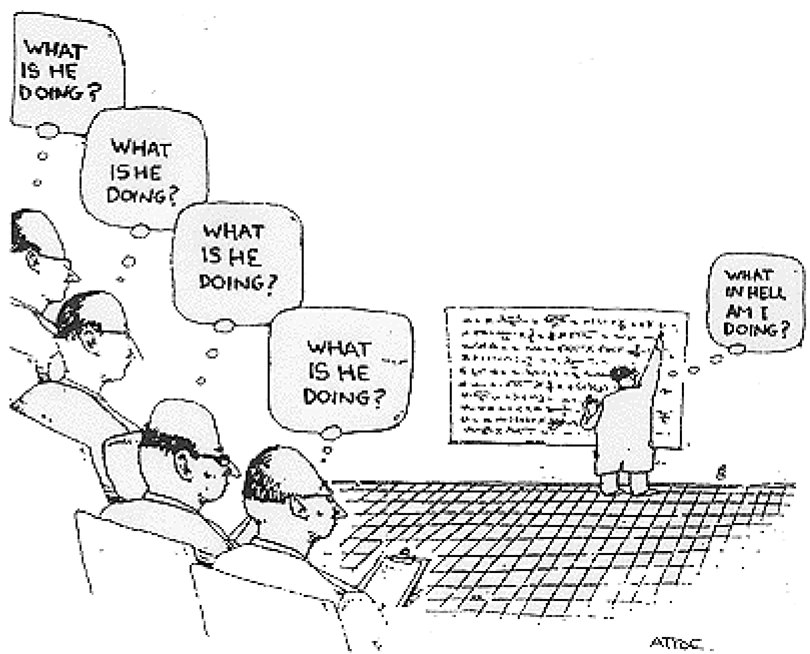
# Hypothesis Testing with *t* Tests

Arlo Clark-Foos

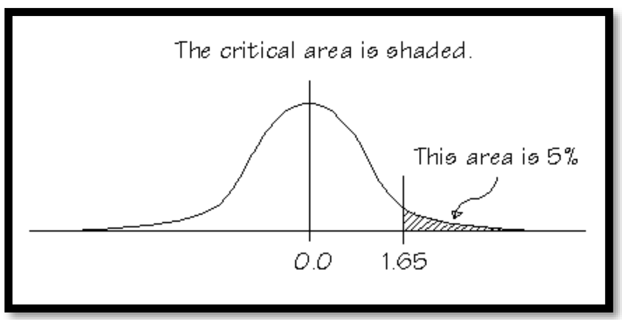


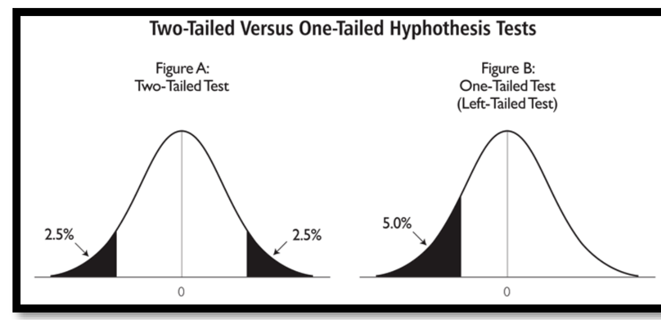
# What have we done so far?

- Hypothesis Testing & Inferential Statistics

alpha levels, cut-offs, p-value

One-tailed vs. Two-tailed tests




  - **Goal**: Estimate likelihood of obtaining sample mean given some known population parameters ($\mu$ *and* $\sigma$)
    - What if we didn't know all of the population parameters?

# The Story of Student's *t*

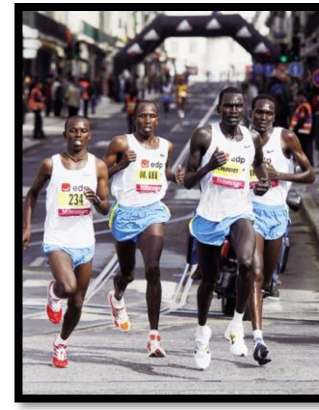


*"Guinness is the best beer available, it does not need advertising as its quality will sell it, and those who do not drink it are to be sympathized with rather than advertised to." --W.S. Gosset (aka "Student")*

- *The pros and cons of ~~beer~~ sampling…*

# Using Samples to Estimate Population Variability

- Acknowledge error
- Smaller samples, less spread



$$s = \sqrt{\frac{\Sigma(X - M)^2}{N \underbrace{-1}_{\text{New!}}}}$$

This correction will affect larger samples less so than it will affect smaller samples.

- *N = 65, N – 1 = 64 (change of 1.5%)*
- *N = 4, N – 1 = 3 (change of 25%)*

# What happened to $\sigma_M$?

- Still there, only used for *z* tests

- We have a new measure of standard deviation for a sample (as opposed to a population): *s*
  - *We need a new measure of standard error based on* *<u>sample</u>* *standard deviation:*

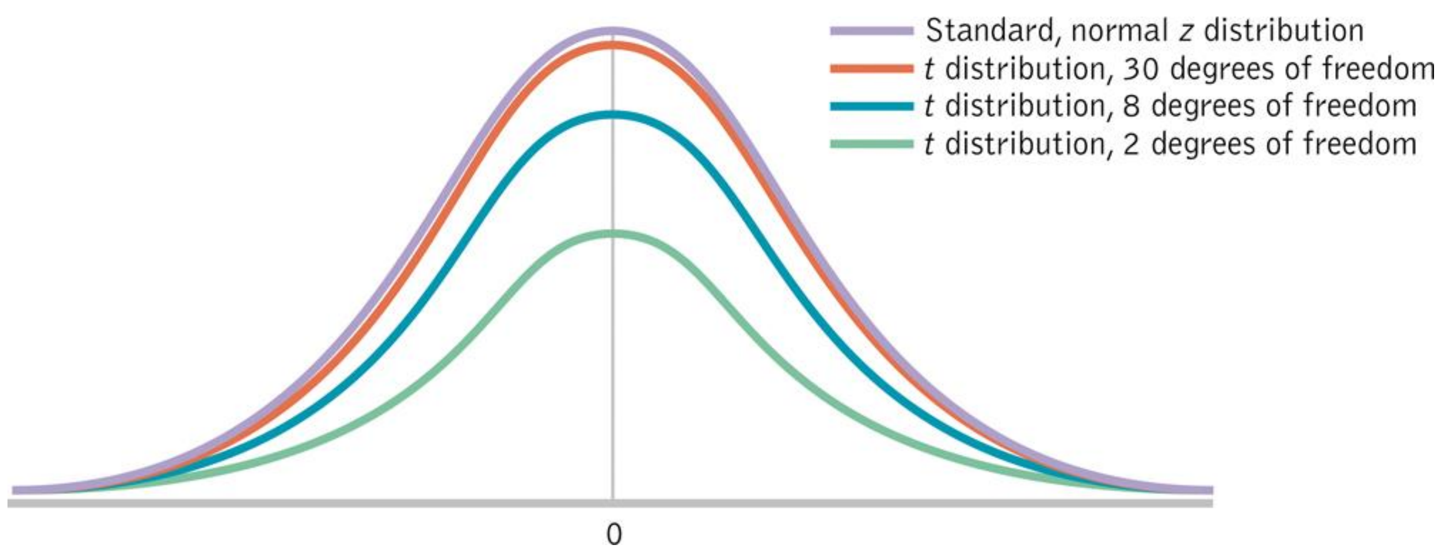
$$s_M = \frac{s}{\sqrt{N}}$$

  - *Wait, what happened to "N-1"?*
  - *We already did that when we calculated s, don't correct again!*

# Student's *t* Statistic

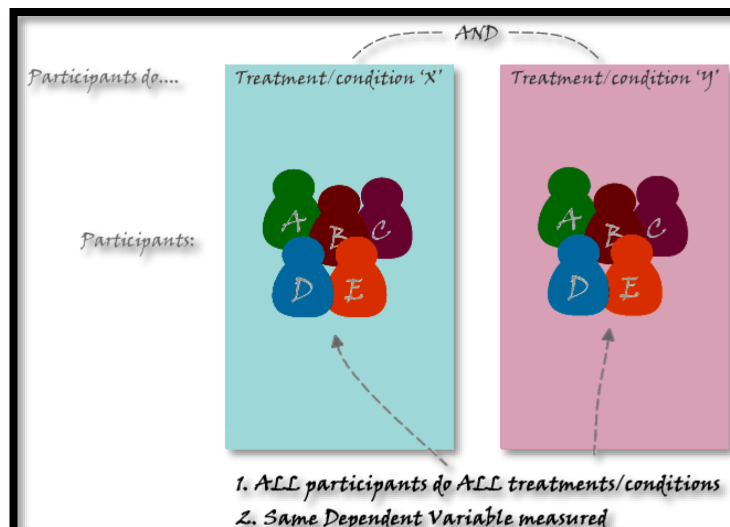
$$t = \frac{(M - \mu_M)}{s_M}$$

*Indicates the distance of a sample mean from a population mean in standard errors (like standard deviations)*



# The many flavors of $t_{ea}$


- Single sample *t*
  - *One sample, compared with <u>known</u> population mean*
  - *Goal: Is our sample different from population?*
- Independent Samples *t*
  - *Different (independent) samples of participants experience each level of IV*
  - *Are our samples from different populations?*
- Paired/Dependent Samples
  - *Same or related (dependent) samples of participants experience each level of IV*
  - *Are our samples from different populations?*



# Degrees of Freedom

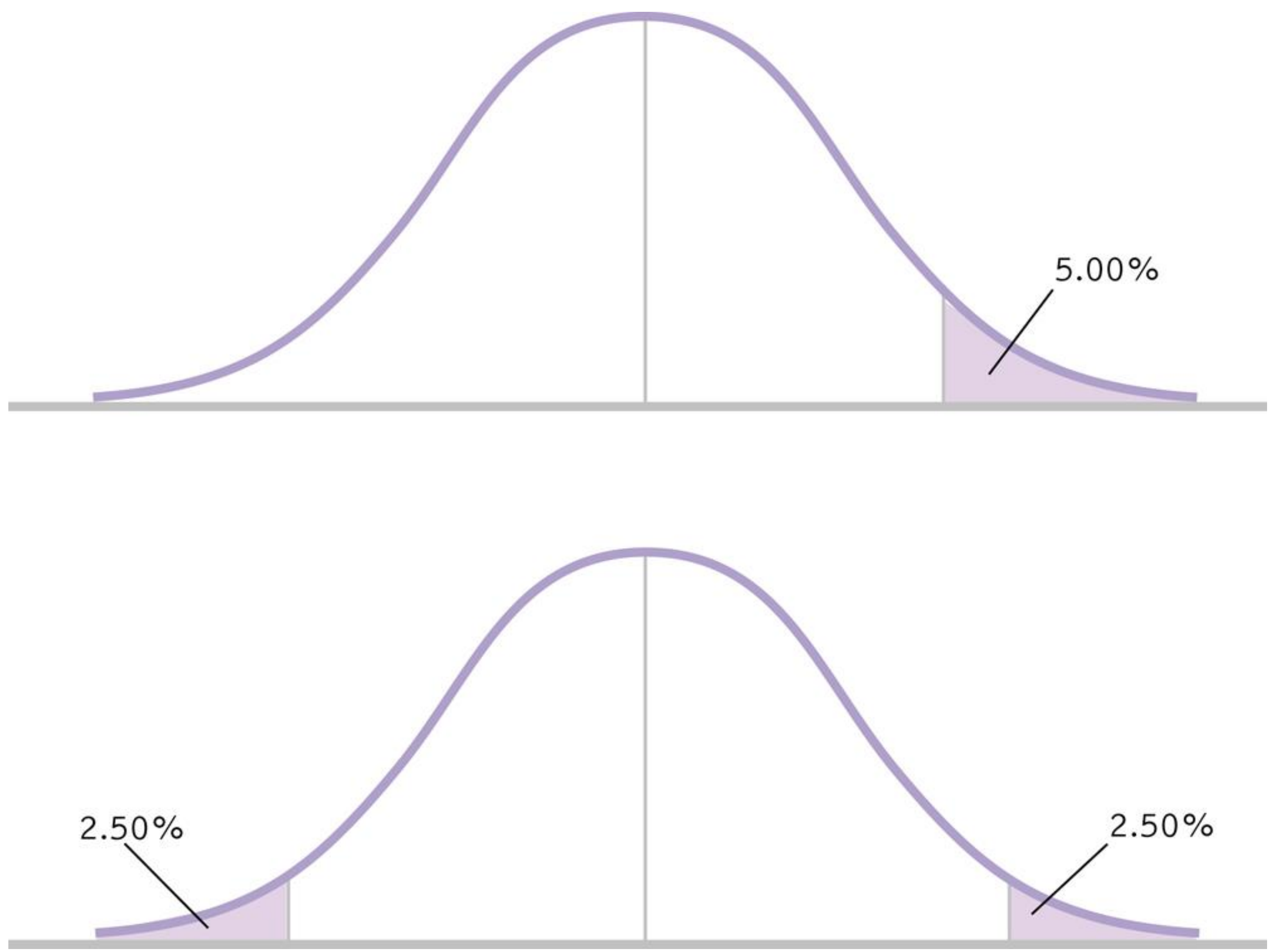
- *Necessary when making estimates…*
- *The number of scores that are free to vary when estimating a population parameter from a sample*
  - $df = N - 1$ (for a Single-Sample $t$ Test)

Example: I decide to ask 6 people how often they floss their teeth and record their average = 2 (times per week)

- Eventual goal: Estimate population parameters (population variability).
- How many scores are free to vary and can still produce an average of 2?

| | | |
|---|---|---|
| 3 | Free | 2 |
| 5 | Free | 1 |
| 1 | Free | 0 |
| 0 | Free | 0 |
| 2 | Free | 0 |
| **1** | **LOCKED** | **9** |
| **Average = 2** | | **Average = 2** |

# One Tailed vs. Two Tailed Tests



# Six Steps for Hypothesis Testing

1. Identify
2. State the hypotheses
3. Characteristics of the comparison distribution
4. Critical values
5. Calculate
6. Decide

*old hat*

# Single-Sample *t* Test: Attendance in Therapy Sessions

- Our Counseling center on campus is concerned that most students requiring therapy do not take advantage of their services. Right now students attend only 4.6 sessions in a given year! Administrators are considering having patients sign a contract stating they will attend at least 10 sessions in an academic year.
- Question: Does signing the contract actually increase participation/attendance?

- We had 5 patients sign the contract and we counted the number of times they attended therapy sessions

| Number of Attended Therapy Sessions |
|---|
| 6 |
| 6 |
| 12 |
| 7 |
| 8 |

# Single-Sample *t* Test: Attendance in Therapy Sessions

1. Identify
   - Populations:
     - Pop 1: All clients who sign contract
     - Pop 2: All clients who do not sign contract
   - Distribution:
     - One Sample mean: Distribution of means
   - Test & Assumptions: Population mean is known but not standard deviation → single-sample *t* test
     1. Data are interval
     2. Probably not random selection
     3. Sample size of 5 is less than 30, therefore distribution might not be normal

# Single-Sample *t* Test: Attendance in Therapy Sessions

2. State the null and research hypotheses

$H_0$: Clients who sign the contract will attend the same number of sessions as those who do not sign the contract.

$H_1$: Clients who sign the contract will attend a different number of sessions than those who do not sign the contract.

# Single-Sample *t* Test: Attendance in Therapy Sessions

3. Determine characteristics of comparison distribution (distribution of sample means)
   - Population: $\mu_M = \mu = 4.6$times
   - Sample: $M$ = _7.8_times, $s$ = _2.490_, $s_M$ = _1.114_

| # of Sessions ($X$) | $X$-$M$ | $(X$-$M)^2$ |
|---|---|---|
| 6 | -1.8 | 3.24 |
| 6 | -1.8 | 3.24 |
| 12 | -4.2 | 17.64 |
| 7 | -0.8 | 0.64 |
| 8 | 0.2 | 0.04 |
| $M_X = 7.8$ | | $SS_X = 24.8$ |

$$s = \sqrt{\frac{\Sigma(X-M)^2}{N-1}} = \sqrt{\frac{24.8}{5-1}} = 2.490$$

$$s_M = \frac{s}{\sqrt{N}} = \frac{2.490}{\sqrt{5}} = 1.114$$

# Single-Sample $t$ Test: Attendance in Therapy Sessions

$\mu_M = 4.6, s_M = 1.114, M = 7.8, N = 5, df = 4$

4. Determine critical value (cutoffs)
   - In Behavioral Sciences, we use $p = .05$ (5%)
   - Our hypothesis ("Clients who sign the contract will attend a different number of sessions than those who do not sign the contract.") is ***nondirectional*** so our hypothesis test is ***two-tailed***.

Significance level = $\alpha$

| Degrees of Freedom | .005 (1-tail) .01 (2-tails) | .01 (1-tail) .02 (2-tails) | .025 (1-tail) .05 (2-tails) | .05 (1-tail) .10 (2-tails) | .10 (1-tail) .20 (2-tails) | .25 (1-tail) .50 (2-tails) |
|---|---|---|---|---|---|---|
| 1 | 63.657 | 31.821 | 12.706 | 6.314 | 3.078 | 1.000 |
| 2 | 9.925 | 6.965 | 4.303 | 2.920 | 1.886 | .816 |
| 3 | 5.841 | 4.541 | 3.182 | 2.353 | 1.638 | .765 |
| 4 | 4.604 | 3.747 | $\boxed{2.776}$ | 2.132 | 1.533 | .741 |
| 5 | 4.032 | 3.365 | 2.571 | 2.015 | 1.476 | .727 |
| 6 | 3.707 | 3.143 | 2.447 | 1.943 | 1.440 | .718 |
| 7 | 3.500 | 2.998 | 2.365 | 1.895 | 1.415 | .711 |

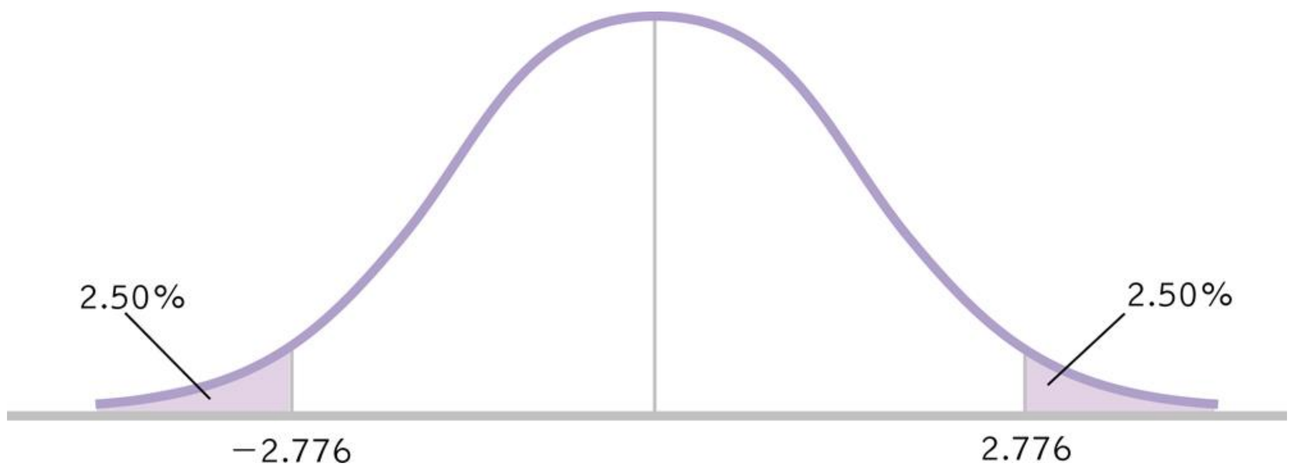
$df = 4 \rightarrow$

# Single-Sample $t$ Test: Attendance in Therapy Sessions

$\mu_M = 4.6, s_M = 1.114, M = 7.8, N = 5, df = 4$

4. Determine critical value (cutoffs)

$t_{crit} = \pm 2.76$



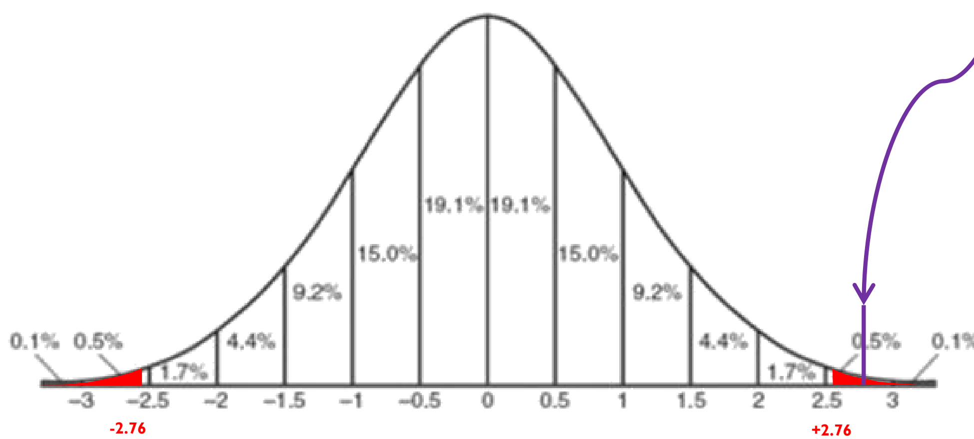
# Single-Sample $t$ Test: Attendance in Therapy Sessions

$\mu_M = 4.6, s_M = 1.114, M = 7.8, N = 5, df = 4$

5. Calculate the test statistic

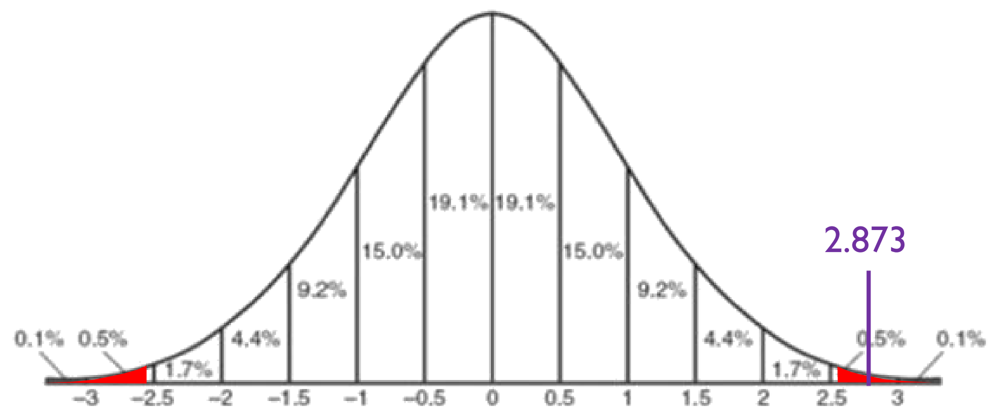
$$t = \frac{(M - \mu_M)}{s_M} = \frac{(7.8 - 4.6)}{1.114} = 2.873$$

6. Make a decision



# Single-Sample *t* Test: Attendance in Therapy Sessions

$\mu_M = 4.6, s_M = 1.114, M = 7.8, N = 5, df = 4$




6. Make a decision

   $t = 2.873 > t_{crit} = \pm 2.776$, *reject the null hypothesis*

   Clients who sign a contract will attend more sessions than those who do not sign a contract,

   $t(4) = 2.87, p < .05.$

# Reporting Results in APA Format

1. Write the symbol for the test statistic (e.g., $z$ or $t$)
2. Write the degrees of freedom in parentheses
3. Write an equal sign and then the value of the test statistic (2 decimal places)
4. Write a comma and then whether the $p$ value associated with the test statistic was less than or greater than the cutoff $p$ value of .05 (or report exact $p$ value).

$$t(4) = 2.87, p < .05$$

# Another example?
# A Tale of a Tail

The citizens of several Georgia towns are worried that a fiberglass insulation plant is polluting the local environment. A doctor at one hospital, *Our Sister of the Failing Mercy*, noted 3 babies born with a coccycx (tailbone) in the past year. Digging deeper she read a report in *Human Pathology* (Dao & Netsky, 1984) that between 1884 and 1984 only 23 babies were born with tails. She believes her hospital data to be unusual but she also writes to doctors at 8 other hospitals to determine how often they had seen this birth defect in the past year (these data are below).

| Hospital | # of Tails |
|---|---|
| North Shore | 2 |
| Town Center | 0 |
| St. Mary | 0 |
| University | 1 |
| Failing Mercy | 3 |
| South Central | 1 |
| Oakmont | 0 |
| Bellevue | 1 |
| East Valley | 0 |
| | $\sum = 8$ |
| | $N = 8$ |
| | $M = 1$ |

*"between 1884 and 1984 only 23 babies were born with tails…"*
*23 babies in 100 years ($\mu$ ) = .23*

# Rolling Kegs for Fun and Profit

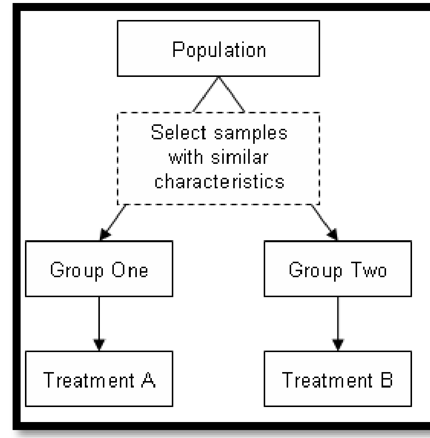
- Making the perfect handmade cask
  - Not too big (too much beer, lower profits)
  - Not too small (not enough beer, fewer clients)

Stella Cunliffe

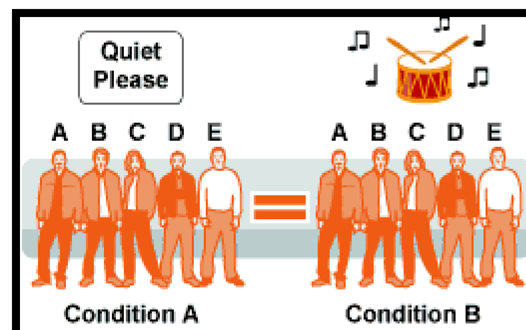
- Expectation: High rejection of poorly sized casks
- Data: Little-to-no rejections…why?
  - Visit workstation for employee making rejections.
  - Conditions must be equal in order to make a fair comparison.
    - *"how impossible it is to find human beings without biases, without prejudices, and without the delightful idiosyncrasies which make them so fascinating."* (Cunliffe, 1976, p.5)

# Studies with Two Samples

- Independent
  - What is it?
  - Pros
  - Cons




- Paired (Dependent)
  - What is it?
  - Pros
  - Cons




- Hypothetical Beer Tasting Experiment
  - What is the ideal design for this study?

Two Related (dependent) samples

# PAIRED SAMPLES *t* TEST

# Paired-Samples *t* Test

$$t = \frac{(M - \mu_M)}{s_M}$$

- *Used to compare 2 means for a within-groups design, a situation in which every participant is in both samples (paired/dependent)*

- New Terminology
    - Distribution of Mean Differences
    - Difference Scores: $X_1 - Y_1, X_2 - Y_2, \ldots$

- Let's walk through an example…

# Six Steps for Hypothesis Testing

1. Identify
2. State the hypotheses
3. Characteristics of the comparison distribution
4. Critical values
5. Calculate
6. Decide

*old hat*

# Paired Samples *t* Test:
# Does Studying in the Exam Room Help?

*I have a debate with a research assistant about context effects in studying for exams. She believes that she does far better when she studies for an exam in the same room as she later takes the exam. I told her that I could see it either hurting or helping students. We agreed to have a group of 5 participants complete two highly similar math exams. Half of these participants studied for and completed the exams in the same room while the other half were in different rooms, order counterbalanced*. Data for SAME and DIFFERENT rooms are below.*

| SAME (*X*) | DIFFERENT (*Y*) | Difference Score X-Y |
|---|---|---|
| 122 | 111 | -11 |
| 131 | 116 | -15 |
| 127 | 113 | -14 |
| 123 | 119 | -4 |
| 132 | 121 | -11 |

$M$ = -11

# Paired Samples *t* Test: Does Studying in the Exam Room Help?

1. Identify
   - Populations:
     - Pop 1: Exam grades when studying and testing are in the <u>same</u> room.
     - Pop 2: Exam grades when studying and testing are in <u>different</u> rooms.
   - Distribution:
     - Mean of Difference Scores: Distribution of Mean Differences
   - Test & Assumptions: One group of participants that is studied at two time points, paired-samples t test
     1. Data are interval
     2. Probably not random selection
     3. Sample size of 5 is less than 30, therefore distribution might not be normal

# Paired Samples *t* Test: Does Studying in the Exam Room Help?

2. State the null and research hypotheses

$H_0$: Studying and testing in the same room will result in the same grade as studying and testing in different rooms.

$H_1$: Studying and testing in the same room will result in a different grade than studying and testing in different rooms.

# Paired Samples $t$ Test: Does Studying in the Exam Room Help?

3. Determine characteristics of comparison distribution (distribution of mean differences)
   - Population: $\mu_M = 0$ (i.e., no mean difference)
   - Sample(s): $M = 11$, $s = 4.301$, $s_M = 1.923$

| Difference Score X-Y | *Deviation Score (Score - Mean)* | Squared Deviation *(Score - Mean)*$^2$ |
|---|---|---|
| -11 | 0 | 0 |
| -15 | -4 | 16 |
| -14 | -3 | 9 |
| -4 | 7 | 49 |
| -11 | 0 | 0 |
| $M = -11$ | | $SS_X = 74$ |

$$s = \sqrt{\frac{\Sigma(X-M)^2}{N-1}} = \sqrt{\frac{74}{5-1}} = 4.301$$

$$s_M = \frac{s}{\sqrt{N}} = \frac{4.301}{\sqrt{5}} = 1.923$$

# Paired Samples *t* Test: Does Studying in the Exam Room Help?

$\mu_M = 0, s_M = 1.923, M = -11, N = 5, df = 4$

4. Determine critical value (cutoffs)
   - In Behavioral Sciences, we use $p = .05$ (5%)
   - Our hypothesis ("Studying and testing in the same room will result in a different grade than studying and testing in different rooms.") is ***nondirectional*** so our hypothesis test is ***two-tailed***.

Significance level = $\alpha$

| Degrees of Freedom | .005 (1-tail) .01 (2-tails) | .01 (1-tail) .02 (2-tails) | .025 (1-tail) .05 (2-tails) | .05 (1-tail) .10 (2-tails) | .10 (1-tail) .20 (2-tails) | .25 (1-tail) .50 (2-tails) |
|---|---|---|---|---|---|---|
| 1 | 63.657 | 31.821 | 12.706 | 6.314 | 3.078 | 1.000 |
| 2 | 9.925 | 6.965 | 4.303 | 2.920 | 1.886 | .816 |
| 3 | 5.841 | 4.541 | 3.182 | 2.353 | 1.638 | .765 |
| 4 | 4.604 | 3.747 | 2.776 | 2.132 | 1.533 | .741 |
| 5 | 4.032 | 3.365 | 2.571 | 2.015 | 1.476 | .727 |
| 6 | 3.707 | 3.143 | 2.447 | 1.943 | 1.440 | .718 |
| 7 | 3.500 | 2.998 | 2.365 | 1.895 | 1.415 | .711 |

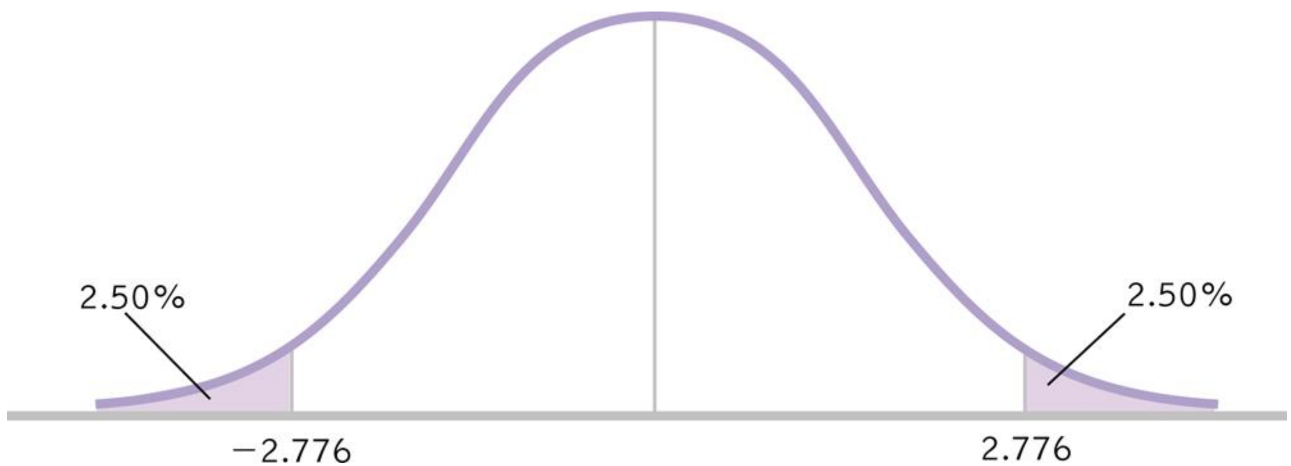
$df = 4 \longrightarrow$

# Paired Samples $t$ Test: Does Studying in the Exam Room Help?

$\mu_M = 0, s_M = 1.923, M = -11, N = 5, df = 4$

4. Determine critical value (cutoffs)

$t_{crit} = \pm 2.76$



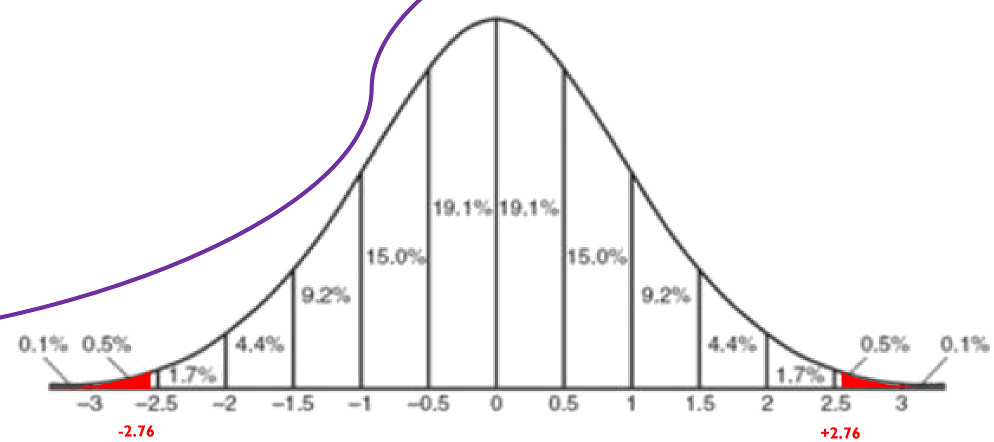
# Paired Samples $t$ Test: Does Studying in the Exam Room Help?

$\mu_M = 0$, $s_M = 1.923$, $M = -11$, $N = 5$, $df = 4$

5. Calculate the test statistic

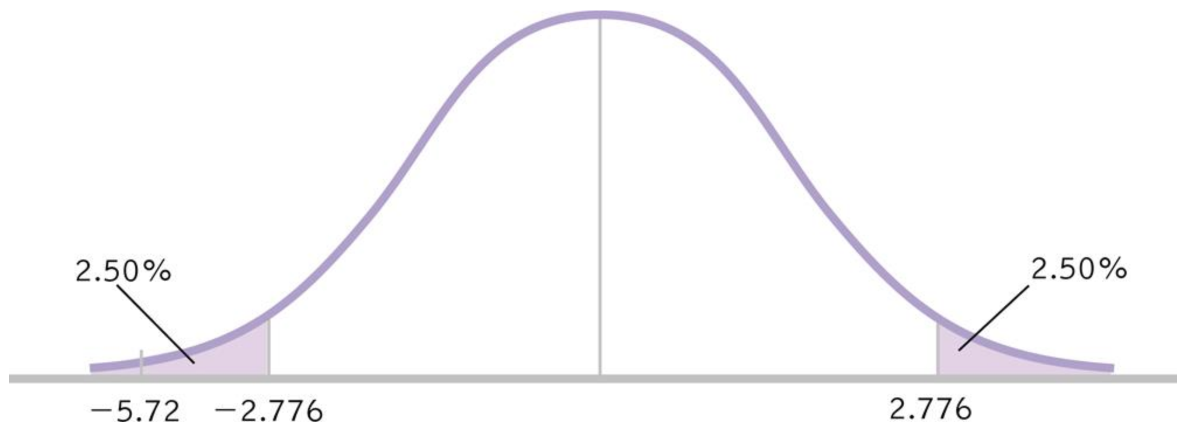
$$t = \frac{(M - \mu_M)}{s_M} = \frac{(-11 - 0)}{1.923} = -5.720$$

6. Make a decision



# Paired Samples *t* Test: Does Studying in the Exam Room Help?

$\mu_M = 0, s_M = 1.923, M = 11, N = 5, df = 4$




6. Make a decision

   $t = -5.720 > t_{crit} = \pm 2.776$, *reject the null hypothesis*

   People studying and testing in different rooms performed worse than … in the same rooms, $t(4) = 5.72, p < .05$.

# Paired Samples: Ideal Design?

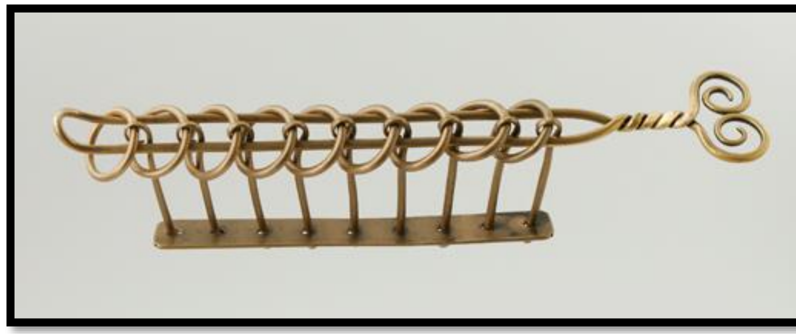
- Given the simple calculations and low cost (time and participants), why wouldn't we always use paired/dependent/within subjects designs?

- Back to the hypothetical beer tasting experiment…

Two Unrelated (unpaired) samples

# INDEPENDENT SAMPLES *t* TEST

# My Father's "Chinese Lock"



- Does teaching someone a strategy decrease the amount of time it takes to solve the puzzle?

- How do you design this study?
  - Paired Samples option #1: No Strategy → Strategy
  - Paired Samples option #2: Strategy → No Strategy

# Independent Samples *t* Test

- *Used to compare 2 means for a between-groups design, a situation in which each participant is assigned to only one condition.*

$$t = \frac{\left[(M_X - M_Y) - (\mu_X - \mu_Y)\right]}{s_{Difference}} = \frac{(M_X - M_Y)}{s_{Difference}}$$

- *New Statistics & Terminology:*
  - *Distribution of Differences Between Means*
  - $df_X$, $df_Y$, $df_{Total}$
  - *Pooled Variance*
  - *Standard Error of the Difference*

# Six Steps for Hypothesis Testing

1. Identify
2. State the hypotheses
3. Characteristics of the comparison distribution
4. Critical values
5. Calculate
6. Decide

*old hat*

# Independent Samples *t* Test: Gender Differences in Humor Appreciation


*I tend to believe that very few differences exist between males and females in cognitive abilities but there is some evidence that there are gender differences in, for example, humor appreciation.*

# Independent Samples *t* Test: Gender Differences in Humor Appreciation

*In this hypothetical study we ask: what percentage of cartoons do men and woman consider funny? We recruited 9 people from the psychology subject pool and asked them to view a cartoon. After the cartoon, each participant gave us a humor rating of the cartoon, from 0-100 (100 being the funniest possible). Here are those data.*


| Women (X) | Men (Y) |
|---|---|
| 84 | 88 |
| 97 | 90 |
| 58 | 52 |
| 90 | 97 |
| | 86 |

# Independent Samples *t* Test: Gender Differences in Humor Appreciation

1. Identify
   - Populations:
     - Pop 1: Women exposed to the cartoon
     - Pop 2: Men exposed to the cartoon
   - Distribution:
     - Difference Between Means: Distribution of <u>Differences Between Means</u>
       - Not Distribution of <u>Mean Differences</u>
   - Test & Assumptions: One group of participants that is studied at two time points, paired-samples t test
     1. Data are interval
     2. Random selection
     3. Sample size of 9 is less than 30, therefore distribution might not be normal

# Independent Samples *t* Test: Gender Differences in Humor Appreciation

2. State the null and research hypotheses


$H_0$: Women will categorize the same number of cartoons as funny as will men.

$H_1$: Women will categorize a different number of cartoons funny than will men.

# Independent Samples $t$ Test: Gender Differences in Humor Appreciation

3. Determine characteristics of comparison distribution (distribution of differences between means)
   - Population: $\mu_1 = \mu_2$ (i.e., no difference between means)

$$t = \frac{\left[\left(M_X - M_Y\right) - \left(\mu_X - \mu_Y\right)\right]}{s_{Difference}} = \frac{\left(M_X - M_Y\right)}{s_{Difference}}$$

What the @#$% ?

# Independent Samples *t* Test: Gender Differences in Humor Appreciation

3. Determine characteristics of comparison distribution

$$S_{Difference}$$

- Standard Error of the Difference:
  a) Calculate variance for each sample
  b) Pool variances, accounting for sample size
  c) Convert from squared standard deviation to squared standard error
  d) Add the two variances
  e) Take square root to get estimated standard error for distribution of differences between means.

# Calculating $s_{Difference}$

a) Calculate variance ($s^2$) for each sample

| Women (X) | X-M | (X-M)² |
|---|---|---|
| 84 | 1.75 | 3.063 |
| 97 | 14.75 | 217.563 |
| 58 | -24.25 | 588.063 |
| 90 | 7.75 | 60.063 |

$M_X$ = 82.25 $SS_X$ = 868.752

$$s_X^2 = \frac{\Sigma(X-M)^2}{N-1} = \frac{868.752}{4-1} = 289.584$$

| Men (Y) | Y-M | (Y-M)² |
|---|---|---|
| 88 | 5.4 | 29.16 |
| 90 | 11.4 | 129.96 |
| 52 | -30.6 | 936.36 |
| 97 | 14.4 | 207.36 |
| 86 | 3.4 | 11.56 |

$M_Y$ = 82.6 $SS_Y$ = 1314.4

$$s_Y^2 = \frac{\Sigma(Y-M)^2}{N-1} = \frac{1314.4}{5-1} = 328.6$$

# Independent Samples *t* Test

- *New* Types of Samples Means…
- *New* Degrees of Freedom

| Women (X) |
|---|
| 84 |
| 97 |
| 58 |
| 90 |
| $N$ = 4 |

$$df_X = N - 1 = 4 - 1 = 3$$

$$df_X = 3$$

$$df_Y = N - 1 = 5 - 1 = 4$$

$$df_Y = 4$$

$$df_{Total} = df_X + df_Y = 3 + 4 = 7$$

$$df_{Total} = 7$$

| Men (Y) |
|---|
| 88 |
| 90 |
| 52 |
| 97 |
| 86 |
| $N$ = 5 |

# Pooled Variance

b) Pool variances, accounting for sample size

- *Weighted average of the two estimates of variance – one from each sample – that are calculated when conducting an independent samples t test.*

$$s_{Pooled}^2 = \left(\frac{df_X}{df_{Total}}\right) s_X^2 + \left(\frac{df_Y}{df_{Total}}\right) s_Y^2$$

$$s_{Pooled}^2 = \left(\frac{3}{7}\right) 289.584 + \left(\frac{4}{7}\right) 328.6$$

$$s_{Pooled}^2 = 124.107 + 187.771 = 311.878$$

# Independent Samples *t* Test: Gender Differences in Humor Appreciation

c) Convert from squared standard deviation to squared standard error

$$s^2_{Pooled} = 311.878$$

$$s^2_{M_X} = \frac{s^2_{Pooled}}{N} = \frac{311.878}{4} = 77.970 \qquad s^2_{M_Y} = \frac{s^2_{Pooled}}{N} = \frac{311.878}{5} = 62.376$$

d) Add the two variances

$$s^2_{Difference} = s^2_{M_X} + s^2_{M_Y} = 77.970 + 62.376 = 140.346$$

e) Take square root to get estimated standard error for distribution of differences between means.

$$s_{Difference} = \sqrt{s^2_{Difference}} = \sqrt{140.346} = 11.847$$

# Independent Samples *t* Test: Gender Differences in Humor Appreciation

4. Determine critical value (cutoffs)
   - In Behavioral Sciences, we use $p$ = .05 (5%)
   - Our hypothesis ("Women will categorize a different number of cartoons funny than will men.") is <u>*nondirectional*</u> so our hypothesis test is <u>*two-tailed*</u>.

$t = \pm 2.365$

Significance level = $\alpha$

$df_{Total} = 7$

| Degrees of Freedom | .005 (1-tail) .01 (2-tails) | .01 (1-tail) .02 (2-tails) | .025 (1-tail) .05 (2-tails) | .05 (1-tail) .10 (2-tails) | .10 (1-tail) .20 (2-tails) | .25 (1-tail) .50 (2-tails) |
|---|---|---|---|---|---|---|
| 1 | 63.657 | 31.821 | 12.706 | 6.314 | 3.078 | 1.000 |
| 2 | 9.925 | 6.965 | 4.303 | 2.920 | 1.886 | .816 |
| 3 | 5.841 | 4.541 | 3.182 | 2.353 | 1.638 | .765 |
| 4 | 4.604 | 3.747 | 2.776 | 2.132 | 1.533 | .741 |
| 5 | 4.032 | 3.365 | 2.571 | 2.015 | 1.476 | .727 |
| 6 | 3.707 | 3.143 | 2.447 | 1.943 | 1.440 | .718 |
| 7 | 3.500 | 2.998 | 2.365 | 1.895 | 1.415 | .711 |
| 8 | 3.355 | 2.896 | 2.306 | 1.860 | 1.397 | .706 |
| 9 | 3.250 | 2.821 | 2.262 | 1.833 | 1.383 | .703 |
| 10 | 3.169 | 2.764 | 2.228 | 1.812 | 1.372 | .700 |
| 11 | 3.106 | 2.718 | 2.201 | 1.796 | 1.363 | .697 |

# Independent Samples *t* Test: Gender Differences in Humor Appreciation

4. Determine critical values
   - $df_{Total} = 7$ $p = .05$ $t = \pm 2.365$

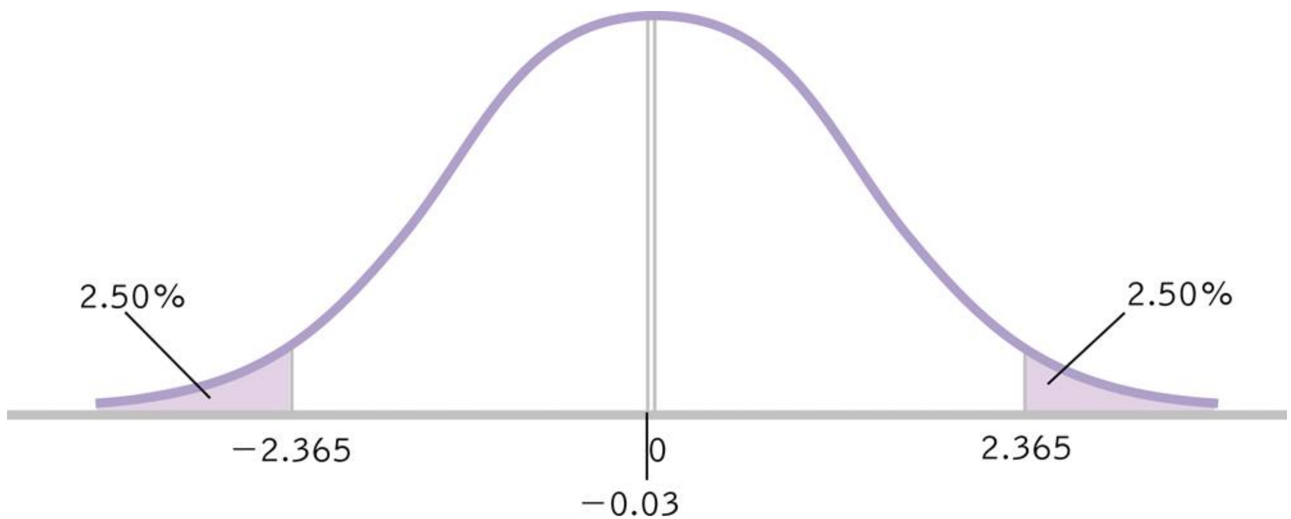
5. Calculate a test statistic

$$t = \frac{\left[(M_X - M_Y) - (\mu_X - \mu_Y)\right]}{s_{Difference}} = \frac{(M_X - M_Y)}{s_{Difference}}$$

$$t = \frac{(82.25 - 82.6)}{11.847} = -.03$$

# Independent Samples *t* Test: Gender Differences in Humor Appreciation

6. Make a decision




- Fail to reject null hypothesis
  - Men and women find cartoons equally humorous, $t(7) = .03, p > .05$

# Summary

**TABLE 9-2. HYPOTHESIS TESTS AND THEIR DISTRIBUTIONS**

We must consider the appropriate comparison distribution when we choose which hypothesis test we will use.

| HYPOTHESIS TEST | NUMBER OF SAMPLES | COMPARISON DISTRIBUTION |
|---|---|---|
| *z* test | one | Distribution of means |
| Single-sample *t* test | one | Distribution of means |
| Paired-samples *t* test | two (same participants) | Distribution of mean difference scores |
| Independent-samples *t* test | two (different participants) | Distribution of differences between means |