Biostatistics I: Hypothesis testing

Continuous data: Two-sample (independent) tests

Eleni-Rosalina Andrinopoulou

Department of Biostatistics, Erasmus Medical Center

e.andrinopoulou@erasmusmc.nl

@erandrinopoulou

Erasmus MC
University Medical Center Rotterdam

In this Section

- Two-sample t-test (independent samples)
- Two-sample Wilcoxon rank sum test (independent samples)
- Examples

Two-sample t-test (independent samples): Theory

Assumptions

- The dependent variables must be continuous
- The observations are independent
- The dependent variables are approximately normally distributed
- The dependent variables do not contain any outliers

Two-sample t-test (independent samples): Theory

Scenario

Is the mean BMI of the students from group 1 different from the mean BMI of the students of group 2?

Connection with linear regression

$y_i = \beta_0 + \beta_1 x_i + \epsilon_i$, where x_i indicates whether a patient was in group 1 or in group 2

$H_0 : \beta_1 = 0$
$H_1 : \beta_1 \neq 0$

Two-sample t-test (independent samples): Theory

Scenario

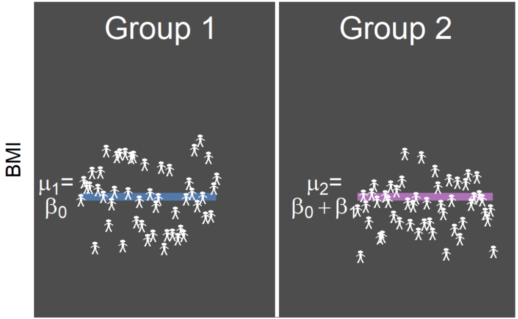
Is the mean BMI of the students from group 1 different from the mean BMI of the students of group 2?

Alternatively

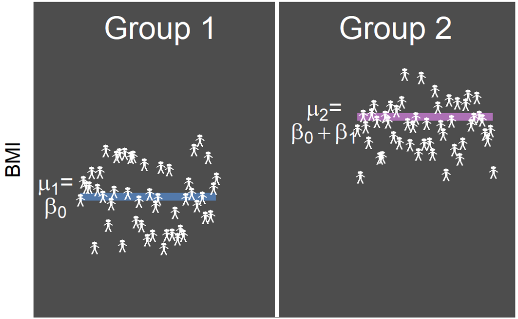
$H_0 : \mu_1 = \mu_2$
$H_1 : \mu_1 \neq \mu_2$

Two-sample t-test (independent samples): Theory

Null hypothesis



Alternative hypothesis



$$\mu_1 = \mu_2$$
$$\beta_1 = 0$$

$$\mu_1 \neq \mu_2$$
$$\beta_1 \neq 0$$

Two-sample t-test (independent samples): Theory

Test statistic

If the variances of the two groups are equal we use the t-statistic:

$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{sd^2(x)}{n_1} + \frac{sd^2(x)}{n_2}}}$, where

$$sd^2(x) = \frac{\sum (x_1 - \bar{x}_1)^2 + \sum (x_2 - \bar{x}_2)^2}{n_1 + n_2 - 2}$$

If the variances of the two groups being compared are *not* equal we use the Welch t-statistic:

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{sd^2(x_1)}{n_1} + \frac{sd^2(x_2)}{n_2}}}$$

- Sample mean of group 1, 2: $\bar{x}_1, \bar{x}_2$
- Standard deviation of group 1, 2: $sd(x_1), sd(x_2)$
- Number of subjects in group 1, 2: n_1, n_2

Two-sample t-test (independent samples): Theory

Sampling distribution

- t-distribution with:
  - If the variance of the two groups are equal: $df = n_1 + n_2 - 2$
  - If the variance of the two groups are *not* equal: $df = \frac{\left[\frac{sd^2(x_1)}{n_1} + \frac{sd^2(x_2)}{n_2}\right]^2}{\frac{[sd^2(x_1)/n_1]^2}{n_1 - 1} + \frac{[sd^2(x_2)/n_2]^2}{n_2 - 1}}$
- Critical values and p-value

Type I error

- Normally $\alpha = 0.05$

Draw conclusions

- Compare test statistic (t) with the critical values or the p-value with α

Two-sample t-test (independent samples): Application

Scenario

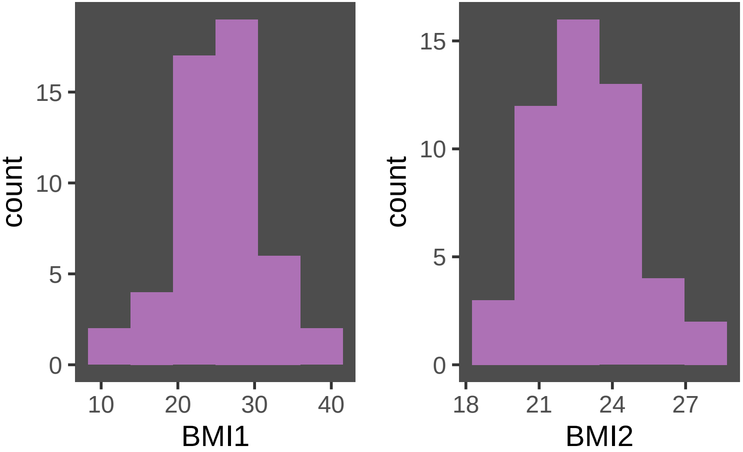
Is the mean BMI of the students from group 1 different from the mean BMI of the students of group 2?

Hypothesis

$H_0 : \mu_1 = \mu_2$
$H_1 : \mu_1 \neq \mu_2$

Two-sample t-test (independent samples): Application

Collect and visualize data



Two-sample t-test (independent samples): Application

Hypothesis

$H_0 : \mu_1 = \mu_2$
$H_1 : \mu_1 \neq \mu_2$

Test statistic

Let's assume that:

- Sample mean of group 1: $\bar{x}_1 = 24$
- Sample mean of group 2: $\bar{x}_2 = 23$
- Standard deviation of group 1: $sd(x_1) = 6$
- Standard deviation of group 2: $sd(x_2) = 2$
- Number of subjects in group 1: $n_1 = 50$
- Number of subjects in group 2: $n_2 = 50$

Two-sample t-test (independent samples): Application

We first test homogeneity of variances:

$H_0 : \frac{variance_1}{variance_2} = 1$

$H_1 : \frac{variance_1}{variance_2} \neq 1$

F statistic: $\frac{highest\ variance}{lowest\ variance} = 9$

DF: $n_1 - 1 = 50 - 1 = 49$, $n_2 - 1 = 50 - 1 = 49$

If $\alpha = 0.05$ and two-tailed test, get critical value in R:

```
qf(p = 0.05, df1 = 49, df2 = 49, lower.tail = FALSE)
[1] 1.607289
```

$\frac{highest\ variance}{lowest\ variance} = 9 > 1.61 \Rightarrow H_0$ is rejected

Two-sample t-test (independent samples): Application

Test statistic

Not equal variances:

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{sd^2(x_1)}{n_1} + \frac{sd^2(x_2)}{n_2}}} = \frac{24 - 23}{\sqrt{\frac{36}{50} + \frac{4}{50}}} = 1.12$$

Degrees of freedom

$$df = \frac{\left[\frac{sd^2(x_1)}{n_1} + \frac{sd^2(x_2)}{n_2}\right]^2}{\frac{[sd^2(x_1)/n_1]^2}{n_1 - 1} + \frac{[sd^2(x_2)/n_2]^2}{n_2 - 1}} = \frac{\left(\frac{36}{50} + \frac{4}{50}\right)^2}{\frac{(36/50)^2}{49} + \frac{(4/50)^2}{49}} = 59.76$$

Type I error

$\alpha = 0.05$

Two-sample t-test (independent samples): Application

Critical values

Using R we get the critical values from the *t*-distribution:

critical value$_{\alpha/2}$ = critical value$_{0.05/2}$

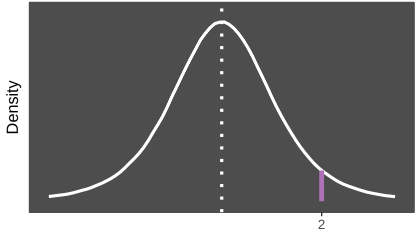
```
qt(p = 0.05/2, 59.76, lower.tail = FALSE)
```

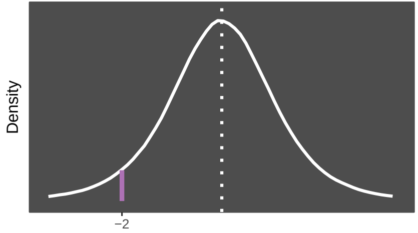
```
[1] 2.000463
```

-critical value$_{\alpha/2}$ = -critical value$_{0.05/2}$

```
qt(p = 0.05/2, 59.76, lower.tail = TRUE)
```

```
[1] -2.000463
```





Two-sample t-test (independent samples): Application

Draw conclusions

We reject the H_0 if:

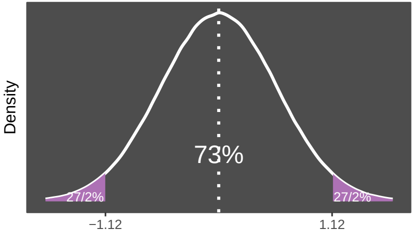
- t > critical value$_{\alpha/2}$ or t < - critical value$_{\alpha/2}$

We have $1.12 < 2 \Rightarrow$ we do not reject the H_0

Using R we obtain the p-value from the t-distribution:

```r
2 * pt(q = 1.12, df = 59.76,
       lower.tail = FALSE)
```

```
[1] 0.2671949
```



Two-sample Wilcoxon rank sum test: Theory

Assumptions

- Population distribution is symmetric
- The observations are independent of one another

Two-sample Wilcoxon rank sum test: Theory

Scenario

Is the distribution of the score values of the students in group 1 different from the distribution of the score values of the students in group 2?

Connection with linear regression

$rank(y_i) = \beta_0 + \beta_1 x_i + \epsilon_i$, where x_i indicates whether a patient was in group 1 or in group 2

$H_0 : \beta_1 = 0$
$H_1 : \beta_1 \neq 0$

Alternatively

H_0 : the distributions of both populations are equal
H_1 : the distributions are not equal

Two-sample Wilcoxon rank sum test: Theory

Test statistic

- Calculate the ranks for the two groups (r_1 and r_2)
- Obtain the sum of those ranks $R_1 = \sum r_1$, $R_2 = \sum r_2$
- Calculate $U_1 = n_1 n_2 + \frac{n_1(n_1+1)}{2} - R_1$ and $U_2 = n_1 n_2 + \frac{n_2(n_2+1)}{2} - R_2$
- The test statistic (U) is the minimum of U_1 and U_2

If **one-tailed**: use either U_1 or U_2 for the test statistic (U) depending on the direction of the alternative hypothesis

Two-sample Wilcoxon rank sum test: Theory

Sampling distribution

For large sample size: we can use the normal approximation, that is, U is normally distributed

$$\mu_U = \frac{n_1 n_2}{2}$$

$$\sigma_U = \sqrt{\frac{n_1 n_2 (n_2 + n_1 + 1)}{12}}$$

If there are ties in ranks, we should use:

$$\sigma_U = \sqrt{\frac{n_1 n_2}{12}\left[(n+1) - \sum_{i=1}^{K} \frac{t_i^2 - t_i}{n(n-1)}\right]}$$

where $n = n_1 + n_2$ and t_i is the number of subjects sharing the rank i. K is the number of ranks

$$z = \frac{|min(U_1, U_2) - \mu_U| - 1/2}{\sigma_U}$$

Two-sample Wilcoxon rank sum test: Theory

Sampling distribution

For small sample size: we can use the exact distribution

- Get critical values and p-value

Type I error

- Normally α = 0.05

Draw conclusions

- Compare test statistic with the critical values or the p-value with α

Two-sample Wilcoxon rank sum test: Application

Scenario

Is the distribution of the score values of the students in group 1 different from the distribution of the score values of the students in group 2?

Hypothesis

H_0 : the distributions of both populations are equal
H_1 : the distributions are not equal

Two-sample Wilcoxon rank sum test: Application

Collect and visualize data

variable	value	rank
x	9.75508	1
x	11.10491	4
x	10.69730	3
y	14.71926	5
y	15.79611	6
y	10.15486	2

Test statistic

$R_1 = 1 + 4 + 3 = 8$

$U_1 = n_1 n_2 + \frac{n_1(n_1+1)}{2} - R_1 = 7$

$R_2 = 5 + 6 + 2 = 13$

$U_2 = n_1 n_2 + \frac{n_2(n_2+1)}{2} - R_2 = 2$

Type I error

$\alpha = 0.05$

Hypothesis

H_0 : the distributions of both populations are equal

H_1 : the distributions are not equal

Two-sample Wilcoxon rank sum test: Application

Critical values

Using `R` we get the critical values from the exact distribution:
low critical value$_{\alpha/2}$ = low critical value$_{0.05/2}$

```
qwilcox(p = 0.05/2, m = 3, n = 3, lower.tail = TRUE)
```

```
[1] 0
```

high critical value$_{\alpha/2}$ = high critical value$_{0.05/2}$

```
qwilcox(p = 0.05/2, m = 3, n = 3, lower.tail = FALSE)
```

```
[1] 9
```

Two-sample Wilcoxon rank sum test: Application

Draw conclusions

We reject the H_0 if:

- $U_1 >$ high critical value$_{\alpha/2}$ and $U_2 <$ low critical value$_{\alpha/2}$

We have $7 < 9$ and $2 > 0 \Rightarrow$ we do *not* reject the H_0

Two-sample Wilcoxon rank sum test: Application

Draw conclusions

Using R we obtain the p-value from the exact distribution:

$p - value = 2 * Pr(U <= 2):$

```r
2 * pwilcox(q = 2, m = 3, n = 3, lower.tail = TRUE)
```

```
[1] 0.4
```

or

$p - value = 2 * Pr(U >= 7) = 2 * (1 - Pr(U < 7)):$

```r
2 * (1 - pwilcox(q = 7 - 1, m = 3, n = 3, lower.tail = TRUE))
```

```
[1] 0.4
```

```r
2 * pwilcox(q = 7 - 1, m = 3, n = 3, lower.tail = FALSE)
```

```
[1] 0.4
```