

Comparing Two Means: Paired and Independent Samples

Textbook Sections 5.2, 5.3

Emile Latour, Nicky Wakim, Meike Niederhausen

February 23, 2026



Learning Objectives

By the end of today's lecture, you will be able to:

1. Distinguish between paired and independent samples
2. Perform paired t-tests for dependent data in R
3. Construct and interpret confidence intervals for paired differences
4. Conduct two-sample t-tests for independent groups
5. Choose the appropriate test based on study design



Roadmap for Today

Part 1: Paired Data

- What makes data “paired”?
- Examples and study designs
- The vegetarian diet example
- Paired t-tests in R

Part 2: The Math Behind Paired Tests

- Population parameters vs. sample statistics
- Test statistics and p-values
- Confidence intervals for differences

Part 3: Independent Samples

- What makes samples “independent”?
- The caffeine and finger tapping example
- Two-sample t-tests in R

Part 4: Choosing the Right Test

- Paired vs. independent: key differences
- Study design implications
- Common mistakes to avoid



CI's and hypothesis testing for different scenarios:

Day	Section	Population parameter	Symbol	Point estimate	Symbol
9	5.1	Population mean	μ	Sample mean	\bar{x}
10	5.2	Population mean of paired differences	μ_d or δ	Sample mean of paired differences	\bar{x}_d
10	5.3	Differences in population means	$\mu_1 - \mu_2$	Differences in sample means	$\bar{x}_1 - \bar{x}_2$
13	8.1	Population proportion	p	Sample proportions	\hat{p}
14	8.2	Differences in population proportions	$p_1 - p_2$	Differences in sample proportions	$\hat{p}_1 - \hat{p}_2$



Paired Data



Where are we in the course?

We've been building up our inference toolkit:

So far:

- Single-sample mean: μ
 - CI: $\bar{x} \pm t^* \times \frac{s}{\sqrt{n}}$
 - Test: Compare \bar{x} to hypothesized μ_0

Today:

- Mean difference from **paired** samples: μ_d or δ
- Difference in means from **independent** samples: $\mu_1 - \mu_2$

Why this matters: The study design determines which test we use!



Reminder: The six steps of hypothesis testing

Before we dive into examples, let's review our standard framework:

1. **State hypotheses** (H_0 and H_A)
2. **Set significance level** (usually $\alpha = 0.05$)
3. **Check assumptions** (independence, normality/large n)
4. **Calculate test statistic**

- Paired: $t = \frac{\bar{x}_d - 0}{s_d / \sqrt{n}}$
- Independent: $t = \frac{(\bar{x}_1 - \bar{x}_2) - 0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$

5. **Find p-value** (probability of seeing this or more extreme)
6. **Make conclusion** (reject or fail to reject H_0 , with context)

Today we'll apply this framework to both paired and independent samples!



What are paired data?

Definition: Paired Data

Paired data occur when two sets of observations are uniquely matched so that an observation in one set corresponds to exactly one observation in the other.

Common scenarios:

1. Before-and-after studies (longitudinal data)

- Measure the same people at two time points
- Example: Blood pressure before and after medication

2. Matched pairs

- Twin studies: One twin gets treatment, other gets control
- Matched controls: Pair people based on age, sex, etc.

3. Natural pairs

- Left eye vs. right eye
- Parent-child pairs



Examples of paired designs

Example 1: Swimmer performance

- Competitive swimmers tested twice
- Once wearing wetsuit
- Once wearing regular swimsuit
- Compare maximum speed

Example 2: Drug effectiveness

- Patients tested before treatment
- Same patients tested after treatment
- Compare blood glucose levels

Why paired?

Why paired?

- Same swimmer in both conditions
- Controls for individual differences in ability

- Same patient at two time points
- Natural “before vs. after” comparison

Key insight: Pairing reduces variability because we’re comparing each person to themselves!



Today's example: Vegetarian diet and cholesterol

Research Question

Can a vegetarian diet reduce cholesterol levels?

Study design:

- 43 non-vegetarian adults enrolled
- Instructed to adopt a vegetarian diet
- Cholesterol measured **before** and **after** diet
- Follow-up period: 3 months

Why is this paired data?

- Same individuals measured twice
- Each person serves as their own control
- We can calculate: After - Before for each person



Loading and exploring the cholesterol data

```
1 library(dplyr)
2 library(here)
3 library(readr)
4
5 # Load the data
6 chol <- read_csv(here("data",
7                      "chol213_n40.csv"))
8
9 # Take a look at the structure
10 glimpse(chol)
```

Rows: 43

Columns: 2

\$ Before <dbl> 195, 145, 205, 159, 244, 166, 250, 236, 192, 224, 238, 197, 169...
\$ After <dbl> 146, 155, 178, 146, 208, 147, 202, 215, 184, 208, 206, 169, 182...

What do we have?

- **Before**: Cholesterol level before diet (mg/dL)
- **After**: Cholesterol level after diet (mg/dL)
- Each row is one person



Calculate the paired differences

For paired data, we create a new variable: the difference

```
1 # Calculate difference: After - Before
2 chol <- chol %>%
3   mutate(DiffChol = After - Before)
4
5 # Look at first few differences
6 chol

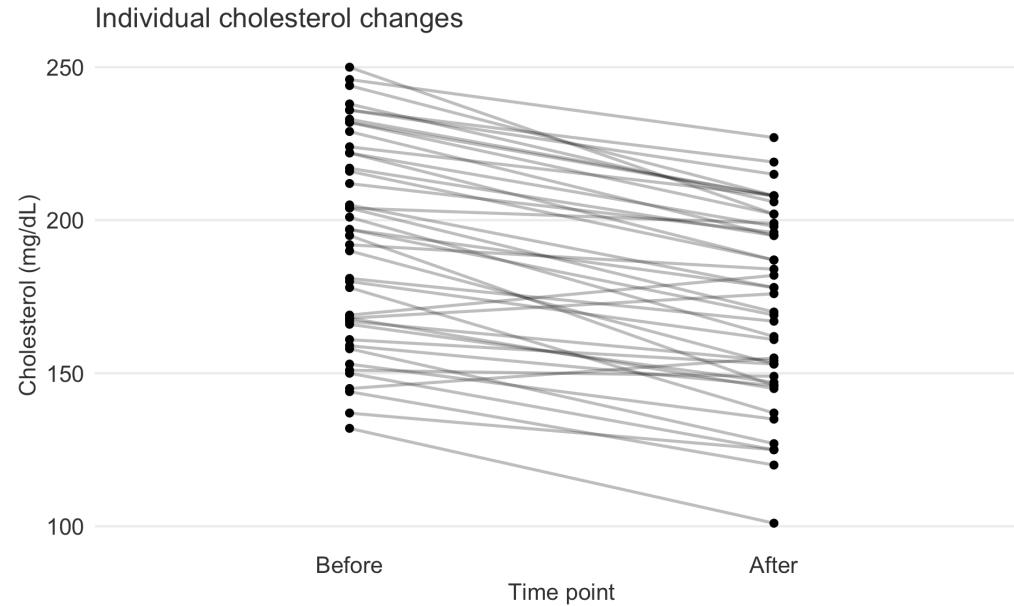
# A tibble: 43 × 3
  Before After DiffChol
  <dbl> <dbl>    <dbl>
1 195   146     -49
2 145   155      10
3 205   178     -27
4 159   146     -13
5 244   208     -36
6 166   147     -19
7 250   202     -48
8 236   215     -21
9 192   184      -8
10 224   208     -16
# i 33 more rows
```

Interpretation of differences:

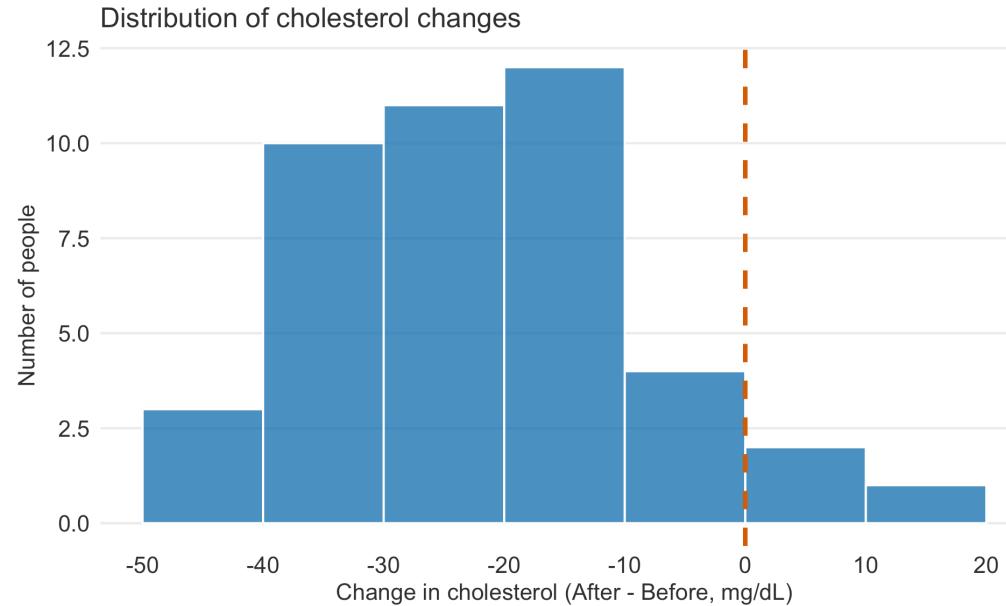
- Negative values: cholesterol decreased
- Positive values: cholesterol increased
- Zero: no change



Visualizing the paired differences



Each line represents one person's change



Most people decreased (negative values)



Summary statistics for paired data

```
1 # Summary statistics for differences
2 chol %>%
3   get_summary_stats(DiffChol, type = "common") %>%
4   gt() %>%
5   tab_options(table.font.size = 40)
```

variable	n	min	max	median	iqr	mean	sd	se	ci
DiffChol	43	-49	13	-23	16	-21.767	13.89	2.118	4.275

What we see:

- Mean difference: $\bar{x}_d = 21.767$ mg/dL
- On average, cholesterol decreased by about 22 mg/dL
- But is this difference statistically significant?

To answer this, we need a hypothesis test!



Notation for paired data

Population Parameters vs. Sample Statistics

Population (what we want to know)

- Mean difference: δ (delta) or μ_d
- Standard deviation: σ_d
- Sample size: N

Sample (what we observe)

- Sample mean difference: \bar{x}_d
- Sample standard deviation: s_d
- Sample size: n

Key insight: Once we calculate the differences, this becomes a one-sample problem!

- We have one number per person (the difference)
- We test if \bar{x}_d is significantly different from 0



Hypothesis test for paired data

Research question: Is there evidence that cholesterol changed after the vegetarian diet?

Step 1: State hypotheses

$$H_0 : \delta = 0$$

$$H_A : \delta \neq 0$$

In words:

- H_0 : The population mean difference in cholesterol is zero (no change)
- H_A : The population mean difference in cholesterol is not zero (there is a change)

Step 2: Set significance level

- Use $\alpha = 0.05$



Check the assumptions

Step 3: Step 1: Check the assumptions

- The assumptions to run a hypothesis test on a sample are:
 - **Independent pairs:** Each pair is independent from all other pairs,
 - **Approximately normal sample or big n:** the distribution of the sample should be approximately normal, or the sample size should be at least 30
- In our example, we would check the assumptions with a statement:
 - The pairs of observations are independent from each other and the number of pairs in our sample is 43. Thus, we can use CLT to approximate the sampling distribution.



The test statistic for paired data

Step 4: Calculate test statistic

The test statistic for paired data is:

$$t = \frac{\bar{x}_d - 0}{\frac{s_d}{\sqrt{n}}}$$

This looks familiar! It's the same formula as for a one-sample t-test.

```
1 # Calculate components
2 n <- nrow(chol)
3 xbar_d <- mean(chol$DiffChol)
4 s_d <- sd(chol$DiffChol)
5 SE <- s_d / sqrt(n)
6
7 # Calculate t-statistic
8 t_stat <- (xbar_d - 0) / SE
9 t_stat
[1] -10.27603
```

```
1 # Calculate the p-value
2 2 * pt(abs(t_stat), df = n - 1, lower.tail = FALSE)
[1] 4.945625e-13
```



Running a paired t-test in R (1/5)

Step 5: Find a p-value

Option 1: Use the difference variable

```
1 # Test the differences directly  
2 t.test(x = chol$DiffChol, mu = 0)
```

One Sample t-test

```
data: chol$DiffChol  
t = -10.276, df = 42, p-value = 4.946e-13  
alternative hypothesis: true mean is not equal to 0  
95 percent confidence interval:  
-26.04229 -17.49259  
sample estimates:  
mean of x  
-21.76744
```



Running a paired t-test in R (2/5)

Option 2: Use the `paired = TRUE` argument

```
1 # Let R calculate differences for us  
2 t.test(x = chol$After, y = chol$Before, paired = TRUE)
```

Paired t-test

```
data: chol$After and chol$Before  
t = -10.276, df = 42, p-value = 4.946e-13  
alternative hypothesis: true mean difference is not equal to 0  
95 percent confidence interval:  
 -26.04229 -17.49259  
sample estimates:  
mean difference  
 -21.76744
```



Running a paired t-test in R (3/5)

`broom::tidy()`: Use the `tidy()` function in the `broom` package with either Option 1 or Option 2

```
1 # The argument conf.int = TRUE gives a confidence interval (default is 95%)
2
3 t.test(x = chol$After, y = chol$Before, paired = TRUE) %>%
4   broom::tidy(conf.int = TRUE)

# A tibble: 1 × 8
  estimate statistic p.value parameter conf.low conf.high method    alternative
    <dbl>      <dbl>    <dbl>      <dbl>     <dbl>     <dbl> <chr>     <chr>
1    -21.8     -10.3 4.95e-13       42     -26.0     -17.5 Paired t... two.sided
```



Running a paired t-test in R (4/5)

Option 3: Use `rstatix` package`

- Requires that the data are in *long* format which means that
 - all of the outcome values are in one column and
 - another column indicates which group the values are from

```
1 chol_long
# A tibble: 86 × 4
  DiffChol ID    Time   Cholesterol
  <dbl>   <fct> <fct>      <dbl>
1     -49  1    Before     195
2     -49  1    After      146
3      10  2    Before     145
4      10  2    After      155
5     -27  3    Before     205
6     -27  3    After      178
7     -13  4    Before     159
8     -13  4    After      146
9     -36  5    Before     244
10    -36  5    After      208
# i 76 more rows
```



Running a paired t-test in R (5/5)

Option 3: Use `rstatix` package

```
1 library(rstatix)
2
3 t_test(data = chol_long,
4         Cholesterol ~ Time,
5         paired = TRUE,
6         detailed = TRUE)

# A tibble: 1 × 13
#>   estimate .y.    group1 group2     n1     n2 statistic      p    df conf.low
#>   <dbl> <chr> <chr> <chr> <int> <int>     <dbl> <dbl> <dbl> <dbl>
#> 1 21.8  Choleste... Before After     43     43     10.3 4.95e-13    42    17.5
#> # i 3 more variables: conf.high <dbl>, method <chr>, alternative <chr>
```



Interpreting the paired t-test output

```
1 # Save and tidy the results
2 chol_test <- t.test(x = chol$DiffChol,
3                      mu = 0)
4
5 chol_test %>%
6   tidy(conf.int = TRUE) %>%
7   gt() %>%
8   tab_options(table.font.size = 40)
```

estimate	statistic	p.value	parameter	conf.low	conf.high	method	alternative
-21.76744	-10.27603	4.945625e-13		42	-26.04229	-17.49259	One Sample t-test two.sided

What we see:

- t-statistic = -10.28
- p-value < 0.001 (very small!)
- 95% CI for δ : (-26.04, -17.49)

Step 6: Make a conclusion

- Since $p\text{-value} < \alpha = 0.05$, we reject H_0 .
- **Conclusion:** There is sufficient evidence that cholesterol levels changed after the vegetarian diet ($p < 0.001$).



Writing a complete conclusion

Best Practice: Include Key Numbers

A good conclusion includes:

1. Decision about H_0 (reject or fail to reject)
2. Context (what was measured)
3. Effect size (mean difference)
4. Confidence interval
5. P-value

Our conclusion:

After adopting a vegetarian diet, cholesterol levels decreased by an average of 21.77 mg/dL (95% CI: 17.49 to 26.04 mg/dL lower), which is significantly different from zero ($t = -10.28$, $df = 42$, $p < 0.001$, Paired t-test).



One-sided vs. two-sided tests

Two-sided test (what we just did):

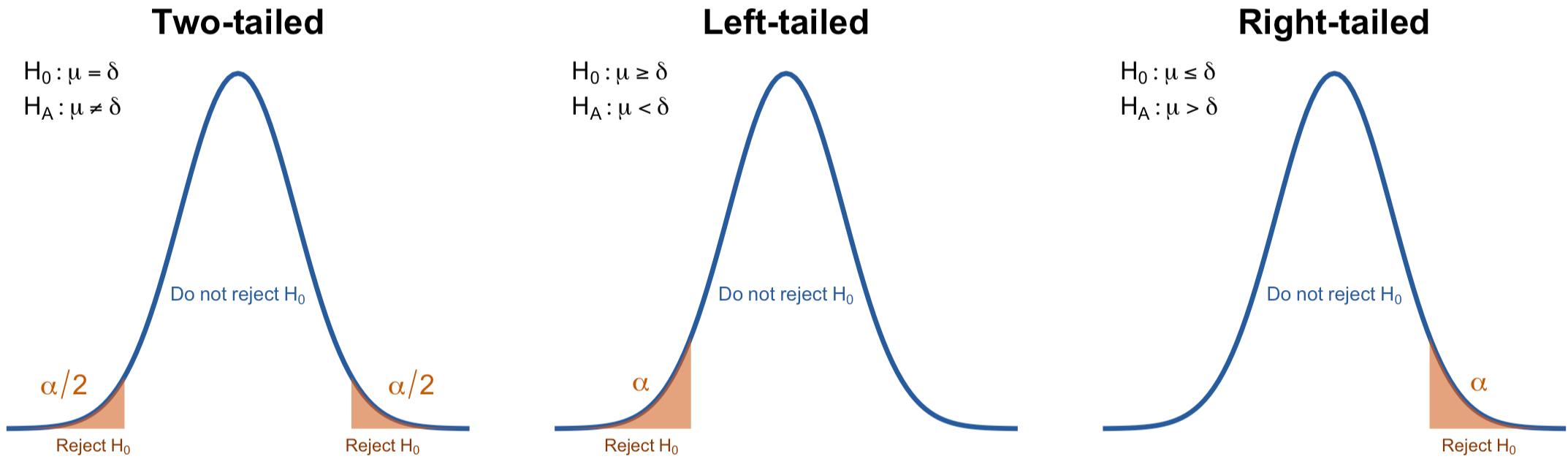
- $H_A : \delta \neq 0$ (different from zero)
- Appropriate when we don't know direction

One-sided test (if we have directional prediction):

- $H_A : \delta < 0$ (decrease)
- $H_A : \delta > 0$ (increase)



Visual: One-sided vs. two-sided



One-sided example

Example: If we specifically want to test if diet *decreased* cholesterol:

$$H_0 : \delta \geq 0$$

$$H_A : \delta < 0$$

In words:

- H_0 : The population mean difference in cholesterol greater than or equal to zero
- H_A : The population mean difference in cholesterol is less zero (there is a change)

```
1 t.test(x = chol$DiffChol, mu = 0, alternative = "less") %>%
2   broom::tidy(conf.int = TRUE) %>%
3   gt() %>%
4   tab_options(table.font.size = 40)
```

estimate	statistic	p.value	parameter	conf.low	conf.high	method	alternative
-21.76744	-10.27603	2.472813e-13	42	-Inf	-18.20461	One Sample t-test	less

Notice: For a one-sided test, the p-value changes and the confidence interval is one-sided.



Independent Samples



What are independent samples?

Definition: Independent Samples

Independent samples occur when individuals in one group are completely unrelated to individuals in the other group. There is no natural pairing or matching.

Key characteristics:

- Different people in each group
- No before/after measurements
- No natural matching
- Typically: different sample sizes are possible

Common scenarios:

- Treatment vs. control groups (randomized trials)
- Case vs. control studies
- Men vs. women
- Exposed vs. unexposed



Paired vs. Independent: The key difference

Paired Data

Structure:

- Same people measured twice
- OR matched pairs
- Same sample size for both conditions

Analysis:

- Calculate differences
- One number per person/pair
- One-sample t-test on differences

Example:

Weight before and after diet program (same 50 people)

Independent Samples

Structure:

- Different people in each group
- No natural pairing
- Can have different sample sizes

Analysis:

- Compare two separate means
- Two numbers: \bar{x}_1 and \bar{x}_2
- Two-sample t-test

Example:

Weight of 50 men vs. 50 women (different people)



Today's example: Caffeine and finger tapping

Research Question

Does caffeine increase finger tapping speed?

Study design:

- 70 college students trained to tap fingers rapidly
- Randomly assigned to two groups:
 - **Control:** Decaffeinated coffee
 - **Caffeine:** Coffee with ~200mg caffeine
- Double-blind design
- After 2 hours, tested finger taps per minute

Why independent samples?

- Different students in each group
 - Each person tested only once
-  No pairing or matching

Loading the caffeine data

```
1 # Load the data
2 CaffTaps <- read_csv(here("data",
3                           "CaffeineTaps_n35.csv"))
4
5 # Check structure
6 glimpse(CaffTaps)
```

Rows: 70

Columns: 2

\$ Taps <dbl> 246, 248, 250, 252, 248, 250, 246, 248, 245, 250, 242, 245, 244, ...
\$ Group <chr> "Caffeine", "Caffeine", "Caffeine", "Caffeine", "Caffeine", "Caf...

What we have:

- **Taps**: Finger taps per minute
- **Group**: Caffeine or NoCaffeine
- Each row is one person
- 35 people per group



Exploring the data by group

```
1 # Summary statistics by group
2 CaffTaps %>%
3   group_by(Group) %>%
4   get_summary_stats(Taps, type = "mean_sd") %>%
5   gt() %>%
6   tab_options(table.font.size = 40)
```

Group	variable	n	mean	sd
Caffeine	Taps	35	248.114	2.621
NoCaffeine	Taps	35	244.514	2.318

What we see:

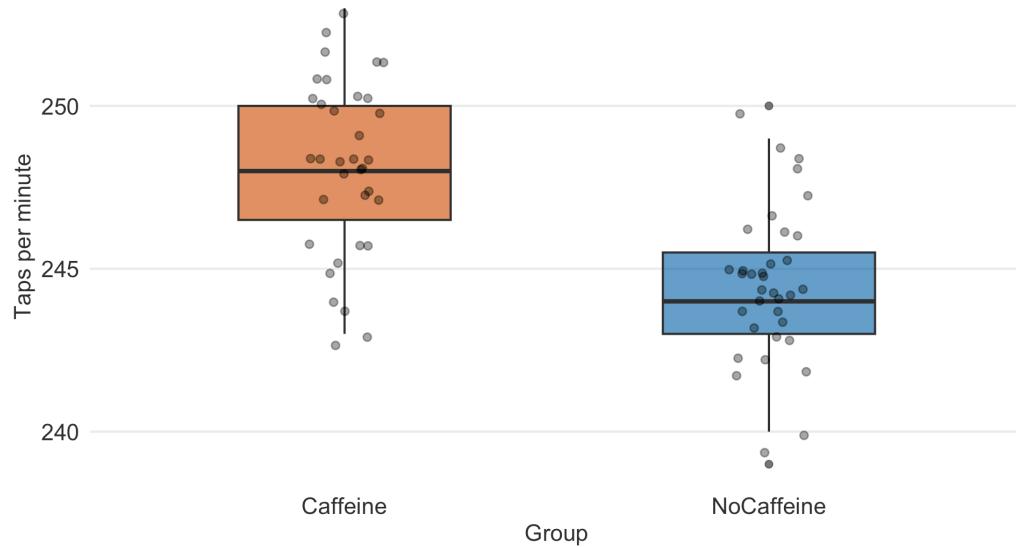
- Caffeine group: mean = 248.1 taps/min
- Control group: mean = 244.5 taps/min
- Difference = 3.6 taps/min

Question: Is this difference statistically significant?



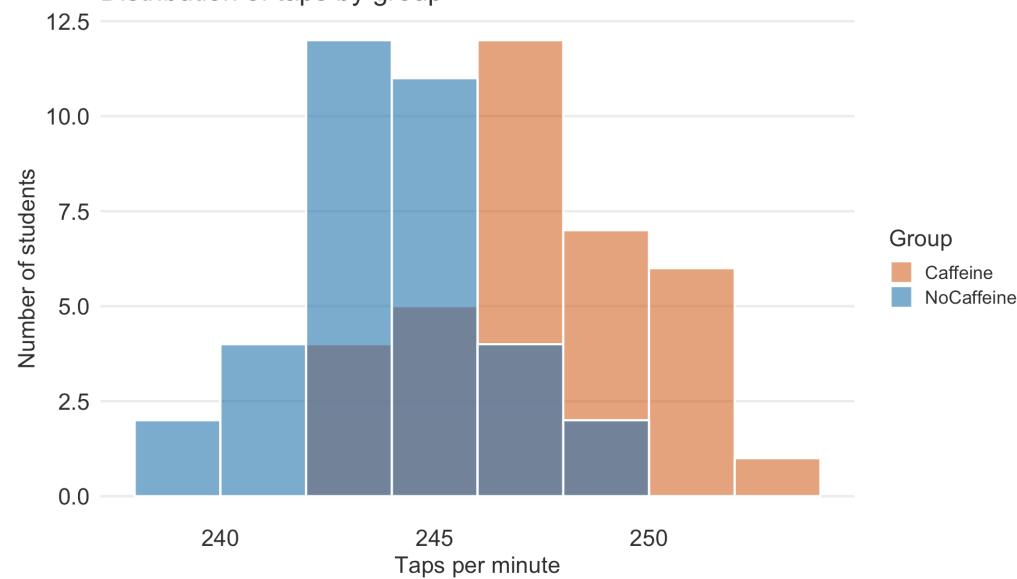
Visualizing independent samples

Finger tapping by group



Box plots with individual points

Distribution of taps by group



Overlapping histograms

Note: We cannot calculate 35 paired differences - these are different people!



Notation for two independent samples

Population Parameters vs. Sample Statistics

Population

- Group 1 mean: μ_1
- Group 2 mean: μ_2
- Difference: $\mu_1 - \mu_2$
- Group 1 SD: σ_1
- Group 2 SD: σ_2

Sample

- Group 1 mean: \bar{x}_1
- Group 2 mean: \bar{x}_2
- Difference: $\bar{x}_1 - \bar{x}_2$
- Group 1 SD: s_1
- Group 2 SD: s_2

Key difference from paired data:

- Two separate groups with potentially different SDs
- Cannot reduce to a single set of differences
- Need a different standard error formula!



Hypothesis test for two independent samples (1/2)

Caffeine example: Does caffeine increase finger tapping speed?

Null hypothesis (no effect):

$$H_0 : \mu_{\text{caffeine}} = \mu_{\text{control}} \iff \mu_{\text{caffeine}} - \mu_{\text{control}} = 0$$

Alternative hypothesis (caffeine increases tapping):

$$H_A : \mu_{\text{caffeine}} > \mu_{\text{control}} \iff \mu_{\text{caffeine}} - \mu_{\text{control}} > 0$$

Why one-sided? We specifically predicted caffeine would *increase* tapping, not just “make it different”



Hypothesis test for two independent samples (2/2)

General form

For two independent samples, we always test $H_0 : \mu_1 - \mu_2 = 0$ against ONE of:

- $H_A : \mu_1 - \mu_2 \neq 0$ (two-sided)
- $H_A : \mu_1 - \mu_2 > 0$ (one-sided upper)
- $H_A : \mu_1 - \mu_2 < 0$ (one-sided lower)



Example: Does caffeine increase finger tapping speed?

Step 1: State hypotheses

$$H_0 : \mu_1 - \mu_2 = 0$$

$$H_A : \mu_1 - \mu_2 > 0$$

Where: μ_1 = mean for Caffeine group, μ_2 = mean for Control group

Step 2: Set significance level

- Use $\alpha = 0.05$



Example: Does caffeine increase finger tapping speed?

Step 3: Calculate test statistic

The test statistic for two independent samples:

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - 0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

Notice the different SE formula!

- For paired: $SE = \frac{s_d}{\sqrt{n}}$
- For independent: $SE = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$

Degrees of freedom: By default, `t.test()` uses the Welch two-sample t-test (does not assume equal variances), so `df` is computed with the Welch–Satterthwaite formula and can be non-integer.



Check the assumptions

Step 3: Check the assumptions

- The assumptions to run a hypothesis test on a sample are:
 - **Independent observations:** Each observation from both samples is independent from all other observations
 - **Approximately normal sample or big n:** the distribution of *each sample* should be approximately normal, or the sample size of *each sample* should be at least 30
- In our example, we would check the assumptions with a statement:
 - The observations are independent from each other. Each caffeine group (aka sample) has 35 individuals. Thus, we can use CLT to approximate the sampling distribution for each sample.



Running a two-sample t-test in R

Steps 4 and 5: Calculate test statistic and p-value

Using formula notation:

```
1 # Test using formula: outcome ~ group
2 t.test(Taps ~ Group,
3        data = CaffTaps,
4        alternative = "greater")
```

Welch Two Sample t-test

```
data: Taps by Group
t = 6.0867, df = 67.002, p-value = 3.133e-08
alternative hypothesis: true difference in means between group Caffeine and group NoCaffeine is greater than 0
95 percent confidence interval:
 2.613502      Inf
sample estimates:
 mean in group Caffeine mean in group NoCaffeine
          248.1143           244.5143
```

Note: R automatically determines which group is “group 1” alphabetically (Caffeine comes before NoCaffeine)



Using rstatix

```
1 library(rstatix)
2
3 t_test(data = CaffTaps,
4         Taps ~ Group,
5         detailed = TRUE)
# A tibble: 1 × 15
#>   estimate estimate1 estimate2 .y. group1 group2    n1    n2 statistic      p
#>   <dbl>     <dbl>     <dbl> <chr> <chr> <chr> <int> <int>     <dbl>    <dbl>
#> 1 3.60      248.     245. Taps  Caffe... NoCaf... 35     35      6.09 6.27e-8
#> # i 5 more variables: df <dbl>, conf.low <dbl>, conf.high <dbl>, method <chr>,
#> # alternative <chr>
```



Interpreting the two-sample t-test

Steps 6: make a conclusion

```
1 # Save and tidy results
2 caff_test <- t.test(Taps ~ Group,
3                      data = CaffTaps,
4                      alternative = "greater")
5
6 tidy(caff_test,
7       conf.int = TRUE) %>%
8     gt() %>%
9     tab_options(table.font.size = 30)
```

estimate	estimate1	estimate2	statistic	p.value	parameter	conf.low	conf.high	method	alternative
3.6	248.1143	244.5143	6.086677	3.132816e-08	67.00222	2.613502	Inf	Welch Two Sample t-test	greater

What we see:

- Mean difference: $\bar{x}_1 - \bar{x}_2 = 3.6$ taps/min
- t-statistic = 6.09
- p-value = < 0.001
- One-sided 95% CI: (2.6, Inf)

Conclusion: There is strong evidence that caffeine increases finger tapping speed ($p < 0.001$).



Complete conclusion for two-sample test

Best Practice: Report Both Groups

For two-sample tests, report:

1. Means and SDs for both groups
2. Difference in means
3. Test statistic and df
4. P-value
5. Confidence interval for difference

Our conclusion:

Mean (SD) tapping speed among students who consumed caffeine was 248.1 taps/min compared to 244.5 taps/min among the control group. Taps/min was significantly higher in the caffeine group ($t = 6.09$, $df = 67.00$, $p < 0.001$, one-sided test). On average, students who consumed caffeine tapped 3.6 taps/min faster than those in the control group (95% CI: at least 2.6 taps/min faster)



Choosing the Right Test



One rule to remember

The decision rule

Ask one question:

Can I calculate a meaningful difference for each individual (or pair)?

- Yes → Paired t-test
- No → Two-sample t-test

This is determined by **study design**, not by the values in the data.

Key idea:

Pairing is about how the data were collected, not how they are stored or analyzed.



Examples: Paired vs. Independent

Scenario	Data Structure	Test
Blood pressure before and after medication	Same people, two measurements	Paired t-test
Weight loss: Diet A vs. Diet B	Different people in each group	Two-sample t-test
Left eye vs. right eye vision	Same people	Paired t-test
Test scores: Men vs. Women	Different people	Two-sample t-test
Cholesterol: Twin 1 vs. Twin 2	Matched pairs	Paired t-test
Pain score: Treatment vs. Placebo (RCT)	Different people	Two-sample t-test

Shortcut:

If you can draw a line connecting two measurements, it's paired.



Common mistakes to avoid

Mistakes we see all the time

1. Using a paired test for independent data

- Example: Testing men vs. women with `paired = TRUE`
- Why it's wrong: R pairs observations by row order, not by any real relationship
- Result: Meaningless inference

2. Treating paired data as independent

- Example: Analyzing before/after as two separate groups
- Why it's wrong: Ignores within-person comparison
- Result: Larger SE and reduced power

Bottom line:

You don't get to choose the test based on convenience — the design chooses it for you.



Checking your work: Does it make sense?

After running your test, ask:

1. Does the sample size match my study design?

- Paired: Should see $n =$ number of pairs
- Independent: Should see $n_1 + n_2 =$ total people

2. Do the means make sense?

- Check that group means match your data summaries

3. Is the SE reasonable?

- Paired tests usually have smaller SE (less variability)
- Independent tests have larger SE (more variability)

4. Does the conclusion answer the research question?

- Make sure you're interpreting the right comparison



Summary: Paired vs. Independent

Paired Data:

- Same people or matched pairs
- Calculate differences first
- Use: `t.test(differences, mu = 0)` or `t.test(x, y, paired = TRUE)`
- SE: $\frac{s_d}{\sqrt{n}}$
- df: $n - 1$

Independent Samples:

- Different people in each group
- Compare two means directly
- Use: `t.test(outcome ~ group, data = ...)`
- SE: $\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$
- df: $n_1 + n_2 - 2$ (approximately)



Key takeaways

1. Study design determines analysis

- Paired or independent is about how data were collected
- Cannot change after the fact!

2. Paired tests are more powerful

- When appropriate, pairing reduces variability
- Smaller SE → easier to detect true effects

3. Both tests follow same logic

- State hypotheses
- Calculate test statistic
- Find p-value
- Make conclusion

4. Always check assumptions

- Independence (within or between pairs)
- Approximate normality (or large n)



Looking ahead

Next time:

- Midterm review

