NRES 776 Lecture 7

Hypothesis test II

Sunny Tseng

Our schedule today

- Announcement (5 min)
- Three ways to conduct hypothesis test (25 min)
- Recap and project I (15 min)
- Wrap up (5 min)

Announcement

- Recording
- Let’s take this lecture easy…

Artwork by @allison_horst

Hypothesis test: two hypotheses

Hypothesis test is a decision making process. We use it to translate a question of interest into a mathematical form. Then, we can make inference to population parameters.

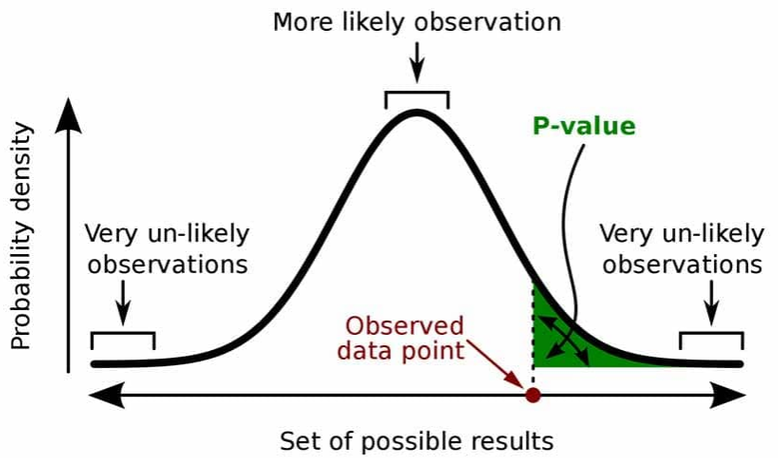
- You can formulate question into two mutual exclusive hypotheses:
  - H_0: **Null hypothesis** (equal to, no effect, the same as, etc.)
  - H_a: **Alternative hypothesis** (not equal to, different, etc.)

Research question	H_0	H_a
Does tooth flossing affect the number of cavities?	Tooth flossing has **no effect** on the number of cavities. $\mu_{floss} = \mu_{nofloss}$	Tooth flossing has **effect** on the number of cavities. $\mu_{floss} \neq \mu_{nofloss}$

- We usually use $=$ in null hypothesis
- Always use **parameters** (e.g., μ, β) when writing out hypothesis

Put everything together

1. **Set up H_0 and H_a**
   - e.g., $H_0 : \mu = 180, H_a : \mu \neq 180$
2. **Determine statistical significance level α**
   - usually be 0.05 or 0.1
3. **Calculate the test statistics**
   - If doing t-test, calculate $t_{statistics} = \frac{\bar{x} - \mu}{s/\sqrt{n}}$
   - If doing z-test, calculate $z_{statistics} = \frac{\bar{x} - \mu}{s/\sqrt{n}}$
4. **Find the p-value corresponded to the test statistics**
   - p-value $= P(t \geq t_{statistics})$
   - The probability of a value larger than test statistics
5. **Compare α and p-value and draw conclusion**
   - $p - value \geq \alpha$: fail to reject H_o
   - $p - value < \alpha$: reject H_0



A **p-value** (shaded green area) is the probability of an observed (or more extreme) result assuming that the null hypothesis is true.

JPRF black bear example

The JPRF manager claim that the population mean (μ) of the black bear weight is 390 kg. Is this claim valid?

You, as a researcher, came to the forest and measured the weight from 25 black bears in JPRF forest, and got their averaged weight as 400 kg, with standard deviation **40** kg.

- $n = 25$
- $\bar{x} = 400$
- $s = 40$

Image from: https://www.flickr.com/photos/archer10/37223656595

JPRF black bear example

1. Set up H_0 and H_a

- $H_0 : \mu = 390$: the population mean of the black bear weight is 390 kg.
- $H_a : \mu \neq 390$, the population mean of the black bear weight is not 390 kg.

2. Determine statistical significance level α

- We set the significance level α as 0.1

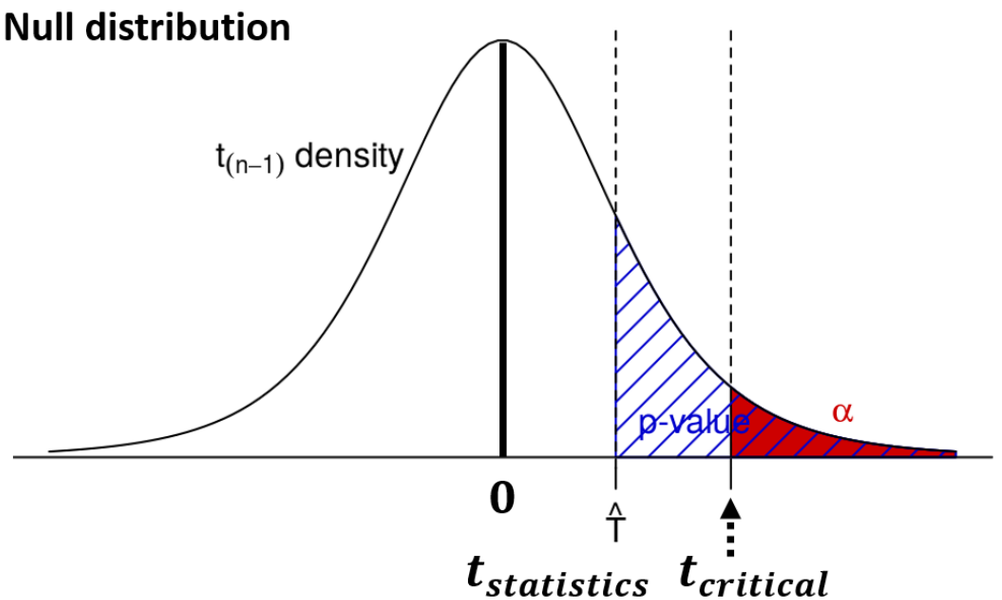
3. Calculate the test statistics

- We are using t-statistics (because n is small)
- $t_{statistics} = \frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{400 - 390}{40/\sqrt{25}} = 1.25$

JPRF black bear example

4. Find the p-value for test statistics

- Using t-table, or R, or online calculator
- Find the p-value for $t_{statistics} = 1.25$
  - $df = n - 1 = 24$
  - two-tailed
- $p - value = 0.22$



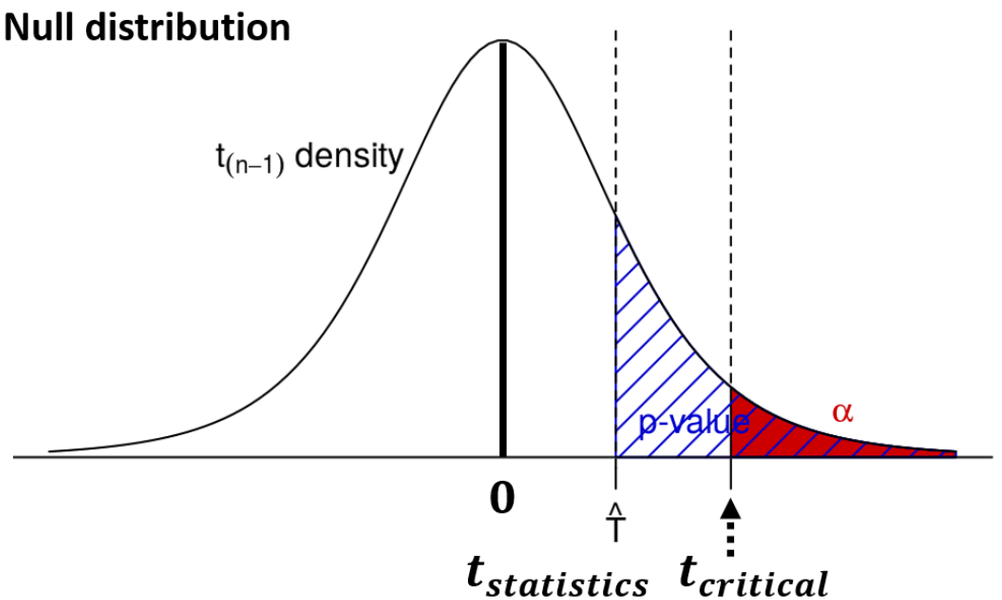
5. Compare α and p-value and draw conclusion

- $p - value > \alpha$, we fail to reject H_0.
- Thus, we support the statement that the population mean of the black bear weight is equal to 390.

JPRF black bear example (other methods)

4.1 Find the critical statistics for α

- Using t-table, or R, or online calculator
- Find the $t_{critical}$ corresponded with α
  - $df = n - 1 = 24$
  - two-tailed
- $t_{ctitical} = 1.711$



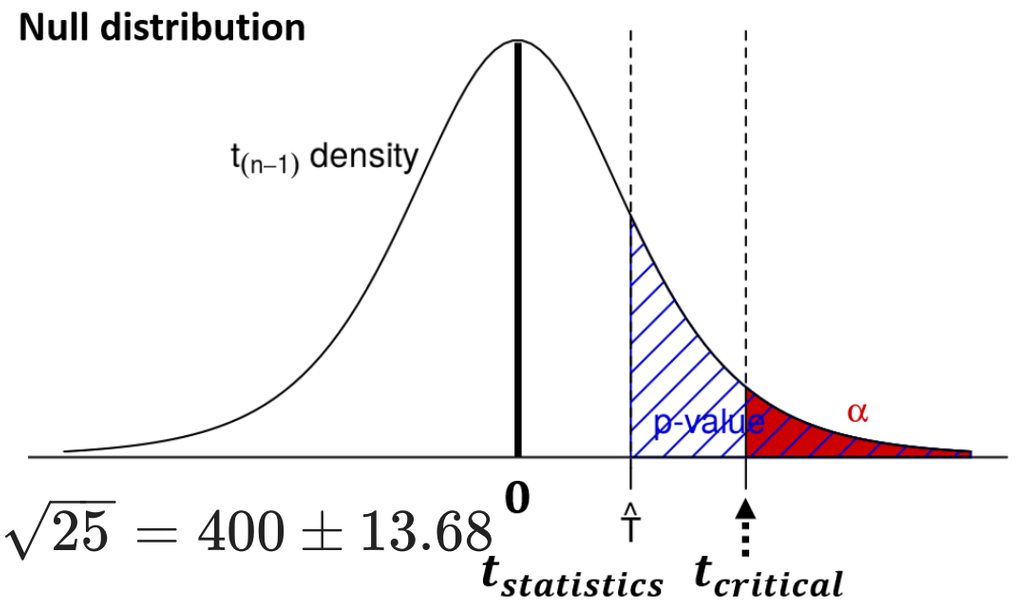
5.1 Compare $t_{statistics}$ and $t_{critical}$ and draw conclusion

- $t_{statistics} \leq t_{critical}$, we fail to reject H_0.
- Thus, we support the statement that the population mean of the black bear weight is equal to 390.

JPRF black bear example (other methods)

4.2 Find upper and lower limits

- Using t-table, or R, or online calculator
- Find the $t_{critical}$ corresponded with $\alpha = 0.1$
  - $df = n - 1 = 24$
  - two-tailed
- $t_{ctitical} = 1.711$
- $CI = \bar{x} \pm t_{critical} * \frac{s}{\sqrt{n}} = 400 \pm 1.711 * 40/\sqrt{25} = 400 \pm 13.68$



5.2 Compare CI with μ and draw conclusion

- $\mu = 390$, which is within (386.3, 413.7), we fail to reject H_0.
- Thus, we support the statement that the population mean of the black bear weight is equal to 390.

Three ways to do hypothesis test

A. compare p-value and α

- $p-value \geq \alpha$: fail to reject H_0
- $p-value < \alpha$: reject H_0

B. compare statistical value vs critical value

- $t_{statistics} < t_{criticle}$: fail to reject H_0
- $t_{statistics} \geq t_{criticle}$: reject H_0

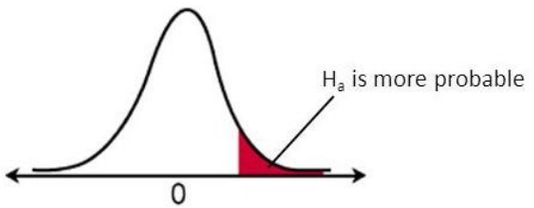
C. compare confidence interval

- If the confidence interval covers μ: fail to reject H_0
- Otherwise: reject H_0

One-tailed or two-tailed test?

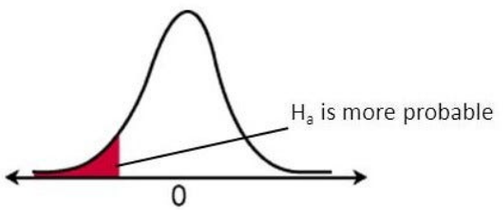
How to tell

- Based on the research question
- If the H_0 is written with $=$, then it's a two-tailed test
- If the H_0 is written with $>$ or $<$, then it's a one-tailed test



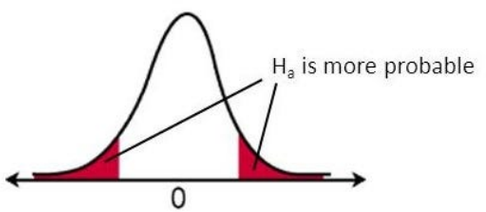
Right-tail test

H_a: $\mu >$ value



Left-tail test

H_a: $\mu <$ value



Two-tail test

H_a: $\mu \neq$ value

So what

- This would influence how you look up your p-value

- Two-tailed test, when looking up the p-value, the probability needs to multiply by 2.
- Two-tailed test, when looking up the $t_{critical}$, the probability is $\alpha/2$.

Gum thickness: a two-tailed t-test

The story

A manufacturer claims that the thickness of the spearmint gum it produces is 7.5 one-hundredths.

A quality control specialist regularly checks this claim. On one production run, he took a random sample of *n* = 10 pieces of gum and measured their thickness.

He obtained:

Thicknesses of 10 Pieces of Gum				
7.65	7.60	7.65	7.70	7.55
7.55	7.40	7.40	7.50	7.50

$$n = 10, \bar{x} = 7.55, s = 0.1027, \alpha = 0.05$$

Base on this sample, is the claim from the manufacturer valid?

Gum thickness

To solve this

1. Set up H_0 and H_a

- $H_0 : \mu = 7.5$: the mean of the spearmint gum thickness equals 7.5
- $H_a : \mu \neq 7.5$, the mean of the spearmint gum thickness doesn't equal 7.5

2. Determine statistical significance level α

- We set the significance level α as 0.05

3. Calculate the test statistics

- We are using t-statistics (because n is small)
- $t_{statistics} = \frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{7.55 - 7.5}{0.1027/\sqrt{10}} = 1.54$

Gum thickness

To solve this

4. Find the p-value for test statistics

- Using t-table, or R, or online calculator
- Find the p-value for $t_{statistics} = 1.54$
  - $df = n - 1 = 9$
  - **two-tailed**
- $p - value = 0.158$

5. Compare α and p-value and draw conclusion

- $p - value > \alpha$, we fail to reject H_0.
- Thus, we support the statement that the mean of the spearmint gum thickness equals 7.5.

Gum thickness

Can you try the CI method?

$$n = 10, \bar{x} = 7.55, s = 0.1027, \alpha = 0.05$$

$$CI = \bar{x} \pm t_{\frac{\alpha}{2}, n-1} * \frac{s}{\sqrt{n}}$$

$$CI = 7.55 \pm 2.262 * \frac{0.1027}{\sqrt{10}} = (7.476, 7.623)$$

What we learned in the past month

Statistics

- Population vs sample, descriptive statistics
- Types of variables, probability distribution
- Sampling distribution, central limit theorem and confidence interval
- Hypothesis test

R coding

- Basic R (R & RStudio, object, function, data type)
- Data visualization and data I/O
- Data wrangling

Project I

Submission on Moodle (Oct.19)

- 2 pages report “NRES776_*firstname_lastname*_project_1.**pdf**” (23:59)
- 5 mins oral presentation “NRES776_*firstname_lastname*_project_1.ppt” (8:00)

Content

- Research background (e.g., efficient audio analysis methods is needed in avian study)
- Hypothesis (e.g., frequency features in audio can be used to identify sound sources)
- Methods (e.g., use logistic modelling, a type of GMM, to identify significant features, or variables)
- Potential contribution (i.e., how your results be helpful for science/public/or your audience)
- They doesn’t need to be in this order

Project I

Statistical methods

- Multiple linear regression
- Generalized linear models
- Mixed-effect models
- Multivariate analysis
- If none of these can be applied to your data, or you are not sure, come have a chat with me or Lisa. The goal of this project is to help you practice the methods we learned in this course, which doesn’t cover all analysis methods.

Encouraged

- Data visualization would be fun
- Funding proposal for a lay audience
- Feel free to be creative

Wrap up

Before we meet again

- Review the two lectures (important)
- Review and practice with the R coding (also important)
- Maybe use what we learned to explore your own data

Next time

- Oct. 19 @ 8 am lab, project I presentation, **In Person**, I will be in Prince George 
- Before that, enjoy the lecture and lab with Lisa 😀