

Why p -values don't mean what you think they mean.

STAT 587 (Engineering)
Iowa State University

October 2, 2020

p -values

From the ASA statement on p -values:

a p -value is the probability under a specified statistical model that a statistical summary of the data would be equal to or more extreme than its observed value.

Do you have a random sample from your population? If not, then you cannot make a statistically valid statement about the population.

t-test

Let $Y_i \stackrel{\text{ind}}{\sim} N(\mu, \sigma^2)$ with hypotheses

$$H_0 : \mu = \mu_0 \quad \text{versus} \quad H_A : \mu \neq \mu_0.$$

Calculate

$$t = \frac{\bar{y} - \mu_0}{s/\sqrt{n}}$$

which has a t_{n-1} distribution if H_0 is true and the associated p -value

$$p\text{-value} = 2P(T \leq -|t|)$$

where $T \sim t_{n-1}$.

You set a significance level α and **reject H_0** if

$$p\text{-value} < \alpha.$$

ASA interpretation

From the ASA:

a p-value can indicate how incompatible the data are with a specified statistical model

If you **reject** H_0 , the data are incompatible with the model associated with the null hypothesis.
In our example,

$$H_0 : Y_i \overset{ind}{\sim} N(\mu_0, \sigma^2).$$

So perhaps

- the data are not independent,
- the data are not normal,
- the variance is not constant,
- the mean is not μ_0 , or
- you got unlucky.

Context matters

From the ASA statement:

Scientific conclusions and business or policy decisions should not be based only on whether a p -value passes a specific threshold.

Imagine these scenarios all with p -value = 0.05:

- a small-scale agricultural field trial,
- an extrasensory perception experiment,
- a large-scale clinical trial, or
- a 30,000 gene screening for disease progression.

A simple model

Let $Y \sim N(\mu, 1)$ and $H_0 : \mu = 0$ vs $H_A : \mu \neq 0$.

You observe p -value, $p = 0.05$. What does it mean?

Bayes rule:

$$P(H_0|p = 0.05) = \frac{P(p = 0.05|H_0)P(H_0)}{P(p = 0.05|H_0)P(H_0) + P(p = 0.05|H_A)P(H_A)}$$

To calculate this we need

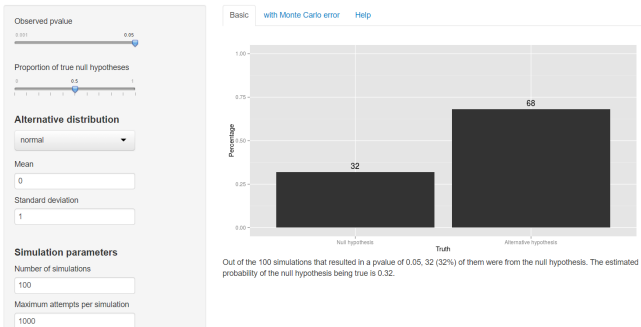
- $P(H_0) = 1 - P(H_A)$ and
- $P(p = 0.05|H_A)$:
 - distribution for μ when H_A is true.

p-value shiny app

```
install.packages("shiny")  
shiny::runGitHub("jarad/pvalue")
```

<https://jaradniemi.shinyapps.io/pvalue/>

Null hypothesis probability conditional on an observed pvalue



Summary

- Random sample \rightarrow population
- Model assumptions
- Context matters
- Error rate is likely much larger than significance level