2 / 44

LECTURE 08: t-tests

ENVS475: Exp. Design and Analysis

Spring 2023

1 / 44

outline

1) t-test overview

2) One-sample t-test

3) Null hypothesis testing

4) Two-sample t-test

5) Paired t-test

2 / 44

t-test overview

- Is my sample different?
  - Value of interest
  - Two samples different?
- General formula of t-test:

$$t_{statistic} = \frac{difference}{variation}$$

- No difference: $t_{statistic} = 0$
- There is a difference: $t_{statistic} \neq 0$

History of t-test

- t-test developed by William Sealy Gossett
- Worked at Guinness, wanted to test quality of stout
- Very small sample sizes ($n \leq 4$)
- Uses the t-distribution, which can be thought of as a small-sample-size version of a standard normal distribution
- Published under pseudonym *Student* because company policy prevented employees from publishing scientific papers

Types of t-tests

1) One Sample

- Is my sample different from some value of interest?

2) Two-sample

- Are two populations different from one another?

3) Paired t-test

- Two measurements on same experimental unit

one sample t-test

Context

- We want to know if a population mean (μ) differs from some value μ_0
- Formally, we want to test the following hypotheses:

$$H_0 : \mu - \mu_0 = 0$$

$$H_a : \mu - \mu_0 \neq 0$$

- **Note** that this is a two-tailed test
- There is also a one-tailed alternative
- This class takes a linear model approach
- All hypotheses in this course will be two-tailed

one sample: motivating example

- We just got a new piece of equipment, and want to ensure that the measurement error is 0

- We take 10 measurements on a known quantity to get a sample to test

- data:

    - 0.5, -0.2, 0.2, 0.1, 0.1, 0.2, 0.2, -0.1, 0.1, -0.1

- Estimate of average measurement error(μ) = 0.1

- μ will almost never exactly equal μ_0 due to sampling variability.

- So how do we answer our original question?

7 / 44

NULL HYPOTHESIS TESTING

8 / 44

NHT

> Formal approach to assessing if sample data represents a
> statistically significant difference, or is due to sampling error.

- requires two hypotheses:
  - The null, H_0, which is no difference/relationship/effect
  - Alternative, H_A, which is that there is some difference/relationship/effect
  - Hypotheses refer to the **populations**

One sample t-test procedure

1) Draw a random sample from a population

2) Calculate the mean and standard error of the mean (SEM)

3) Compute a $t_{statistic}$

$$t_{statistic} = \frac{\bar{y} - \mu_0}{SEM}$$

4) Compare the $t_{statistic}$ with a *critical value* ($t_{critical}$

- If $|t_{statistic}| > |t_{critical}|$ we **Reject the NULL hypothesis**
- If $|t_{statistic}| < |t_{critical}|$ we **FAIL to Reject the NULL hypothesis** (FTR)

5) Use $t_{statistic}$ to compute a p_{value}

- If $p < \alpha$: REJECT
- If $p > \alpha$: FTR

One sample t-test procedure

1) Data: 0.5, -0.2, 0.2, 0.1, 0.1, 0.2, 0.2, -0.1, 0.1, -0.1

2) mean: 0.1

- SEM: 0.063

3) Compute a $t_{statistic}$

$$t_{statistic} = \frac{\bar{y} - \mu_0}{SEM}$$

$$t_{statistic} = \frac{0.1 - 0}{0.063} = 1.581$$

4) Compare the $t_{statistic}$ to $t_{critical}$

basic idea

Reject the null hypothesis if $t_{statistic}$ is larger than would be expected under the null hypothesis

This will happen if:

- The sample mean is far from 0

- And/or the SE is small

null hypothesis testing

Question:

> How do we know the likely values of t under the null hypothesis?

Answer:

> Theory says that the test statistic will follow a t-distribution with $n - 1$ degrees of freedom, **if the null hypothesis is true**

THE t-DISTRIBUTION

- t-distribution is a small-sample-size version of standard normal distribution

- with fewer samples, there is more uncertainty

- this uncertainty is represented in "fat tails"
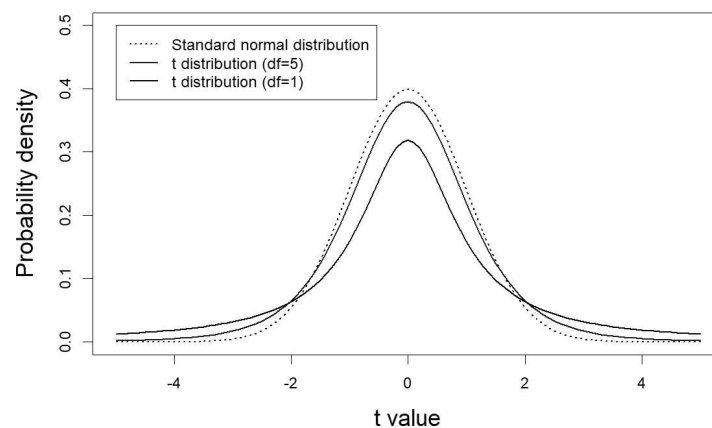
calculate $t_{critical}$ values

- There's always a chance that our data is far from the expected value by chance

- This is called the false-positive, or type I error rate, and is set as an α value

- By convention, α usually equals 0.05, and most statistical software (including R) uses it as a default

- This can vary, and you will occasionally see papers that set α value to 0.1, 0r 0.01

- for this class, we will leave $\alpha = 0.05$

- split α into both tails of distribution

15 / 44

CRITICAL VALUES FOR $\alpha = 0.05$

t distribution with df=9
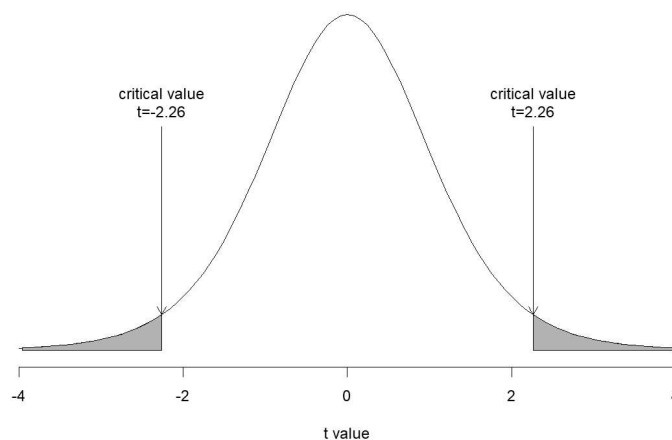
16 / 44

t-critical value in R

- Remember the `qnorm`, `dnorm`, and `pnorm` functions?

- Have equivalents for the t-distribution: `qt`, `dt`, and `pt`

- If we want to know what **value** is at a certain **quantile** of our data, we use the `qt()` function

- critical value = `qt(0.025, df = 9)` = -2.26

- We can also use the `pt()` function to calculate a p-value by hand

- `pt(t_stat, df = 9, lower.tail = FALSE) * 2` = 0.148

   - Use the absolute values to avoid confusion between signs
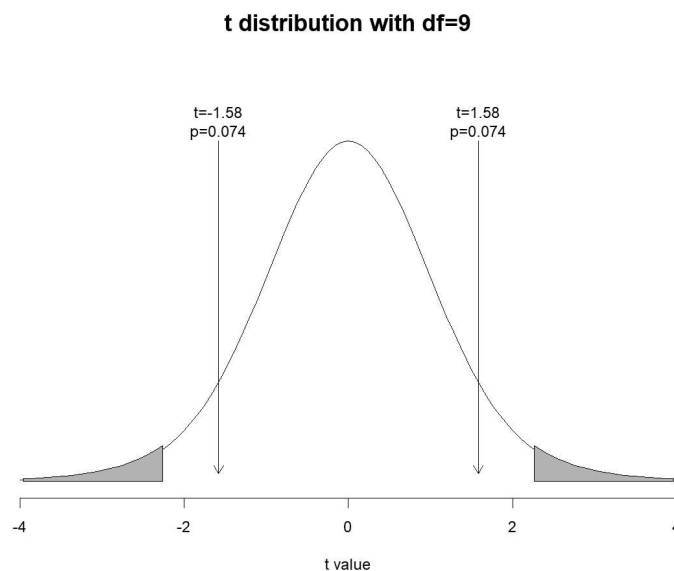
   - multiply value by 2 since we have 2-tails

p-values

t distribution with df=9

MORE ON p-VALUES

A p-value tells us how likely the null hypothesis is, given your observations

Our conclusion must be to either reject or "fail to reject" the null hypothesis

A p-value does not tell us how much evidence there is in favor of a particular difference in means

one-tailed vs. two-tailed tests

t distribution with df=9
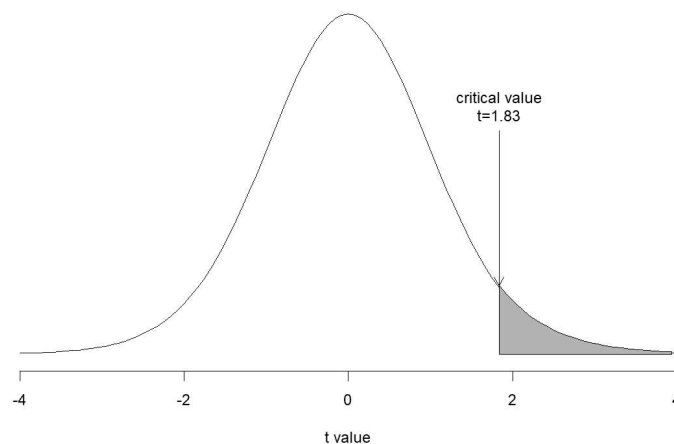
more on degrees of freedom

The degrees of freedom for a calculation on a set of numbers is the number of elements in the set (i.e., how many numbers there are) minus the number of different things you must know about the set in order to complete the calculation

Example:

Consider a set of n = 5 numbers. In the absence of any information about them, all ♠ve are free to be any value. However, if you are also told that the sum of the set is 20, then only 4 of the numbers are free to be anything, but the fifth is constrained by your knowledge that the sum must be 20. Hence, $df = n - 1 = 4$

Big t-statistics = REJECT null

- What makes a big $t_{statistic}$?

- Large difference (big numerator)
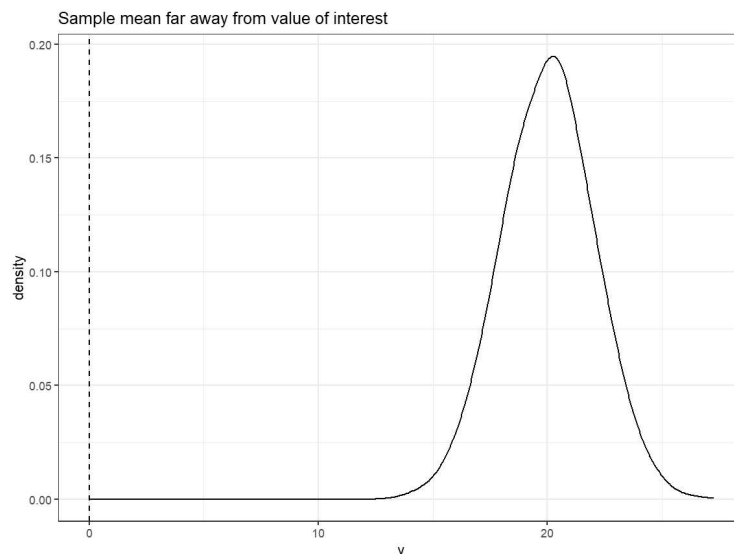
Sample mean far away from value of interest

Big t-statistics = REJECT null

- What makes a big $t_{statistic}$?

- Little variation in data (small numerator)
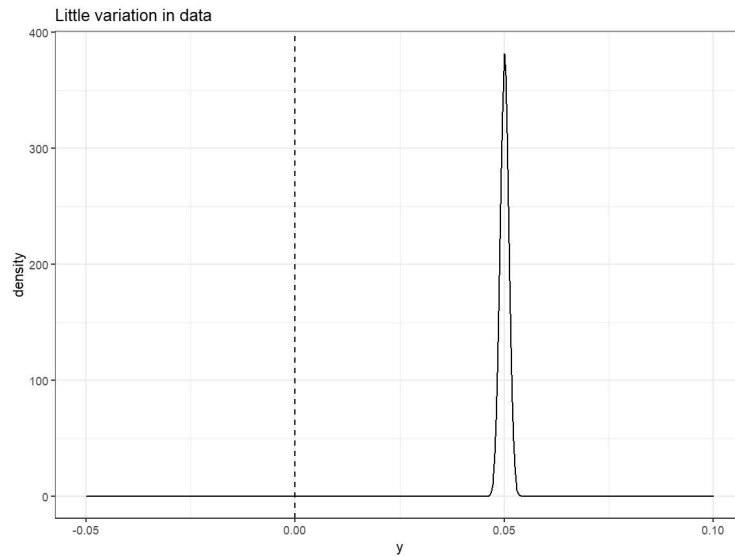
simple linear model:

Intercept-only

$$y_i = \beta_0 + \epsilon_i$$

$$\epsilon_i \sim N(0, \sigma)$$

- Where $\beta_0 = Intercept = \mu$

- Equation becomes:

$$y_i = \mu + \epsilon_i$$

Intercept-only model

- use `lm()` function to create intercept-only model

```
df_one <- data.frame(y = c(0.5, -0.2,  0.2,  0.1,  0.1,  0.2, 0.2, -0

lm0 <- lm(y ~ 1, data = df_one)

summary(lm0)
```

```
##
## Call:
## lm(formula = y ~ 1, data = df_one)
##
## Residuals:
##    Min     1Q Median     3Q    Max
##  -0.30  -0.15   0.00   0.10   0.40
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.10000    0.06325   1.581    0.148
##
## Residual standard error: 0.2 on 9 degrees of freedom
```

one sample interpretation

Based on our data, we fail to reject the null hypothesis that our measurement error is equal to 0 ($t_{statistic} = 1.58, ~p = 0.15, ~df = 9$).

Remember that our example here was we were assessing a new piece of equipment. We fail to reject the null, which means that our device is accurate and measures the quantities within the expected amount of error.

TWO-SAMPLE t-TEST

TWO-SAMPLE t-TEST

Concept

- We want to determine if two population means differ

- The null hypothesis is: $H_0 : \mu_1 = \mu_2$

- Note that this is the same as: $H_0 : \mu_1 - \mu_2 = 0$

- The two-tailed alternative hypothesis is $H_a : \mu_1 \neq \mu_2$

- Note that this is the same as: $H_0 : \mu_1 - \mu_2 \neq 0$

- Appropriate when:

    - The two samples, one from each population, are independent
    - Both populations are (approximately) normally distributed
    - The population variances are unknown but are the same for both populations

procedure

1) Draw two random samples from two populations

2) Compute the standard error of the difference in means:

$$SEDM = \sqrt{SEM_1^2 + SEM_2^2}$$

3) Compute the t statistic:

$$t = \frac{\bar{y}_1 - \bar{y}_2}{SEDM}$$

4) If t is more extreme than the critical values, reject the null hypothesis

5) Critical value is based on $t_{\alpha,df} = n_1 + n_2 - 2$

worked example

Question:

- Is there a difference in the density of trees at low and high elevations?

Hypothesis:

- $H_0 : \mu_{low} - \mu_{high} = 0$
- $H_A : \mu_{low} - \mu_{high} \neq 0$

Field procedure:

- $n = 10$ plots are sampled using randomly located belt transects 100m long $\times$ 10m wide at both high and low elevations
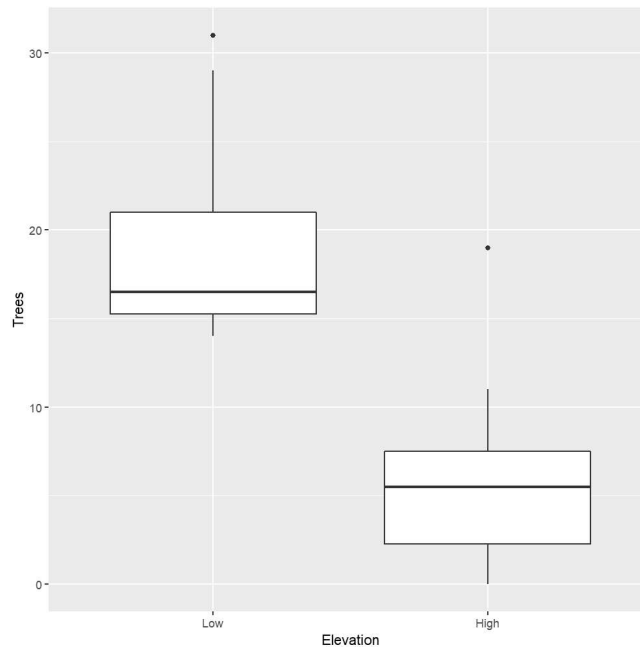
Data:

```
Low elevation: 16, 14, 18,        High elevation: 2, 11, 6, 8,
17, 29, 31, 14, 16, 22, 15        0, 3, 19, 1, 6, 5
```

worked example

worked example

Low elevation   High elevation

16, 14, 18, 17, 29,   2, 11, 6, 8, 0,
31, 14, 16, 22, 15   3, 19, 1, 6, 5

- Mean of low group: $\bar{y}_L = 20.2$

- Mean of high group: $\bar{y}_H = 6.1$

- Standard deviation of low group: $s_L = 6.03$

- Standard deviation of high group: $s_H = 5.63$

- Standard error of difference in means $SEDM_1 = 2.61$

- Test statistic: $t = (20.2 - 6.1)/2.61 = 5.4$

- Critical value: $t_{0.975, df=10+10-2} = 2.1$, p-value: < 0.001

simple linear model:

Categorical predictor

$$y_i = \beta_0 + \beta_1 * x_i + \epsilon_i$$

$$\epsilon_i \sim N(0, \sigma)$$

- Where $\hat{y}_{high} = \beta_0$

- And $\hat{y}_{low} = \beta_0 + \beta_1$

33 / 44

Categorical predictor

- use `lm()` function to create intercept-only model

```
lm1 <- lm(Trees ~ Elevation, data = df)
summary(lm1)
```

```
##
## Call:
## lm(formula = Trees ~ Elevation, data = df)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -6.100 -4.125 -1.700  2.125 12.900
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)     19.200      1.866  10.291 5.73e-09 ***
## ElevationHigh  -13.100      2.638  -4.965    1e-04 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 5.9 on 18 degrees of freedom
## Multiple R-squared:  0.578,    Adjusted R-squared:  0.5545
## F-statistic: 24.65 on 1 and 18 DF,  p-value: 0.0001001
```

34 / 44

NHT in linear models

- Linear models have a null hypothesis for each coefficient (row) in the output table

Intercept (β_0)

- $H_0 : \beta_0 = 0$

    - Intercept is just the estimated mean for one of the groups
    - Often not actually relevant to the research question

Regression coeficient (β_1)

- $H_0 : \beta_1 = 0$

    - β_1 is the estimated difference between means
    - Here, this is the question of interest
    - Because we have a low p-value, we can reject the null hypothesis and accept the alternative that the densities do differ based on elevation

35 / 44

PAIRED *t*-TEST

36 / 44

PAIRED *t*-TEST

Context

- Used when two measurements are taken on each experimental unit

    - Measure before and after a drug treatment

    - What is the effect of mining effluent on paired streams?

- Problem can be analyzed by taking differences of each pair and then conducting a one-sample *t*-test

PAIRED *t*-TEST

PAIRED *t*-TEST

Context

- Used when two measurements are taken on each experimental unit

- Problem can be analyzed by taking differences of each pair and then conducting a one-sample *t*-test

- Examples:
  - Does a new pharmaceutical decrease blood cholesterol?
  - Is small mammal density higher before or after the use of prescribed fire?

39 / 44

motivation

Does Pharmaceutical lower LDL levels?
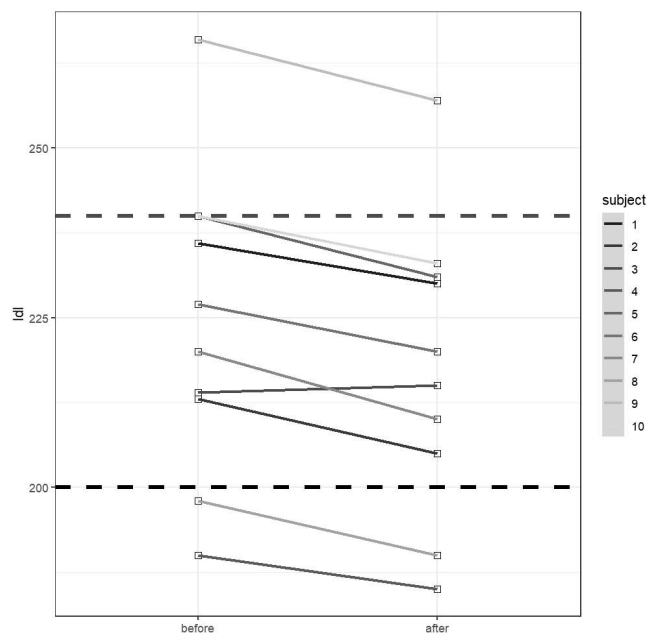
40 / 44

worked example

Hypotheses (μ_d is the mean difference)

- $H_0 : \mu_d = 0$
- $H_a : \mu_d \neq 0$

t-statistic

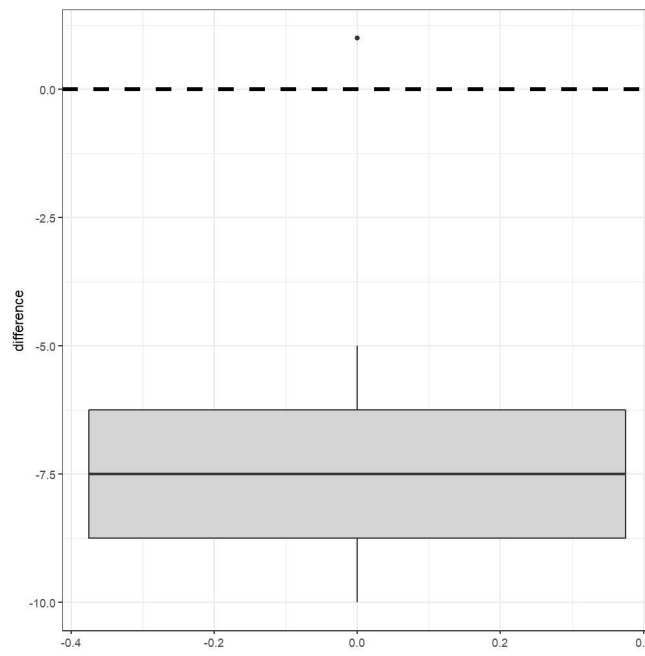$$t = \frac{\bar{y}_{Diff} - 0}{SED}$$

worked example

Calculations

- Mean differences: $\bar{y}_d = -6.8$

- Standard deviation of differences: $s_d = 3.12$

- Standard error of mean differences: $SEM_d = 0.987$

- Test statistic: $t = -6.8/9.82 = -6.893$, Critical value: $t_{0.95,10} = -2.26$

- $p < 0.001$

looking ahead

This week: t-tests and Null Hypothesis Testing lab and HW

Reading: Hector Ch. 6, 9

Take Home Exam: Assigned Monday after Spring Break, Due Friday after Spring Break

Aknowledgements This lecture was largely based on Clark Rushings NR6750 course materials