Independent Samples t-test
(part 1)

Lecture 10

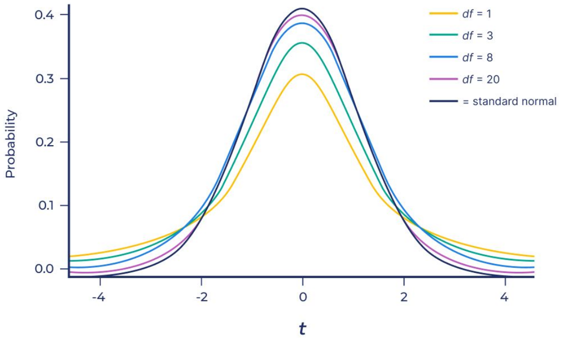
Emma Ning, M.A.

From our last lecture...

- One-sample t-test

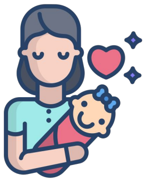
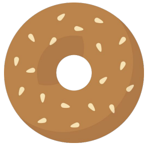
$$t = \frac{M - \mu}{s_M} = \frac{M - \mu}{\frac{s}{\sqrt{n}}}$$

- t-distribution & Effect size



$$d = \frac{M - \mu}{s}$$

- Type 1 and Type 2 errors

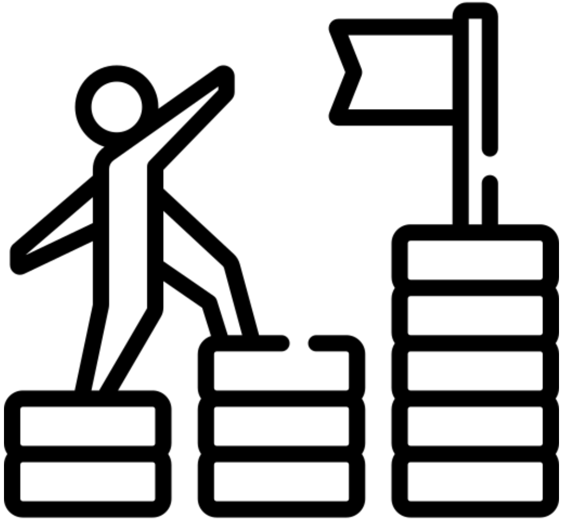



Well… we said last time we need one-sample t-test instead of z-test because we don’t know population parameters.

We used <u>sample SD</u> instead of population SD to estimate SE.

But what about that population mean in the numerator?

$$t = \frac{M - \boxed{\mu}}{s_M} = \frac{M - \boxed{\mu}}{\frac{s}{\sqrt{n}}}$$



Independent samples t-test

z-test

One-sample t-test

More realistic

TODAY'S PLAN

01

The Independent Samples Design

02

NHST Steps for Independent Samples t-test

03

Worked Example

04

Wrap-Up

Learning objectives

- **Differentiate** between a **between-subjects** and **within-subjects** design
- **Understand** and can **explain** why we use **pooled variance** in our calculation of estimated standard error for an independent samples t test
- **Calculate** the t statistic, degrees of freedom, and effect size for an independent samples t test
- **Interpret** and **write** the results of an independent samples t test in APA style with plain English explanation

01

The Independent Samples Design

Recap: The Two Tests We Learned

$$z = \frac{M - \mu}{\boxed{\sigma}_M} = \frac{M - \mu}{\frac{\boxed{\sigma}}{\sqrt{n}}}$$

$$t = \frac{M - \mu}{\boxed{s}_M} = \frac{M - \mu}{\frac{\boxed{s}}{\sqrt{n}}}$$

	z-test	One-sample t-test
μ	✅	✅
σ	✅	❌

Recap: The Two Tests We Learned

Both of them are asking the question "Is my *__sample__* weird compared to the *__population__*?"

But in real life, we almost never know the population standard deviation—so we use the t-test instead of the z-test.

$$z = \frac{M - \mu}{\boxed{\sigma}_M} = \frac{M - \mu}{\frac{\boxed{\sigma}}{\sqrt{n}}}$$

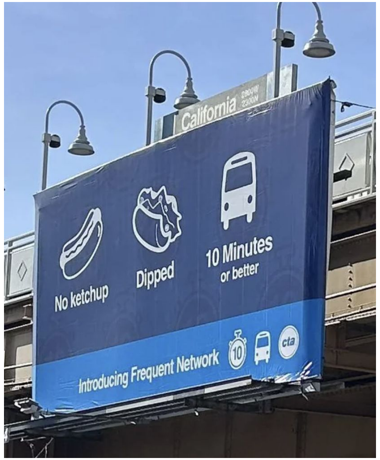
$$t = \frac{M - \mu}{\boxed{s}_M} = \frac{M - \mu}{\frac{\boxed{s}}{\sqrt{n}}}$$

	z-test	One-sample t-test
μ	✅	✅
σ	✅	❌

Tell me about your interests!

Think of a research question you're genuinely curious about — from your lab, your life, or just something random that interests you.

– Could we answer it with the kinds of tests we've learned so far?
– Or would we need something new?



Is the average wait time for the Blue Line longer than the city's posted estimate of 10 minutes?

Frequent Network

Is the average wait time for the #8 bus longer than #54 bus? (#8 is not on frequent network, #54 is)

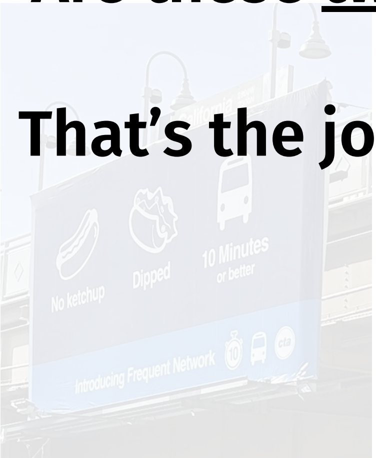
Very often, we don't know anything about the population at all. That's when the question becomes: Are these <u>two groups</u> different from each other?

That's the job of the independent samples t-test.

Tell me about your interests!

Think of a research question you're genuinely curious about — from your lab, your life, or just something random that interests you.

– Could we answer it with the kinds of tests we've learned so far?
– Or would we need something new?



Is the average wait time for the Blue Line longer than the city's posted estimate of 10 minutes?

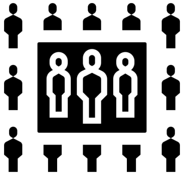
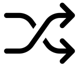
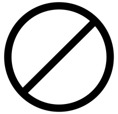
[Frequent Network]

Is the average wait time for the #8 bus longer than #54 bus? (#8 is not on frequent network, #54 is)

What is Independent Samples t-test

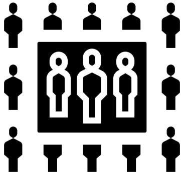
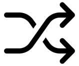
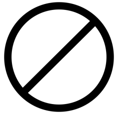
It is used to compare the means of two groups.
Suppose we are a pharma company, and we want to see if our drug improves sleep meaningfully.

We take our sample and randomly assign participants to either a placebo group or a drug group.


Avg hours slept 6.50 8.50

What is Independent Samples t-test


Avg hours slept	6.50	8.50

The t-test helps us answer:

Do people in the drug group sleep more, on average, than those in the placebo group—and is that difference large enough that it's unlikely to be due to chance?

Now, let's think with the knowledge we already have.

We can already reason out how to compute for the independent samples t-test.

This is where understanding wins 100x over memorization.

Intuition: Independent Samples t-test Formula

$$t = \frac{M - \boxed{\mu}}{\boxed{s}_M} = \frac{M - \boxed{\mu}}{\frac{\boxed{s}}{\sqrt{n}}}$$

Last class, we used sample SD to estimate SE (denominator).

Now instead of comparing to the population mean, we are comparing two groups (i.e., samples).

$$t = \frac{M_1 - M_2}{\text{some sort of SE}}$$

Intuition: Independent Samples t-test Formula

Signal over noise

$$t = \frac{M_1 - M_2}{\text{some sort of SE}}$$

Numerator:
Effect caused by our **treatment/experimental manipulation**
(e.g., different in sleep between control vs. drug group)

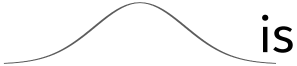
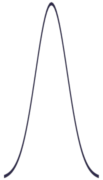
Denominator:
Represents chance variation—specifically, **sampling variability**. It tells us how much we would expect sample means (or mean differences) to vary from one sample to another

What is in the denominator as SE?

$$t = \frac{M_1 - M_2}{\text{some sort of SE}}$$

$$t = \frac{M - \mu}{s_M} = \frac{M - \mu}{\frac{\boxed{s}}{\sqrt{n}}}$$

We now have 2 groups. So should we use the sample SD of group 1 or group 2?

And, what if group 1  is more variable than group 2  ?

We can take into account the variability of both groups by calculating **pooled** variance/SD/SE.

Pooled Variance

$$S_p^{\,2} = \frac{SS_1 + SS_2}{df_1 + df_2}$$

The pooled variance is our best single **estimate of the population variance**.

$$t = \frac{M_1 - M_2}{\boxed{\text{some sort of SE}}}$$

This $S_p^{\,2}$ is partly involved in calculating the denominator! But not entirely...

$$t = \frac{M - \mu}{s_M} = \frac{M - \mu}{\frac{s}{\boxed{\sqrt{n}}}}$$

What about sample size n?

Estimated Standard Error

(for the Difference in Means)

$$S_{M1 - M2} = \sqrt{\frac{S_p^2}{n_1} + \frac{S_p^2}{n_2}}$$

After we have our best estimate of the population variance, S_p^2, we use that to estimate the **sampling distribution** of the difference between two means (M1 - M2).

This is our independent-samples t-test formula

Group 1
Sample Mean

Group 2
Sample Mean

$$t = \frac{\boxed{M_1} - \boxed{M_2}}{s_{M_1 - M_2}} = \frac{\boxed{M_1} - \boxed{M_2}}{\sqrt{\frac{s_p^2}{n_1} + \frac{s_p^2}{n_2}}}$$

Pooled SE

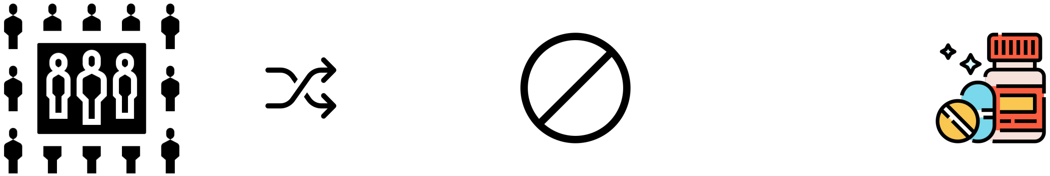
Every single step of this is a nod to our soup analogy

We are using our sample to infer about the population

If we only have sample statistics (without knowing μ or σ), we use our sample statistics to estimate them

And all t-tests operate based on the signal-to-noise logic

Why is it called Independent Samples t-test



Avg hours slept 6.50 8.50

If you’re in the placebo group, you can’t also be in the drug group. That means the two groups consist of entirely different people—so we call them **independent**.

Independent samples t-test is the same thing as “two-sample t-test”.

Paired Samples t-test

Independent samples t-test is just one type of t-test.
There is also paired samples t-test.

Here, we measure the **same** group of people twice—like before and after they take a drug—to see if there's any change.

Pre-drug  Post-drug

Recap: The Three Tests We Learned

	z-test	One-sample t-test	Independent Samples t-test
μ	✅	✅	❌
σ	✅	❌	❌
M	✅	✅	✅
s	✅	✅	✅

Compares 2 **samples**/groups

Compares 1 sample to **population**

02

NHST Steps for Independent Samples t-test

NHST STEPS

1 State the **null** and **alternative hypotheses**.

2 Set your **cutoff score** (find t_{crit}). *Use a **t-table**!*

3 Calculate your **test statistic** (i.e., t-statistic) <u>and **effect size**</u>.

4 Make your **decision**. *Reject the null* or *Fail to reject the null*

5 **Interpret** your results (in APA style).

NHST STEPS

1 State the **null** and **alternative hypotheses**.

2 Set your **cutoff score** (fi... *Use a **t-table**!*

3 Calculate your **test statistic** (i.e., t-statistic) and **effect size**.

4 Make your **de...** Fail to reject the null

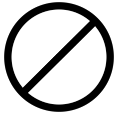
5 **Interpret** your results (in APA style).

The steps are EXACTLY the same as one-sample t-test!

You just write and interpret it differently.

Plus a few minor things about hypothesis, df and effect size calculation.

Basic Logic



Avg hours slept | 6.50 ⟷ Difference of 2 ⟷ 8.50

The t-test helps us answer:

Do people in the drug group sleep more, on average, than those in the placebo group—and is that difference large enough that it's unlikely to be due to chance?

TWO POSSIBILITIES

The 2 difference between groups is due to
sampling variability (not a real effect).

or

The 2 difference between groups is due to our
IV (not sampling error).

Hypotheses

Null Hypothesis

H_0

There is **no difference** in [**DV**] between the Sample 1 and Sample 2

$\mu_1 = \mu_2$

$\mu_1 - \mu_2 = 0$

Alternative Hypothesis

H_A

There is **a difference** in [**DV**] between the Sample 1 and Sample 2

$\mu_1 \neq \mu_2$

$\mu_1 - \mu_2 \neq 0$

Note: This is a two-tailed hypothesis example

Degrees of freedom

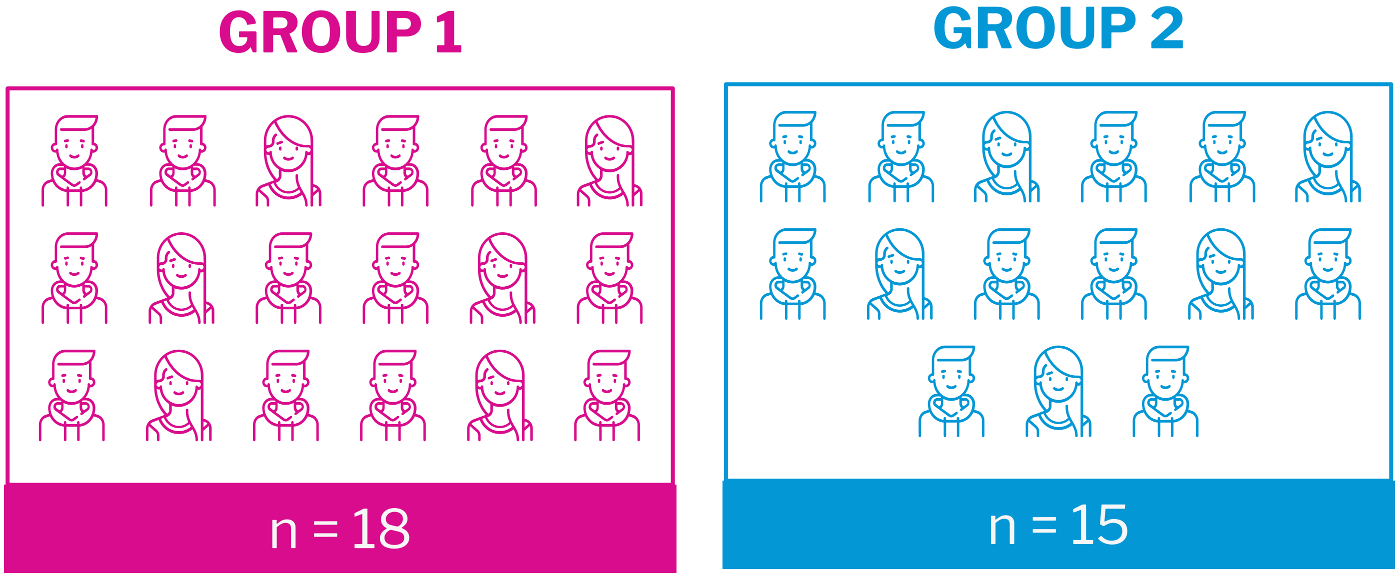
$$S_p^2 = \frac{SS_1 + SS_2}{df_1 + df_2}$$

$$df_1 = n - 1$$

$$df_2 = n - 1$$

So, our overall degrees of freedom would be $df_1 + df_2$, which is $n - 2$

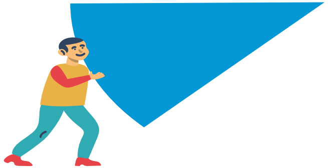
EXAMPLE



total $n = 18 + 15 = \mathbf{33}$

$df = 33 - 2 = \mathbf{31}$

also: $df_1 + df_2$
$= 17 + 14$
$= 31$



EFFECT SIZE

tells us **"how much"** of a difference there is between the samples

Cohen's *d*

Interpretation:

- **very small** (d < .2)
- **small** (d = .20 - .49)
- **medium** (d = .50 - .79)
- **large** (d > .80)

$$d = \frac{M_1 - M_2}{\sqrt{s_p^2}}$$

the **number of standard deviations** the means differ by

Cohen's d demo

03

Worked Example

Researchers at Instragram are interested in comparing social media usage rates between two generations: **Millennials** and **Gen Z** to plan political messaging. The collect data from 42 Millennials and 38 Gen Z-ers and want to compare the average daily Instagram usage (in mins) between the two generations.

Gen Z	Millennials
M = **88**	M = **75**

$$s_p^2 = 255$$

$$s_{M1 - M2} = 3.58$$

STEP 0 : Annotate Your Problem

sample 1 mean	sample 2 mean	total sample size	degrees of freedom	pooled variance	estimated standard error
M_1	M_2	n	df	s_p^2	s_{M1-M2}
75	88	80	?	255	3.58

NHST STEPS

1 State the **null** and **alternative hypotheses**.

2 Set your **cutoff score** (find t_{crit}). *Use a **t-table**!*

3 Calculate your **test statistic** (i.e., t-statistic) <u>and **effect size**</u>.

4 Make your **decision**. *Reject the null* or *Fail to reject the null*

5 **Interpret** your results (in APA style).

STEP 1: State Hypotheses

Null Hypothesis

H_0

There is **no difference** in Instagram usage between Millenials and Gen Z.

$\mu_1 = \mu_2$

Alternative Hypothesis

H_A

There **is a difference** in Instagram usage between Millenials and Gen Z.

$\mu_1 \neq \mu_2$

NHST STEPS

1. State the **null** and **alternative hypotheses**.

2. Set your **cutoff score** (find t_{crit}). *Use a **t-table**!*

3. Calculate your **test statistic** (i.e., t-statistic) <u>and **effect size**</u>.

4. Make your **decision**. *Reject the null* or *Fail to reject the null*

5. **Interpret** your results (in APA style).

STEP 2: Cutoff Score (t_{crit})

	Proportion (a) in One tail				
	.05	.025	.01	.005	.0005
	Proportion (a) in Two tails combined				
df	.10	.05	.02	.01	.001
18	1.734	2.101	2.552	2.878	3.922
19	1.729	2.093	2.539	2.861	3.883
20	1.725	2.086	2.528	2.845	3.850
21	1.721	2.080	2.518	2.831	3.819
22	1.717	2.074	2.508	2.819	3.792
23	1.714	2.069	2.500	2.807	3.768
24	1.711	2.064	2.492	2.797	3.745
25	1.708	2.060	2.485	2.787	3.725
26	1.706	2.056	2.479	2.779	3.707
27	1.703	2.052	2.473	2.771	3.689
28	1.701	2.048	2.467	2.763	3.674
29	1.699	2.045	2.462	2.756	3.660
30	1.697	2.042	2.457	2.750	3.646
40	1.684	2.021	2.423	2.704	3.551
60	1.671	2.000	2.390	2.660	3.460
120	1.658	1.980	2.358	2.617	3.373
∞	1.645	1.960	2.326	2.576	3.290

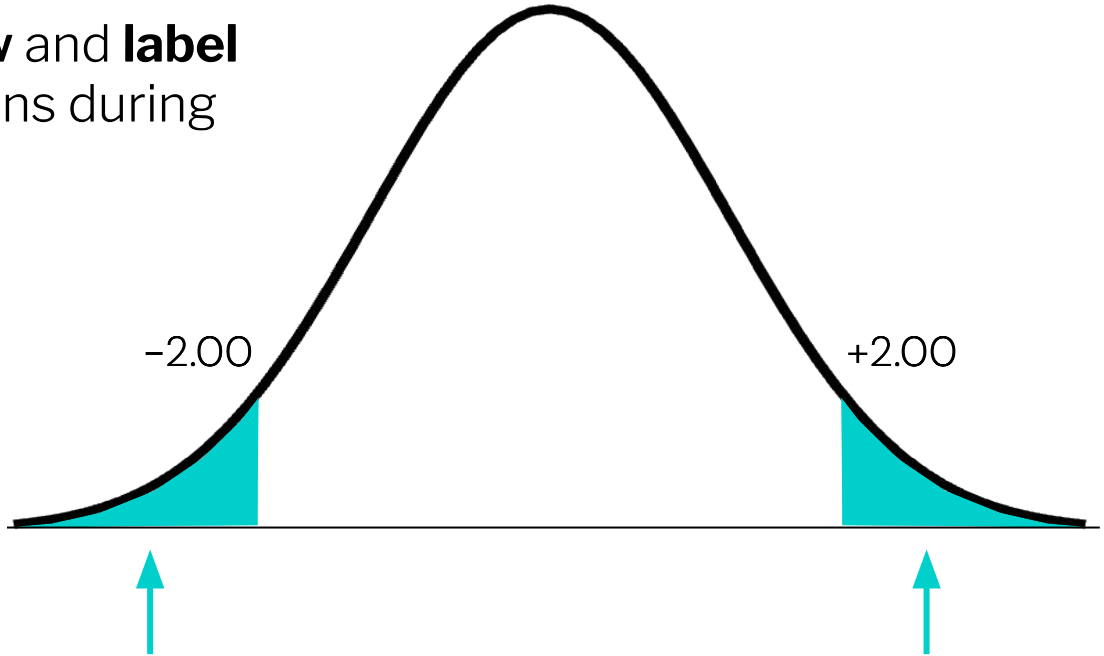
What we need to know:

- df = 78 (42 +38 - 2)
- $\alpha = 0.05$
- two-tailed test

Our t_{crit} is **±2.00**

STEP 2: Find Cutoff Score (t_{crit})

Always **draw** and **label** critical regions during this step!



$t_{crit} = \pm\mathbf{2.00}$

NHST STEPS

1 State the **null** and **alternative hypotheses**.

2 Set your **cutoff score** (find t_{crit}). *Use a **t-table**!*

3 Calculate your **test statistic** (i.e., t-statistic) <u>and **effect size**</u>.

4 Make your **decision**. *Reject the null* or *Fail to reject the null*

5 **Interpret** your results (in APA style).

STEP 3: Calculate *t* Statistic + Effect Size

sample 1 mean	sample 2 mean	total sample size	degrees of freedom	pooled variance	estimated standard error
M_1	M_2	n	df	s_p^2	s_{M1-M2}
75	88	80	78	255	3.58

3A Compute you **t statistic**.

$$t = \frac{M_1 - M_2}{s_{M1 - M2}} = \frac{88 - 75}{3.58} = \boxed{3.63}$$

3B Compute your ***effect size***.

$$d = \frac{M_1 - M_2}{\sqrt{s_p^2}} = \frac{88 - 75}{\sqrt{255}} = \boxed{0.81}$$

NHST STEPS

1 State the **null** and **alternative hypotheses**.

2 Set your **cutoff score** (find t_{crit}). *Use a* ***t-table****!*

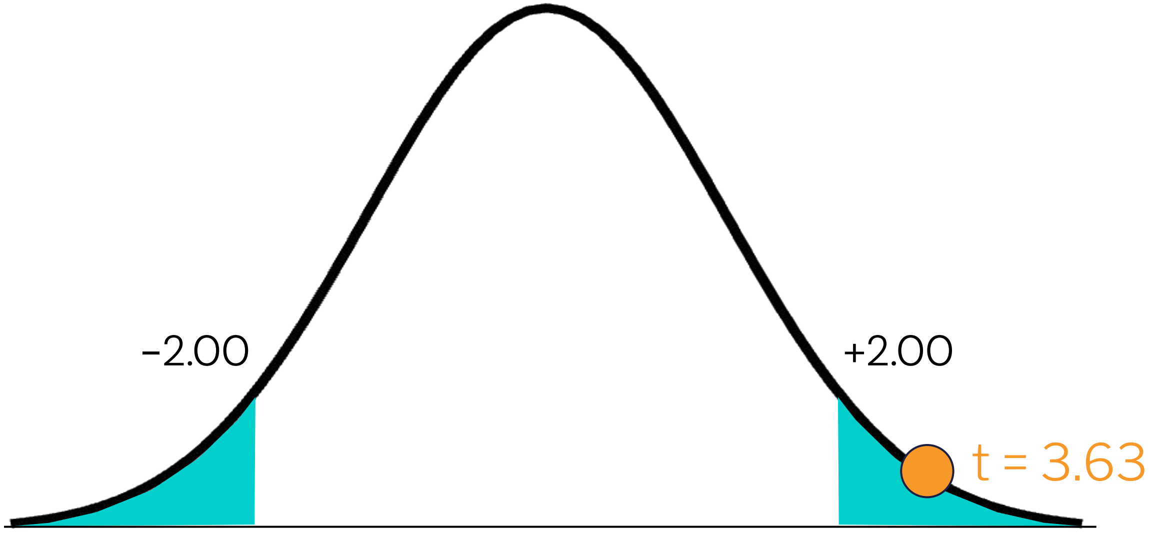
3 Calculate your **test statistic** (i.e., t-statistic) and **effect size**.

4 Make your **decision**. *Reject the null* or *Fail to reject the null*

5 **Interpret** your results (in APA style).

STEP 4 : Make Our Decision About Null

t_{crit} = ±**2.00**



Our t statistic is in the critical region, so we **reject the null hypothesis**.

NHST STEPS

1 State the **null** and **alternative hypotheses**.

2 Set your **cutoff score** (find t_{crit}). *Use a **t-table**!*

3 Calculate your **test statistic** (i.e., t-statistic) <u>and **effect size**</u>.

4 Make your **decision**. *Reject the null* or *Fail to reject the null*

5 **Interpret** your results (in APA style).

STEP 5: Write results in APA style.

Results from an **independent samples t-test** indicated that Gen Z **(*M* = 88)** used Instagram **significantly more** than Millennials **(*M* = 75)**, ***t*(78) = 3.63, *p* < .05, *d* = 0.81**, with a large effect size.

Important to report the direction!

$$\underbrace{t}_{\text{t-test}}(\overbrace{78}^{\text{df}}) = \underbrace{3.63}_{\text{t-statistic}},\ \overbrace{p}^{\text{p-value}} < \underbrace{.05}_{\alpha},\ \overbrace{d = 0.81}^{\text{Cohen's d}}$$

ICA 10

A team of researchers at a university mental health center is studying stress levels among undergraduate and graduate students. They collect data from 45 undergraduate students and 40 graduate students, measuring their average daily stress levels using a scale ranging from 0 to 100.

Undergraduates	Graduate Students
M_1 = **56**	M_2 = **61**

$s_p^2 = 148$ $s_{M1 - M2} = 2.65$

Conduct an independent samples t test and **show each step** on the board (α = 0.05, two-tailed test).

	Proportion (a) in <u>One</u> tail				
	.05	.025	.01	.005	.0005
	Proportion (a) in <u>Two tails</u> combined				
df	.10	.05	.02	.01	.001
18	1.734	2.101	2.552	2.878	3.922
19	1.729	2.093	2.539	2.861	3.883
20	1.725	2.086	2.528	2.845	3.850
21	1.721	2.080	2.518	2.831	3.819
22	1.717	2.074	2.508	2.819	3.792
23	1.714	2.069	2.500	2.807	3.768
24	1.711	2.064	2.492	2.797	3.745
25	1.708	2.060	2.485	2.787	3.725
26	1.706	2.056	2.479	2.779	3.707
27	1.703	2.052	2.473	2.771	3.689
28	1.701	2.048	2.467	2.763	3.674
29	1.699	2.045	2.462	2.756	3.660
30	1.697	2.042	2.457	2.750	3.646
40	1.684	2.021	2.423	2.704	3.551
60	1.671	2.000	2.390	2.660	3.460
120	1.658	1.980	2.358	2.617	3.373
∞	1.645	1.960	2.326	2.576	3.290

CALCULATIONS

3A Compute your **t statistic**.

$$t = \frac{M_1 - M_2}{s_{M1 - M2}} = \frac{56 - 61}{2.65} = \boxed{1.89}$$

3B Compute your ***effect size***.

$$d = \frac{M_1 - M_2}{\sqrt{s_p^2}} = \frac{56 - 61}{\sqrt{148}} = \boxed{0.41}$$

APA STYLE SENTENCE

Results from an independent samples t-test indicated that there was **not a significant difference** in **stress levels** between **undergraduates** (*M* = 56) and **graduate students** (*M* = 61), ***t*(83) = 1.89, *p* > .05, *d* = 0.40**.

df p-value Cohen's d

***t*(83) = 1.89, *p* > .05, *d* = 0.40**

t-test t-statistic α

04

Wrap Up

Key Takeaways

- An independent samples t-test compares the means between **two separate (independent) groups**; this is a **between-subjects design**.

- We must "pool" the variance between the two samples and use this **pooled variance** (s_p^2) to estimate our standard error.

- All the tests we learned fall under the framework of NHST, and their steps are very similar, with the main difference being:
  1) working with sample statistics rather than population;
  2) what we are comparing (two groups? paired?)