Statistical Probability and Models

C. Ryan Campbell

Duke University

c.ryan.campbell@duke.edu

28 Nov 2017

Project Updated

- Still Grading Rough Drafts - Done by Thursday
- Presentations are next week
- Present what you have, should be more polished than the rough draft
- 20 minutes per group presentation
- 5min intro, 5min per member
- Order?

Overview

This Week's Goals

- Understand p-values & adjustments
- Learn about Maximum Likelihood
- Learn about Bayes Theorem & Bayesian Probability

Probability

- “Odds” that some event occurs
- Bounded from 0 to 1
- Usually expressed as a fraction or percent
- Often using the notation: Pr(event) or P(event)

Or

- Probabilities of multiple events can be combined
- “Or” condition
- Probability either thing happens: A or B
- when A and B are independent and mutually exclusive:
- Pr(A or B) = Pr(A) + Pr(B)

Or

- Probabilities of multiple events can be combined
- “Or” condition
- when A and B are independent and not exclusive:
- Pr(A or B) = Pr(A) + Pr(B) - Pr(A & B)

And

- Probabilities of multiple events both occuring can be combined
- “And” condition
- Probability both things happen, A & B
- When A & B are independent:
- Pr(A & B) = Pr(A) * Pr(B)
- “And” is commutative:
- Pr(B & A) = Pr(A & B)

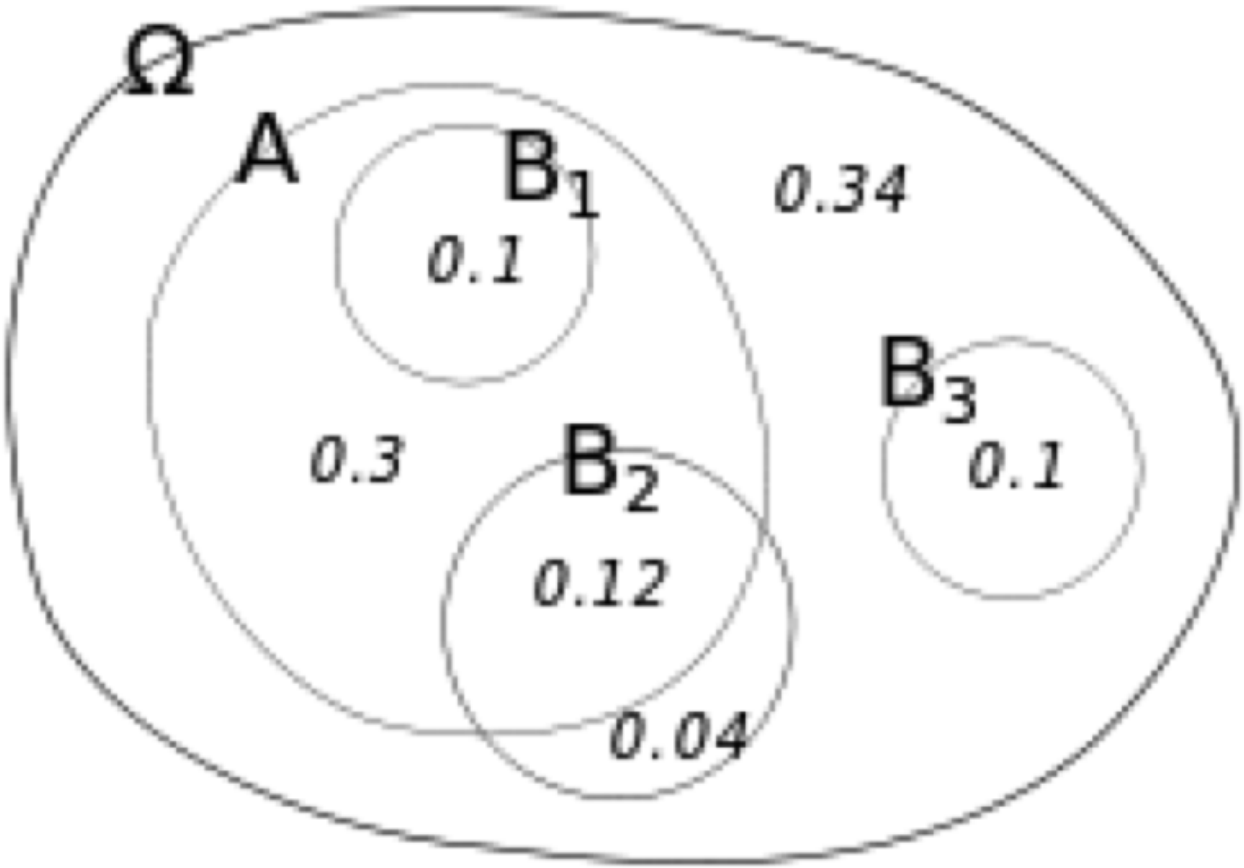
Conditional Probabilities

- Probabilities of event A given event B
- Probability of A if we know B has occured
- When B happens, how likely is it that A happens
- Numerator = Pr(A & B)
- Denominator = Pr(B)

$$\frac{\Pr(A \,\&\, B)}{\Pr(B)}$$

Some Quick Math

$\Pr(B_1 \text{ given } A) = ?$
$\Pr(B_2 \text{ given } A) = ?$
$\Pr(A \text{ given } B_2) = ?$



P-Value

- What is a P-Value?

Hypothesis Testing

- Need the context of Hypothesis testing for p-values to have meaning
- H0: Group A = Group B
- H1: Group A *does not equal* Group B

Hypothesis Testing

- We assume H0 is true
- Compare Group A and Group B
- The p-value measures how likely it is the differences between A and B are due to chance
- A lower p-value gives more power to reject H0

Let's Try an Example

- Run a comparison of means in R
- Compare two random sets of data with a t-test, using

```
> rnorm()
> t.test()
```

- mean 0, stdev 1, n 10
- What is your p-value?

Let's Try an Example

- Run a comparison of means in R
- Compare two random sets of data with a t-test, using

```
> rnorm()
> t.test()
```

  mean of 0

  stdev of 1

  n of 10

- What is your p-value?

Let's Try an Example

- Is H0 true or false?
- Did anyone get a p-value suggesting otherwise?
- Why?

Let's Try an Example

- What is the distribution of p-values over many tests?
- How many tests is a single DESeq analysis?
- Use a for loop and R to replicate the example, but on a DESeq scale

Let's Try an Example

Use a for loop and R to replicate the example, but on a DESeq scale

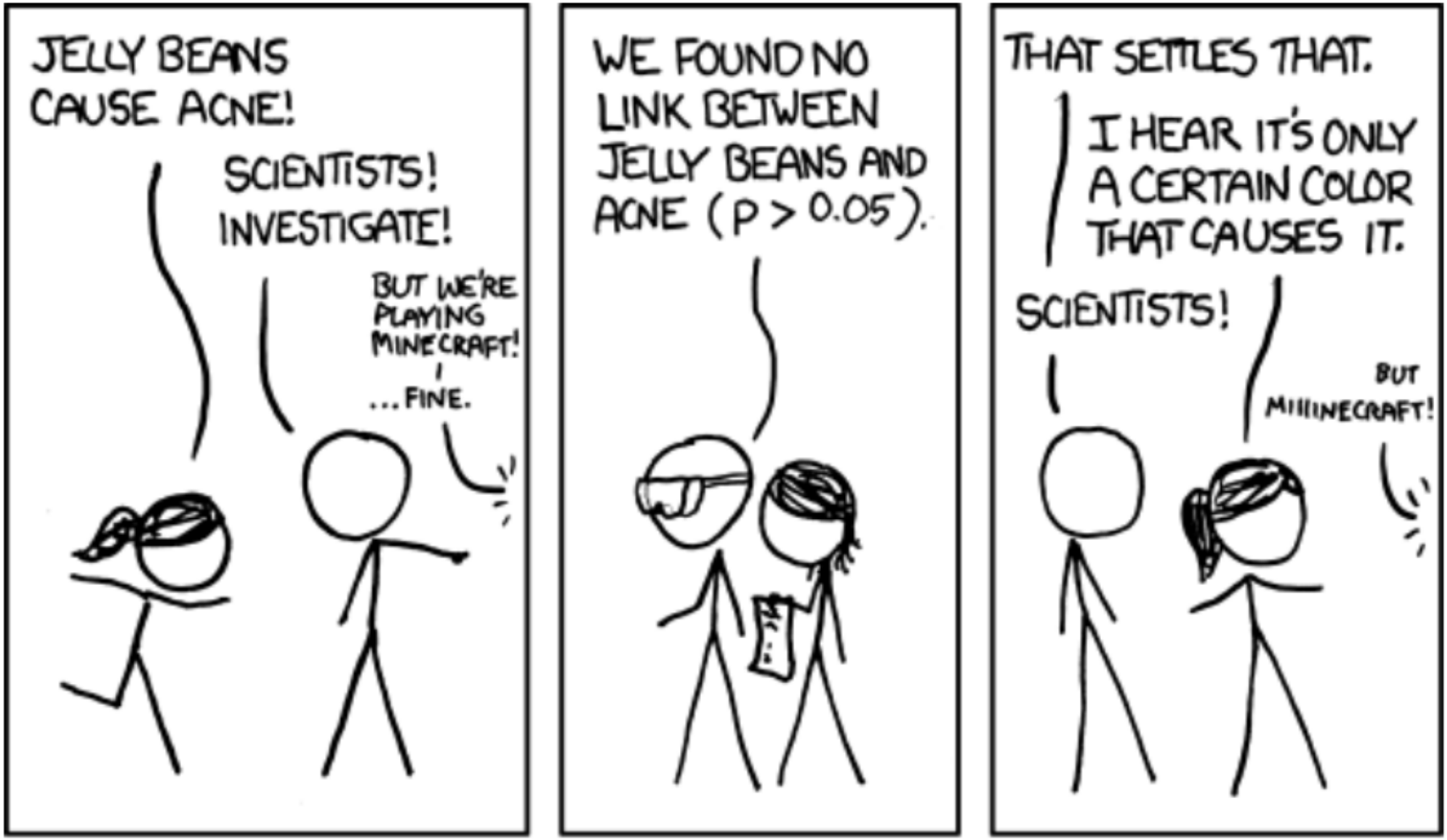
```
for (n in 1:high number) {
generate p-values
}
```

How do you save all your p-values?

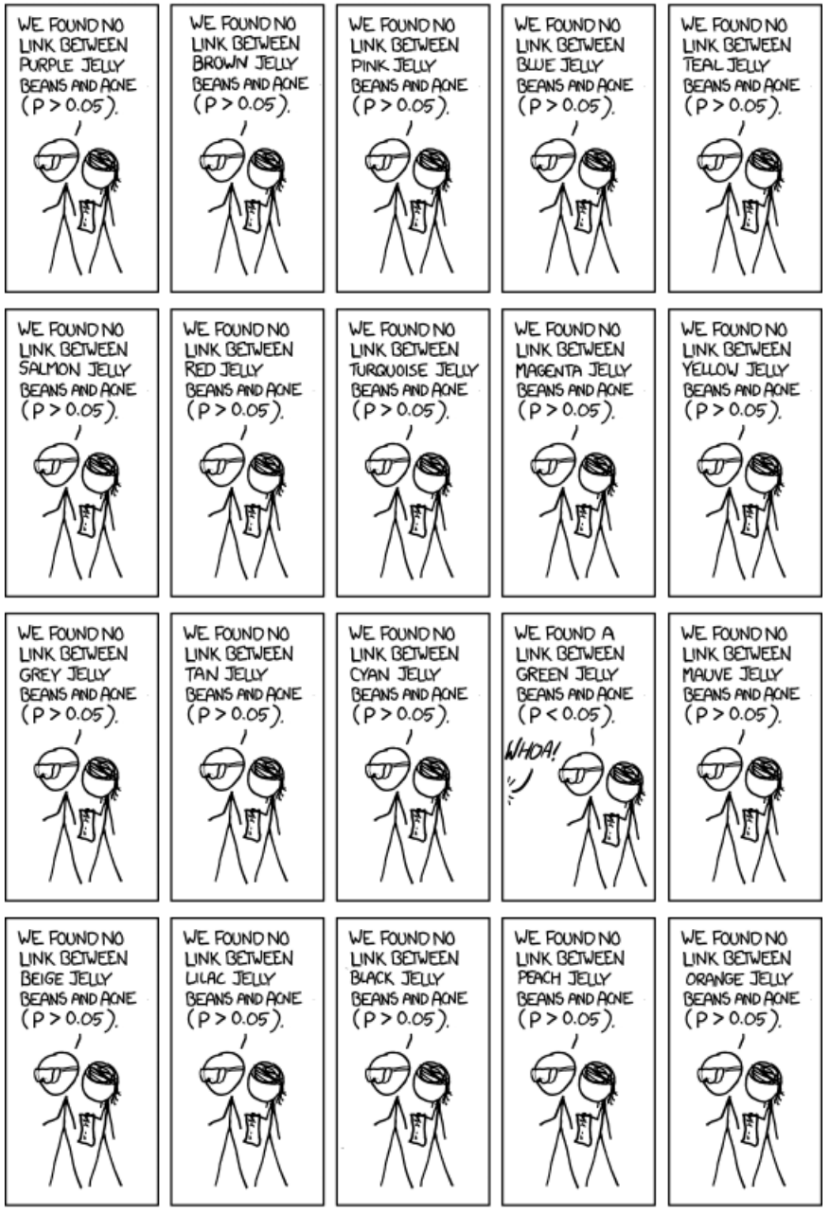
Let's Try an Example

- Visualize the distribution of p-values
- What do you see?

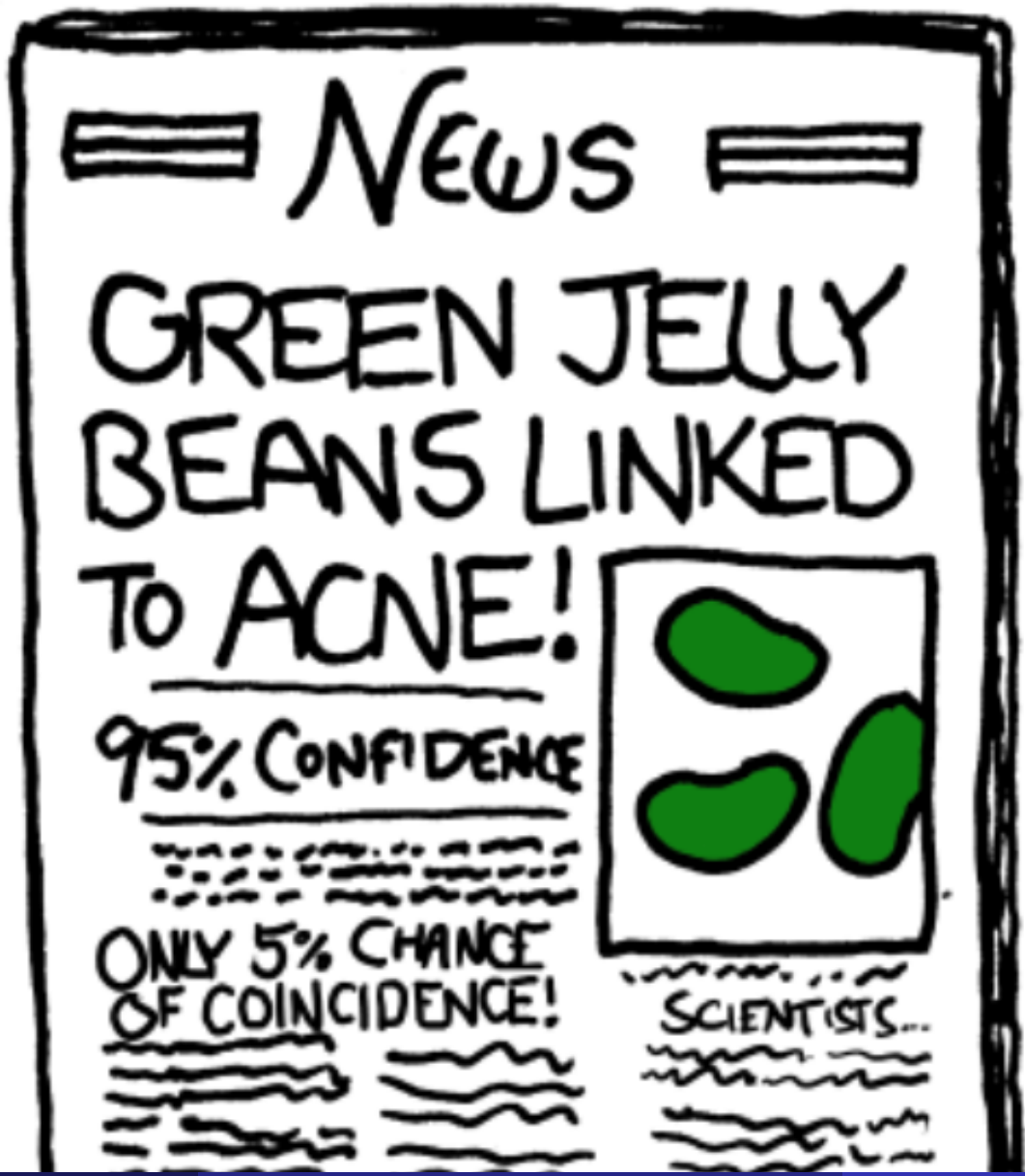
Multiple Tests: xkcd



Multiple Tests: xkcd



Multiple Tests: xkcd



Multiple Tests

- If you conduct 10,000 t-tests
- AND the null hypothesis is true
- How many will yield a p-value capable of rejecting the null
- at alpha = 0.05?
- at alpha = 0.01?

Multiple Tests: Solutions

- What do we do?
- How do you take this into account?
- Bonferroni corrections
- Adjusted P-Values
- Permutation tests

Bonferroni Correction

- Very simple approach
- Combine your alpha level (0.05)
- With the number of tests (10,000)
- 0.05/10,000 is the altered p-value level
- per-test alpha = 5×10^{-6}

Bonferroni Correction

- Pros and Cons:
- + Easy to apply
- + Flexible for number of tests
- - Overly Conservative (will accept false nulls)
- - Assumes independence

Adjusted P-Value

- More statistically complex
- Aims to reduce FDR - False Discovery Rate
- False Discovery - all of the random samples we drew earlier with T-tests that fell below 0.05
- This is why you'll want to rely on padj from DESeq

Adjusted P-Value: Example

- Benjamini-Hochberg procedure
- Sort results by p-value
- Assign each test a BH score of:
- (rank/N of tests) * acceptable FDR

Adjusted P-Value: Example

- For the “best” results (lowest P-Values) you should see:
- p-value less than BHscore
- Find the worst p-value that is still less than its BHscore
- All tests below that are significant

Adjusted P-Value: Example

- Find the worst p-value that is still less than its BHscore
- That test and all tests below that are significant:

(1/25) * .25 = 0.01:

Variable	P Value	Rank	(I/m)Q
Depression	0.001	1	0.01
Family History	0.008	2	0.02
Obesity	0.039	3	0.03
Other health	0.041	4	0.04
Children	**0.042**	**5**	**0.05**
Divorce	0.060	6	0.06
Death of Spouse	0.074	7	0.07
Limited income	0.205	8	0.08

Permutation Test

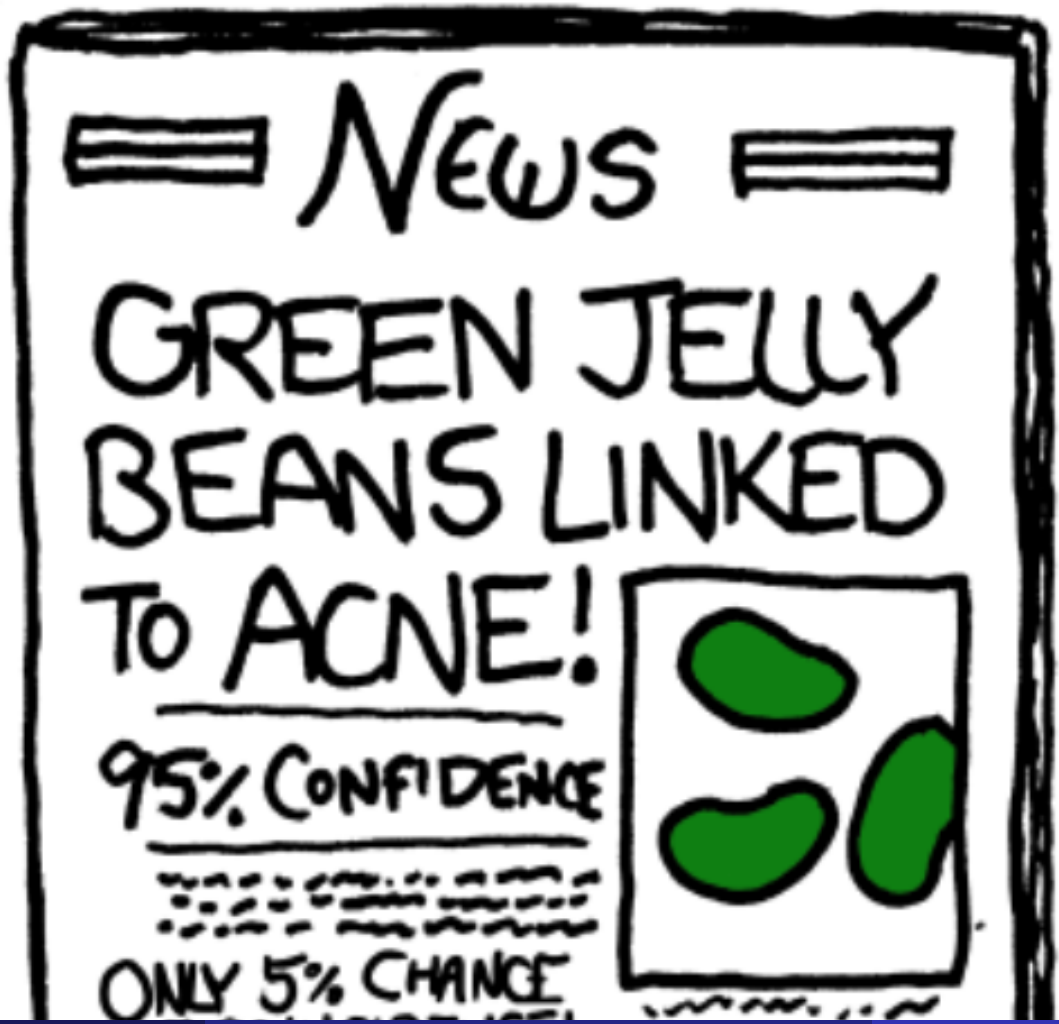
- A form of resampling your data
- (other forms are Jackknifing or Bootstrapping)
- You re-label your samples (control v. experimental)
- Rerun the analysis
- See where the actual test's p-value falls on the range of p-values this produces

Permutation Test

- Two possible outcomes:
- 1) The instance with the correct labels falls in the middle of the distribution
- 2) The instance with the correct labels is a significant outlier
- What do each of these mean?

Multiple Tests: xkcd

How would each correctional method have dealt with our jelly bean problem?



The End