PSY 3307

t-test Combined

Jonathan A. Pedroza PhD

Cal Poly Pomona

2021-10-12

What is a One-Sample t-test

- It's pretty similar to a z-test
  - t-test used more often in behavioral research
- z-test requires we know population standard deviation
  - often not possible in behavioral research
- uses unbiased estimators (N - 1 formulas)
- computes something like the z-score for our sample mean
  - t-score

One-Sample t-test

- parametric test for when the population standard deviation is unknown
- still compares the sample mean to the population mean

Steps to One-Sample t-test

1. Statistical Hypotheses
   - what is the population mean and is your sample mean different from that population mean
   - H0: sample mean equals the population mean
   - H1: sample mean is different from the population mean
2. Select an alpha
3. Check assumptions
   - Outcome needs to be continuous (interval or ratio scale)
   - Population score forms a normal distribution
   - variability of raw score population is estimated from the sample

Steps to One-Sample t-test

- All we need to know is the t critical value and if the t obtained value is within the regions of rejection

Steps to a z-test/One-Sample t-test

- get population standard deviation (z-test)
- get estimated standard deviation (t-test)
- get the standard error (SE) of the mean (z-test)
- get the **estimated** SE (t-test)
- calculate the score by subtracting the population mean from the sample mean and dividing by the SE
    - either obtained z or t value

Changes between the z-test and t-test

$$S_X^2 = \frac{\Sigma X^2 - \frac{(\Sigma X)^2}{N}}{N-1}$$

$$S_x = \sqrt{\frac{\Sigma X^2 - \frac{(\Sigma X)^2}{N}}{N-1}}$$

$$\sigma_{\overline{X}} = \frac{\sigma_X}{\sqrt{N}}$$

$$z_{obt} = \frac{\overline{X} - \mu}{\sigma_{\overline{X}}}$$

$$S_{\overline{X}} = \frac{S_X}{\sqrt{N}}$$

$$t_{obt} = \frac{\overline{X} - \mu}{S_{\overline{X}}}$$

Small Change in Formulas

- SE calculation will start to look slightly different as it will use the variance squared

- Due to future formulas using slightly different notation, we will adopt that for our SE

$$S_{\overline{X}} = \sqrt{\frac{S_x^2}{N}}$$

Example

```
set.seed(092221)

numbers = rnorm(10, mean = 5, sd = 1.2)

numbers
```

```
##  [1] 4.770670 5.271329 6.899812 5.598090 4.219022 6.828034 6.046497 2.92124
##  [9] 4.993870 5.948126
```

1. Calculate the Variance
2. Calculate the SE
3. Compute t

```r
# population mean is 10

4.770670 + 5.271329 + 6.899812 + 5.598090 + 4.219022 + 6.828034 + 6.0464
```

```
## [1] 53.4967
```

```r
# 53.50

53.50/10
```

```
## [1] 5.35
```

```r
# 5.35

4.770670^2 + 5.271329^2 + 6.899812^2 + 5.598090^2 + 4.219022^2 + 6.82803
```

```
## [1] 299.3272
```

```r
# 299.33

53.50^2
```

```
## [1] 2862.25
```

```r
# 2862.25
```

```r
2862.25/10
```

```
## [1] 286.225
```

```r
# 286.23

299.33 - 286.23
```

```
## [1] 13.1
```

```r
# 13.1

13.1/9
```

```
## [1] 1.455556
```

```r
# 1.46 variance
```

```r
# se is...
1.46/10
```

```
## [1] 0.146
```

```r
sqrt(1.46/10)
```

```
## [1] 0.3820995
```

```r
sqrt(.146)
```

```
## [1] 0.3820995
```

```r
# compute t
(5.35 - 10)/.38
```

```
## [1] -12.23684
```

```r
# t value of -12.24
```

t-distribution & Degrees of Freedom (df)

- we will now be working with the t-distribution
    - this also means we'll be working with a t-table
- **t-distribution** is the sampling distribution of all values of t when samples of a particular size (differing N size) are selected from the raw score population in the null hypothesis

t-distribution & Degrees of Freedom (df)

- higher values on the t-distribution are to the right of the population mean, lower values to the left of the population mean
- t-tests also have regions of rejection
- doesn't always represent a perfectly normal distribution
    - dependent on N value
    - larger the sample the more normal the distribution looks
- the different shapes are important because our regions of rejection will look different dependent on the sample size

t-distribution & Degrees of Freedom (df)

- the distribution changes based on the sample size, which then means that the 5% of the regions of rejection and critical value change
- remember to be conservative about estimating variance and SD, we have been using `N - 1`
- the name of that is the **degrees of freedom** or df
    - number of scores in a sample that reflect the variability in the population
    - determines shape of sampling distribution when estimating standard deviation for the population

t-distribution & Degrees of Freedom (df)

- since the df is the sample size - 1, the larger the df, the closer to resembling a normal distribution our data becomes
  - df of 120+ is the same as a z-distribution

Using the t-table

- the t-table is different from the z-table
- has df, \alpha = .05 and \alpha = .01
    - this is dependent on our sample size - 1, and what our alpha is `a priori`

t-table

- we need to figure out our `t critical value`
- we need our sample size, and a decision on what alpha we want to use (.05 or .01)
- since not all df are listed, if your df is between two values, a statistically significant finding is a t-value larger than the larger df and smaller than the smaller df

Examples

- sample size = 200
  - alpha = .05
- sample size = 90
  - alpha = .05
- sample size = 37
  - alpha = .01

t-test Interpretation

- If a statistically significant finding is found
  - your sample is significantly different from the population in whatever the outcome was

One-tailed one-sample t-test

- if you know if your sample will do better or worse than the population, you'd use a one-tailed test
- Example: you know that your sample will get higher grades than the population

Confidence Intervals

- **point estimation** is a way to estimate a point where you think the population's outcome value will be
  - this is why we can't say we're certain mu is a specific number and have to say *around* that number
- **interval estimation** is when we state that mu will fall within a range of values
  - margin of error, such as getting an exam and stating that the average test score was 84 plus or minus 3 points
  - due to sampling error
- **confidence intervals** are a range of values which we are certain our value falls within
  - when we say *around* a value, we are saying that we got one value but we are certain it is within a range of values
  - around 84 points on an exam, but we are certain the correct value is between 80 and 87

Confidence Intervals

- We're choosing a range of values that are not significantly different from our sample mean
- we compute confidence intervals after we have a statistically significant finding
- It is often stated as:
    - We got a statistically significant finding where our sample scored ***points compared to the population's score*** ; t(df) = t-value, p-value
    - Example: $t(31) = 4.7$, $p = .037$

```r
coffee <- read_csv('https://raw.githubusercontent.com/rfordatascience/t
  mutate(species = as.factor(species),
         process = recode(processing_method, "Washed / Wet" = "washed",
                          "Semi-washed / Semi-pulped" = "not_washed",
                          "Pulped natural / honey" = "not_washed",
                          "Other" = "not_washed",
                          "Natural / Dry" = "not_washed",
                          "NA" = NA_character_),
         process = as.factor(process),
         species = as.factor(species),
         country_of_origin = as.factor(country_of_origin),
         variety = as.factor(variety)) %>%
  drop_na(process, color)
```

```
##
## -- Column specification ----------------------------------------------------
## cols(
##   .default = col_character(),
##   total_cup_points = col_double(),
##   number_of_bags = col_double(),
##   aroma = col_double(),
##   flavor = col_double(),
##   aftertaste = col_double(),
##   acidity = col_double(),
##   body = col_double(),
##   balance = col_double(),
##   uniformity = col_double(),
```

```r
psych::describe(coffee$total_cup_points, na.rm = TRUE)
```

```
##    vars    n  mean   sd median trimmed  mad   min   max range  skew kurtosi
## X1    1 1071 82.03 2.67  82.42    82.3 1.85 59.83 90.58 30.75 -2.11    10.5
##      se
## X1 0.08
```

```r
# mean is 82.03
# SE is .08
# sample size is 1071
```

```r
t.test(coffee$total_cup_points, mu = 85) #conf int only works for two ta
```

```
##
##      One Sample t-test
##
## data:  coffee$total_cup_points
## t = -36.39, df = 1070, p-value < 0.00000000000000022
## alternative hypothesis: true mean is not equal to 85
## 95 percent confidence interval:
##  81.87385 82.19373
## sample estimates:
## mean of x
##  82.03379
```

```r
t.test(coffee$total_cup_points, mu = 85, alternative = "less")
```

```
##
##      One Sample t-test
##
## data:  coffee$total_cup_points
## t = -36.39, df = 1070, p-value < 0.00000000000000022
## alternative hypothesis: true mean is less than 85
## 95 percent confidence interval:
##       -Inf 82.16798
## sample estimates:
## mean of x
##  82.03379
```

```r
t.test(coffee$total_cup_points, mu = 85, alternative = "greater")
```

```
##
##      One Sample t-test
##
## data:  coffee$total_cup_points
## t = -36.39, df = 1070, p-value = 1
## alternative hypothesis: true mean is greater than 85
## 95 percent confidence interval:
##  81.8996     Inf
## sample estimates:
## mean of x
##  82.03379
```

Confidence Interval Calculations

$$(s_x)(-t_{crit}) + \overline{X} \;\leq\; \mu \;\leq\; (s_x)(t_{crit}) + \overline{X}$$

```
# t critical value is 1.96 since we have such a large sample and df

# mu = 85
# sample mean = 82.03
# SE = .08
# df = 1070

# lower
.08*-1.96 + 82.03
```

```
## [1] 81.8732
```

```
# 81.8732

# higher
.08*1.96 + 82.03
```

```
## [1] 82.1868
```

```r
t.test(coffee$total_cup_points, mu = 85)
```

```
##
##      One Sample t-test
##
## data:  coffee$total_cup_points
## t = -36.39, df = 1070, p-value < 0.0000000000000022
## alternative hypothesis: true mean is not equal to 85
## 95 percent confidence interval:
##  81.87385 82.19373
## sample estimates:
## mean of x
##  82.03379
```

$t(1070) = -36.39$, $p < .05$, 95% CI [81.87, 82.19]

Our one-sample t-test comparing a sample of coffee ratings ($M = 82.03$, $SD = 2.67$) to the population of coffee ratings ($M = 85$) showed evidence of a statistically significant difference. Specifically, the sample's average coffee rating was significantly lower than the population's average coffee rating; $t(1070) = -36.39$, $p < .05$, 95% CI [81.87, 82.19]. We are 95% certain that the actual sample mean is between 81.87 and 82.19.

Independent-samples t-test

Between & Within Designs

- Experiments can be broken down into two different types of designs
- **Between-subject/group** design is when you are interested in comparing two (for now) or more groups on an outcome variable
- **Within-subject/group** design is when you have the same participants but you test them twice (either with two different variables or two different time points)

Two tests we are talking about

- **independent samples t-test** is when there are two groups of participants are separated into two different conditions to compare based on that condition
    - comparing the physical activity levels (DV) of sexes (Condition 1 = Male, Condition 2 = Female)
    - parametric test
- **paired-samples t-test** is when there are two experimental conditions that the same participants take part in
    - interested in two variables in the same sample of participants
    - can be the same variable and two different time points
    - bmi levels before an experiment and after the experiment for all participants
    - parametric test

Independent Samples t-test

- JP note: probably the most often used t-test
- because it is a parametric test, it has assumptions
- Assumptions are
  - DV is normally distributed interval/ratio scores
  - populations have homogeneous variance
  - not a true assumption but something important to note is that your groups should be equal in n (condition) size

Homogeneity of Variance

- **homogeneity of variance** is when the variances of the populations represented in a study have "equal" variances
- in order to test that the variances are equal, we can look at it through visuals
    - however, a better option is to use the Levene's test

Independent samples t-test

- hypotheses are now focused on the differences between the two groups/conditions

$$H0 : \mu_1 - \mu_2 = 0$$

H0: There will be no difference in DV scores between group 1 and group 2.

- both samples/groups represent the population

$$H1 : \mu_1 - \mu_2 \neq 0$$

H1: There will be differences in DV scores between group 1 and group 2.

- the groups represent a different population or don't represent the current population

t-distribution for independent samples t-test

- we are interested in the difference between our group/sample means
- we have two samples from one raw score population
- **sampling distribution of differences between means** show all differences between two means that occur when random samples are drawn from a population of scores
- the mean of the sampling distribution is zero because both sample means will equal the population mean of the raw score population

Independent samples t-test

- determines the probability of obtaining our difference between our means when H0 is true
- Term changes
  - N is now the full sample size
  - n is the size of each group/sample/condition
  - so for each group/sample/condition, we have an n

Performing the indepdendent samples t-test

$$s_x^2 = \frac{\Sigma X^2 - \frac{(\Sigma X)^2}{N}}{N-1}$$

```
male_scores = c(4, 6, 2, 3, 5, 1, 2, 4, 3, 5)
female_scores = c(4, 6, 9, 6, 5, 8, 2, 5, 3, 7)

male_scores
```

```
##  [1] 4 6 2 3 5 1 2 4 3 5
```

```
female_scores
```

```
##  [1] 4 6 9 6 5 8 2 5 3 7
```

First we'll calculate the variance

```r
# male sum
4+6+2+3+5+1+2+4+3+5
```

```
## [1] 35
```

```r
# sum is 35
```

$$s^2_{x_1} = \frac{\Sigma X^2 - \frac{(35)^2}{N}}{N-1}$$

```r
# female sum
4+6+9+6+5+8+2+5+3+7
```

```
## [1] 55
```

```r
# sum 55
```

$$s_{x_2}^2 = \frac{\Sigma X^2 - \frac{(55)^2}{N}}{N-1}$$

```r
35/10
```

```
## [1] 3.5
```

```r
# male mean 3.5
```

```r
55/10
```

```
## [1] 5.5
```

```r
# female mean 5.5
```

```r
# male sum of squared Xs
4^2+6^2+2^2+3^2+5^2+1^2+2^2+4^2+3^2+5^2
```

```
## [1] 145
```

```r
# 145
```

$$s_{x_1}^2 = \frac{145 - \frac{(35)^2}{10}}{10 - 1}$$

```r
# female sum of squared Xs
4^2+6^2+9^2+6^2+5^2+8^2+2^2+5^2+3^2+7^2
```

```
## [1] 345
```

```r
# female 345
```

$$s_{x_2}^2 = \frac{345 - \frac{(55)^2}{10}}{10-1}$$

```r
# male sum of X squared and divided by N
35^2
```

```
## [1] 1225
```

```r
1225/10
```

```
## [1] 122.5
```

```r
# 122.5
```

$$s_{x_1}^2 = \frac{145 - \frac{1225}{10}}{10 - 1}$$

```r
# female sum of X squared and divided by N
55^2
```

```
## [1] 3025
```

```r
3025/10
```

```
## [1] 302.5
```

```r
# 302.5
```

$$s_{x_2}^2 = \frac{345 - \frac{302.5}{10}}{10 - 1}$$

```r
# male variance calculations
(145 - 122.5)/(10-1)
```

```
## [1] 2.5
```

```r
# variance is 2.5
```

$$s^2_{x_1} = \frac{145 - 122.5}{10 - 1}$$

```r
# female variance calculations
(345 - 302.5)/(10 - 1)
```

```
## [1] 4.722222
```

```r
# variance is 4.72
```

$$s_{x_2}^2 = \frac{345 - 302.5}{10 - 1}$$

```r
sd(male_scores)^2
```

```
## [1] 2.5
```

```r
sd(female_scores)^2
```

```
## [1] 4.722222
```

New Terms

- **pooled variance** is the weighted average variance of the groups'/samples' variances in a independent samples t-test
- **standard error of the difference** is the estimated standard deviation of the sampling distribution of differences between the means

Now we can calculate the pooled variance n1 = 10 n2 = 10 variance of group 1 = 2.5 variance of group 2 = 4.72

$$S^2_{pool} = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{(n_1 - 1) + (n_2 - 1)}$$

```r
# start with the numerator
(10 - 1)*2.5 + (10 - 1)*4.72
```

```
## [1] 64.98
```

```r
# numerator is 64.98

# denominator
(10 - 1) + (10 - 1)
```

```
## [1] 18
```

```r
# denominator is 18
```

$$S^2_{pool} = \frac{(10-1)2.5 + (10-1)4.72}{(10-1)+(10-1)}$$

```r
9*2.5 + 9*4.72
```

```
## [1] 64.98
```

```r
# 64.98

9+9
```

```
## [1] 18
```

```r
# 18
```

$$S^2_{pool} = \frac{(9)2.5 + (9)4.72}{9+9}$$

```r
64.98/18
```

```
## [1] 3.61
```

```r
# pooled variance is 3.61
```

$$S^2_{pool} = \frac{64.98}{18}$$

$$S^2_{pool} = 3.61$$

Let's calculate for the standard error of the difference

$$S_{\overline{X_1} - \overline{X_2}} = \sqrt{(S^2_{pool})(\frac{1}{n_1} + \frac{1}{n_2})}$$

```r
1/10
```

```
## [1] 0.1
```

```r
3.61*(.1 + .1)
```

```
## [1] 0.722
```

```r
sqrt(.72)
```

```
## [1] 0.8485281
```

$$S_{\overline{X_1}-\overline{X_2}} = \sqrt{(3.61)(\frac{1}{10} + \frac{1}{10})}$$

$$S_{\overline{X_1}-\overline{X_2}} = \sqrt{(3.61)(.1 + .1)}$$

```r
3.61*(.1+.1)
```

```
## [1] 0.722
```

```r
# se is .72
```

$$S_{\overline{X_1}-\overline{X_2}} = \sqrt{.72}$$

```r
sqrt(.72)
```

```
## [1] 0.8485281
```

```r
# se of the difference is .85
```

$$S_{\overline{X_1}-\overline{X_2}} = .85$$

Now we can calculate the independent samples t-test obtained value

$$t_{obt} = \frac{(\overline{X_1} - \overline{X_2}) - (\mu_1 - \mu_2)}{S_{\overline{X_1} - \overline{X_2}}}$$

Note the population mean 1 minus the population mean 2 is what is specified in the null hypothesis, so it will be zero

```
((3.5 - 5.5) - 0)/.85
```

```
## [1] -2.352941
```

```
# t obtained value is -2.35
```

$$t_{obt} = \frac{(3.5 - 5.5) - 0}{.85}$$

Let's now calculate the degrees of freedom

$$df = (n_1 - 1) + (n_2 - 1)$$

```r
(10 - 1) + (10 - 1)
```

```
## [1] 18
```

```r
# df is 18

# t critical is +-2.101
```

$$df = (10 - 1) + (10 - 1)$$

$$df = 18$$

So we get a value of -2.35 and the t-critical value is -2.101

Is there a statistically significant difference between the two groups?

-2.35 > -2.101

Now let's get confidence intervals

$$Lower\ Bound : (\overline{X}_1 - \overline{X}_2) - t_{\alpha/2}\ *\ \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}$$

$$Upper\ Bound : (\overline{X}_1 - \overline{X}_2) + t_{\alpha/2}\ *\ \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}$$

```r
# group 1 mean = 3.5
# group 2 mean = 5.5
# t critical value is 2.101
# n1 = 10
# n2 = 10
# variance of group 1 = 2.5
# variance of group 2 = 4.72

# lower
(3.5 - 5.5) - 2.101 * sqrt((2.5/10) + (4.72/10))
```

```
## [1] -3.785232
```

```r
# -3.79

# upper
(3.5 - 5.5) + 2.101 * sqrt((2.5/10) + (4.72/10))
```

```
## [1] -0.214768
```

```r
# -.21
```

Effect Sizes

- Reminder: r effect sizes are .1 = small, .3 = medium, .5 = large
- Reminder: cohen's d effect sizes are .2 = small, .5 = medium, .8 = large
- these are both measures of the strength of a relationship
    - better than simply using p value alone
- cohen's d can never be negative so the value you get is the absolute value (.e.g., its always positive)
- if unequal sample sizes in groups/conditions then you'll use `Hedges' g`
    - same formula and will be roughly the same once sample sizes get larger than 20 (N = 20)

$$d = \frac{(\overline{X_1} - \overline{X_2})}{\sqrt{S^2_{pool}}}$$

Small sample sizes use the following formula (samples under 50)

$$d = \frac{(\overline{X_1} - \overline{X_2})}{\sqrt{S^2_{pool}}} * (\frac{N-3}{N-2.25}) * \sqrt{\frac{N-2}{N}}$$

```r
(3.5 - 5.5)/sqrt(3.61)
```

```
## [1] -1.052632
```

```r
# each step below
3.5 - 5.5
```

```
## [1] -2
```

```r
-2/sqrt(3.61)
```

```
## [1] -1.052632
```

```r
# cohen's d is .92 or the number of standard deviations between the mea
```

$$d = \frac{(3.5 - 5.5)}{\sqrt{3.61}}$$

```r
(3.5 - 5.5)/sqrt(3.61)
```

```
## [1] -1.052632
```

```r
(10-3/10 -2.25)
```

```
## [1] 7.45
```

```r
sqrt((10-2)/10)
```

```
## [1] 0.8944272
```

```r
-1.05*7.45*.89
```

```
## [1] -6.962025
```

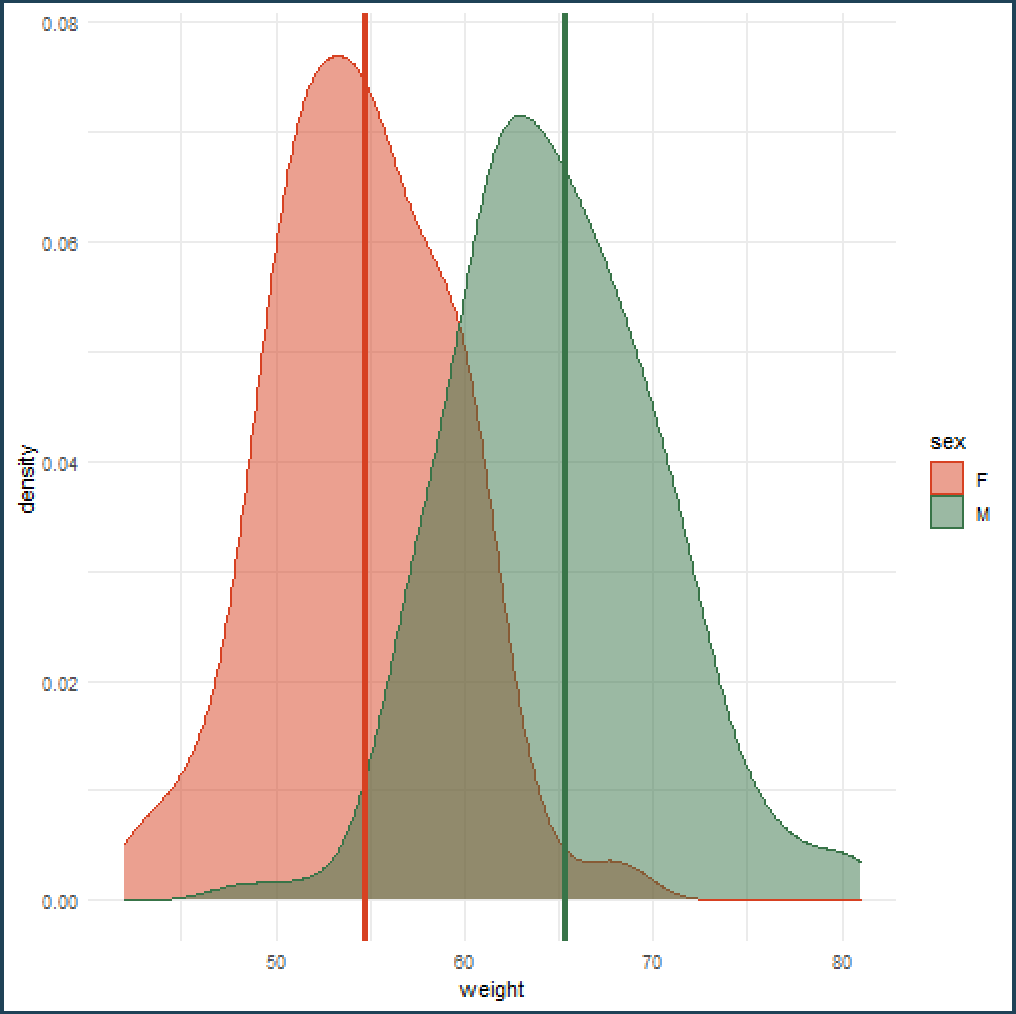
$$d = \frac{(3.5 - 5.5)}{\sqrt{3.61}} * \left(\frac{10 - 3}{10 - 2.25}\right) * \sqrt{\frac{10 - 2}{10}}$$

Steps for independent samples t-test

1. Get the means of both groups/samples
2. Get the variances of both groups/samples
3. Get the group/sample sizes (n)
4. Get the pooled variance by getting the groups'/samples' variances averaged
5. Get the standard error of the differences
6. Calculate the t-obtained value
7. Get the degrees of freedom
8. Calculate the confidence intervals
9. Get the effect size

Independent samples t-test Example

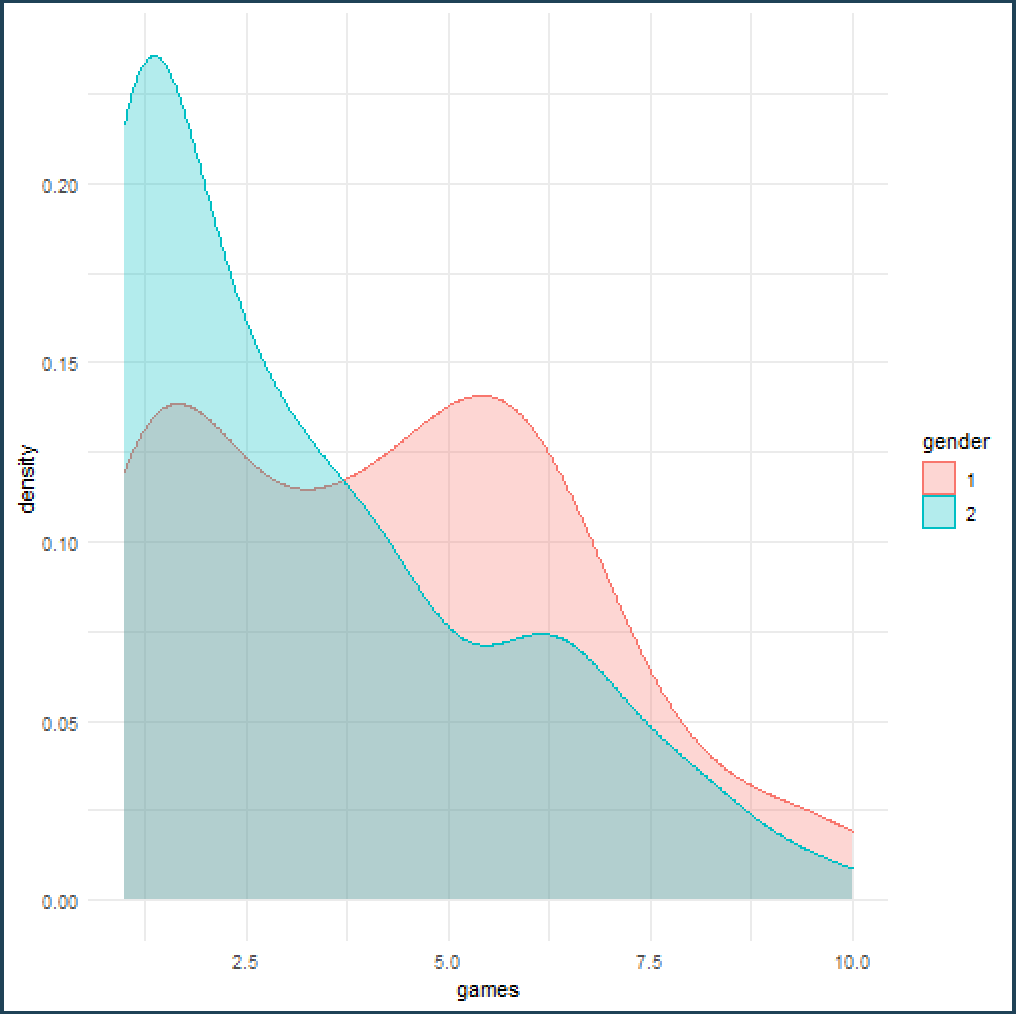
```
##
##  Descriptive statistics by group
## group: F
##    vars   n  mean   sd median trimmed  mad min max range skew kurtosis   se
## X1    1 200 54.52 4.83     54   54.52 4.45  42  69    27 0.07     0.02 0.34
## ------------------------------------------------------------
## group: M
##    vars   n  mean   sd median trimmed  mad min max range skew kurtosis   se
## X1    1 200 64.87 5.44   64.5   64.71 5.19  48  81    33 0.25     0.28 0.38
```


0.08
0.06
0.04
0.02
0.00
density
50
60
70
80
weight
sex
F
M

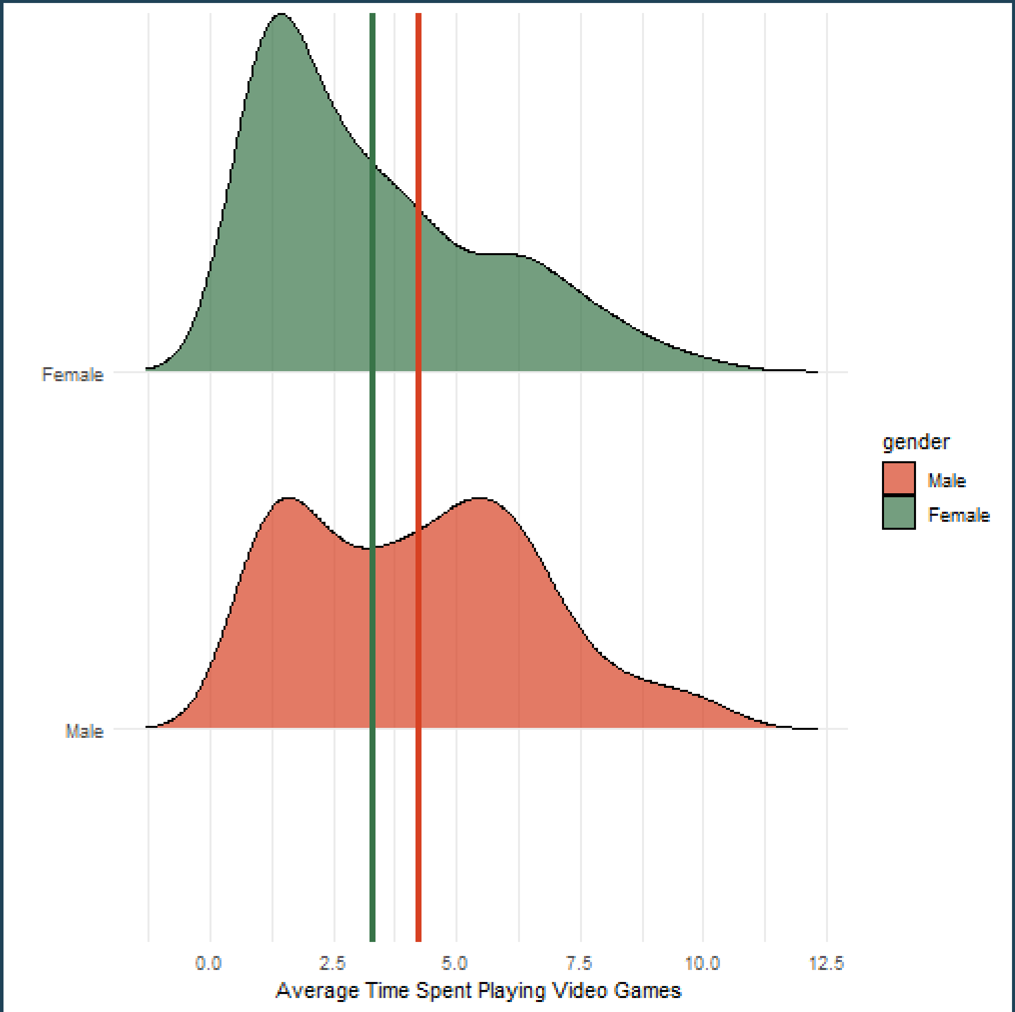
```
## Levene's Test for Homogeneity of Variance (center = median)
##        Df F value Pr(>F)
## group   1   1.972  0.161
##       398

##
##     Two Sample t-test
##
## data:  weight by sex
## t = -20.116, df = 398, p-value < 0.00000000000000022
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -11.361529  -9.338471
## sample estimates:
## mean in group F mean in group M
##           54.52           64.87
```

Real life independent samples t-test

```
## # A tibble: 2 x 2
##   gender games
##   <fct>  <dbl>
## 1 1       4.23
## 2 2       3.30
```


density
0.00
0.05
0.10
0.15
0.20
2.5
5.0
7.5
10.0
games
gender
1
2

```
## Picking joint bandwidth of 0.763
```



```
## Levene's Test for Homogeneity of Variance (center = median)
##        Df F value Pr(>F)
## group   1  1.3319 0.2492
##       370

##
##     Two Sample t-test
##
## data:  games by gender
## t = 3.5171, df = 370, p-value = 0.0004906
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  0.4066502 1.4379978
## sample estimates:
## mean in group 1 mean in group 2
##        4.227011        3.304688
```

```
##    vars   n mean   sd median trimmed  mad min max range skew kurtosis   se
## X1    1 372 3.59 2.38      3    3.34 2.97   1  10     9 0.69    -0.52 0.12
```

One-tail independent-samples t-test

- used when we are confident that the direction of the relationship
- if you think one group will have a higher score/value then you would make that statement in your alternative hypothesis

Steps for one-tailed test

- Variable(Groups) = Intervention(Got Intervention/Control)
- Outcome = Fruit & Vegetable Intake

1. decide which sample/group score you think will be larger
    - I think that the intervention group will eat more fruits and vegetables
2. decide which condition/group to subtract from the other
    - `intervention group - control group`
3. decide whether the difference will be positive or negative
    - a positive value should be returned from the previous equation

1. create hypotheses
    - The intervention group will eat the same or less fruits and vegetables than the control group
    - H0: mu(intervention group) - mu(control group) ≤ 0
    - The intervention group will eat more fruits and vegetables than the control group
    - H1: mu(intervention group) - mu(control group) > 0

1. locate regions of rejection
    - since we are interested in a positive difference in our hypothesis, we'll only be looking at the upper tail of the distribution
2. conduct your independent samples t-test
    - make sure your groups/samples are subtracted the same way as your hypotheses
    - `intervention group - control group`

Paired Samples t-test

- within-subjects/groups design
- also called related-samples t-test
  - each participant gets measured in each condition
- an example would be an intervention for weight loss where everyone's weight is measured before (first time point to measure) the intervention and after (second time point to measure)
- **matched-samples design** is when a participant in one condition is matched with a participant in the other condition
- **repeated-measures design** is when each participant gets measured twice
  - can be more for more rigorous statistics, but not the ones we'll learn about in this class

Why we use paired-samples t-tests

- pretest/posttest design
  - before everyone gets the intervention/experiment (pretest or first time point) and then after the intervention/experiment (posttest or second time point)

- we subtract every person's before/after score from the before/score score to get a `difference` score
- you can subtract whatever score you want from the other (before - after or after - before)
- **mean differences** the mean of the differences between the paired scores
  - repesented as dbar
- add all the values of differences for before/after to get a sum and then divide by the number of participants (each person has a before and after score or the number of difference scores)

$$Sample\ mean\ difference = \overline{D}$$

$$Population\ mean\ difference = \mu_D$$

- now because we have a sample mean from a sample, we can now perform a one-sample t-test

Hypotheses

- our null hypothesis is that there will be no change in scores on both occasions so there should be a difference of zero

$$H_0 : \mu_D = 0$$

- our alternative hypothesis is that either our before or after scores should be higher
  - our hypothesis supports either a positive or negative change

$$H_1 : \mu_D \neq 0$$

- similar to the independent samples t-test, our sampling distribution is of mean differences

Steps to paired-samples t-test

Calculate the estimated variance of the difference scores

$$S_D^2 = \frac{\Sigma D^2 - \frac{(\Sigma D)^2}{N}}{N - 1}$$

Note the notation changes from X to D because we are working with differences in scores and not means

Calculate the standard error of the mean differences

$$S_{\overline{D}} = \sqrt{\frac{S_D^2}{N}}$$

Find the obtained t-value

$$t_{obt} = \frac{\overline{D} - \mu_D}{S_{\overline{D}}}$$

The population difference will be zero unless you are testing for a nonzero difference

Calculate the degrees of freedom

$$df = N - 1$$

Calculate effect size

$$d = \frac{\overline{D}}{\sqrt{S_D^2}}$$

OR

$$r_{pb}^2 = \sqrt{\frac{(t_{obt})^2}{(t_{obt})^2 + df}}$$

Discuss Significance

- Is the difference between time point 1 and time point 2 statistically significant?

Example

```
set.seed(100521)
t1 = rnorm(10, mean = 10, sd = 1.5)

t2 = rnorm(10, mean = 5.8, sd = 4)

df <- data.frame(participant = 1:10,
                 t1 = round(t1, 2),
                 t2 = round(t2, 2),
                 score_difference = c("_____________", "_____________",
                                    "_____________", "_____________", "
                 score_difference_squared = c("_____________", "________
                                    "_____________", "_____________", "
```

```r
df
```

```
##    participant    t1    t2 score_difference score_difference_squared
## 1            1  8.32 12.75    _____________            _____________
## 2            2 13.96 10.54    _____________            _____________
## 3            3  8.17  7.51    _____________            _____________
## 4            4  8.40  4.32    _____________            _____________
## 5            5  8.90  9.95    _____________            _____________
## 6            6  7.58  3.64    _____________            _____________
## 7            7  9.67  6.45    _____________            _____________
## 8            8 11.51  3.40    _____________            _____________
## 9            9 10.13  3.43    _____________            _____________
## 10          10 11.49 13.77    _____________            _____________
```

```r
df$t2
```

```
##  [1] 12.75 10.54  7.51  4.32  9.95  3.64  6.45  3.40  3.43 13.77
```

```r
df$t1
```

```
##  [1]  8.32 13.96  8.17  8.40  8.90  7.58  9.67 11.51 10.13 11.49
```

```r
12.75 - 8.32
```

```
## [1] 4.43
```

```r
10.54 - 13.96
```

```
## [1] -3.42
```

```r
7.51 - 8.17
```

```
## [1] -0.66
```

```r
4.32 - 8.40
```

```
## [1] -4.08
```

```r
9.95 - 8.90
```

```
## [1] 1.05
```

```r
3.64 - 7.58
```

```
## [1] -3.94
```

```r
6.45 - 9.67
```

```
## [1] -3.22
```

```r
3.40 - 11.51
```

```
## [1] -8.11
```

```r
3.43 - 10.13
```

```
## [1] -6.7
```

```r
13.77 - 11.49
```

```
## [1] 2.28
```

To get your dataframe to include the numbers you just included, you need to run this again. Then it should fill in the blanks with the numbers you provided.

```
df <- data.frame(participant = 1:10,
  t1 = round(t1, 2),
                 t2 = round(t2, 2),
                 score_difference = c(4.43, -3.42, -.66, -4.08, 1.05, -3
                 score_difference_squared = c("_____________", "_______
                                    "_____________", "_____________", "
```

```r
df$score_difference
```

```
##  [1]  4.43 -3.42 -0.66 -4.08  1.05 -3.94 -3.22 -8.11 -6.70  2.28
```

```r
(4.43)^2
```

```
## [1] 19.6249
```

```r
(-3.42)^2
```

```
## [1] 11.6964
```

```r
(-.66)^2
```

```
## [1] 0.4356
```

```r
(-4.08)^2
```

```
## [1] 16.6464
```

```r
(1.05)^2
```

```
## [1] 1.1025
```

```r
(-3.94)^2
```

```
## [1] 15.5236
```

```r
(-3.22)^2
```

```
## [1] 10.3684
```

```r
(-8.11)^2
```

```
## [1] 65.7721
```

```r
(-6.7)^2
```

```
## [1] 44.89
```

```r
(2.28)^2
```

```
## [1] 5.1984
```

```r
df <- data.frame(participant = 1:10,
  t1 = round(t1, 2),
                 t2 = round(t2, 2),
                 score_difference = c(4.43, -3.42, -.66, -4.08, 1.05, -3
                 score_difference_squared = c(19.62, 11.70, .44, 16.64,

df
```

```
##    participant    t1    t2 score_difference score_difference_squared
## 1            1  8.32 12.75             4.43                    19.62
## 2            2 13.96 10.54            -3.42                    11.70
## 3            3  8.17  7.51            -0.66                     0.44
## 4            4  8.40  4.32            -4.08                    16.64
## 5            5  8.90  9.95             1.05                     1.10
## 6            6  7.58  3.64            -3.94                    15.52
## 7            7  9.67  6.45            -3.22                    10.37
## 8            8 11.51  3.40            -8.11                    65.77
## 9            9 10.13  3.43            -6.70                    44.89
## 10          10 11.49 13.77             2.28                     5.20
```

```r
df$score_difference
```

```
##  [1]  4.43 -3.42 -0.66 -4.08  1.05 -3.94 -3.22 -8.11 -6.70  2.28
```

```r
4.43 + (-3.42) + (-0.66) + (-4.08) + 1.05 + (-3.94) + (-3.22) + (-8.11)
```

```
## [1] -22.37
```

```r
# sum difference is -22.37

-22.37/10
```

```
## [1] -2.237
```

```r
# mean difference is -2.24
```

Calculate the estimated variance of the difference scores

$$s_D^2 = \frac{\Sigma D^2 - \frac{(\Sigma D)^2}{N}}{N-1}$$

$$s_D^2 = \frac{\Sigma D^2 - \frac{(-22.37)^2}{10}}{10-1}$$

```
19.62+ 11.70+ .44+ 16.64+ 1.10+ 15.52+ 10.37+ 65.77+ 44.89+ 5.20
```

```
## [1] 191.25
```

$$s_D^2 = \frac{191.25 - \frac{(-22.37)^2}{10}}{10 - 1}$$

```r
(-22.37)^2
```

```
## [1] 500.4169
```

$$s_D^2 = \frac{191.25 - \frac{500.42}{10}}{10 - 1}$$

```r
500.42/10
```

```
## [1] 50.042
```

$$s_D^2 = \frac{191.25 - 50.04}{9}$$

```r
191.25 - 50.04
```

```
## [1] 141.21
```

$$s_D^2 = \frac{141.21}{9}$$

```r
141.21/9
```

```
## [1] 15.69
```

$$s_D^2 = 15.69$$

Calculate the standard error of the mean differences

$$S_{\overline{D}} = \sqrt{\frac{S_D^2}{N}}$$

$$S_{\overline{D}} = \sqrt{\frac{15.69}{10}}$$

```r
15.69/10
```

```
## [1] 1.569
```

$$S_{\overline{D}} = \sqrt{1.57}$$

```r
sqrt(1.57)
```

```
## [1] 1.252996
```

$$S_{\overline{D}} = 1.25$$

Find the obtained t-value

$$t_{obt} = \frac{\overline{D} - \mu_D}{S_{\overline{D}}}$$

$$t_{obt} = \frac{-2.24 - 0}{1.25}$$

```r
-2.24 - 0
```

```
## [1] -2.24
```

$$t_{obt} = \frac{-2.24}{1.25}$$

```r
-2.24/1.25
```

```
## [1] -1.792
```

$$t_{obt} = -1.79$$

Calculate the degrees of freedom

$$df = N - 1$$

```r
10 - 1
```

```
## [1] 9
```

$$df = 9$$

Calculate effect size

$$d = \frac{\overline{D}}{\sqrt{S_D^2}}$$

```r
mean(df$score_difference)
```

```
## [1] -2.237
```

```r
sd(df$score_difference)^2
```

```
## [1] 15.69073
```

```r
-2.24
```

```
## [1] -2.24
```

```r
15.69
```

```
## [1] 15.69
```

$$d = \frac{-2.24}{\sqrt{15.69}}$$

```r
sqrt(15.69)
```

```
## [1] 3.96106
```

```r
sd(df$score_difference)
```

```
## [1] 3.961153
```

$$d = \frac{-2.24}{3.96}$$

```r
-2.24/3.96
```

```
## [1] -0.5656566
```

$$d = -.57$$

so really this means

$$d = .57$$

$$r_{pb}^2 = \sqrt{\frac{(t_{obt})^2}{(t_{obt})^2 + df}}$$

$$r_{pb}^2 = \sqrt{\frac{(-1.79)^2}{(-1.79)^2 + df}}$$

```r
(-1.79)^2
```

```
## [1] 3.2041
```

$$r_{pb}^2 = \sqrt{\frac{3.20}{3.20 + 9}}$$

```r
3.20 + 9
```

```
## [1] 12.2
```

$$r_{pb}^2 = \sqrt{\frac{3.20}{12.20}}$$

```r
3.2/12.2
```

```
## [1] 0.2622951
```

$$r^2_{pb} = \sqrt{.26}$$

```r
sqrt(.26)
```

```
## [1] 0.509902
```

$$r_{pb}^2 = .51$$

Discuss Significance

t-obtained value of -1.79

t-critical value of

One-tailed paired-samples t-test

- choose whether you think your `after` score would be lower/higher than the `before` score
- have your hypotheses reflect that
- if you think after scores should be higher than before (learning intervention) then the difference scores should be positive
    - because you subtracted before scores from after scores

$$H_a : \mu_D > 0$$

$$H_0 : \mu_D \leq 0$$

Reporting

- report in a similar style to all other t-test
  - t(df) = t value, p value
- you'll also report the means of the before and after scores
  - you won't report the difference between the two means

Effect Sizes

- **effect size** is the amount of influence that changing the conditions of the IV has on the DV
- **cohen's d** is a measure of effect size in two-sample studies that reflects the magnitude of difference
  - small = .2
  - medium = .5
  - large = .8
- larger effect size means stronger the strength of the association/relationship between the IV and DV

Effect Size using Proportion of Variance Accounted For

- used to determine how consistently scores change
- **proportion of variance accounted for** is the proportion of differences in DV scores assocaited with changing the conditions of the IV
  - effect size using **squared point-biserial correlation coefficient**, which indicates the proprotion of variance in DV scores that is accounted fro by IV variable in two-sample studies, are below
  - small = .09
  - moderate = .10-.25
  - large = .25