Five-minute review

1. Does a small p-value measure how large a difference is?

2. n=10: df?

Write your answers before we discuss them.

Comparing independent proportions

What we will learn

Test equality of two proportions using a pooled standard error.

Why does it matter?

A difference between sample proportions may reflect chance. The equality test uses a pooled estimate of the common null proportion.

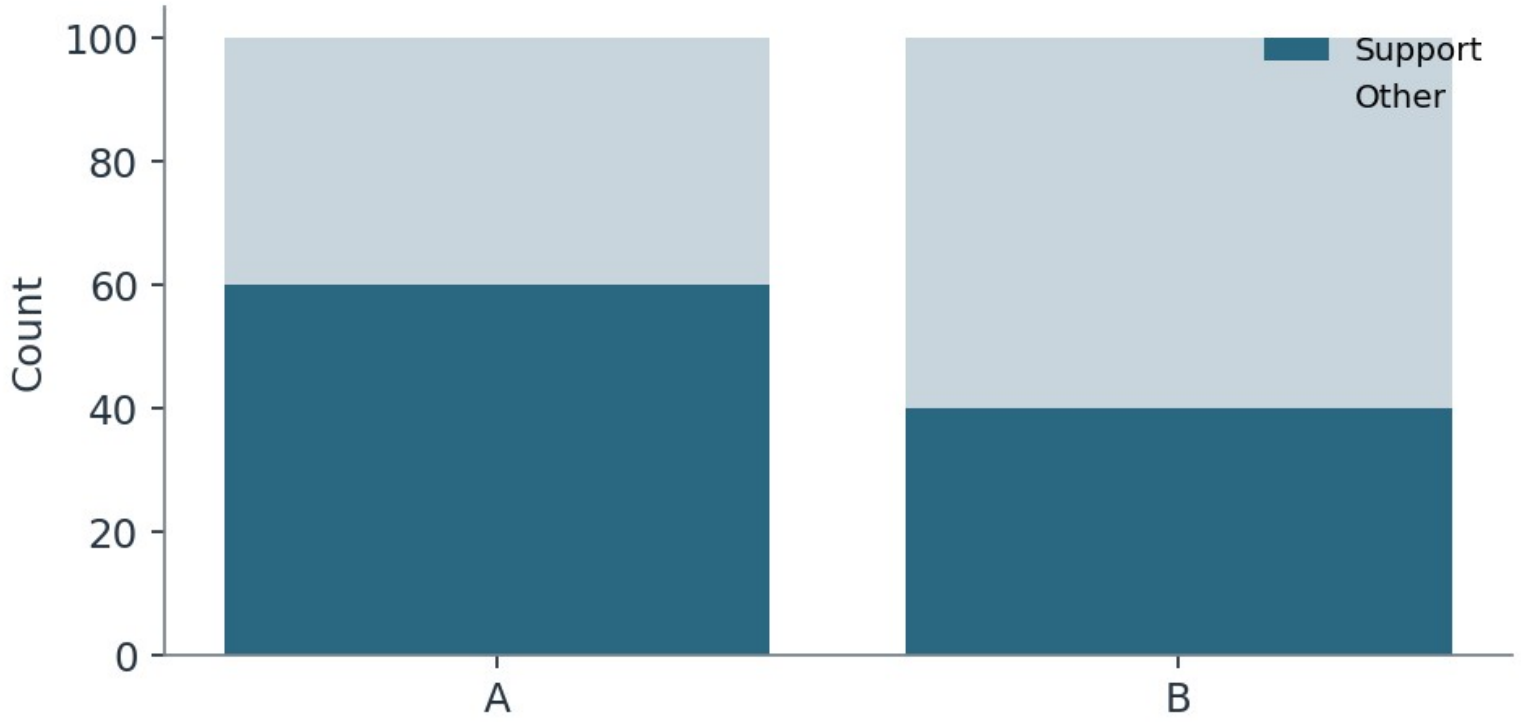
Definitions: Comparing independent proportions

Independent samples: observations in one group are not paired with those in another.

Difference of proportions: $p_1 - p_2$, estimated by $\hat{p}_1 - \hat{p}_2$.

Pooled proportion: combined successes divided by combined sample size, used for the equality-null test.

Illustration and interpretation



Notice the pattern

A difference between sample proportions may reflect chance. The equality test uses a pooled estimate of the common null proportion.

Illustrative data / model

Formulas

$$\hat{p} = \frac{x_1 + x_2}{n_1 + n_2}$$

$$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})(\frac{1}{n_1} + \frac{1}{n_2})}}$$

Important ideas and notation

Let $\hat{p}_1 = x_1/n_1$, $\hat{p}_2 = x_2/n_2$. Under $H_0: p_1 = p_2$, pooled $\hat{p} = (x_1 + x_2)/(n_1 + n_2)$.

$z = (\hat{p}_1 - \hat{p}_2)/\sqrt{[\hat{p}(1-\hat{p})(1/n_1 + 1/n_2)]}$. Check independent random groups, independent observations and at least 10 expected successes/failures per group under H_0.

Use chosen tail and α. A confidence interval for the difference uses an unpooled SE; do not substitute it into this pooled test.

Example 8.1.1

Group A $\frac{60}{100}$; B $\frac{40}{100}$; $H_a: p_1 \neq p_2$, $\alpha = .05$. Test whether the two population proportions differ.

Solution:

Pooled=.5

SE=√(.25×.02)=.07071

z=2.828

p≈.00468

reject

evidence proportions differ.

Example 8.1.2

A $\frac{30}{100}$, B $\frac{20}{100}$. Right-tail $\alpha=.05$. Test whether the first population proportion exceeds the second.

Solution:

Pooled=.25

SE≈.06124

z≈1.633

p≈.0512

fail to reject.

Example 8.1.3 • Comparing two independent surveys

Find the pooled sample proportion.

Group	Support	Other	Total
A	60	40	100
B	40	60	100

Illustrative classroom dataset

Example 8.1.3 • Solution

Find the pooled sample proportion.

Solution:

$$\frac{60+40}{100+100}=0.50.$$

Individual practice

1. A 45/100, B 55/100; pooled?

2. For same, direction of z?

3. A two-sided test p=.02 at α=.05?

Group activity

A 75/150, B 50/150; test $p_1 > p_2$ at .05.

Show your calculation or graph. Write a sentence explaining the result in context.

Check your answers

1. .50.

2. Negative.

3. Reject.

Group solution

Solution:

$\text{Pooled} = \frac{125}{300} = .4167$

$SE \approx .05693$

$z \approx 2.928$

$p \approx .0017$

reject.

Compare your conclusion with the question. Check your units and explain what the numbers represent.

Exit ticket • 5 minutes

1. Why pool proportions for equality test?

2. Group A has 42 successes in 60 trials; B has 24 in 40. Find both sample proportions and the pooled estimate.

Show your reasoning. Submit before leaving.

Five-minute review

1. Why pool proportions for equality test?

2. A 45/100, B 55/100; pooled?

Write your answers before we discuss them.

Comparing independent means

What we will learn

Choose an independent two-sample test and interpret mean differences.

Why does it matter?

When group membership differs, use an independent-samples method. Welch's t test does not require equal population variances.

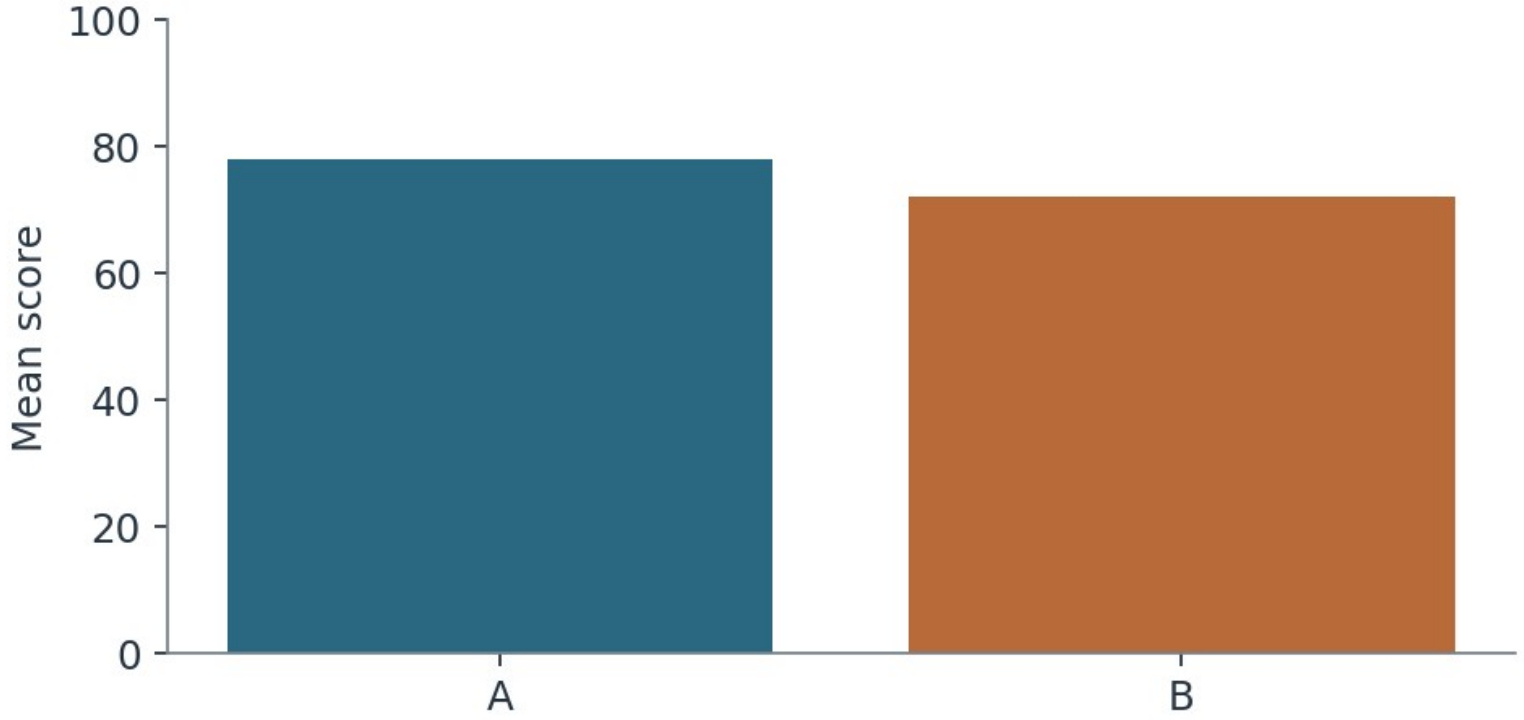
Definitions: Comparing independent means

Difference of means: $\mu_1 - \mu_2$, estimated by $\bar{x}_1 - \bar{x}_2$.

Welch t test: an independent two-mean test that does not assume equal population variances.

Pooled t test: a two-mean test using a shared variance estimate when equal variances are justified.

Illustration and interpretation



Notice the pattern

When group membership differs, use an independent-samples method. Welch's t test does not require equal population variances.

Illustrative data / model

Formulas

$$t = \frac{\bar{x}_1 - \bar{x}_2 - \Delta_0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

$$df \approx \frac{(a+b)^2}{\frac{a^2}{n_1 - 1} + \frac{b^2}{n_2 - 1}}, \quad a = \frac{s_1^2}{n_1},\ b = \frac{s_2^2}{n_2}$$

Important ideas and notation

For known σ_1,σ_2, $z=(\bar{x}_1-\bar{x}_2-\Delta_0)/\sqrt{(\sigma_1^2/n_1+\sigma_2^2/n_2)}$. Unknown variances: Welch t uses $s_1^2/n_1+s_2^2/n_2$ under the square root.

Welch degrees of freedom $\approx (a+b)^2/[a^2/(n_1-1)+b^2/(n_2-1)]$, where $a=s_1^2/n_1$, $b=s_2^2/n_2$. Prefer Welch unless equal variances are justified.

Pooled t uses $s_p^2=[(n_1-1)s_1^2+(n_2-1)s_2^2]/(n_1+n_2-2)$, $SE=s_p\sqrt{(1/n_1+1/n_2)}$, $df=n_1+n_2-2$. Require independent groups, normality or adequate n.

Example 8.2.1

Independent $n_1=n_2=25$, means 78,72, $s_1=s_2=10$; two-sided $\alpha=.05$. Test whether the two population means differ.

Solution:

Welch $SE=\sqrt{8}\approx 2.828$

$t\approx 2.121, df=48$

$p\approx .0391$

reject

evidence means differ.

Example 8.2.2

Known $\sigma_1=\sigma_2=8$, $n_1=n_2=32$, means 45,43. Test whether the two means differ at $\alpha=.05$.

Solution:

SE=2

z=1

two-sided p≈.3173

fail to reject at .05.

Example 8.2.3 • Comparing two independent class samples

What is the observed difference in means (A−B)?

Group	n	Mean	Sample SD
A	25	78	10
B	25	72	10

Illustrative classroom dataset

Example 8.2.3 • Solution

What is the observed difference in means (A−B)?

Solution:

6 points.

Individual practice

1. $n_1=n_2=16, s_1=s_2=4$: SE?

2. Means 20 and 18, SE=1: statistic?

3. Should paired pre/post scores use this independent test?

Group activity

Two independent samples: $n_1=n_2=16$, means 25,22, SDs 4,4; two-sided $\alpha=.05$.

Show your calculation or graph. Write a sentence explaining the result in context.

Check your answers

1. $\sqrt{2} \approx 1.414$.

2. 2.

3. No.

Group solution

Solution:

$SE \approx 1.414$

$t \approx 2.121, df = 30$

$p \approx .0423$

reject.

Compare your conclusion with the question. Check your units and explain what the numbers represent.

Exit ticket • 5 minutes

1. When is pooled t appropriate?

2. Two independent groups have $n_1=n_2=25$, means 60 and 55, and $s_1=s_2=5$. Calculate the Welch t statistic for equality.

Show your reasoning. Submit before leaving.

Five-minute review

1. When is pooled t appropriate?

2. $n_1=n_2=16, s_1=s_2=4$: SE?

Write your answers before we discuss them.

Paired means and matched data

What we will learn

Test a mean difference with matched observations.

Why does it matter?

Measure the same person twice and analyze each person's change. The pairs, not the individual measurements, determine the sample size.

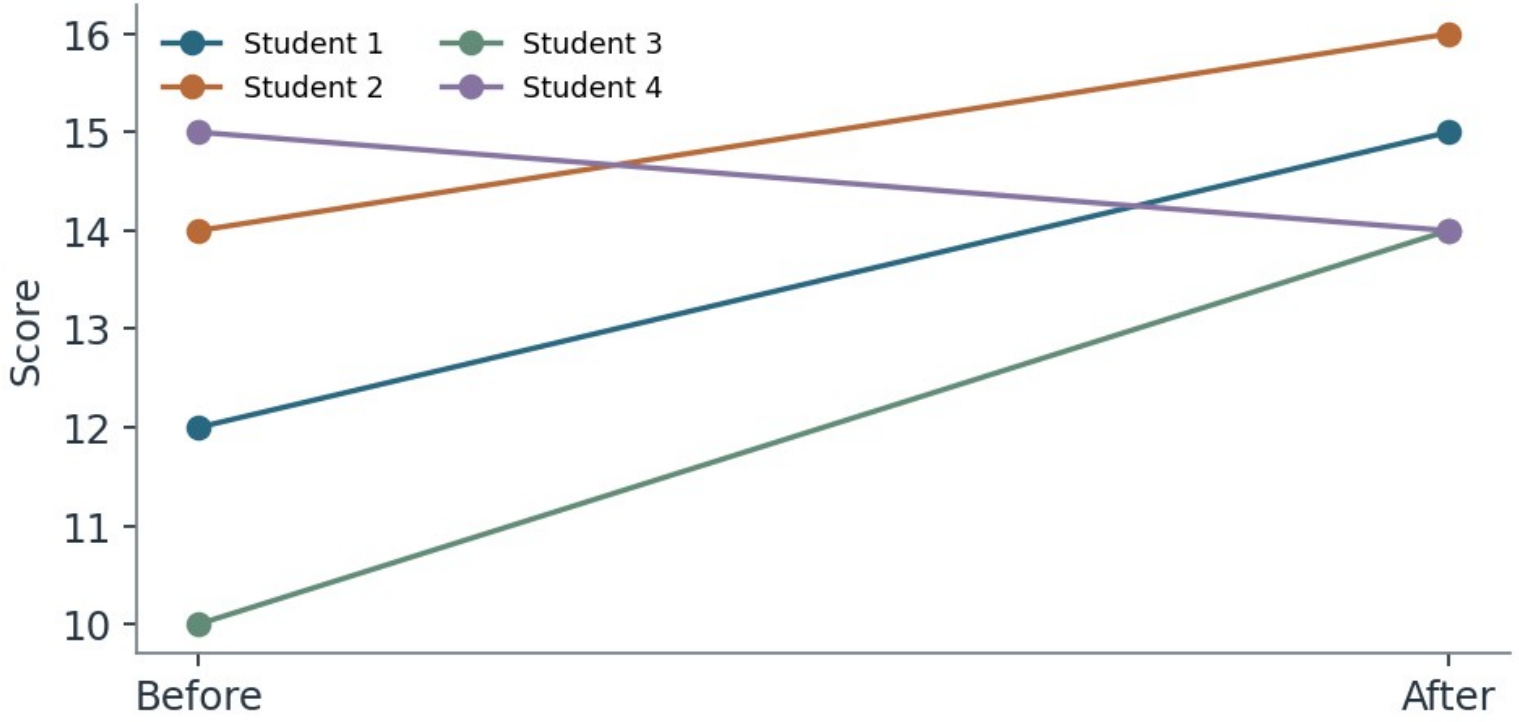
Definitions: Paired means and matched data

Paired data: each observation in one condition has a meaningful match in the other condition.

Difference score: the difference within each pair, using a consistent direction.

Paired t test: a one-sample t test of the mean difference scores.

Illustration and interpretation



Notice the pattern

Measure the same person twice and analyze each person's change. The pairs, not the individual measurements, determine the sample size.

Illustrative data / model

Formulas

$$d_i = x_{i,\text{after}} - x_{i,\text{before}}$$

$$t = \frac{\bar{d}}{\frac{s_d}{\sqrt{n}}}, \quad df = n - 1$$

Important ideas and notation

Paired observations have meaningful one-to-one matching. Compute d_i=after−before (or another declared direction). Test H_0:μd=0.

Paired t = $\bar{d}/(sd/\sqrt{n})$, df=n−1, with n pairs. Require independent pairs and differences approximately normal for small n.

The relevant variability is the SD of the differences, not the two separate SDs. A paired test does not alone rule out confounding in a before/after study.

Example 8.3.1

Nine pairs have mean difference 3, SD 3. $H_a: \mu d > 0$, $\alpha = .05$. Test whether the population mean difference is positive.

Solution:

$t = 3/(3/3) = 3, df = 8$

critical 1.860

$p \approx .0085$

reject

evidence positive mean difference.

Example 8.3.2

Before 10,12,14; after 12,13,17. Calculate the difference scores, their mean, sample SD and paired t statistic.

Solution:

Differences 2,1,3

mean 2

sample SD=1

$t = 2\sqrt{3} \approx 3.464$, $df = 2$.

Example 8.3.3 • Before and after practice

Find the average change for these four students.

Student	Before	After	Change
1	12	15	3
2	14	16	2
3	10	14	4
4	15	14	−1

Illustrative classroom dataset

Example 8.3.3 • Solution

Find the average change for these four students.

Solution:

$$\frac{3+2+4-1}{4} = 2 \text{ points.}$$

Individual practice

1. 12 pairs: df?

2. Differences 2,4,6: mean and SD?

3. If subtraction direction reverses, what changes?

Group activity

n=16 pairs,đ=1,sd=4; two-sided α=.05.

Show your calculation or graph. Write a sentence explaining the result in context.

Check your answers

1. 11.

2. 4 and 2.

3. Sign of mean difference and t; two-sided p-value unchanged.

Group solution

Solution:

SE=1

t=1,df=15

p≈.3332

fail to reject

insufficient evidence of mean change.

Compare your conclusion with the question. Check your units and explain what the numbers represent.

Exit ticket • 5 minutes

1. What is sample size in a paired study?

2. Three people have before values 8, 10, 12 and after values 9, 13, 14. Find differences and their mean.

Show your reasoning. Submit before leaving.