

STAT 113

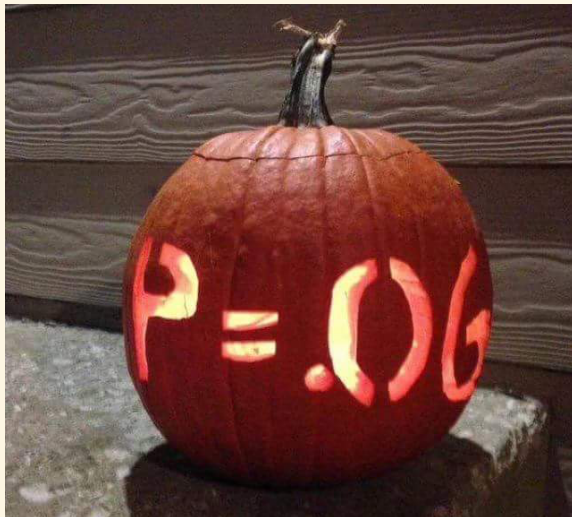
Tests and Intervals

Colin Reimer Dawson

Oberlin College

October 31, 2017

Happy Halloween!



Outline

Tests and CIs

Test for a Single Population Mean

When does the procedure fail?

- Consider a procedure to create a 95% confidence interval for a mean, μ , using $\bar{x} \pm 2SE$.
- How often will we miss μ ? **We will miss 5% of the time.**
- *What* will cause us to miss? **We will miss when $|\bar{x} - \mu| > 2SE$**

Confidence Level and Significance Level

- In a two-tailed test for μ with $\alpha = 0.05$, we reject H_0 if the sample mean, \bar{x} falls in the outer 5% of the H_0 sampling distribution of possible \bar{x} values.
- How far away do μ and \bar{x} need to be for this to happen? The outer 5% of \bar{x} values are at least 2SE from μ
- When H_0 is true, how often will we incorrectly reject it? We will make a Type I Error 5% (α) of the times that H_0 is true.
- *What* will cause a Type I Error? when the sample statistic happens to be more than 2SE from the population parameter

Confidence Level and Significance Level

- Suppose $H_0 : \mu = 80$ is true. What proportion of possible random samples (datasets) yield a CI that contains 80?
- Since the CI contains μ 95% of the time and $\mu = 80$, the CI will contain 80 95% of the time
- So if our CI *doesn't* contain 80, there are two possibilities...
- Either $\mu \neq 80$ after all, or we happened into an unlucky sample where $|\mu - \bar{x}| > 2SE$
- These are the two cases when we *reject* H_0 (first is true positive, second is false discovery/Type I Error).

Another Way to Interpret CIs

We can think of the $(1 - \alpha)$ Confidence Interval as the set of null parameter values that would *not* be rejected using a test with significance level α .

Example: Police Body Cameras

`http://bwc.thelab.dc.gov/results.html`

Outline

Tests and CIs

Test for a Single Population Mean

Hypothesis Testing for a Single Population Mean

Hypothesis Testing for a Single Population Mean

If the sampling distribution is symmetric and bell-shaped, then we can reject H_0 at $\alpha = 0.05$ if $|\mu_0 - \bar{x}| > 2SE$ (μ_0 is what H_0 says μ is). Estimate SE using a bootstrap distribution.

More generally, estimate the randomization distribution by shifting the bootstrap distribution to be centered at μ_0 , and compute P -values as for other tests.

Alternatively, construct a $(1 - \alpha)$ CI, and reject if μ_0 lies outside it.

Mercury in Fish

Mercury in Fish

The US Food and Drug Administration has a limit for mercury content in fish of 1.0 ppm (parts per million). In Canada, the limit is 0.5 ppm.

- (a) Suppose we want to estimate the mean mercury content of fish in Florida lakes, using a random sample of 53 fish.
- (b) Suppose we want to ascertain whether the evidence supports the assertion that fish from this population would be considered safe to eat in (1) the U.S. and (2) Canada.

- StatKey

Body Temperature

Does healthy human body temperature really average 98.6? A random sample of 50 healthy adults participated in a study to find out. Each person had their temperature taken, and recorded.

1. What are the null and alternative hypotheses?
2. What statistic do we want to use from the sample?
3. Use StatKey (Bootstrap CI for a mean) to construct a 99% confidence interval for the parameter of interest.
4. Shift the bootstrap distribution to be centered at the null parameter value to approximate a randomization distribution, and compute a P -value.
5. Now go to “Randomization Test for a Single Mean” and generate the shifted bootstrap distribution directly. Does the P -value agree with yours?