M6S2 - P-values

Professor Jarad Niemi

STAT 226 - Iowa State University

October 30, 2018

Outline

- Review of statistical hypotheses
  - Null vs alternative
  - One-sided vs two-sided
- Pvalues
  - test statistic
  - as or more extreme
  - interpretation

Statistical hypotheses

Most statistical hypotheses are statements about a population parameters.

For example, for a population mean μ, we could have the following null hypothesis with a two-sided alternative hypothesis:

$$H_0 : \mu = 0 \quad \text{versus} \quad H_a : \mu \neq 0$$

Or we could have the following null hypothesis with a one-sided alternative

$$H_0 : \mu = 98.6 \quad \text{versus} \quad H_a : \mu > 98.6$$

or, equivalently

$$H_0 : \mu lle 98.6 \quad \text{versus} \quad H_a : \mu > 98.6$$

P-values

Definition

A test statistic is a summary statistic that you use to make a statement about a hypothesis. A p-value is the (frequency) probability of obtaining a test statistic as or more extreme than you observed if the null hypothesis (model) is true.

We will discuss the following phrases one at a time

- if the null hypothesis (model) is true,
- test statistic,
- as or more extreme than you observed, and
- (frequency) probability.

Null hypothesis (model)

Recall that we have a null hypothesis, e.g.

$$H_0 : \mu = m_0$$

for some known value m_0, e.g. 0. But we also have statistical assumptions, e.g.

$$X_i \stackrel{iid}{\sim} N(\mu, \sigma^2).$$

Thus, the statement if the null hypothesis (model) is true means that we assume

$$X_i \stackrel{iid}{\sim} N(m_0, \sigma^2).$$

ACT scores example

The mean composite score on the ACT among the students at Iowa State University is 24. We wish to know whether the average composite ACT score for business majors is different from the average for the University. We sample 51 business majors and calculate an average score of 26 with a standard deviation of 4.38.

Let X_i be the composite ACT score for student i who is a business major at Iowa State University with $E[X_i] = \mu$.

What is the null hypothesis? The null hypothesis is

$$H_0 : \mu = 24.$$

What is the null hypothesis model?

$$X_i \stackrel{iid}{\sim} N(24, \sigma^2).$$

Test statistic

Let $X_i \stackrel{iid}{\sim} N(\mu, \sigma^2)$. The following are all summary statistics:

- sample mean ($\overline{X}$),
- sample median (Q2),
- sample standard deviation (S),
- sample variance (S^2),
- min, max, range, Q1, Q3, interquartile range, etc.

The test statistic ... you observed is just the actual value you calculate from your sample, e.g. the observed sample mean ($\overline{x}$), the observed sample standard deviation (s), etc.

We will be primarily interested in the t-statistic:

$$t = \frac{\overline{x} - m_0}{s/\sqrt{n}}.$$

ACT scores example

The mean composite score on the ACT among the students at Iowa State University is 24. We wish to know whether the average composite ACT score for business majors is different from the average for the University. We sample 51 business majors and calculate an average score of 26 with a standard deviation of 4.38.

What is the observed sample mean?

$$\overline{x} = 26$$

What is the observed sample standard deviation?

$$s = 4$$

What is the t-statistic when the null hypothesis is true?

$$t = \frac{26 - 24}{4.38/\sqrt{51}} \approx 3.261$$

As or more extreme than you observed

When you collect data and assume the null hypothesis is true, i.e. $H_0 : \mu = m_0$, you calculate the t-statistic using the formula

$$t = \frac{\overline{x} - m_0}{s/\sqrt{n}}.$$

This is what you observe. If

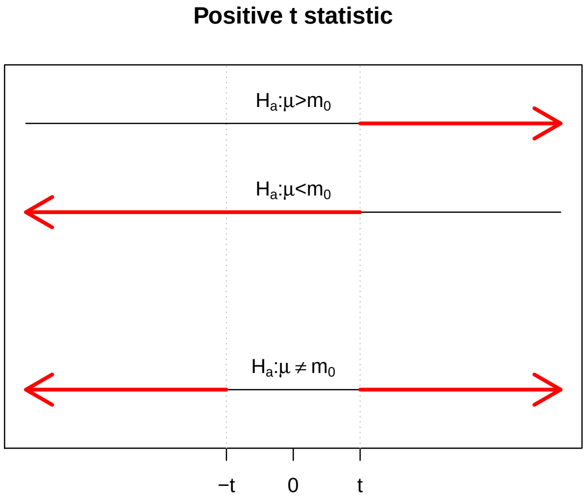
- $\mu = m_0$ then it is likely that $t \approx 0$,
- $\mu > m_0$ then it is likely that $t > 0$, and
- $\mu < m_0$ then it is likely that $t < 0$.

The phrase as or more extreme means away from the null hypothesis and toward the alternative.

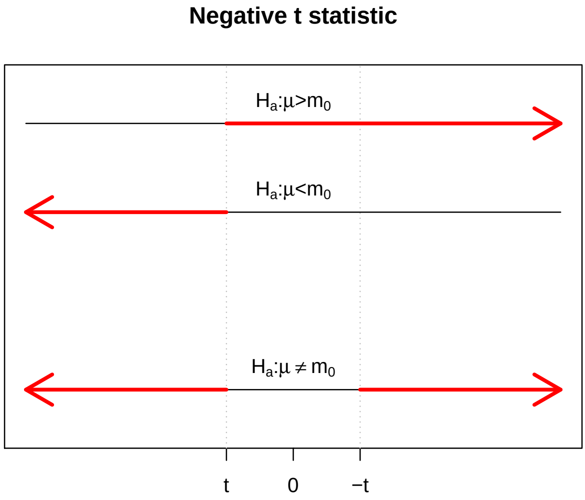
Thus the as or more extreme regions are

- $H_a : \mu > m_0$ implies the region $T_{n-1} > t$,
- $H_a : \mu < m_0$ implies the region $T_{n-1} < t$, and
- $H_a : \mu \neq m_0$ implies the region $T_{n-1} < -|t|$ or $T_{n-1} > |t|$.

As or more extreme than you observed (graphically)



As or more extreme than you observed (graphically)



Sampling distribution of the t-statistic

Recall that if $X_i \stackrel{iid}{\sim} N(\mu, \sigma^2)$, then

$$T_{n-1} = \frac{\overline{X} - \mu}{S/\sqrt{n}} \sim t_{n-1}$$

i.e. T_{n-1} has a t distribution with $n-1$ degrees of freedom.

If the null hypothesis, $H_0 : \mu = m_0$ is true, then

$$T_{n-1} = \frac{\overline{X} - m_0}{S/\sqrt{n}} \sim t_{n-1}.$$

Recall that for random variables, we can calculate probabilities such as the following by calculating areas under the pdf.

- $P(T_5 > 2.015) = 0.05$
- $P(T_{18} > 3.197) = 0.0025$
- $P(T_{26} < -1.315) = P(T_{26} > 1.315) = 0.10$ (by symmetry).

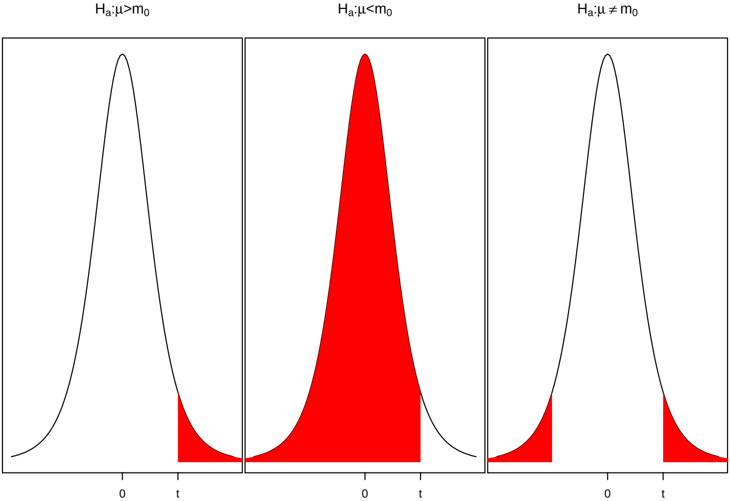
Probability

The (frequency) probability of being as or more extreme than you observed is just the areas under the pdf of a t-distribution with $n-1$ degrees of freedom for the as or more extreme than you observed regions.

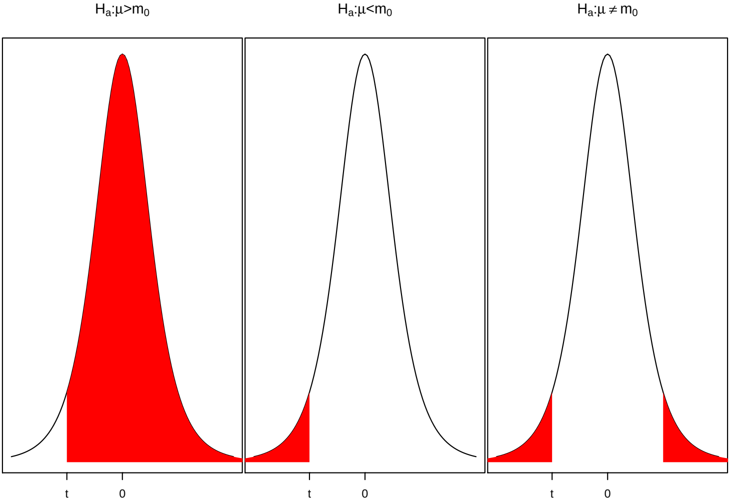
In particular if you observe the t-statistic t and have n observations, then these are the probability calculations associated with each alternative hypothesis:

Alternative hypothesis	Probability
$H_a : \mu > m_0$	$P(T_{n-1} > t)$
$H_a : \mu < m_0$	$P(T_{n-1} < t)$
$H_a : \mu \neq m_0$	$P(T_{n-1} < -\lvert t\rvert \text{ or } T_{n-1} > \lvert t\rvert)$ $= P(T_{n-1} < -\lvert t\rvert) + P(T_{n-1} > \lvert t\rvert)$

Probability (graphically) - positive t



Probability (graphically) - negative t



Calculating probabilities using the t table

Since the t table is constucted for areas to the right, i.e. probabilities such as $P(T_{n-1} > t)$, we need to convert all our probability statements to only have a $>$ sign.

Using symmetry properties of the t distribution, we have

Alternative hypothesis	Probability
$H_a : \mu > m_0$	$P(T_{n-1} > t)$
$H_a : \mu < m_0$	$P(T_{n-1} < t)$ $= P(T_{n-1} > -t)$
$H_a : \mu \neq m_0$	$P(T_{n-1} < -\lvert t\rvert \text{ or } T_{n-1} > \lvert t\rvert)$ $= P(T_{n-1} < -\lvert t\rvert) + P(T_{n-1} > \lvert t\rvert)$ $= 2P(T_{n-1} > \lvert t\rvert)$

P-values for $H_0 : \mu = m_0$

Definition

A p-value is the (frequency) probability of obtaining a test statistic as or more extreme than you observed if the null hypothesis (model) is true.

So for the null hypothesis $H_0 : \mu = m_0$, calculate

$$t = \frac{\overline{x} - m_0}{s/\sqrt{n}}$$

and find the appropriate probability:

- $H_a : \mu \neq m_0$ implies p-value $= 2P(T_{n-1} > |t|)$,
- $H_a : \mu < m_0$ implies p-value $= P(T_{n-1} > -t)$, and
- $H_a : \mu > m_0$ implies p-value $= P(T_{n-1} > t)$.

ACT scores example

The mean composite score on the ACT among the students at Iowa State University is 24. We wish to know whether the average composite ACT score for business majors is different from the average for the University. We sample 51 business majors and calculate an average score of 26 with a standard deviation of 4.38.

Let X_i be the composite ACT score for student i who is a business major at Iowa State University. Assume $X_i \stackrel{iid}{\sim} N(\mu, \sigma^2)$.

- Null hypothesis $H_0 : \mu = 24$
- Alternative hypothesis $H_a : \mu \neq 24$
- t-statistic:

$$t = \frac{26 - 24}{4.38/\sqrt{51}} \approx 3.261$$

- p-value:

$$\begin{aligned} 2P(T_{n-1} > |t|) \quad &= 2P(T_{50} > 3.261) \\ &= 2 \cdot 0.001 \\ &= 0.002 \end{aligned}$$