1 Basic Ideas

Should We Experiment or Should We Merely Observe?

Example: Does talking on a cell-phone cause brain cancer?

(a) Conduct an **observational study** (i.e., take a **sample** survey) among humans.

- What are the drawbacks to this observational study (**sample** survey)?

(b) Conduct an **experiment** with humans.

- What are the drawbacks to this experiment?

(c) Conduct an **experiment** with mice.

- What are the drawbacks to this experiment?

Example: University of Michigan, May 19, 2003, news report by three social scientists: "Student drug testing not effective in reducing drug use."

www.umich.edu/news/Releases/2003/May03/r051903.html

Study was from 1998–2001, based on 722 secondary schools (497 high schools and 225 middle schools). Students were asked if they used marijuana in the past 12 months, and a school administrator was asked about the drug testing policy. Overall, schools testing for drugs were virtually similar to schools not testing for drugs in

terms of marijuana use among students. The authors implied that drug testing is a waste of money.

(a) What are the *explanatory* and *response* variables?

(b) Was this study *observational* or an *experiment*.

(c) Is the authors' implication valid?

1.1 Sampling

This section focuses on **sampling** (i.e., collecting observational data) rather than **experiments**.

Some Basic Terms in Statistics

The **population** is the total set of subjects in which we are interested.

A **sample** is the subset of the population for whom we have (or plan to have) data.

A Sample Survey and a Census

Example: Suppose we are interested in the population proportion of American adults who support the President's foreign policy in Ukraine?

Ideally, take a _____.

What would be the *disadvantages*?

What would be the *advantages*?

Instead, take a _____.

Definition: A **sample survey** selects a sample of people from a population and interviews them to collect data.

What would be the *advantages*?

What would be the *disadvantages*?

Types of Sampling Methods

Definition: A **simple random sample** of n subjects from a population is one in which each possible sample of that size has the same chance of being selected.

Definition: A **stratified random sample** divides the population into groups called **strata**, and then selects a simple random sample from each stratum.

Example: Suppose that a university is known to be 60% female and 40% male, and a survey is to be conducted related to the abortion issue. Enough funding (or time) is available to sample 100 students.

(a) How would a *simple random sample* be taken?

(b) How would a *stratified random sample* be taken?

(c) Which sample is better and why?

Definition: Divide the population into a large number of **clusters**. Select a simple random sample of the clusters. All elements within each *sampled* cluster are sampled, to form a **cluster random sample**.

Example: You have one week to estimate the average annual church donation of Baptist members in Rhode Island.

(a) How would a *simple random sample* be taken?

(b) How would a *cluster random sample* be taken?

(c) Which sample is better and why?

Definition: Sample of Convenience

What are some examples?

Statistics and Parameters

A **parameter** is a numerical summary of the *population*.

A **statistic** is a numerical summary of a *sample* taken from the population.

1.4 Bias in Studies

This section also focuses on **sampling** (i.e., collecting observational data) rather than **experiments**.

What are Good Ways and Poor Ways to Sample?

Definition: (again!) A **simple random sample** of n subjects from a population is one in which each possible sample of that size has the same chance of

being selected.

Definition: The **sampling frame** is the list of subjects in the population from which the sample is taken.

Example: Sample only American senior citizens when estimating how American adults feel about Social Security issues.

What are the *population* and *sampling frame*?

Would the results of this poll be valid?

□

Types of Bias in Sample Surveys

(1) Sampling bias occurs from using nonrandom samples or having undercoverage.

(2) Nonresponse bias occurs when some sampled subjects cannot be reached or refuse to participate or fail to answer some questions.

(3) Response bias occurs when the subject gives an incorrect response (perhaps lying), or the question wording or the way the interviewer asks the questions is confusing or misleading. (*Response bias* includes *social acceptability bias* and *leading question bias,* as listed in the textbook.)

Example: For a JMU student research project a few years ago, the researchers asked questions similar to the following.

"Do you smoke marijuana?"

"Do you think JMU students smoke marijuana?"

□

Example: The *Literary Digest* Poll *(Know this example in detail, although you need not memorize the numbers.)*

Franklin Roosevelt vs. Alfred Landon, Election of 1936.

Since 1916, the *Literary Digest* correctly picked the Presidents.

Digest mailed questionnaires to 10 million people, whose names were from country club membership lists, phone books, and automobile registrations.

George Gallup, polling 50,000 people, predicted *Digest's* results in advance.

3rd party candidates were excluded in the numbers below.

	Roosevelt's percentage
The election result	62
Digest's prediction	43
Gallup's prediction of *Digest*	44
Gallup's prediction of election	56

□

Example: Thomas Dewey vs. Harry Truman, Election of 1948. *(Know this example in detail, although you need not memorize the numbers.)*

candidates	Crossley	Gallup	Roper	results
Harry Truman	45	44	38	50
Thomas Dewey	50	50	53	45
Strom Thurmond	2	2	5	3
George Wallace	3	4	4	2

A Gallup Poll interviewer in St. Louis was required to interview 13 subjects of whom

- 6 live in suburbs, 7 in central city
- 7 men, 6 women

- AND additional criteria based on age, race, monthly rent.

□

Do we trust the results of **volunteer response samples**, such as internet polls? Why or why not?

Remark: A sample is said to be **representative** of the population, if the survey is **unbiased,** even if the sample size is small.

1.3 Design of Experiments

The three principles of experimental design are

(1) Control

(2) Randomize

(3) Use enough subjects

Definition: In an experiment, the **subjects** are the **experimental units.**

Example: *(Know this example in detail.)* Diethylstibestrol (DES) is an artificial hormone, and in the 1940s was believed to prevent miscarriages.

The rate of miscarriages among pregnant women was known to be some fixed number (for non-DES-users).

In five large studies, pregnant women volunteered to try DES, and the rate of miscarriages for these DES-users was lower than for the non-DES-users.

In each of these five studies, the researchers concluded that DES lowers the rate of miscarriages.

Were the conclusions of these researchers valid?

Even in the late 1960s, doctors were prescribing the drug to 50,000 women each year.

DES was banned for use on pregnant women in 1971. □

Ideally, a **double-blinded** experiment is best.

When is a double-blinded experiment not possible (or not ethical)?

Example: Salk Vaccine Field Trial *(Know this example in detail, although you need not memorize the numbers.)*

In 1916 polio epidemic in United States.

In 1950s Jonas Salk had a promising "vaccine," which worked well in laboratory (i.e., the vaccine seemed safe and produced antibodies against polio).

What now? Test whether or not vaccine works.

(a) (hypothetical) Test vaccine on a small sample of children (e.g., 10 children).

If successful on them, mass distribute the vaccine.

(b) (hypothetical) Offer vaccine to a large number of children.

Typically, not everyone will accept the vaccine.

We have two groups: **treatment** (those who accepted the vaccine) and **control** (those who declined the vaccine).

What is the **explanatory variable**?

What is the **response variable**?

Would this study be considered a valid experiment?

(c) (real data) The National Foundation for Infantile Paralysis (NFIP) proposed vaccinating all grade 2 children (if consent was given), and leaving grades 1 and 3 for control.

Would this study be considered a valid experiment?

(d) (real data) Randomized control.

Offer many children the ability to participate in the experiment, but do not tell them if they are given treatment or placebo.

Also, do not tell doctors or nurses.

Statistics tell us that with large enough samples, **randomized, controlled** experiments determine whether or not a **treatment** (e.g., drug, vaccine) works.

If the treatment group has a higher success rate than the placebo group, we need to decide if this was due to chance or due to a successful treatment.

Typically, we require overwhelming evidence that the treatment was successful before marketing the new treatment.

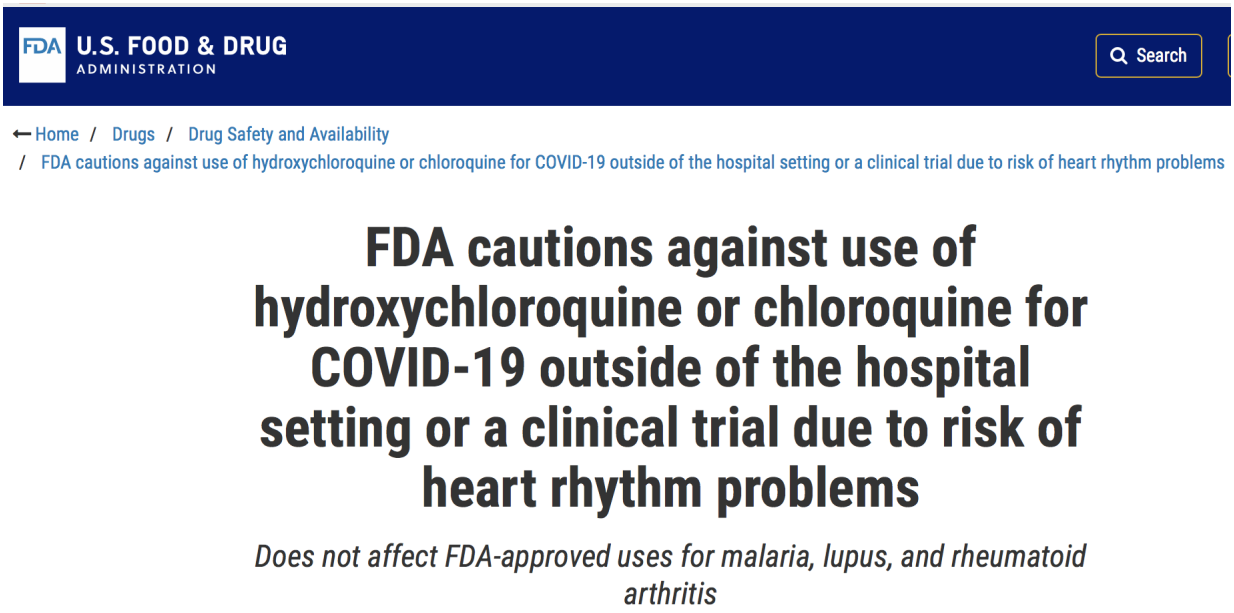
Salk vaccine trial of 1954

Rate of polio per 100,000

Randomized controlled double-blinded experiment			The NFIP study		
	size	rate		size	rate
treatment (high hyg.)	200,000	28	Grade 2 (vaccine, high hyg.)	225,000	25
control (high hyg.)	200,000	71	Grades 1 & 3 (control, average hyg.)	725,000	54
no consent (low hyg.)	350,000	46	Grade 2 (no consent, low hyg.)	125,000	44

□

Example: Should hydroxychloroquine be used to treat COVID-19?



July 1, 2020 Update: A summary of the FDA review of safety issues with the use of hydroxychloroquine and chloroquine to treat hospitalized patients with COVID-19 is now available. This includes reports of serious heart rhythm problems and other safety issues, including blood and lymph system disorders, kidney injuries, and liver problems and failure.

June 15, 2020 Update: Based on ongoing analysis and emerging scientific data, FDA has revoked the emergency use authorization (EUA) to use hydroxychloroquine and chloroquine to treat COVID-19 in certain hospitalized patients when a clinical trial is unavailable or participation is not feasible. We made this determination based on recent results from a large, randomized clinical trial in hospitalized patients that found these medicines showed no benefit for decreasing the likelihood of death or speeding recovery. This outcome was consistent with other new data, including those showing the suggested dosing for these medicines are unlikely to kill or inhibit the virus that causes COVID-19. As a result, we determined that the legal criteria for the EUA are no longer met. Please refer to the Revocation of the EUA Letter and FAQs on the Revocation of the EUA for Hydroxychloroquine Sulfate and Chloroquine Phosphate for more information.

□

Multifactor Experiments

Often we are interested in more than one *explanatory variable* (i.e., factor).

Suppose we are testing two different treatments for cancer. What might be another factor of interest?

Experiments: Randomized or Matched?

An **experiment** may use:

(a) **completely randomized design** – All experimental units are randomly assigned to *treatment* or *control*.

(b) **matched pairs** – Each individual (or pairs of similar individuals) is given both the *treatment* and the *control*.

Give an example of a medical experiment using *matched pairs*.

When is a medical experiment using *matched pairs* not possible?

Which is better, and why: *completely randomized design* or *matched pairs*?

Recall: An **experiment** has a *treatment* and a *control*, whereas an **observational study** (sample survey) consists of just polling.

1.2 Types of Data

Variables may be **quantitative (numerical)** or **qualitative (categorical)**.

Examples of **quantitative** variables are:

Examples of **qualitative** variables are:

Two types of **quantitative** variables are **discrete** and **continuous**.

1. **Discrete variable** takes values which are distinct numbers with gaps.

2. **Continuous variable** takes any value in an interval.

Levels of Measurement for Qualitative Data

Qualitative data

A **nominal scale** classifies data into distinct categories in which no ranking (ordering) is implied.

An **ordinal scale** classifies data into distinct categories in which ranking is implied.