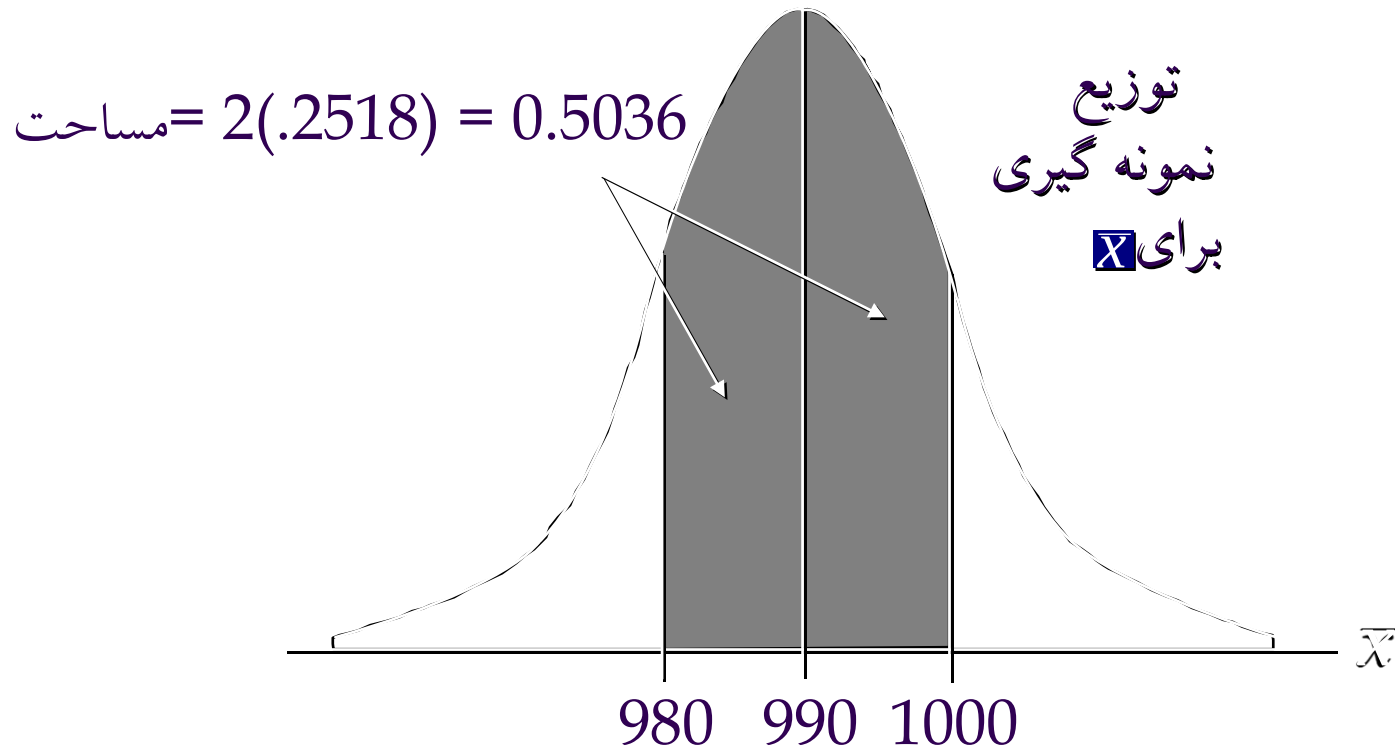
با استفاده از جدول احتمال نرمال استاندارد داریم

$$z = 10/14.6 = 0.68, \text{ مساحت می شود } = (2)(.2518) = 0.5036$$

با احتمال 0.5036 میانگین نمونه از میانگین واقعی جمعیت 10 تا بیشتر یا کمتر خواهد بود .

مثال: کالج

■ توزیع نمونه گیری $\bar{X}$ برای نمرات امتحانی ورودی دانش آموزان



توزیع نمونه گیری برای $\bar{p}$

■ توزیع نمونه گیری $\bar{p}$ ، توزیع احتمال تمام مقادیر ممکن نسبت نمونه است.

■ مقدار مورد انتظار برای

$$E(\bar{p}) = p$$

که:

نسبت جمعیت $p =$

توزیع نمونه گیری برای $\bar{p}$

■ انحراف معیار $\bar{p}$

جامعه نا متناهی

$$\sigma_{\bar{p}} = \sqrt{\frac{p(1-p)}{n}}$$

جامعه متناهی

$$\sigma_{\bar{p}} = \sqrt{\frac{p(1-p)}{n}} \sqrt{\frac{N-n}{N-1}}$$

• $\sigma_{\bar{p}}$ انحراف معیار میانگین نسبت نمونه به عنوان خطای استاندارد برای نسبت نامیده می شود.

توزیع نمونه گیری برای $\bar{p}$

■ هر گاه حجم نمونه بزرگ باشد توزیع نمونه گیری $\bar{p}$ را میتوان به یک توزیع احتمال نرمال تقریب کرد.

■ حجم نمونه را بزرگ فرض می کنیم هر گاه این شرایط برقرار باشد:

$$np \geq 5$$

و

$$n(1 - p) \geq 5$$

توزیع نمونه گیری برای $\bar{p}$

- برای مقادیر p نزدیک به 0.5، نمونه با حجمی به کوچکی 10 می تواند دارای یک تقریب نرمال باشد.
- برای مقادیر p که بسیار کوچک (نزدیک به 0) یا بسیار بزرگ (نزدیک به 1) هستند، نیازمند به نمونه های بسیار بزرگ هستیم.

مثال: کالج

■ توزیع نمونه گیری $\bar{p}$ برای وضعیت اقامت در خوابگاه

توزیع احتمال نرمال یک تقریب قابل قبول است زیرا:

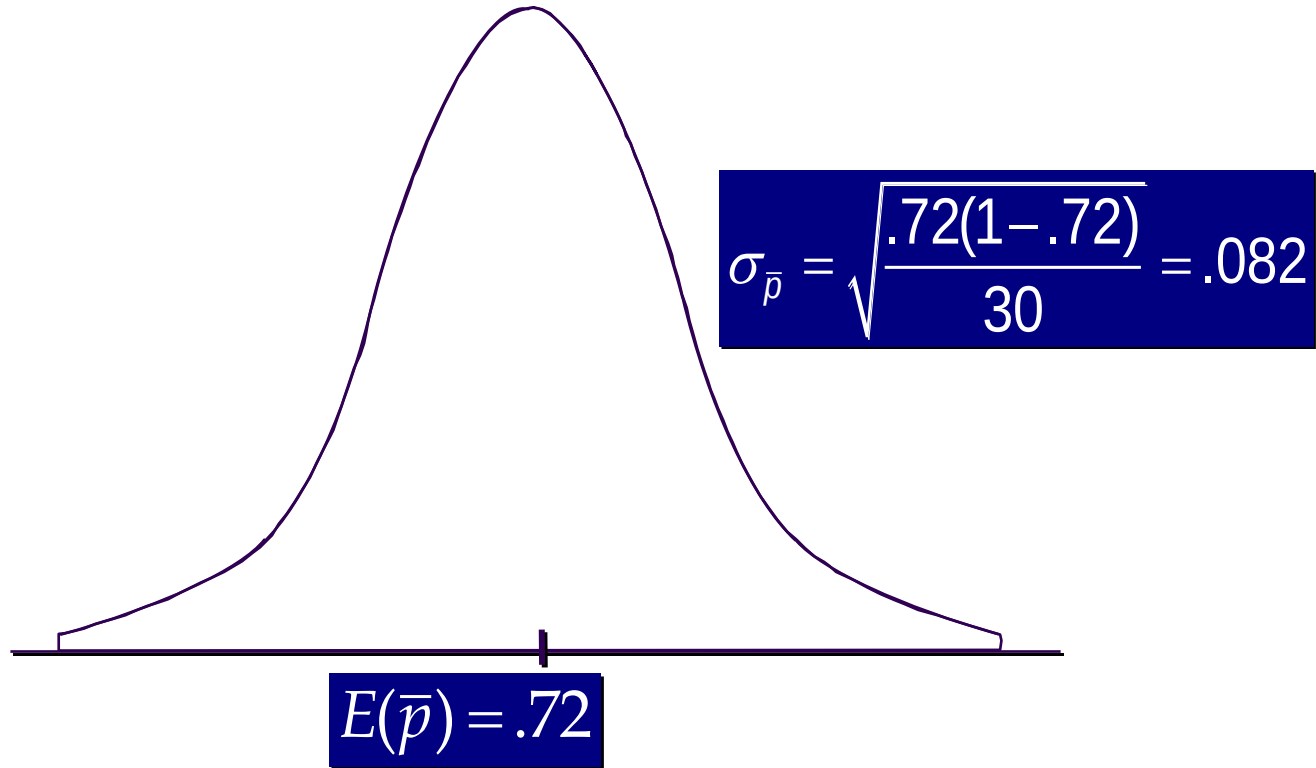
$$np = 30(.72) = 21.6 \geq 5$$

و

$$n(1 - p) = 30(.28) = 8.4 \geq 5$$

مثال: کالج

■ توزیع نمونه گیری برای وضعیت اقامت در خوابگاه



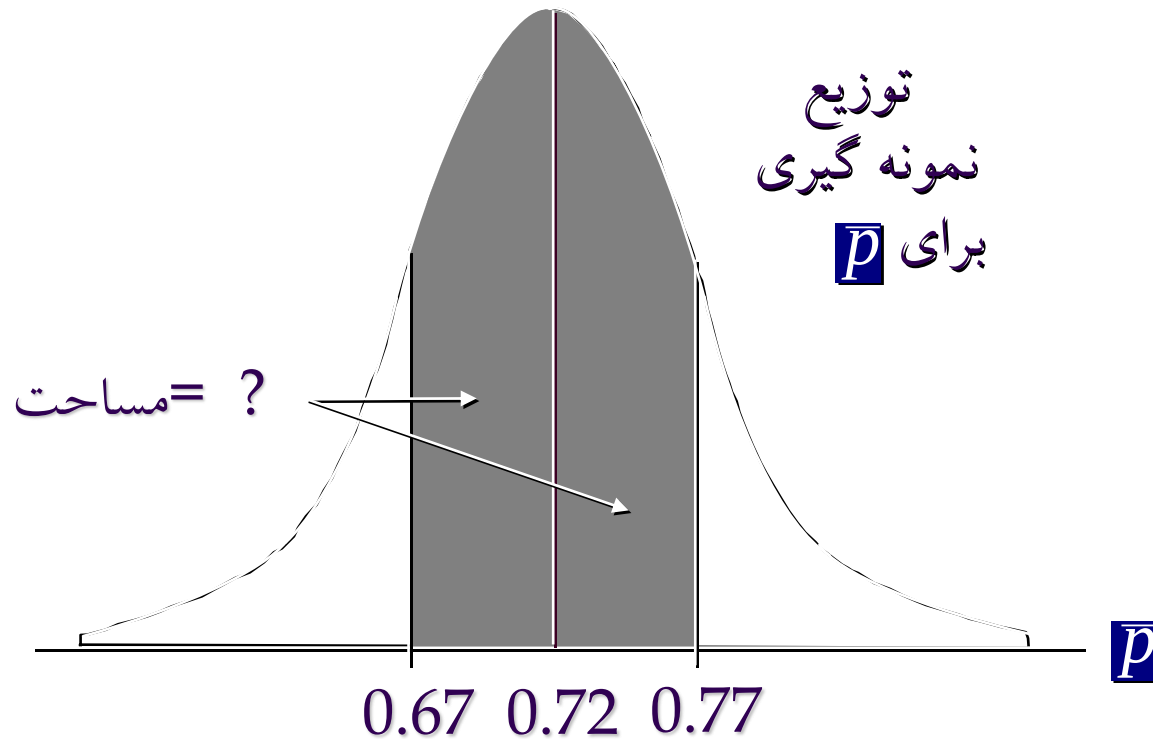
مثال: کالج

■ توزیع نمونه گیری $\bar{p}$ برای وضعیت اقامت در خوابگاه احتمال اینکه نسبت یک نمونه تصادفی ساده شامل 30 متقاضی، یک برآورد برای نسبت جامعه متقاضی های خوابگاه دانشجویی فراهم کند که از نسبت واقعی جامعه به اندازه 0.5 کمتر یا بیشتر باشد؟

به عبارت دیگر احتمال اینکه $\bar{p}$ بین 0.67 و 0.77 قرار بگیرد، چقدر است؟

مثال: کالج

■ توزیع نمونه گیری $\bar{p}$ برای وضعیت اقامت در خوابگاه



مثال: کالج

■ توزیع نمونه گیری $\bar{p}$ برای وضعیت اقامت در خوابگاه

برای

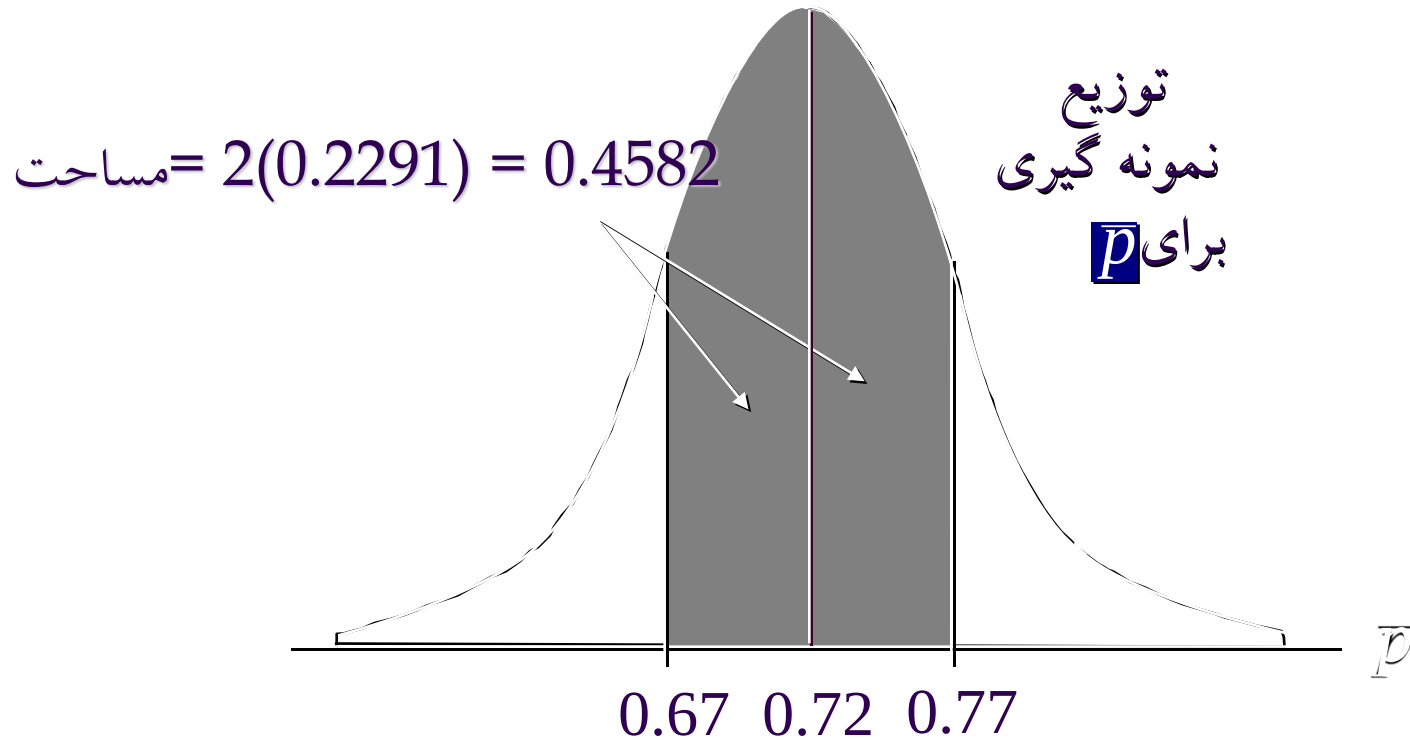
$$z = 0.05 / 0.082 = 0.61$$

$$\text{مساحت} = (0.2291)(2) = 0.4582.$$

با احتمال 0.4582 نسبت نمونه، از نسبت واقعی جامعه به میزان 0.5 کمتر یا بیشتر است .

مثال: کالج

■ توزیع نمونه گیری $\bar{p}$ برای وضعیت اقامت .



روش های نمونه گیری

■ نمونه گیری تصادفی طبقه بندی شده

■ نمونه گیری خوشه ای

■ نمونه گیری سیستماتیک

■ نمونه گیری متناسب

■ نمونه گیری با نظر شخصی

نمونه گیری تصادفی طبقه بندی شده

- در ابتدا جامعه به گروه هایی از واحدها تقسیم می شوند که طبقه نامیده می شود .
- هر واحد درون جامعه متعلق به یکی و فقط یکی از طبقه ها است.
- بهترین نتیجه زمانی بدست می آید که واحد درون هر طبقه تا حد امکان یکسان باشد. (به عبارت دیگر، گروه همگن باشد)
- یک نمونه تصادفی ساده از هر طبقه گرفته می شود.
- فرمول ها، برای دسته بندی نتایج نمونه طبقاتی، برای برآورد پارامتر جامعه، قابل استفاده می باشند.

نمونه گیری تصادفی طبقه بندی شده

- مزیت: اگر طبقات همگن باشد، این روش به اندازه نمونه گیری تصادفی ساده دقیق است اما با یک حجم نمونه کوچکتر.
- مثال: پایه و اساس برای طبقه بندی می تواند: رشته تحصیلی، موقعیت، سن، نوع صنعت و ... باشد .

نمونه گیری خوشه ای

- در ابتدا جامعه را به گروه هایی از واحدهای قابل تفکیک که خوشه نامیده می شود تقسیم می کنیم.
 - به صورت مطلوب هر خوشه یک نماینده از کوچکترین مقیاس تشریح شده در جامعه است. (به عبارت دیگر، گروه ناهمگن)
 - سپس یک نمونه تصادفی ساده از خوشه ها می گیریم.
 - تمام واحدهای خوشه ها که انتخاب شده اند سرشماری می شود.
- ادامه دارد...



نمونه گیری خوشه ای

- مزیت: نزدیکی واحدها می تواند تاثیر گذار باشد. (به عبارت دیگر، مشاهدات نمونه می توانند در یک زمان کوتاه بدست بیایند)
- معایب: این روش، عموماً مستلزم یک نمونه کلی با حجم بزرگتری از نمونه گیری تصادفی ساده یا نمونه گیری تصادفی طبقه بندی شده است.

- مثال : در کاربرد ابتدایی، نمونه گیری در سطحی است که در آن خوشه ها بلوک های شهری یا دیگر مساحت های تعریف شده هستند .

نمونه گیری سیستماتیک

■ اگر یک نمونه به حجم n از جامعه ای شامل N واحد مورد نظر باشد، ممکن است یک واحد را برای هر n/N واحد درون جامعه به عنوان نمونه بگیریم.

■ به طور تصادفی یکی از اولین واحد n/N را از لیست جامعه انتخاب می کنیم.

■ سپس ما هر واحد n/N ام را که از لیست جامعه پیروی می کند انتخاب می کنیم.

■ این روش ویژگی های یک روش نمونه گیری تصادفی ساده را دارد، به خصوص اگر لیست واحدهای جامعه به صورت تصادفی مرتب شده باشند.

ادامه دارد ...



نمونه گیری سیستماتیک

- مزیت: اگر از نمونه گیری سیستماتیک استفاده شود نمونه معمولا راحت تر از نمونه تصادفی ساده تشخیص داده می شود.
- مثال: انتخاب 100 امین لیست از یک دفترچه تلفن بعد از انتخاب تصادفی اولین لیست.

نمونه گیری متناسب

- این یک تکنیک نمونه گیری غیر محتمل است. موارد شامل نمونه هایی هستند که احتمال انتخاب شدن آنها نا معلوم است.
- تشخیص ابتدایی نمونه به وسیله تناسب صورت می گیرد.
- مزیت: انتخاب نمونه و جمع آوری داده ها نسبتاً آسان است.
- معایب: تعیین اینکه نمونه، چگونه نماینده جامعه است غیر ممکن است.
- مثال: یک پرفسور تحقیقی را انجام می دهد که از دانشجویهای داوطلب برای تشکیل یک نمونه استفاده می کند.

نمونه گیری با نظر شخصی

■ شخصی با آگاهی بیشتر در موضوع مورد مطالعه، واحدهایی از جامعه را انتخاب می کند که احساس می کند نماینده بهتری برای جامعه است.

■ این یک تکنیک نمونه گیری غیر محتمل است.

■ مزیت: این یک روش نسبتاً آسان برای انتخاب یک نمونه است.

■ معایب: کیفیت نتایج نمونه وابسته به قضاوت شخص انتخاب کننده نمونه است.

■ مثال: یک خبرنگار نمونه های سه تایی یا چهار تایی از سناتور ها می گیرد تا قضاوت کند که آنها منعکس کننده عقیده مجلس سنا هستند.

پایان

