Strength of Evidence + Practice

Grinnell College

November 11, 2024

Review

Last week we introduced the idea of the **null distribution**

If we were to collect many samples of $\overline{X}$, the null distribution refers to the distribution of statistics

$$t = \frac{\overline{X} - \mu_0}{\hat{\sigma}/\sqrt{n}}$$

when $H_0 : \mu = \mu_0$, i.e., the null hypothesis is true

Consider the t

Consider the pieces of a t-statistic

$$t = \frac{\overline{x} - \mu_0}{\hat{\sigma}/\sqrt{n}}$$

1. $\overline{x} - \mu_0$ indicates the distance between my observed data and my null hypothesis, though we cannot use this alone as it does not include a degree of "certainty" associated with $\overline{x}$
2. $\hat{\sigma}$ is my estimate of the population's standard deviation. When this is large, there will be more uncertainty in my estimate of $\overline{x}$
3. n represents the number of observations in my sample – the more observations I have, the more confidence I will have in my estimate

We should have a sense of how each of these components impact my t-statistic

Consider the t

$$t = \frac{\overline{x} - \mu_0}{\hat{\sigma}/\sqrt{n}}$$

We can think, then, of the t-statistic as being a measure of evidence *against* the null hypothesis

If:

1. $\overline{x}$ is far from μ and
2. Our certainty in $\overline{x}$ is high (i.e., low $\hat{\sigma}$ or large n)

then our statistic t will be larger. A larger t statistic is less likely than a smaller one

What we consider "large" will depend on the t-distribution

t-distribution

When the null hypothesis is true,

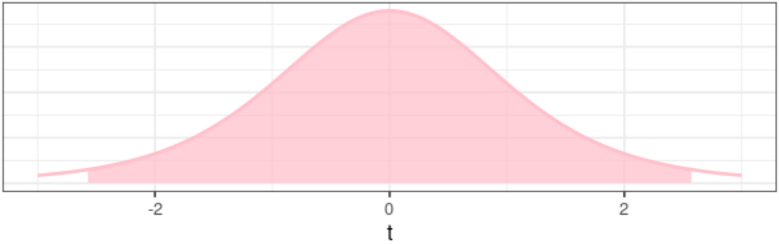
$$t = \frac{\overline{x} - \mu_0}{\hat{\sigma}/\sqrt{n}}$$

follows a t-distribution with $n - 1$ degrees of freedom

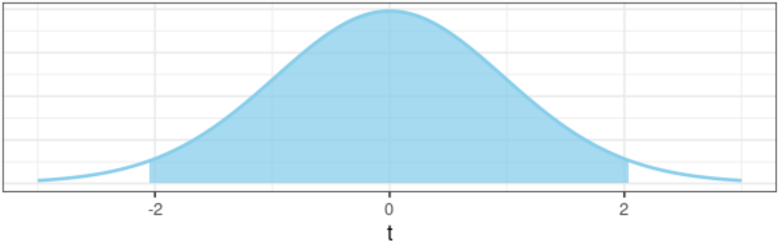
The degrees of freedom tells us, relatively speaking, what values are considered "large"

$t = 2.1$ may be considered "large" when $df = 30$ but not when $df = 5$

95% of a t-distribution with df = 5



95% of a t-distribution with df = 30



In summary

The process goes like this:

1. Assume our null hypothesis $H_0 : \mu = \mu_0$ is true
2. Compute a t-statistic with our observed data
3. Ask: is this t-statistic "large"?
   - This will depend on the degrees of freedom
   - It will also depend on what range we consider acceptable, i.e., 80%, 95%, 99%, etc.,
4. (After the exam) If we know our t statistic and we know our degrees of freedom, we can find the probability of observing our data if the null is true. This probability is our **p-value**

Question 4 Null Distribution

Suppose $\overline{x} = 24$, $\hat{\sigma} = 5$ and $n = 25$

Question:

1. Would $\overline{x}$ be in a 95% CI around $\mu_0 = 20$?
2. What about $\mu_0 = 27$?
3. Which null hypothesis does our sample provide more evidence *against*?

Q4

$$t = \frac{24 - \mu_0}{5/\sqrt{25}}$$

We see that $\hat{\sigma}/\sqrt{n} = 1$. From our sheet we find $C = 2.0639$ for $df = 24$

1. If $\mu_0 = 20$, a 95% CI around this value would be

$$20 \pm 2.0639 \times 1 = (17.936, 22.064)$$

2. If $\mu_0 = 27$, a 95% CI around the null would be

$$27 \pm 2.0639 \times 1 = (24.936, 29.064)$$

Q4

Neither of these intervals contain $\overline{x}$ so our observed data provides evidence against each. To see which hypothesis our data provides *more* evidence against, we can compute t-statistics

1. $$t = \frac{24 - 20}{5/\sqrt{25}} = 4$$

2. $$t = \frac{24 - 27}{5/\sqrt{25}} = -3$$

Because these come from the same distribution (i.e., $df = 24$ in both cases), we can compare these directly and see that we have more evidence against the hypothesis that $\mu_0 = 20$

Question 5

Scenario 1
In a sample with $n = 20$, we find that $\hat{\sigma} = 2$ and $\overline{x} = 15.916$. Our null hypothesis is $\mu_0 = 15$
Scenario 2
In a sample with $n = 40$, we find that $\hat{\sigma} = 5$ and $\overline{x} = 126.62$. Our null hypothesis is $\mu_0 = 125$

1. Which of these two scenarios shows the greatest difference between the observed and hypothesized mean?
2. Which of these scenarios do you think offers more compelling evidence that the null hypothesis is false?

Q5

Scenario 1

$$t = \frac{15.916 - 15}{2/\sqrt{20}} = \frac{0.916}{0.447} = 2.048$$

Scenario 2

$$t = \frac{126.62 - 120}{5/\sqrt{40}} = \frac{1.62}{0.791} = 2.049$$

Question:

1. Which shows largest difference between observed and hypothesized mean?
2. Which provides more compelling evidence that the null false?

Exam

- Probability stuff
- CLT – when does it apply and what does it say?
- Sampling distributions
- Confidence intervals and critical values
- t-distribution, t-statistics
- Null hypothesis and p-values