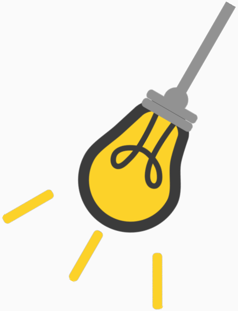
浙江大学爱丁堡大学联合学院
ZJU-UoE Institute

The *t*-test: practical applications and variants

Nicola Romanò - nicola.romano@ed.ac.uk

At the end of this lecture, you should be able to:

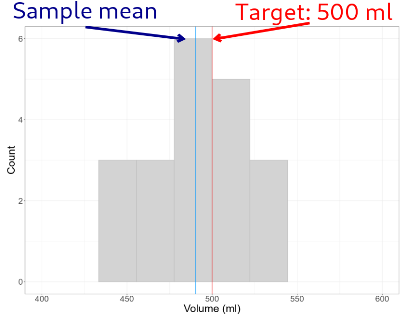
- Describe variants of the t-test
    - 1-tailed vs 2-tailed
    - 1- vs 2-sample
    - Paired vs unpaired
- Choose an appropriate type of t-test for a given problem
- Describe and use non-parametric alternatives to the t-test



Our earlier example

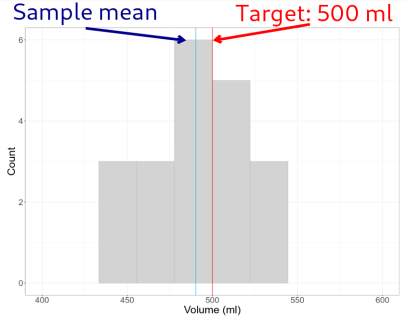
Is the factory filling each bottle with enough Guinness?

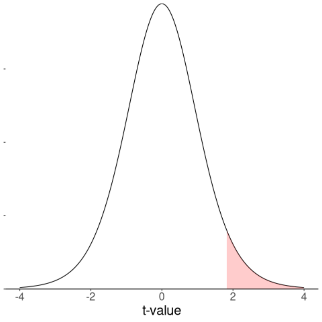
Is the factory filling each bottle with enough Guinness?



Our earlier example

Is the factory filling each bottle with enough Guinness?





Variants of the *t*-test

Variants of the *t*-test

1-tailed vs 2-tailed

One-tailed vs two-tailed t-test

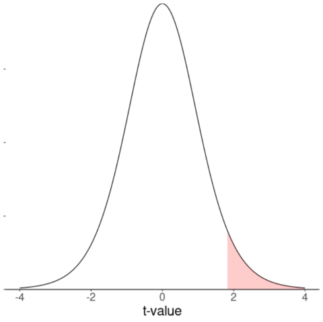
Compare the following two questions:

- Is the factory filling each bottle with **enough** Guinness?
- Is the factory filling each bottle with a volume **different** from 500ml?

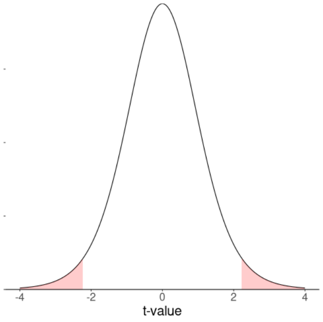
One-tailed vs two-tailed t-test

Compare the following two questions:

- Is the factory filling each bottle with **enough** Guinness?
- Is the factory filling each bottle with a volume **different** from 500ml?

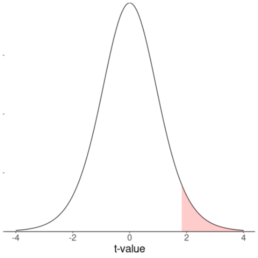


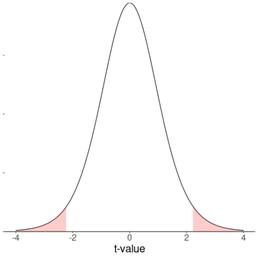
1-tailed - Area of interest is only on one side of the distribution (ex. 5%)



2-tailed - Area of interest is on both sides of the distribution (ex. 2.5% on each side)

Which one should I use?





One-tailed

- If we look for outcomes in one direction only, e.g. A > B (or A < B)
- H_0: A is not greater (or smaller) than B
- H_1: A is greater (or smaller) than B

Two-tailed

- If we look for outcomes in both directions, e.g. A $\neq$ B
- H_0: A is equal to B
- H_1: A is not equal to B

Critical values for 1-tailed and 2-tailed tests

For the same significance level, the critical values for 1-tailed and 2-tailed tests are different!

One tailed Two tailed

d.f.	$t_{.100}$	$t_{.050}$*	$t_{.025}$**	$t_{.010}$	$t_{.005}$	d.f.
1	3.078	6.314	12.706	31.821	63.657	1
2	1.886	2.920	4.303	6.965	9.925	2
3	1.638	2.353	3.182	4.541	5.841	3
4	1.533	2.132	2.776	3.747	4.604	4
5	1.476	2.015	2.571	3.365	4.032	5
6	1.440	1.943	2.447	3.143	3.707	6
7	1.415	1.895	2.365	2.998	3.499	7
8	1.397	1.860	2.306	2.896	3.355	8
9	1.383	1.833	2.262	2.821	3.250	9
10	1.372	1.812	2.228	2.764	3.169	10

Variants of the *t*-test

1-sample vs 2-sample

1-sample vs 2-sample t-test

Compare the following two questions

Is the factory filling each bottle with a volume different from 500ml?

Is the factory in Dublin filling each bottle with a volume different from the factory in Glasgow?

1-sample vs 2-sample t-test

Compare the following two questions

Is the factory filling each bottle with a volume different from 500ml?

Is the factory in Dublin filling each bottle with a volume different from the factory in Glasgow?

difference from the ref. value

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$

sample standard deviation (estim, of the population σ)

standard error of the mean

1-sample vs 2-sample t-test

Compare the following two questions

Is the factory filling each bottle with a volume different from 500ml?

difference from the ref. value

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$

sample standard deviation (estim, of the population σ)

standard error of the mean

Is the factory in Dublin filling each bottle with a volume different from the factory in Glasgow?

diff. of means

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

combined error

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

Equal variances
Student's t-test

Unequal variances
Welch's t-test

Example:

Sample 1: 4, 6, 8, 10

Sample 2: 1, 5, 3, 4

H_0: $\mu_1 = \mu_2$

H_A: $\mu_1 \neq \mu_2$

What should we do to test this hypothesis?

Variants of the *t*-test

Paired vs unpaired

Paired vs unpaired t-test

Paired *t*-test: Same subjects are measured at different times or under different conditions.

Unpaired *t*-test: Different subjects are measured at different times or under different conditions.

Paired vs unpaired t-test

Paired *t*-test: Same subjects are measured at different times or under different conditions.

Unpaired *t*-test: Different subjects are measured at different times or under different conditions.

Paired *t*-test example You want to study the effect of a drug on the blood pressure of 10 patients. You measure the blood pressure of each patient before and after the administration of the drug.

Paired t-test - calculating t statistics

The paired t-test is effectively a 1-sample t-test on the differences between the two measurements (with reference value = 0).

avg. difference between paired measurements

$$t = \frac{\bar{d}}{\left(\frac{s_d}{\sqrt{n}}\right)}$$

standard deviation of the differences

t-test variants in R

Use the `t.test()` function in R to perform a *t*-test.

```
t.test(x, y = NULL,
alternative = c("two.sided", "less", "greater"),
mu = 0, paired = FALSE, var.equal = FALSE,
conf.level = 0.95, ...)
```

t-test variants in R

Use the `t.test()` function in R to perform a *t*-test.

```
t.test(x, y = NULL,
alternative = c("two.sided", "less", "greater"),
mu = 0, paired = FALSE, var.equal = FALSE,
conf.level = 0.95, ...)
```

1 sample

`mu` = reference value

2 sample

Must provide two vectors of values as `x` and `y`

For Welch's t-test`var.equal` = `FALSE` ()

1-tailed

`alternative` = `"greater"` or `"less"`

2-tailed

`alternative` = `"two.sided"` (default)

Paired

`paired` = `TRUE`

Unpaired

`paired` = `FALSE` (default)

How to choose the appropriate type of t-test?

State the hypothesis. What is H_0 and what is H_A?

1. Are the assumptions met?

How to choose the appropriate type of *t*-test?

State the hypothesis. What is H_0 and what is H_A?

1. Are the assumptions met?
    - Samples must be **independent** (not applicable to paired *t*-test)

How to choose the appropriate type of *t*-test?

State the hypothesis. What is H_0 and what is H_A?

1. Are the assumptions met?
    - Samples must be **independent** (not applicable to paired *t*-test)
    - Samples must be **random**

How to choose the appropriate type of *t*-test?

State the hypothesis. What is H_0 and what is H_A?

1. Are the assumptions met?
    - Samples must be **independent** (not applicable to paired *t*-test)
    - Samples must be **random**
    - **Normality** of the sample(s)
      For paired *t*-test, the **differences** must be normally distributed.

How to choose the appropriate type of *t*-test?

State the hypothesis. What is H_0 and what is H_A?

1. Are the assumptions met?
    - Samples must be **independent** (not applicable to paired *t*-test)
    - Samples must be **random**
    - **Normality** of the sample(s)
      For paired *t*-test, the **differences** must be normally distributed.
    - **Homogeneity of variances** (not applicable to one-sample *t*-test)

How to choose the appropriate type of *t*-test?

State the hypothesis. What is H_0 and what is H_A?

1. Are the assumptions met?
    - Samples must be **independent** (not applicable to paired *t*-test)
    - Samples must be **random**
    - **Normality** of the sample(s)
      For paired *t*-test, the **differences** must be normally distributed.
    - **Homogeneity of variances** (not applicable to one-sample *t*-test)
2. What direction are you looking at? 1-tailed or 2-tailed?

How to choose the appropriate type of *t*-test?

State the hypothesis. What is H_0 and what is H_A?

1. Are the assumptions met?
    - Samples must be **independent** (not applicable to paired *t*-test)
    - Samples must be **random**
    - **Normality** of the sample(s)
      For paired *t*-test, the **differences** must be normally distributed.
    - **Homogeneity of variances** (not applicable to one-sample *t*-test)
2. What direction are you looking at? 1-tailed or 2-tailed?
3. Are you comparing 1-sample or 2-samples?

How to choose the appropriate type of *t*-test?

State the hypothesis. What is H_0 and what is H_A?

1. Are the assumptions met?
    - Samples must be **independent** (not applicable to paired *t*-test)
    - Samples must be **random**
    - **Normality** of the sample(s)
      For paired *t*-test, the **differences** must be normally distributed.
    - **Homogeneity of variances** (not applicable to one-sample *t*-test)
2. What direction are you looking at? 1-tailed or 2-tailed?
3. Are you comparing 1-sample or 2-samples?
4. Are the observations paired or unpaired?

What to do when assumptions are not met?

- **Normality of the sample(s)**
  - If the deviation from normality is small, the t-test is robust to it, and we can still use it.
  - If the deviation from normality is large, we can try transforming the data and see if the data become more normal. Examples of transformations include log, square root, inverse, etc.
  - If the data cannot be transformed, we can use a **non-parametric** test instead (see next slides).

What to do when assumptions are not met?

- **Normality of the sample(s)**
  - If the deviation from normality is small, the t-test is robust to it, and we can still use it.
  - If the deviation from normality is large, we can try transforming the data and see if the data become more normal. Examples of transformations include log, square root, inverse, etc.
  - If the data cannot be transformed, we can use a **non-parametric** test instead (see next slides).
- **Homogeneity of variances**
  - We can use Welch's t-test instead of Student's t-test.

A **non-parametric** (or distribution-free) test is a statistical test that does not assume that the data follow a particular distribution.
As alternatives to the t-test, we can use the following non-parametric tests:

- **Wilcoxon signed-rank test** - for paired data or one-sample data
- **Mann-Whitney U test** - for unpaired data

Non-parametric tests

A **non-parametric** (or distribution-free) test is a statistical test that does not assume that the data follow a particular distribution.
As alternatives to the t-test, we can use the following non-parametric tests:

- **Wilcoxon signed-rank test** - for paired data or one-sample data
- **Mann-Whitney U test** - for unpaired data

The idea of these two tests is very similar, and in R you can use the same function `wilcox.test()` to perform both tests.

Wilcoxon: `wilcox.test(x, y, paired = TRUE, ...)`
Wilcoxon, one-sample: `wilcox.test(x, mu = 0, ...)`
Mann-Whitney: `wilcox.test(x, y, paired = FALSE, ...)`

The idea behind non-parametric tests

Wilcoxon and Mann-Whitney tests are based on the **rank** of the observations.
They do not assume a particular distribution of the data, but they do assume that the observations are **independent** and that the **distribution of the observations is the same** (or similar) in the two groups.
They compare medians rather than means.

The idea behind non-parametric tests

Wilcoxon and Mann-Whitney tests are based on the **rank** of the observations.
They do not assume a particular distribution of the data, but they do assume that the observations are **independent** and that the **distribution of the observations is the same** (or similar) in the two groups.
They compare medians rather than means.

Note: they have **less statistical power** than the *t*-test, so they are less likely to detect a difference when there is one. However, they are more robust to deviations from normality and to outliers.

How non-parametric tests work

Mann-Whitney test:

- Combine all observations and rank them from smallest to largest
- Sum the ranks in each group
- Calculate the test statistic U and compare it to the critical value (we're not going to do this by hand, but R will do it for you!)

How non-parametric tests work

Mann-Whitney test:

- Combine all observations and rank them from smallest to largest
- Sum the ranks in each group
- Calculate the test statistic U and compare it to the critical value (we're not going to do this by hand, but R will do it for you!)

Wilcoxon test:

- Calculate the difference between the two observations in each pair
- Rank the differences from smallest to largest
- Sum the ranks of the positive differences and the ranks of the negative differences
- Calculate the test statistic W and compare it to the critical value

Learning objectives

You should now be able to:

- Describe variants of the t-test
  - 1-tailed vs 2-tailed
  - 1- vs 2-sample
  - Paired vs unpaired
- Choose an appropriate type of t-test for a given problem
- Describe and use non-parametric alternatives to the t-test