Introduction

Marius Ionescu

08/30/2011

Numbers vs percentages

Example

Claim: Syracuse schools are better than Hamilton schools because they have twice as many National Scholarship finalists. What do you think?

Numbers vs percentages

Example (Is Colgate worth the money?)

A student compared the number of students who graduated in the last four years at Colgate, Dartmouth and UConn and who earn more than \$75,000 per year. The numbers are

Colgate	Dartmouth	UConn
1011	1240	2500

Should you have enrolled at UConn?

Fact

Percentages are useful for comparing population of different sizes.

Some theory

Fact

We measure something in order to

1. *compare it to something else*
2. *see if it is growing or decreasing*

Fleischmann's Rapid Rise yeast

Example

Fleischmann's Rapid Rise yeast claims that it rises 50% faster. If a pizza recipe says to let the dough rise 1 hour, how long should we let it rise with rapid rise yeast?

HIV testing

Example

Is it a good idea to test everyone for HIV antibodies? Suppose

- people with antibodies test positive 99% of the time.
- people without antibodies test negative 99% of the time
- 1.1 million out of 310 million Americans have HIV antibodies
- everyone was tested.

What fraction of those whose tests are positive are positive?

Some examples of "statistics gone wrong"

Example

USA Today reported that 28% of the people surveyed said they microwave leftovers almost daily, 43% said they microwave leftovers two to four times a week, and 15% said they do it once a week. What is wrong with these numbers?

Trident gum commercial

Example

Recall the commercial: "Four out of five dentists surveyed recommended Trident gum for their patients who chew gum". Do you agree?

Sample of size one

Example

You have probably seen the TV commercial pitting one paper towel brand against another, where one piece of each type of paper towel is used to try to absorb the same amount of red juice. What is wrong with this comparison?

2 minutes

Example

A study conducted a few years back evaluated videotaped sessions of 1265 patient appointments with 59 primary-care physicians and 6 surgeons in Colorado and Oregon. This study found that physicians who not been sued for malpractice spent an average of 18 minutes with each patient, compared to 16 minutes for physicians who had been sued for malpractice. Is two minutes that important?

Definition

Statisticians use the "method of comparison" to show cause and effect through corellation.

Ideal situation: Two groups of subjects identical except for two ways:

1. the treatment
2. the response.

We call the first group the treatment group and the second one the control group.

Fact (Basic Idea)

If the treatment group has a different response, that response must have been caused by the treatment variable.

Experimental design

- Subjects should be assigned to treatment or control at **random**.
- If the treatment group differs from the control group, these differences **confound** with the effects of the treatment.

Example

Suppose that we want to study the effect of partying in college to lung cancer. Smoking is a **confoundly variable**, because those who party are more likely to smoke, and smoking is related to lung cancer.

Placebos; Double blind experiments

Fact

With human subjects, the idea of being treated is very powerful and can confound with actual treatment. Two methods are used to reduce this effect

1. *Placebo: a non-medical treatment, given to the control group so that they do not know which group they are in.*
2. *Double blind experiment: neither the subjects nor the doctors who measure the responses should know who was in the treatment group and who was in the control group.*

Ideal (gold standard) experiment

Fact

Randomized controlled double blind experiment: Take a population, divide it randomly into two groups, compare the results and decide whether chance could have produced the result.

Fact

Some questions don't allow controlled experiments.
Examples:

- *effects of gender on career path*
- *effects of black holes on nearby stars*
- *effects of minority status on income path*
- *effect of diet on health*