Elementary statistics

Michael Ernst

CSE 190p

University of Washington

A dice-rolling game

- Two players each roll a die
- The higher roll wins
  - Goal: roll as high as you can!
- Repeat the game 6 times

Hypotheses regarding Mike's success

- Luck
- Fraud
  - loaded die
  - inaccurate reporting

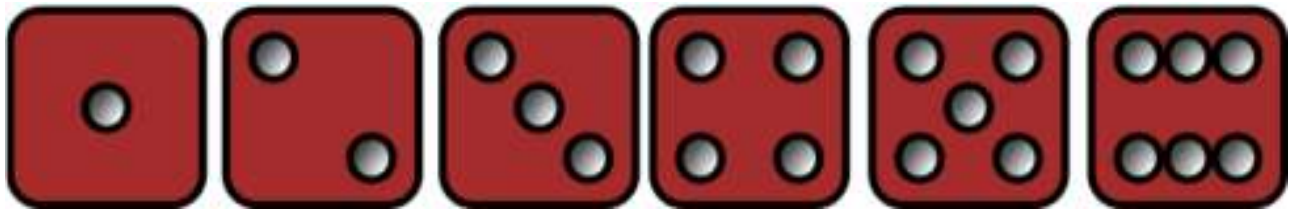
- How likely is luck?
- How do we decide?

Questions that statistics can answer

- I am flipping a coin. Is it fair?
  How confident am I in my answer?
- I have two bags of beans, each containing some black and some white beans. I have a handful of beans. Which bag did the handful come from?
- I have a handful of beans, and a single bag. Did the handful come from that bag?

- Does this drug improve patient outcomes?
- Which website design yields greater revenue?
- Which baseball player should my team draft?
- What premium should an insurer charge?
- Which chemical process leads to the best-tasting beer?

What can happen when you roll a die?



What is the likelihood of each?

A dice-rolling experiment

Game: Roll one die, get paid accordingly:

Roll	1	2	3	4	5	6
Payoff	1 CHF	2 CHF	3 CHF	4 CHF	5 CHF	0 CHF

Player self-reports the die roll and takes the money

– no verification of the actual roll

From “Lies in disguise: An experimental study on cheating”
by Urs Fischbacher and Franziska Heusi

What can happen when you roll two dice?

How likely are you to roll *11 or higher*?

This probability is known as the “p value”.

2 3 4 5 6 7 8 9 10 11 12

How to compute p values

- Via a statistical formula
  - Requires you to make assumptions and know which formula to use
- Computationally (simulation)
  - Run many experiments
  - Count the fraction with a better result
    - Requires a metric/measurement for “better”
  - Requires you to be able to run the experiments

Interpreting p values

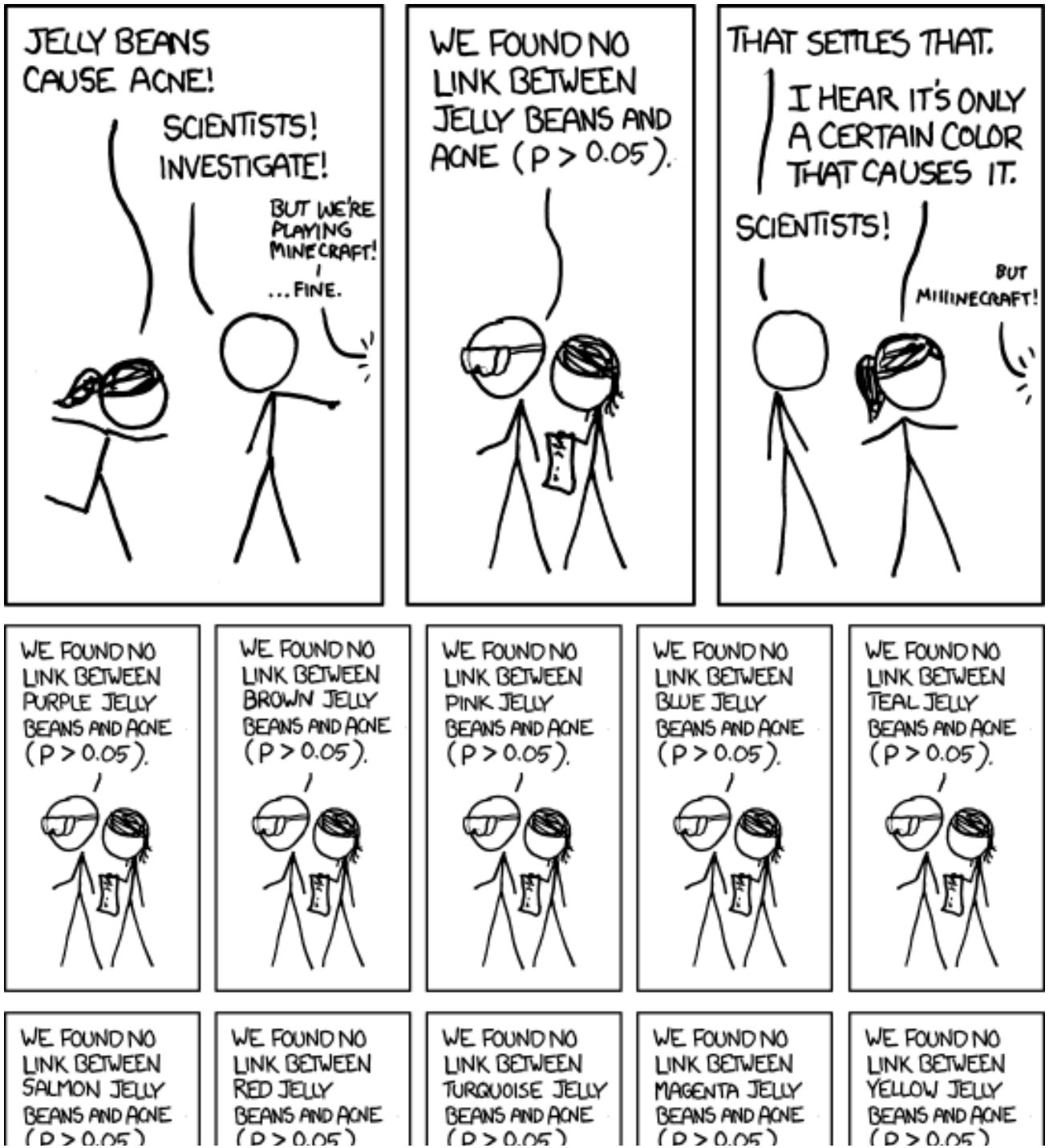
p value of 5% or less = statistically significant

- This is a *convention*; there is nothing magical about 5%

Two types of errors may occur in statistical tests:

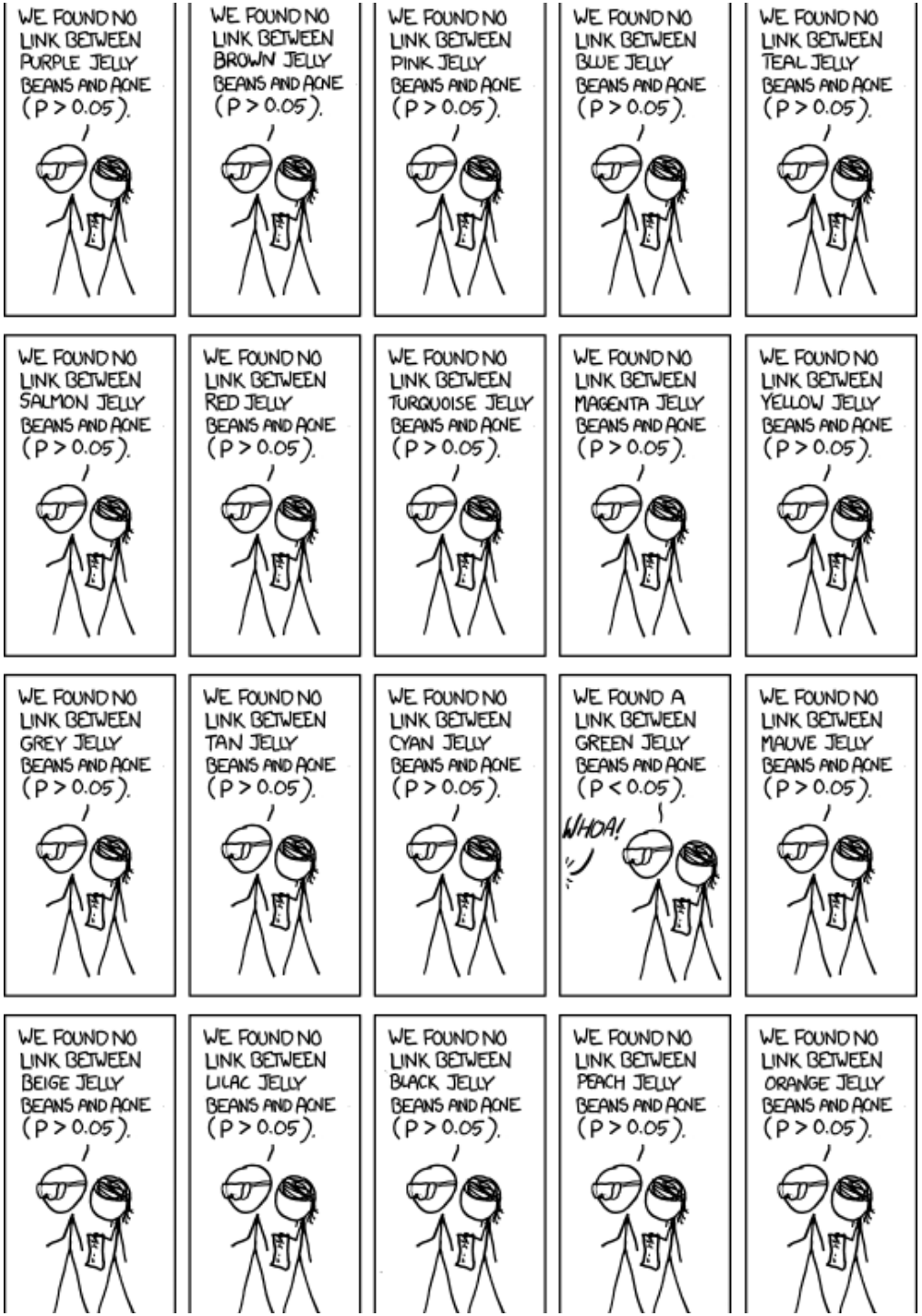
- false positive (or false alarm or Type I error): no real effect, but report an effect (through good/bad luck or coincidence)
  - If no real effect, a false positive occurs about 1 time in 20
  - If there is a real effect, a false positive occurs less often
- false negative (or miss or Type II error): real effect, but report no effect (through good/bad luck or coincidence)
  - The smaller the effect, the more likely a false negative is
  - How many die rolls to detect a die that is only slightly loaded?

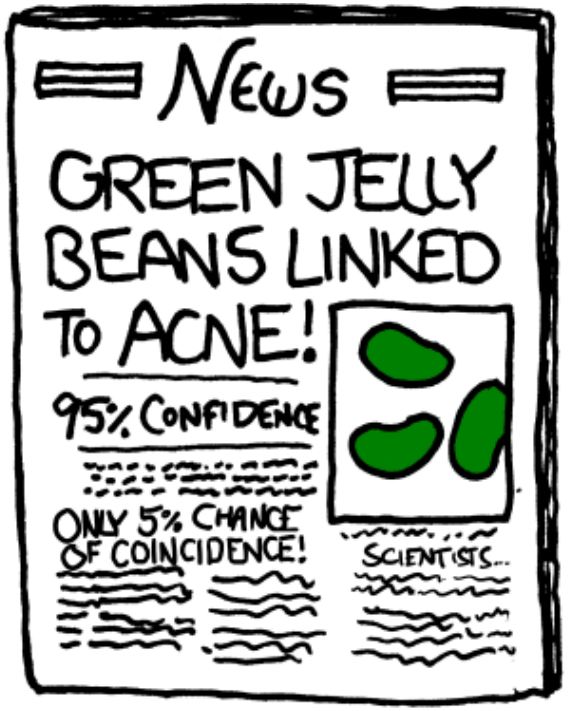
The *larger* the sample, the *less the likelihood* of a false positive or negative



A false positive

http://xkcd.com/882/

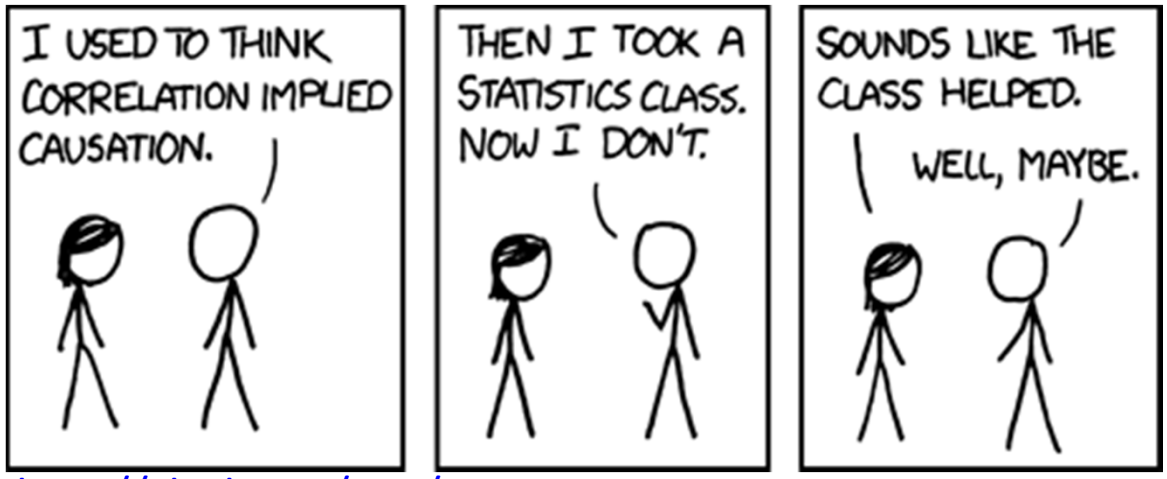




http://xkcd.com/882/

Correlation ≠ causation

Ice cream sales and murder rates are correlated



http://xkcd.com/552/

Statistical significance
≠ practical importance