



## Chapter 1

---

# The Benefits and Risks of Using Statistics

---

# Thought Question 1:

A recent newspaper article concluded that *smoking marijuana* at least three times a week *resulted in lower grades* in college.

How do you think the researchers came to this conclusion? Do you believe it? Is there a more reasonable conclusion?

# Thought Question 2:



**Theory:** *On average, men have lower resting pulse rates than women do.*

How could you go about trying to prove or disprove that? Would it be sufficient to measure the pulse rates of one member of each sex? Two members of each sex? What information about men's and women's pulse rates would help you decide how many people to measure?

# Thought Question 3:



Suppose you were to learn that the *large state university* in a particular state graduated more students who eventually went on to become **millionaires** than any of the *small liberal arts colleges* in the state.

Would that be a fair comparison?

How should the numbers be presented in order to make it a fair comparison?

# Thought Question 4:



In its March 3–5, 1995 issue, *USA Weekend* magazine asked readers to return a survey with a variety of questions about sex and violence on television. Of 65,142 readers who responded, 97% were “very or somewhat concerned about violence on TV” (*USA Weekend*, 2–4 June 1995, p. 5).

Based on this survey, can you conclude that about 97% of U.S. citizens are concerned about violence on TV? Why or why not?

# 1.1 Statistics




**Statistics** is a collection of procedures and principles for gaining and analyzing information in order to help people make decisions when faced with uncertainty.

# Statistics in Your Own Life



**Which route to go to school,  
class, work?**

Try each a few times  
Then choose *best* one.

Sampling routes and comparing them  
 gained and analyzed useful  
information to make a decision.

## Case Study 1.1: *Heart or Hypothalamus?*



Hypothalamus plays role in emotion, yet most people think about the heart. Psychologist Lee Salk studied role of heart in human relationships.

### **Observations:**

- Rhesus monkey at zoo held baby on left, by the heart, on 40 out of 42 occasions.
- Out of 287 human mothers (within 4 days of giving birth), 237 or 83% held babies on left – regardless of handedness.

# Case Study 1.1: *Heart or Hypothalamus?*



## More Observations:

- Out of 438 shoppers leaving store with one bag, exactly half carried bag on the left.
- Dental patients asked to hold a 5-inch ball while teeth worked on. Substantially more than half held it on the left (under stress).

## Salk's conjecture:

Must be survival value to having a newborn placed close to sound of mother's heartbeat.

# Case Study 1.1: *Heart or Hypothalamus?*



## Controlled Experiment:

Baby nursery at NY City hospital.

- 1<sup>st</sup> group of newborns: Continuous sound of human heartbeat over loudspeaker.
- 2<sup>nd</sup> group of newborns: no sound played.

**Response:** weight gain (or loss) after 4 days.

## Results:

Infants treated to heartbeat sound gained more weight (lost less) and cried less.

# 1.2 Detecting Patterns and Relationships



## Obvious or Not?

- Average man taller than average woman.
- Taking aspirin helps prevent heart attacks.
- More likely to buy blue jeans in certain months than others.
- Men have lower resting pulse rates than women.
- Listening to Mozart improves spatial reasoning.

# Statistical Studies



**To conduct a statistical study properly, one must**

1. Get a representative sample.
2. Get a large enough sample.
3. Decide whether the study should be an observational study or a randomized experiment.

**Example: Assess if men have *lower* resting pulse rates than women.**

# 1. Get a representative sample.

**Sample** = those actually studied

**Population** = larger group from which  
sample was chosen

Researchers want to extend results  
beyond just the sample.

## **For pulse rates:**

Convenient sample = members of your class.

→ Not valid, e.g., if entire track team in class.

→ Couldn't extend to other age groups.

## 2. Get a large enough sample.

How large is “*large enough*”?

Depends on how much the responses **vary**.

**More variable** within each group

➡ **larger sample** needed to detect a difference.

**For pulse rates:**

- If all men had rate of 65 and all women had rate of 75, then just need a *few* to see the difference.
- If men’s rate varied from 50 to 80 and women’s rates ranged from 52 to 82, then would take *more* measurements to see a difference.

### 3. An observational Study or Randomized Experiment?



**Observational Study:** merely observe things about our sample.

**Randomized Experiment:** randomly assign participants to one of various treatment groups.

**For pulse rates:**

- Sufficient to measure or “*observe*” both pulse rates and gender of people in sample.
- To learn if aspirin helps prevent heart attacks, a randomized experiment would be better.

# Case Study 1.2: *Does Aspirin Prevent Heart Attacks?*



## Physicians' Health Study (1988)

5-year randomized experiment

22,071 male physicians (40 to 84 years old).

- Group 1: took ordinary *aspirin* tablet every other day.
- Group 2: took *placebo* (looked like aspirin but no active ingredients).

Subjects did not know which group they were assigned (called blinding).

## Case Study 1.2: *Does Aspirin Prevent Heart Attacks?*



Condition	Heart Attack	No Heart Attack	Attacks per 1000
Aspirin	104	10,933	9.42
Placebo	189	10,845	17.13

### Results:

- Heart attack rate in aspirin group only 55% of rate in placebo group.
- Randomization → other factors, such as amount of exercise, expected to be *similar* for both groups.
- Subjects were male physicians. Results may *not* extend to all males, nor to women.

# 1.3 Don't Be Deceived By Improper Use of Statistics



## Example 1

- Business Magazine sent **2000** questionnaires to readers about Republican primary: Chrysler president Lee Iacocca vs. VP George Bush.
- Based on **only first 200 respondents**, reported Iacocca would beat Bush 54% to 47%.
- **Sample not representative of population** of all people likely to vote: those who feel strongly about an issue most likely to respond.

# Improper Use of Statistics



## Example 2

- U.S. EPA study led to report that New Jersey ranked 22<sup>nd</sup> in its release of toxic chemicals.
- **Problem:** wrong units ... ranking was based on **total pounds released**.
- Using appropriate rate of **pounds released *per square mile***, New Jersey ranked 4<sup>th</sup> on the list.
- Figures were misleading until adjusted for size.

# Improper Use of Statistics



## Example 3

- Article headline:  
    “Study: Smoking May Lower Kids’ IQs”
- Headline implication: a causal connection.
- Problem: results based on an observational study (since experiment would be unethical).
- Moral: *Can’t make causal connections from observational studies.*

# Improper Use of Statistics

## **New study confirms too much pot impairs brain**

“More evidence that chronic marijuana smoking impairs mental ability: Researchers at the University of Iowa College of Medicine say a test shows those who smoke seven or more marijuana joints per week had lower math, verbal and memory scores than non-marijuana users. Scores were particularly reduced when marijuana users held a joint’s smoke in their lungs for longer periods.” *San Francisco Examiner*, 13 March 1993, p. D-1.

## **Example 4**

- **Headline implication: a causal connection.**
- **Problem: based on an **observational study****
- **Can conclude *there is a relationship* but *cannot conclude a causal connection.***

## Case Study 1.3: *A Mistaken Accusation of Cheating*



- Two students **accused of cheating on multiple choice medical exam** due to whispering during one part and wrong answers often matched each other.
- Licensing board estimated **odds of having answers as close as these two were < 1 in 10,000**.
- At trial, statistician showed agreement in this exam was no higher than in other exams.
- **Key:** two students were husband and wife, both from India. *Licensing board incorrectly assumed they were just guessing.* Instead, common mistakes were often the right answers for India background.
- Whispering had to do with status of sick child.

# 1.4 Summary and Conclusions



- There are both advantages and dangers of using statistical methods.
- Results of a study are not enough!
- You need to know how data were collected, who was asked, etc.
- Over the next several chapters, you will become an educated consumer of statistical information.