

Comparison of Bayesian and Frequentist Inference

18.05 Spring 2014

Jeremy Orloff and Jonathan Bloom

Concept question

Three different tests are run all with significance level $\alpha = .05$.

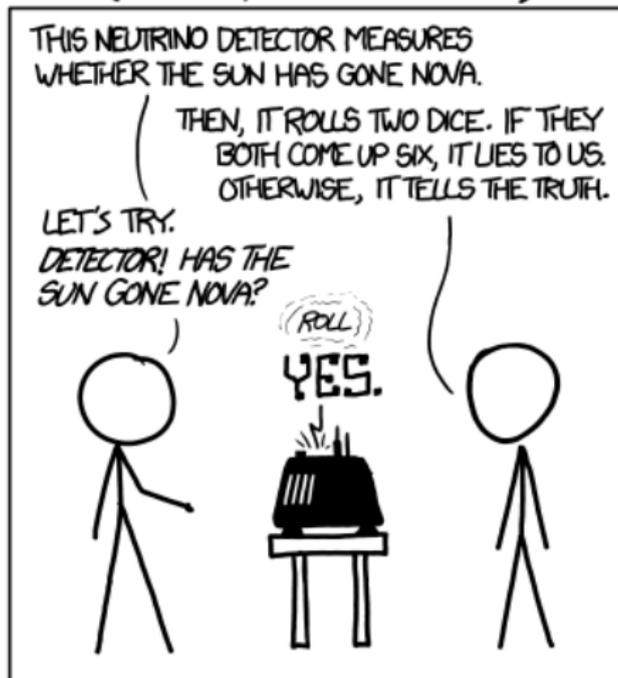
1. Experiment 1: finds $p = .03$ and rejects its null hypothesis H_0 .
2. Experiment 2: finds $p = .049$ and rejects its null hypothesis.
3. Experiment 3: finds $p = .15$ and fails to reject its null hypothesis.

Which result is most likely to be correct?

(Click **4** if you don't know.)

Solution

DID THE SUN JUST EXPLODE?
(IT'S NIGHT, SO WE'RE NOT SURE.)



Courtesy of [xkcd](#). CC-BY-NC.

<http://xkcd.com/1132/>

Board question: chi-square for independence

(From Rice, *Mathematical Statistics and Data Analysis*, 2nd ed. p.489)

Consider the following contingency table of counts:

Education	Married once	Married multiple times	Total
College	550	61	611
No college	681	144	825
Total	1231	205	1436

Are the number of marriages and education level independent?

Test this using a chi-squared test with significance level 0.01.

Concept question: multiple-testing

1. Suppose we use two-sample t -tests at $\alpha = .05$ level to determine whether 6 treatments all have the same recovery time. How many t -tests might we need to run?

- 1) 1 2) 2 3) 6 4) 15 5) 30

Concept question: multiple-testing

1. Suppose we use two-sample t -tests at $\alpha = .05$ level to determine whether 6 treatments all have the same recovery time. How many t -tests might we need to run?

- 1) 1 2) 2 3) 6 4) 15 5) 30

2. In the situation above, assuming all 6 means are the same, what is the probability that we reject at least one of the 15 null hypotheses?

- 1) Less than .05 2) .05 3) .10 4) Greater than .50

Concept question: multiple-testing

1. Suppose we use two-sample t -tests at $\alpha = .05$ level to determine whether 6 treatments all have the same recovery time. How many t -tests might we need to run?

- 1) 1 2) 2 3) 6 4) 15 5) 30

2. In the situation above, assuming all 6 means are the same, what is the probability that we reject at least one of the 15 null hypotheses?

- 1) Less than .05 2) .05 3) .10 4) Greater than .50

Discussion: Recall that there is an F -test that tests if all the means are the same. What is an advantage of using the F -test rather than many two-sample t -tests?

Board question: Stop!

Experiments are run to test a coin that is suspected of being biased towards heads. The significance level is set to $\alpha = .1$

Experiment 1: Toss a coin 5 times. Report the sequence of tosses.

Experiment 2: Toss a coin until the first tails. Report the sequence of tosses.

1. Give the test statistic, null distribution and rejection region for each experiment. List all sequences of tosses that produce a test statistic in the rejection region for each experiment.

2. Suppose the data is *HHHHT*.

(a) Do the significance test for both types of experiment.

(b) Do a Bayesian update starting from a flat prior: $\text{Beta}(1,1)$.

Draw some conclusions about the fairness of coin from your posterior.

(Use R: `pbeta` for computation)

Board question: Stop II

For each of the following experiments (all done with $\alpha = .05$)

(a) Comment on the validity of the claims.

(b) Find the probability of a type I error in each experimental setup.

- 1 By design Peter did 50 trials and computed $p = .04$.
He reports $p = .04$ with $n = 50$ and declares it significant.
- 2 Erika did 50 trials and computed $p = .06$.
Since this was not significant, she then did 50 more trials and computed $p = .04$ based on all 100 trials.
She reports $p = .04$ with $n = 100$ and declares it significant.
- 3 Jerry did 50 trials and computed $p = .06$.
Since this was not significant, he started over and computed $p = .04$ based on the next 50 trials.
He reports $p = .04$ with $n = 50$ and declares it statistically significant.

MIT OpenCourseWare
<http://ocw.mit.edu>

18.05 Introduction to Probability and Statistics

Spring 2014

For information about citing these materials or our Terms of Use, visit: <http://ocw.mit.edu/terms>.