

اصول و روش های آماری در نمونه گیری

دکتر ید... محرابی
 دانشکده بهداشت
 دانشگاه علوم پزشکی شهید بهشتی
ymehrabi@gmail.com

جامعه آماری

مجموعه ای از افراد یا اشیا که در یک یا چند صفت مشترک بوده و با هدف تحقیق مرتبط باشند.

انواع روش های جمع آوری داده ها

Census سرشماری

List of Units

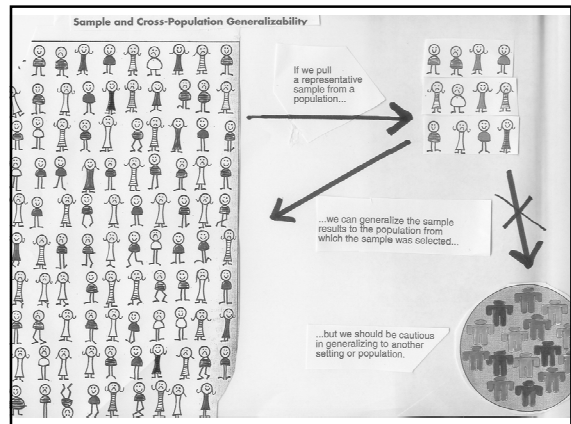
چارچوب نمونه گیری
Sampling Frame

List of Units

Sample

نتایج نمونه

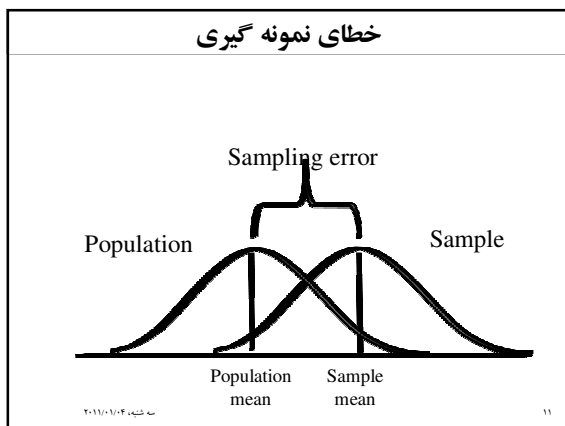
A good sample is
Representative
of the whole
population.



Accuracy and Precision

- Accuracy:
 - How close a particular measure is to the true or correct value.
- Precision:
 - How reproducible or How close replicate measurements become.

Adequate sampling technique helps to ensure that sample measurements become accurate and precise estimate of the population.



برخی خطاهای غیر نمونه گیری

خطای پوشش ناکافی

خطای بی پاسخی
(سوال بدون پاسخ - فرد بدون پاسخ)

خطای اندازه گیری

خطای کدگذاری
خطا در تشکیل فایل کامپیوتری

روشهای متداول نمونه گیری

- نمونه گیری غیر تصادفی (غیر احتمالی)
- نمونه گیری تصادفی (احتمالی)

سه شنبه، ۲۰۱۱/۰۱/۰۴ ۱۳

روشهای نمونه گیری غیر تصادفی

برخی عناصر جامعه آماری شانس انتخاب شدن در نمونه را ندارند

- نمونه گیری آماده
- نمونه گیری سهمیه ای

سه شنبه، ۲۰۱۱/۰۱/۰۴ ۱۴

Quota Sampling, continued...

The problem is that even when we know that a quota sample is representative of the particular characteristics for which quotas have been set, we have no way of knowing if the sample is representative in terms of any other characteristics.

Exhibit 5.8 Quota Sampling

Population
50% male, 50% female
20% white, 30% black

Quota sample
50% male, 50% female

Representative of gender distribution in population, not representative of race distribution.

In Exhibit 5.9, for example, quotas have been set for gender only. Under the circumstances, it's no surprise that the sample is representative of the population only in terms of gender, not in terms of race. Interviewers are only human.

© Free Forge Press, an imprint of Sage Publications, 2004

سه شنبه، ۲۰۱۱/۰۱/۰۴ ۱۵

نمونه گیری تصادفی (احتمالی)

در نمونه گیری احتمالی هر فرد جامعه آماری دارای شانس معین و غیر صفر جهت انتخاب شدن در نمونه می باشد.

لذا بر اساس نتایج حاصل از نمونه می توان با اعتماد قابل اندازه گیری درباره پارامترهای جامعه قضاوت نمود.

سه شنبه، ۲۰۱۱/۰۱/۰۴ ۱۶

انواع روش های نمونه گیری تصادفی (احتمالی)

Simple Random Sampling (SRS)	۱_ نمونه گیری تصادفی ساده
Stratified Sampling	۲_ نمونه گیری طبقه ای
Cluster Sampling	۳_ نمونه گیری خوشه ای
Systematic Sampling	۴_ نمونه گیری سیستماتیک (منظم)
Multistage Sampling	۵_ نمونه گیری چند مرحله ای

سه شنبه، ۲۰۱۱/۰۱/۰۴ ۱۷

نمونه گیری تصادفی ساده (SRS)

در هر مرحله از انتخاب نمونه کلیه افراد:

- ۱_ شانس یکسان جهت انتخاب شدن در نمونه داشته باشند.

- ۲_ نمونه ها مستقل از هم باشند. (انتخاب هر فرد از جامعه بر انتخاب یا عدم انتخاب سایر افراد هیچ تاثیری نداشته باشد).

سه شنبه، ۲۰۱۱/۰۱/۰۴ ۱۸

انتخاب نمونه تصادفی ساده

• با کمک جدول اعداد تصادفی می توان شانس یکسان به افراد جامعه جهت انتخاب در نمونه داد.

• روش استفاده از این جدول بدین ترتیب است که :

بطور تصادفی از نقطه ای از جدول شروع به خواندن و یادداشت کردن اعداد می کنیم تا نمونه مورد نظر تامین گردد.

تعداد ارقام اعداد تصادفی که در جدول خوانده می شود با تعداد ارقام شماره آخرین فرد از جامعه مساوی باشد.

۱۹

جدول ارقام تصادفی

۲۰

جدول ارقام تصادفی

01703	49894	57579	98505	85008	98681	56862	41860
87556	95669	39885	31669	31460	96413	84398	31562
84254	60541	73290	54685	80208	77044	14771	33378
12429	43566	32578	38935	75460	98133	18386	12417
63055	26768	63609	92424	50808	95416	12795	50787
18348	79628	05778	72095	90754	90430	00791	38023
19827	95727	02372	23485	54372	89732	67768	72151
30236	52309	99971	44890	28522	92140	40703	16888
32160	42795	04959	73840	99110	07527	73725	19291
14832	30334	18047	38712	32931	85481	15378	25011
21151	02668	44154	95153	63213	70014	67531	52581
89677	82090	42211	75118	36233	25131	13314	33063
67129	12388	41678	51286	80948	91599	52652	02519
27808	23807	25424	35877	96308	45847	88287	88419
24646	88222	66395	24060	98186	81741	08675	36931
10030	79086	89464	28282	89252	14777	02033	42852
26512	51935	86185	75646	51698	89313	57145	85070
43334	27009	27879	73339	74387	14314	42078	

۲۱

مثال:

می خواهیم نمونه ای به حجم $n = 10$ از هفتاد نفر دانشجویان یک کلاس به صورت تصادفی انتخاب نمایم.

ابتدا به هر یک از دانشجویان شماره ای از ۰۱ تا ۷۰ اختصاص می دهیم.

۰۱ ۰۲ ۰۳ ۶۹ ۷۰

چون شماره آخرین فرد دورقمی است با کمک جدول اعداد تصادفی ده عدد دورقمی انتخاب می نمایم. بدین ترتیب است که ارقام دورقمی بزرگتر از هفتاد در نظر گرفته نمی شود. ارقام تکراری نیز در نظر گرفته نمی شود. به عبارت دیگر:

از یک نقطه تصادفی شروع کرده و ده عدد دورقمی غیر تکراری بین ۰۱ تا ۷۰ پیدا می کنیم.

۲۲

جدول ارقام تصادفی	01703	49894	57579	98505	85008	98681	56862	41860
	87556	95669	39885	31669	31460	96413	84398	31562
	84254	60541	73290	54685	80208	77044	14771	33378
	12429	43566	32578	38935	75460	98133	18386	12417
	63055	26768	63609	92424	50808	95416	12795	50787
	18348	79628	05778	72095	90754	90430	00791	38023
نقطه شروع	19827	95727	02372	23485	54372	89732	67768	72151
	30236	52309	99971	44890	28522	92140	40703	16888
	32160	42795	04959	73840	99110	07527	73725	19291
	14832	30334	18047	38712	32931	85481	15378	25011
	21151	02668	44154	95153	63213	70014	67531	52581
	89677	82090	42211	75118	36233	25131	13314	33063
	67129	12388	41678	51286	80948	91599	52652	02519
	27808	23807	25424	35877	96308	45847	88287	88419
	24646	88222	66395	24060	98186	81741	08675	36931
	10030	79086	89464	28282	89252	14777	02033	42852
	26512	51935	86185	75646	51698	89313	57145	85070
	43334	27009	27879	73339	74387	14314	42078	

شماره افراد انتخاب شده: 48, 28; 52; 29; 21, 40, 70, 31, 68, 32

۲۳

نمونه گیری طبقه ای Stratified Sampling

✓ استفاده از دانش قبلی برای افزایش دقت نمونه

فرد ۱۳، بین اقلی نام نسبی

۲۴

نمونه گیری طبقه ای Stratified Sampling

زمانی قابل استفاده است که پراکندگی بین
طبقات بیشتر از پراکندگی داخل طبقات باشد

سه شنبه، ۲۰۱۱/۰۱/۰۴

۲۵

نمونه گیری خوشه ای

سه شنبه، ۲۰۱۱/۰۱/۰۴

۲۶

نمونه گیری خوشه ای Cluster Sampling

✓ مقرون به صرفه، زمان کمتر،
✓ زمانی قابل استفاده است که پراکندگی بین خوشه
ها کمتر از پراکندگی داخل خوشه ها باشد

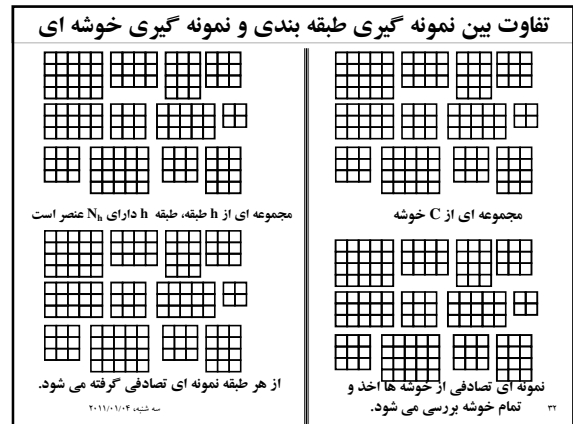
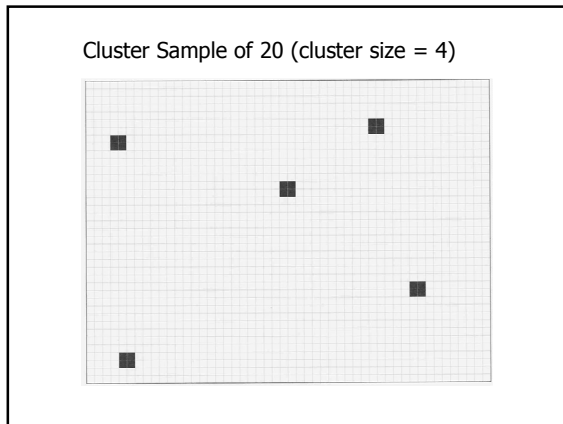
سه شنبه، ۲۰۱۱/۰۱/۰۴

۲۷

Simple Random Sample: $n = 20, N = 2000$

Systematic sample: $n = 20, N = 2000, k = 45$

Stratified sample of 20 from 4 strata

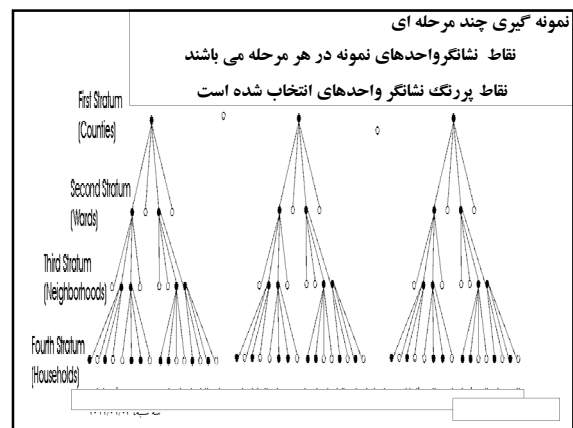


نمونه گیری سیستماتیک (منظم)
Systematic Sampling

حجم جامعه = N حجم نمونه = n فاصله نمونه گیری: $k=N/n$

یک عدد تصادفی بین ۱ و k انتخاب می کنیم مثلاً r . شماره نمونه ها عبارت خواهند بود از:

$r, r+k, r+2k, r+3k, \dots$



تمرین

برای موضوع انتخاب شده در گروه، روش نمونه گیری مناسب را با جزئیات تعیین کنید و حجم نمونه را با اعداد فرضی محاسبه کنید.

حجم نمونه

در مطالعات توصیفی و مقطعی

Sample size in Descriptive and Cross-sectional Studies

اندازه نمونه در پژوهش های توصیفی و مقطعی
Sample size for Descriptive and cross-sectional research

موضوع را در دو حالت مجزا مطرح می کنیم:

۱- اندازه نمونه در مطالعات یک میانگین

۲- اندازه نمونه در مطالعات یک نسبت

سه شنبه، ۲۰۱۱/۰۱/۰۴ ۳۷

فرمول تعیین حجم نمونه
برای برآورد میانگین جامعه

$$n = \frac{z^2 \sigma^2}{d^2}$$

سه شنبه، ۲۰۱۱/۰۱/۰۴ ۳۸

فرمول تعیین حجم نمونه
برای برآورد میانگین جامعه

$$n = \frac{z^2 \sigma^2}{d^2}$$

معمولا انحراف معیار جامعه (σ) معلوم نیست بنابراین انحراف معیار نمونه (از مطالعه مقدماتی یا مطالعات دیگران) به جای آن قرار داده می شود:

$$n = \frac{z^2 s^2}{d^2}$$

سه شنبه، ۲۰۱۱/۰۱/۰۴ ۳۹

$$n = \frac{z^2 s^2}{d^2}$$

n = برآورد تعداد نمونه
 s = برآورد انحراف معیار صفت مورد نظر
 d = حداکثر خطای قابل قبول در برآورد میانگین جامعه
 z = ضریب اطمینان

سه شنبه، ۲۰۱۱/۰۱/۰۴ ۴۰

مثال

برای برآورد میانگین فشارخون سیستولیک (بر حسب میلی متر جیوه) یک جمعیت که انحراف معیار آن طبق مطالعات انجام شده برابر ۱۰ می باشد حداقل تعداد نمونه مورد نیاز با احتمال اشتباه ۰.۰۵ و دقت ۲ میلی متر جیوه چقدر باید باشد؟

سه شنبه، ۲۰۱۱/۰۱/۰۴ ۴۱

مثال

برای برآورد میانگین فشارخون سیستولیک (بر حسب میلی متر جیوه) یک جمعیت که انحراف معیار آن طبق مطالعات انجام شده برابر ۱۰ می باشد حداقل تعداد نمونه مورد نیاز با احتمال اشتباه ۰.۰۵ و دقت ۲ میلی متر جیوه چقدر باید باشد؟

$$n = \frac{z^2 s^2}{d^2} = \frac{(1.96)^2 (10)^2}{(2)^2} = 96$$

سه شنبه، ۲۰۱۱/۰۱/۰۴ ۴۲

مثال:

برای برآورد میانگین اندازه کالری دریافتی دانش آموزان دوره ابتدایی یک شهرستان چه تعداد نمونه لازم است تا با اطمینان ۹۵٪، فاصله بین مقدار برآورد شده با میانگین واقعی کالری دریافتی در جامعه دانش آموزان ابتدایی شهرستان بیش از ۲۰ کالری نباشد؟ در مطالعه ای مشابه انحراف معیار میزان کالری دریافتی دانش آموزان ۷۵ به دست آمده است.

۴۳

مثال:

برای برآورد میانگین اندازه کالری دریافتی دانش آموزان دوره ابتدایی یک شهرستان چه تعداد نمونه لازم است تا با اطمینان ۹۵٪، فاصله بین مقدار برآورد شده با میانگین واقعی کالری دریافتی در جامعه دانش آموزان ابتدایی شهرستان بیش از ۲۰ کالری نباشد؟ در مطالعه ای مشابه انحراف معیار میزان کالری دریافتی دانش آموزان ۷۵ به دست آمده است.

$$n = \frac{z^2 s^2}{d^2} =$$

$$\frac{(1.96)^2 (75)^2}{(20)^2} = 54.02 \approx 55$$

۴۴

فرمول تعیین حجم نمونه

برای برآورد نسبت جامعه (با دقت ثابت)

$$n = \frac{z^2 p(1-p)}{d^2}$$

n= برآورد تعداد نمونه
 p= برآورد نسبت صفت مورد نظر
 d= حداکثر خطای قابل قبول در برآورد نسبت جامعه
 Z= ضریب اطمینان

۴۵

مثال:

فرض کنید می‌خواهیم شیوع یک بیماری را در بین افراد ساکن در یک منطقه برآورد کنیم. بر اساس مطالعات قبلی میزان شیوع بیماری مورد نظر در مناطق مشابه ده درصد بوده است. چه تعداد نمونه لازم است؟ (سطح اطمینان ۹۵٪ و Z=1.96).

۴۶

مثال:

فرض کنید می‌خواهیم شیوع یک بیماری را در بین افراد ساکن در یک منطقه برآورد کنیم. بر اساس مطالعات قبلی میزان شیوع بیماری مورد نظر در مناطق مشابه ده درصد بوده است. چه تعداد نمونه لازم است؟ (سطح اطمینان ۹۵٪ و Z=1.96).

$$n = \frac{z^2 p(1-p)}{d^2}$$

$$= \frac{1.96^2 (0.1)(0.9)}{(0.01)^2} = 3456$$

۴۷

مثال:

مدیریت خدمات بهداشتی در منطقه ای می خواهد برآوردی از نسبت افراد بزرگسال که واکسن هپاتیت دریافت کرده اند را بدست آورد. برای یک نمونه تصادفی چه تعداد افراد باید بررسی شوند تا با اطمینان ۹۵٪ نتایج برآورد در فاصله قابل قبولی از نسبت واقعی جامعه باشد؟

۴۸

مثال:

مدیریت خدمات بهداشتی در منطقه ای می خواهد برآوردی از نسبت افراد بزرگسال که واکسن هپاتیت دریافت کرده اند را بدست آورد. برای یک نمونه تصادفی چه تعداد افراد باید بررسی شوند تا با اطمینان ۹۵٪ نتایج برآورد در فاصله قابل قبولی از نسبت واقعی جامعه باشد؟

$$n = \frac{z^2 p(1-p)}{d^2} = \frac{1.96^2 (0.5)(0.5)}{(0.05)^2} = 384$$

سه شنبه، ۲۰۱۱/۰۱/۰۴

۴۹

مثال:

در مثال قبلی اگر سطح اطمینان را به ۹۹ درصد تغییر دهیم چه تعداد نمونه لازم خواهد بود؟

$$n = \frac{z^2 p(1-p)}{d^2} = \frac{2.57^2 (0.5)(0.5)}{(0.05)^2} = 660$$

سه شنبه، ۲۰۱۱/۰۱/۰۴

۵۰

مقدار Z برای برخی سطوح اطمینان

Z ²	ضریب اطمینان (Z)	سطح خطا	سطح اطمینان
۲.۶۹	۱.۶۴	۰.۱۰	۰.۹۰
۳.۸۴	۱.۹۶	۰.۰۵	۰.۹۵
۶.۶۰	۲.۵۷	۰.۰۱	۰.۹۹

سه شنبه، ۲۰۱۱/۰۱/۰۴

۵۱

Factors Determining Sample Size

- Number of **groups** and subgroups within the sample
- Value of **information** in the study
- **Accuracy** level required in results
- **Cost** of sample
- **Variability** of the population
- Sampling **Method**

سه شنبه، ۲۰۱۱/۰۱/۰۴

۵۲