

Statistics

Unit 6 Sampling

Name _____

Population and sample

The **population** in a statistical study is the entire group of individuals about which we want information.

A **sample** is the part of the population from which we actually collect information. We use information from a sample to draw conclusions about the entire population.

Try This

1. Identify the population and the sample in each of the following settings.
 - a. A furniture maker buys hardwood in large batches. The supplier is supposed to dry wood before shipping. The furniture maker chooses five pieces of wood from each batch and test their moisture content. If any piece exceeds 12% moisture content, the entire batch is sent back.
 - b. Each week, Gallup Poll questions a sample of about 1500 adult US residents to determine national opinion on a wide variety of issues.

How to Sample Badly

Bias

The design of a statistical study shows **bias** if it systematically favors certain outcomes.

Voluntary Response Sample

A **voluntary response sample** consists of people who choose themselves by responding to a general appeal. Voluntary response samples show bias because people with strong opinions (positive or negative) are most likely to respond.

Convenience Sample

Choosing individuals who are easiest to reach results in a **convenience sample**.

Try This

2. For each of the following situations, identify the sampling method used. Then explain how the sampling method could lead to bias.
 - a. A farmer brings a juice company several crates of oranges each week. A company inspector looks at 10 oranges from the top of each crate before deciding whether to buy all the oranges.
 - b. The ABC program *Nightline* once asked whether the United Nations should continue to have its headquarters in the United States. Viewers were invited to call one telephone number to respond “Yes” and another for “No”. There was a charge for calling either number. More than 186,000 callers responded, and 67% said “No.”

How to Sample Well: Random Sampling

Simple Random Sample

A **simple random sample (SRS)** of size n consists of n individuals from the population chosen in such a way that every set of n individuals has an equal chance to be the sample actually selected.

How to Choose an SRS Using a Table of Random Digits

Step 1: Give each member of the population a numerical label of the *same length*.

Step 2: Read consecutive groups of digits of the appropriate length from the table of random digits.

Your sample contains the individuals whose labels you find.

How to Choose an SRS Using a Graphing Calculator

Step 1: Give each member of the population a numerical label (it does not have to be the same length).

Step 2: Use your random integer function on your calculator.

TINspire: Menu, Probability, Random, Integer.

Type in smallest label, highest label, how many you want in the sample.

TI-84: Math, PRB, RandInt

Type in smallest label, highest label, how many you want in the sample.

Try This

3. The school newspaper is planning an article on family-friendly places to stay over spring break at a nearby beach town. The editors intend to call 4 randomly chosen hotels to ask about their amenities for families with children. They have an alphabetized list of all 28 hotels in the town.

Here is the list:

Aloha Kai	Captiva	Palm Tree	Sea Shell
Anchor Down	Casa del Mar	Radisson	Silver Beach
Banana Bay	Coconuts	Ramada	Sunset Beach
Banyan Tree	Diplomat	Sandpiper	Tradewinds
Beach Castle	Holiday Inn	Sea Castle	Tropical Breeze
Best Western	Lime Tree	Sea Club	Tropical Shores
Cabana	Outrigger	Sea Grape	Veranda

a. Use a random digit table to select 4 hotels for the editors to call. (Label them in alphabetical order and start at line 130 in the table.)

b. Use a graphing calculator to select 4 hotels for the editors to call.

Other Sampling Methods

Stratified Random Sample and Strata

To select a **stratified random sample**, first classify the population into groups of similar individuals, called **strata**. Then choose a separate SRS in each stratum and combine these SRSs to form the full sample.

Cluster Sample and Clusters

To take a **cluster sample**, first divide the population into smaller groups. Ideally, these **clusters** should mirror the characteristics of the population. Then choose an SRS of the clusters. *All* individuals in the chosen clusters are included in the sample.

Probability Sample

A **probability sample** gives each member of the population a known chance (greater than zero) to be selected.

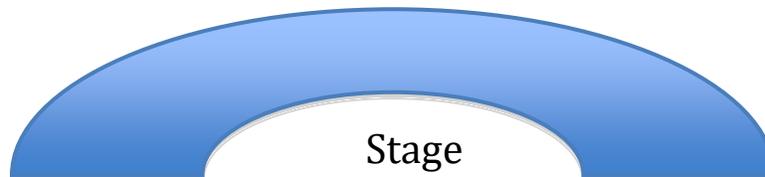
Multistage sample

A combination of two or more sampling methods. Most large scale sample surveys use multistage. Most national surveys are multistage.

Try This

4. The student council wants to conduct a survey during the first five minutes of an all-school assembly in the auditorium about use of the school library. They would like to announce the results of the survey at the end of the assembly. The student council president asks your statistics class to help carry out the survey.

There are 800 students present at the assembly. A map of the auditorium is shown below. Note that students are seated by grade level and that the seats are numbered from 1 to 800.



1	2	3	4	5	6	7	8	9	10
21	22	23	24	25	26	27	28	29	30

11	12	13	14	15	16	17	18	19	20
31	32	33	34	35	36	37	38	39	40

761	762	763	764	765	766	767	768	769	770
781	782	783	784	785	785	787	788	789	790

771	772	773	774	775	776	777	778	779	780
891	792	793	794	795	796	797	798	799	800

9th grade: Seats 601 – 800
11th grade: Seats 201 – 400

10th grade: Seats 401 – 600
12th grade: Seats 1 – 200

Describe how you would use each of the following sampling methods to select 80 students to complete the survey.

- a. Simple random sample
- b. Stratified random sample
- c. Cluster sample

What Can Go Wrong?

Undercoverage

Undercoverage occurs when some groups in the population are left out of the process of choosing the sample.

Nonresponse

Nonresponse occurs when an individual chosen for the sample can't be contacted or refuses to participate.

Response Bias

Response bias occurs when someone gives an incorrect response.

Wording of Questions

Confusing or leading questions can introduce strong bias and changes in wording can greatly change a survey's outcome. Even the order in which questions are asked matters.

Try This

5. Each of the following has a source of error. Label each with the appropriate error.
 - a. The telephone directory is used to choose a sample from.
 - b. An non-church going person is asked how many times they were at church this month, they answer 3 times.
 - c. The person cannot be contacted in five calls.

6. A survey paid for by makers of disposable diapers found that 84% of the sample opposed banning disposable diapers. Here is the actual question:

It is estimated that disposable diapers account for less than 2% of trash in today's landfills. In contrast, beverage containers, third-class mail and yard wastes are estimated to account for about 21% of the trash in landfills. Given this, in your opinion, would it be fair to ban disposable diapers?

Explain how the wording of the question could result in bias. Rewrite the question to eliminate this bias.