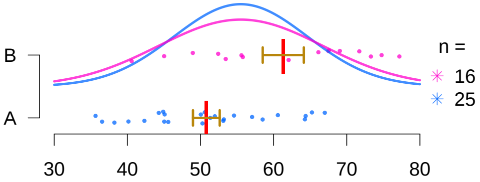
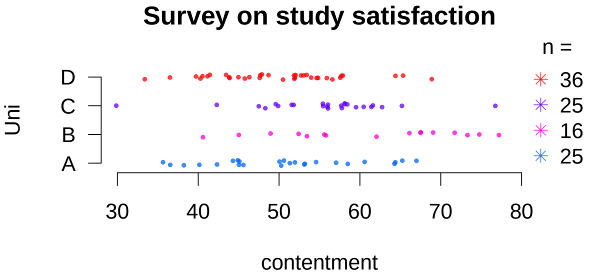
Surrounding the two-sample t-test



For comments please email: M. Messer, michael.messer@tuwien.ac.at

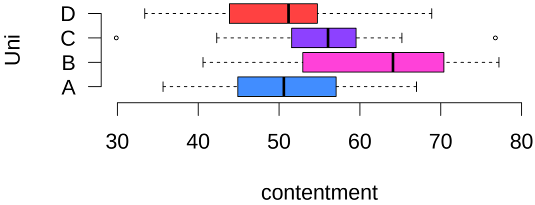
Motivation

At four universities students of a certain study program were interviewed regarding the level of their satisfaction with the study situation. An extensive survey had to be filled out. Subsequently, for every respondent a global value of 'contentment' was evaluated.



Motivation

At four universities students of a certain study program were interviewed regarding the level of their satisfaction with the study situation. An extensive survey had to be filled out. Subsequently, for every respondent a global value of 'contentment' was evaluated.

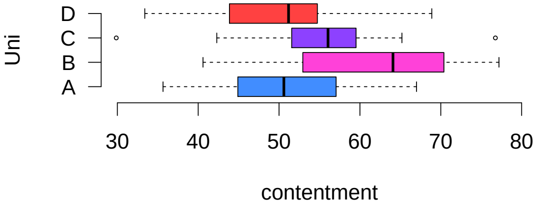


Question?

- The median value of contentment of the respondents from Uni A was about the same as the median of which Uni?

Motivation

At four universities students of a certain study program were interviewed regarding the level of their satisfaction with the study situation. An extensive survey had to be filled out. Subsequently, for every respondent a global value of 'contentment' was evaluated.

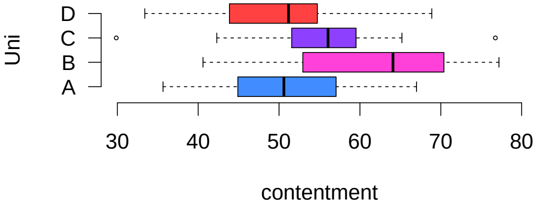


Question?

- The median value of contentment of the respondents from Uni *A* was about the same as the median of which Uni? *D*

Motivation

At four universities students of a certain study program were interviewed regarding the level of their satisfaction with the study situation. An extensive survey had to be filled out. Subsequently, for every respondent a global value of ’contentment’ was evaluated.

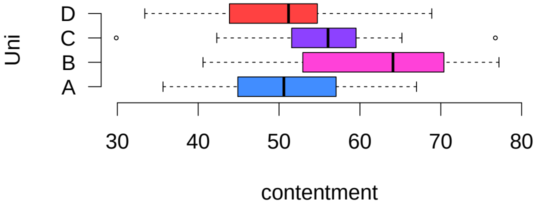


Question?

- The median value of contentment of the respondents from Uni A was about the same as the median of which Uni? D
- The third quartile of the respondants of Uni C was about?

Motivation

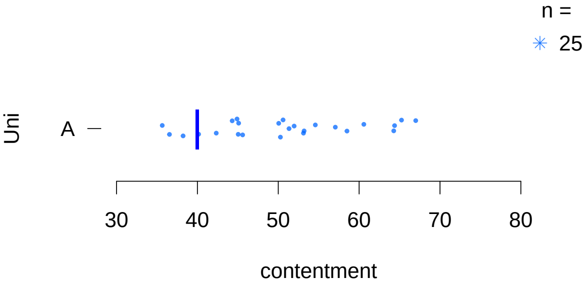
At four universities students of a certain study program were interviewed regarding the level of their satisfaction with the study situation. An extensive survey had to be filled out. Subsequently, for every respondent a global value of 'contentment' was evaluated.



Question?

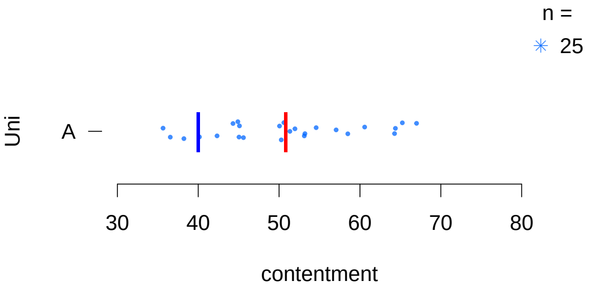
- The median value of contentment of the respondents from Uni *A* was about the same as the median of which Uni? *D*
- The third quartile of the respondants of Uni *C* was about? 60

Reminder: One-sample situation



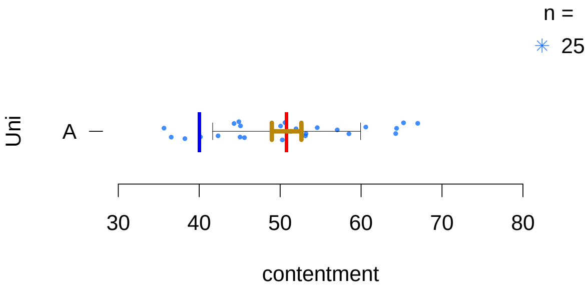
- $H_0 : \mu = \mu_0$

Reminder: One-sample situation



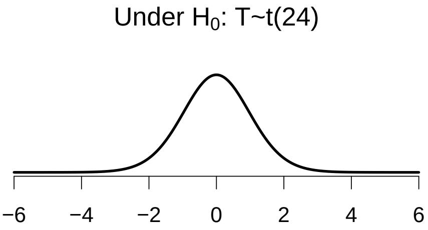
- $H_0 : \mu = \mu_0$
- Mean $\bar{x}$ far away from μ_0?

Reminder: One-sample situation



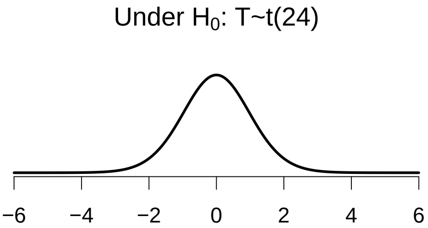
- $H_0 : \mu = \mu_0$
- Mean $\bar{x}$ far away from μ_0?
- Answer: Yes! Many standard errors $sem = s/\sqrt{n}$ away (approx. $6 \cdot sem$)

Reminder: One-sample situation



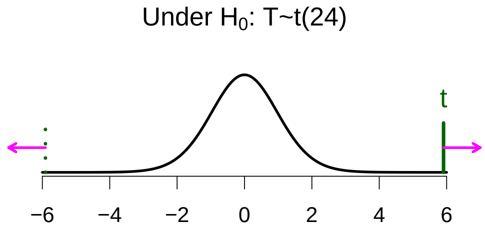
- $H_0 : \mu = \mu_0$
- Mean $\bar{x}$ far away from μ_0?
- Answer: Yes! Many standard errors $sem = s/\sqrt{n}$ away (approx. $6 \cdot sem$)
- In the model $X_1, \ldots, X_n$ i.i.d. RVs with $X_1 \sim N(\mu, \sigma^2)$ it holds under H_0 $T = \frac{\bar{X} - \mu_0}{SEM} \sim t(n-1)$

Reminder: One-sample situation



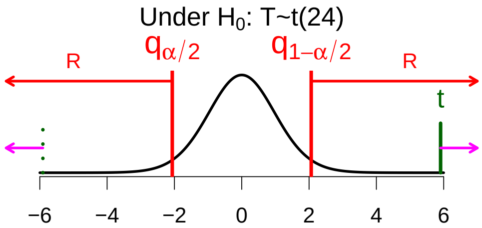
- $H_0 : \mu = \mu_0$
- Mean $\bar{x}$ far away from μ_0?
- Answer: Yes! Many standard errors $sem = s/\sqrt{n}$ away (approx. $6 \cdot sem$)
- In the model $X_1, \ldots, X_n$ i.i.d. RVs with $X_1 \sim N(\mu, \sigma^2)$ it holds under H_0 $T = \frac{\bar{X} - \mu_0}{SEM} \sim t(n-1)$
- The estimated standard deviation of $\bar{X}$ is $1 \cdot SEM$ (not $6 \cdot SEM$)

Reminder: One-sample situation



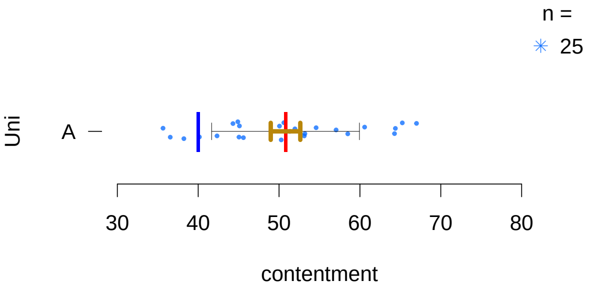
- $H_0 : \mu = \mu_0$
- Mean $\bar{x}$ far away from μ_0?
- Answer: Yes! Many standard errors $sem = s/\sqrt{n}$ away (approx. $6 \cdot sem$)
- In the model $X_1, \ldots, X_n$ i.i.d. RVs with $X_1 \sim N(\mu, \sigma^2)$ it holds under H_0
  $T = \frac{\bar{X} - \mu_0}{SEM} \sim t(n-1)$
- The estimated standard deviation of $\bar{X}$ is $1 \cdot SEM$ (not $6 \cdot SEM$)
- Judge the evaluated data $t = \frac{\bar{x} - \mu_0}{sem}$ according to the $t(n-1)$-distribution

Reminder: One-sample situation



- $H_0 : \mu = \mu_0$
- Mean $\bar{x}$ far away from μ_0?
- Answer: Yes! Many standard errors $sem = s/\sqrt{n}$ away (approx. $6 \cdot sem$)
- In the model $X_1, \ldots, X_n$ i.i.d. RVs with $X_1 \sim N(\mu, \sigma^2)$ it holds under H_0 $T = \frac{\bar{X} - \mu_0}{SEM} \sim t(n-1)$
- The estimated standard deviation of $\bar{X}$ is $1 \cdot SEM$ (not $6 \cdot SEM$)
- Judge the evaluated data $t = \frac{\bar{x} - \mu_0}{sem}$ according to the $t(n-1)$-distribution
- Here: $p < \alpha = 5\%$ resp. $t \in R$, thus reject H_0 on the α level

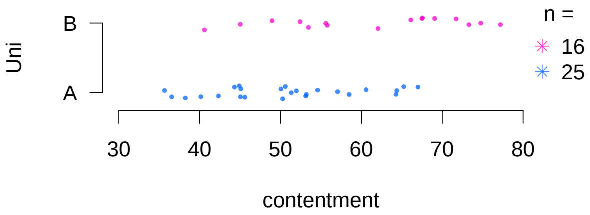
Reminder: One-sample situation



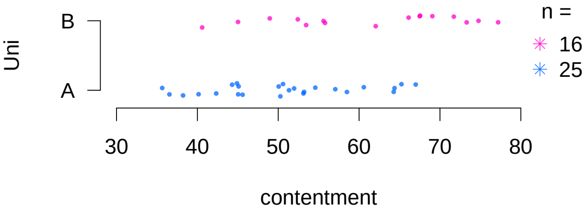
Message:
We judge the discrepancy of $\bar{x}$ and μ_0 in the units 'standard error'

$$\boxed{sem = \frac{s}{\sqrt{n}}}$$

Today: Two-sample situation

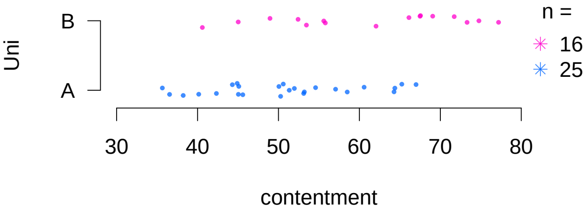


Today: Two-sample situation



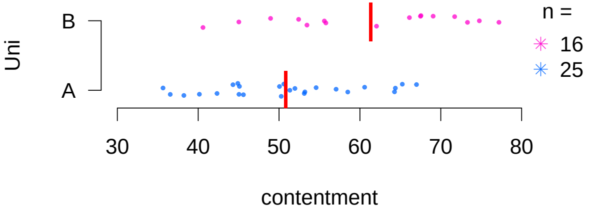
- The satisfaction $(x_1, \ldots, x_{n_1})$ of the students from Uni A ('index 1') has the tendency to be smaller than the satisfaction $(y_1, \ldots, y_{n_2})$ in Uni B ('index 2')

Today: Two-sample situation



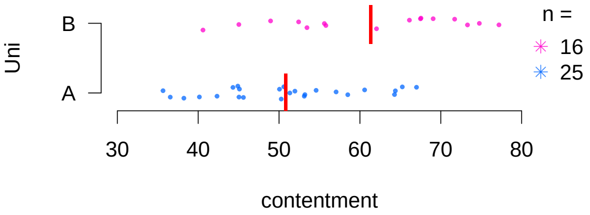
- The satisfaction $(x_1, \ldots, x_{n_1})$ of the students from Uni A ('index 1') has the tendency to be smaller than the satisfaction $(y_1, \ldots, y_{n_2})$ in Uni B ('index 2')
- Can this discrepancy observed in the data be easily explained by chance, if there is actually no difference between the Unis (in the 'populations' of all students of the program)?

Today: Two-sample situation



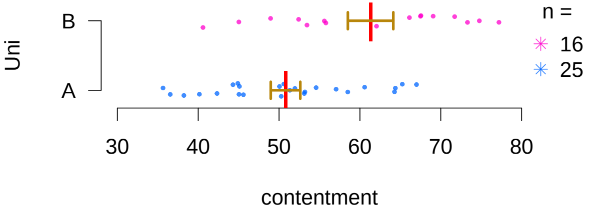
- The satisfaction $(x_1, \ldots, x_{n_1})$ of the students from Uni A ('index 1') has the tendency to be smaller than the satisfaction $(y_1, \ldots, y_{n_2})$ in Uni B ('index 2')
- Can this discrepancy observed in the data be easily explained by chance, if there is actually no difference between the Unis (in the 'populations' of all students of the program)?
- Measure discrepancy through the distance of means $\bar{y}$ and $\bar{x}$

Today: Two-sample situation



- The satisfaction ($x_1, \ldots, x_{n_1}$) of the students from Uni A ('index 1') has the tendency to be smaller than the satisfaction ($y_1, \ldots, y_{n_2}$) in Uni B ('index 2')
- Can this discrepancy observed in the data be easily explained by chance, if there is actually no difference between the Unis (in the 'populations' of all students of the program)?
- Measure discrepancy through the distance of means $\bar{y}$ and $\bar{x}$
- Question: Is this discrepancy large? Respectively, are the means far apart?

Today: Two-sample situation

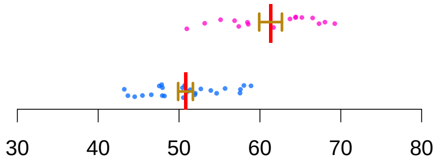


- The satisfaction $(x_1, \ldots, x_{n_1})$ of the students from Uni A ('index 1') has the tendency to be smaller than the satisfaction $(y_1, \ldots, y_{n_2})$ in Uni B ('index 2')
- Can this discrepancy observed in the data be easily explained by chance, if there is actually no difference between the Unis (in the 'populations' of all students of the program)?
- Measure discrepancy through the distance of means $\bar{y}$ and $\bar{x}$
- Question: Is this discrepancy large? Respectively, are the means far apart?
- Answer: This depends on the standard errors sem_y and sem_x

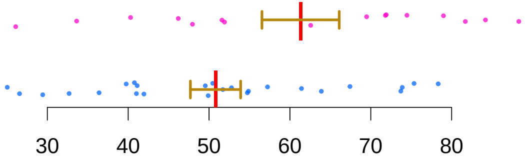
sem_y is the standard error of $\bar{y}$, i.e., based on the data $(y_j)_j$ only, and analogously sem_x is based on the $(x_i)_i$.

Standard errors!

Discrepancy huge



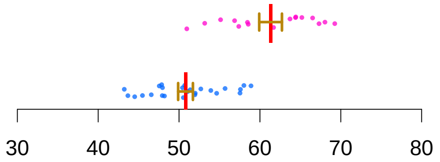
Discrepancy smaller



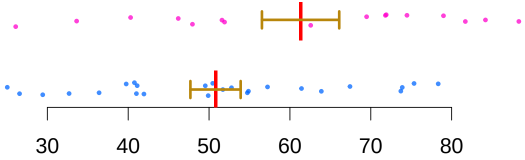
(while the means in the upper and lower graphic coincide)

Standard errors!

Discrepancy huge



Discrepancy smaller



(while the means in the upper and lower graphic coincide)

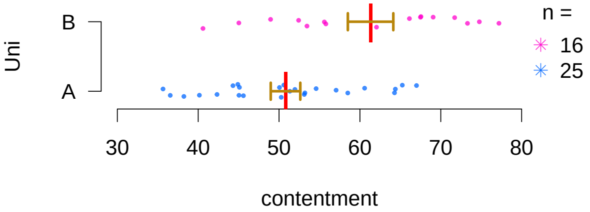
Thus: discrepancy large if the standard errors small

Two-sample t-statistic

The two-sample t-statistic

$$\boxed{t = \frac{\bar{y} - \bar{x}}{\sqrt{sem_y^2 + sem_x^2}}}$$

measures the discrepancy of the means in relation to their standard errors

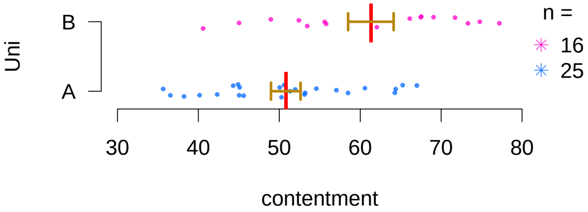


Two-sample t-statistic

The two-sample t-statistic

$$\boxed{t = \frac{\bar{y} - \bar{x}}{\sqrt{sem_y^2 + sem_x^2}}}$$

measures the discrepancy of the means in relation to their standard errors

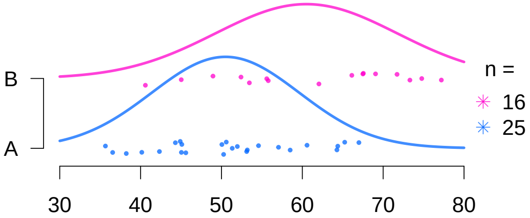


Here: $t \approx 3.1$

$\rightarrow$ Need a model, in order to explicitly judge the value of t

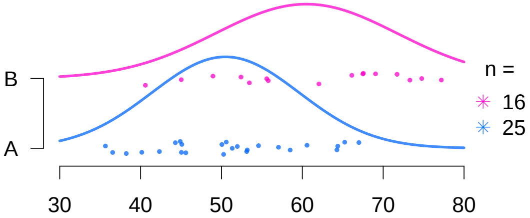
Model

Model: Let $X_1, \dots, X_{n_1}, Y_1, \dots, Y_{n_2}$ be independent RVs with
$X_i \sim N(\mu_1, \sigma_1^2)$ for $i = 1, \dots, n_1$ and
$Y_j \sim N(\mu_2, \sigma_2^2)$ for $j = 1, \dots, n_2$ with $(\mu_1, \sigma_1^2, \mu_2, \sigma_2^2) \in \mathbb{R} \times \mathbb{R}^+ \times \mathbb{R} \times \mathbb{R}^+$



Model

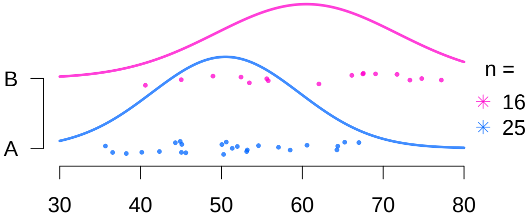
<u>Model</u>: Let $X_1, \dots, X_{n_1}, Y_1, \dots, Y_{n_2}$ be independent RVs with
$X_i \sim N(\mu_1, \sigma_1^2)$ for $i = 1, \dots, n_1$ and
$Y_j \sim N(\mu_2, \sigma_2^2)$ for $j = 1, \dots, n_2$ with $(\mu_1, \sigma_1^2, \mu_2, \sigma_2^2) \in \mathbb{R} \times \mathbb{R}^+ \times \mathbb{R} \times \mathbb{R}^+$



- Independence

Model

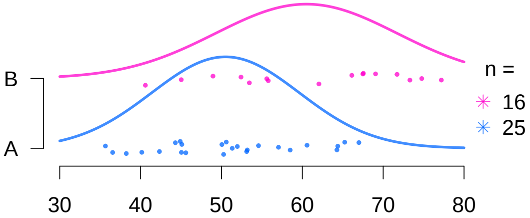
<u>Model</u>: Let $X_1, \dots, X_{n_1}, Y_1, \dots, Y_{n_2}$ be independent RVs with
$X_i \sim N(\mu_1, \sigma_1^2)$ for $i = 1, \dots, n_1$ and
$Y_j \sim N(\mu_2, \sigma_2^2)$ for $j = 1, \dots, n_2$ with $(\mu_1, \sigma_1^2, \mu_2, \sigma_2^2) \in \mathbb{R} \times \mathbb{R}^+ \times \mathbb{R} \times \mathbb{R}^+$



- Independence
- Normal distribution

Model

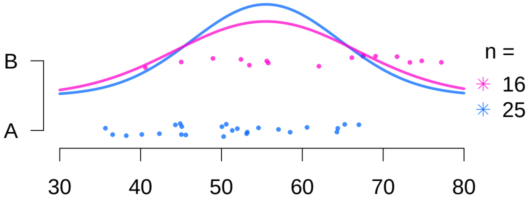
Model: Let $X_1, \ldots, X_{n_1}, Y_1, \ldots, Y_{n_2}$ be independent RVs with
$X_i \sim N(\mu_1, \sigma_1^2)$ for $i = 1, \ldots, n_1$ and
$Y_j \sim N(\mu_2, \sigma_2^2)$ for $j = 1, \ldots, n_2$ with $(\mu_1, \sigma_1^2, \mu_2, \sigma_2^2) \in \mathbb{R} \times \mathbb{R}^+ \times \mathbb{R} \times \mathbb{R}^+$



- Independence
- Normal distribution
- Each group has their own (unknown) parameters
  in particular there is a possible 'shift'

Null hypothesis

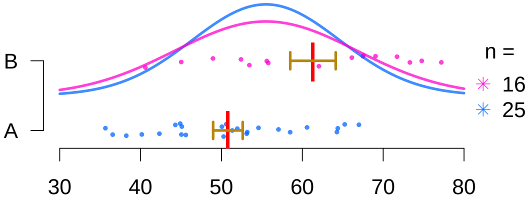
Let $X_1, \ldots, X_{n_1}, Y_1, \ldots, Y_{n_2}$ be independent RVs with
$X_i \sim N(\mu_1, \sigma_1^2)$ for $i = 1, \ldots, n_1$ and
$Y_j \sim N(\mu_2, \sigma_2^2)$ for $j = 1, \ldots, n_2$ with $(\mu_1, \sigma_1^2, \mu_2, \sigma_2^2) \in \mathbb{R} \times \mathbb{R}^+ \times \mathbb{R} \times \mathbb{R}^+$



- Null hypothesis: $H_0 : \mu_1 = \mu_2$
  i.e., distributions are <u>not</u> shifted against each other, 'no difference'

Null hypothesis

Let $X_1, \ldots, X_{n_1}, Y_1, \ldots, Y_{n_2}$ be independent RVs with
$X_i \sim N(\mu_1, \sigma_1^2)$ for $i = 1, \ldots, n_1$ and
$Y_j \sim N(\mu_2, \sigma_2^2)$ for $j = 1, \ldots, n_2$ with $(\mu_1, \sigma_1^2, \mu_2, \sigma_2^2) \in \mathbb{R} \times \mathbb{R}^+ \times \mathbb{R} \times \mathbb{R}^+$



- Null hypothesis: $H_0 : \mu_1 = \mu_2$
  i.e., distributions are <u>not</u> shifted against each other, 'no difference'
- In the model, how unlikely is the observed discrepancy, resp. $t \approx 3.1$, if H_0 holds true, and hence there is no shift?

Distribution of the t-statistic

Let $X_1, \ldots, X_{n_1}, Y_1, \ldots, Y_{n_2}$ be independent RVs with
$X_i \sim N(\mu_1, \sigma_1^2)$ for $i = 1, \ldots, n_1$ and
$Y_j \sim N(\mu_2, \sigma_2^2)$ for $j = 1, \ldots, n_2$ with $(\mu_1, \sigma_1^2, \mu_2, \sigma_2^2) \in \mathbb{R} \times \mathbb{R}^+ \times \mathbb{R} \times \mathbb{R}^+$

Under $H_0 : \mu_1 = \mu_2$ it holds (approx), that

$$T := \frac{\bar{Y} - \bar{X}}{\sqrt{SEM_y^2 + SEM_x^2}} \sim t(\nu)$$

- $t(\nu)$ is the t-distribution with ν degrees of freedom

Distribution of the t-statistic

Let $X_1, \ldots, X_{n_1}, Y_1, \ldots, Y_{n_2}$ be independent RVs with
$X_i \sim N(\mu_1, \sigma_1^2)$ for $i = 1, \ldots, n_1$ and
$Y_j \sim N(\mu_2, \sigma_2^2)$ for $j = 1, \ldots, n_2$ with $(\mu_1, \sigma_1^2, \mu_2, \sigma_2^2) \in \mathbb{R} \times \mathbb{R}^+ \times \mathbb{R} \times \mathbb{R}^+$

Under $H_0 : \mu_1 = \mu_2$ it holds (approx), that

$$T := \frac{\bar{Y} - \bar{X}}{\sqrt{SEM_y^2 + SEM_x^2}} \sim t(\nu)$$

- $t(\nu)$ is the t-distribution with ν degrees of freedom
- ν depends on the sample sizes n_1 und n_2, as well as on the standard errrors SEM_x and SEM_y $\rightarrow$ sufficient if R knows ν

Distribution of the t-statistic

Let $X_1, \dots, X_{n_1}, Y_1, \dots, Y_{n_2}$ be independent RVs with
$X_i \sim N(\mu_1, \sigma_1^2)$ for $i = 1, \dots, n_1$ and
$Y_j \sim N(\mu_2, \sigma_2^2)$ for $j = 1, \dots, n_2$ with $(\mu_1, \sigma_1^2, \mu_2, \sigma_2^2) \in \mathbb{R} \times \mathbb{R}^+ \times \mathbb{R} \times \mathbb{R}^+$

Under $H_0 : \mu_1 = \mu_2$ it holds (approx), that

$$T := \frac{\bar{Y} - \bar{X}}{\sqrt{SEM_y^2 + SEM_x^2}} \sim t(\nu)$$

- $t(\nu)$ is the t-distribution with ν degrees of freedom
- ν depends on the sample sizes n_1 und n_2, as well as on the standard errrors SEM_x and SEM_y $\rightarrow$ sufficient if R knows ν

For completeness: it is

$$\nu = \frac{\left(\frac{\sigma_y^2}{n_2} + \frac{\sigma_x^2}{n_1}\right)^2}{\frac{(\sigma_y^2/n_2)^2}{n_2 - 1} + \frac{(\sigma_x^2/n_1)^2}{n_1 - 1}}$$

while in practice the unknown σ_2^2 and σ_1^2 are estimated using the empirical variances s_y^2 and s_x^2. (Boring... Please do not remember ...)

Distribution of the t-statistic

Let $X_1, \dots, X_{n_1}, Y_1, \dots, Y_{n_2}$ be independent RVs with
$X_i \sim N(\mu_1, \sigma_1^2)$ for $i = 1, \dots, n_1$ and
$Y_j \sim N(\mu_2, \sigma_2^2)$ for $j = 1, \dots, n_2$ with $(\mu_1, \sigma_1^2, \mu_2, \sigma_2^2) \in \mathbb{R} \times \mathbb{R}^+ \times \mathbb{R} \times \mathbb{R}^+$

Under $H_0 : \mu_1 = \mu_2$ it holds (approx), that

$$T := \frac{\bar{Y} - \bar{X}}{\sqrt{SEM_y^2 + SEM_x^2}} \sim t(\nu)$$

- $t(\nu)$ is the t-distribution with ν degrees of freedom
- ν depends on the sample sizes n_1 und n_2, as well as on the standard errrors SEM_x and SEM_y $\rightarrow$ sufficient if R knows ν
- Why is this distribution plausible?
  - Due to independence, standardization of the difference $\bar{Y} - \bar{X}$ yields $\frac{(\bar{Y} - \bar{X}) - (\mu_2 - \mu_1)}{\sqrt{\sigma_2^2/n_2 + \sigma_1^2/n_1}} \sim N(0, 1)$

Distribution of the t-statistic

Let $X_1, \ldots, X_{n_1}, Y_1, \ldots, Y_{n_2}$ be independent RVs with
$X_i \sim N(\mu_1, \sigma_1^2)$ for $i = 1, \ldots, n_1$ and
$Y_j \sim N(\mu_2, \sigma_2^2)$ for $j = 1, \ldots, n_2$ with $(\mu_1, \sigma_1^2, \mu_2, \sigma_2^2) \in \mathbb{R} \times \mathbb{R}^+ \times \mathbb{R} \times \mathbb{R}^+$

Under $H_0 : \mu_1 = \mu_2$ it holds (approx), that

$$T := \frac{\bar{Y} - \bar{X}}{\sqrt{SEM_y^2 + SEM_x^2}} \sim t(\nu)$$

- $t(\nu)$ is the t-distribution with ν degrees of freedom
- ν depends on the sample sizes n_1 und n_2, as well as on the standard errrors SEM_x and SEM_y $\rightarrow$ sufficient if R knows ν
- Why is this distribution plausible?
  - Due to independence, standardization of the difference $\bar{Y} - \bar{X}$ yields $\frac{(\bar{Y}-\bar{X})-(\mu_2-\mu_1)}{\sqrt{\sigma_2^2/n_2+\sigma_1^2/n_1}} \sim N(0,1)$
  - But under H_0 the difference in expectations $\mu_2 - \mu_1$ vanishes

Distribution of the t-statistic

Let $X_1, \ldots, X_{n_1}, Y_1, \ldots, Y_{n_2}$ be independent RVs with
$X_i \sim N(\mu_1, \sigma_1^2)$ for $i = 1, \ldots, n_1$ and
$Y_j \sim N(\mu_2, \sigma_2^2)$ for $j = 1, \ldots, n_2$ with $(\mu_1, \sigma_1^2, \mu_2, \sigma_2^2) \in \mathbb{R} \times \mathbb{R}^+ \times \mathbb{R} \times \mathbb{R}^+$

Under $H_0 : \mu_1 = \mu_2$ it holds (approx), that

$$T := \frac{\bar{Y} - \bar{X}}{\sqrt{SEM_y^2 + SEM_x^2}} \sim t(\nu)$$

- $t(\nu)$ is the t-distribution with ν degrees of freedom
- ν depends on the sample sizes n_1 und n_2, as well as on the standard errrors SEM_x and SEM_y $\rightarrow$ sufficient if `R` knows ν
- Why is this distribution plausible?
  - Due to independence, standardization of the difference $\bar{Y} - \bar{X}$ yields $\frac{(\bar{Y}-\bar{X})-(\mu_2-\mu_1)}{\sqrt{\sigma_2^2/n_2+\sigma_1^2/n_1}} \sim N(0,1)$
  - But under H_0 the difference in expectations $\mu_2 - \mu_1$ vanishes
  - The estimation of σ_1 and σ_2 again yields heavier tails as in $N(0,1)$

Distribution of the t-statistic

Let $X_1, \ldots, X_{n_1}, Y_1, \ldots, Y_{n_2}$ be independent RVs with
$X_i \sim N(\mu_1, \sigma_1^2)$ for $i = 1, \ldots, n_1$ and
$Y_j \sim N(\mu_2, \sigma_2^2)$ for $j = 1, \ldots, n_2$ with $(\mu_1, \sigma_1^2, \mu_2, \sigma_2^2) \in \mathbb{R} \times \mathbb{R}^+ \times \mathbb{R} \times \mathbb{R}^+$

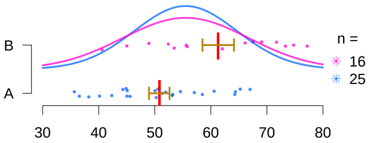
Under $H_0 : \mu_1 = \mu_2$ it holds (approx), that

$$T := \frac{\bar{Y} - \bar{X}}{\sqrt{SEM_y^2 + SEM_x^2}} \sim t(\nu)$$

- $t(\nu)$ is the t-distribution with ν degrees of freedom
- ν depends on the sample sizes n_1 und n_2, as well as on the standard errrors SEM_x and SEM_y $\to$ sufficient if R knows ν
- Why is this distribution plausible?
  - Due to independence, standardization of the difference $\bar{Y} - \bar{X}$ yields $\frac{(\bar{Y}-\bar{X})-(\mu_2-\mu_1)}{\sqrt{\sigma_2^2/n_2+\sigma_1^2/n_1}} \sim N(0,1)$
  - But under H_0 the difference in expectations $\mu_2 - \mu_1$ vanishes
  - The estimation of σ_1 and σ_2 again yields heavier tails as in $N(0,1)$
- Rough approximation: If H_0 holds true, then $|t| \approx 1$ is a typical value, while $|t| \gtrapprox 3$ barely happens (as T is approx $N(0,1)$ distributed)

Judgement of the discrepancy

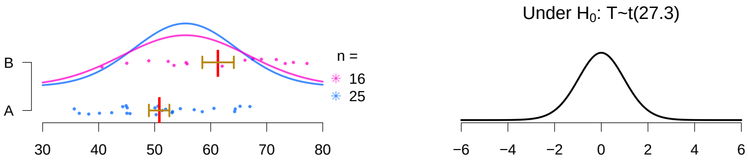
How unlikely is the observed discrepancy, resp. $t \approx 3.1$, if H_0 holds true, i.e., if there is no difference in the population means?



'Business as usual'

Judgement of the discrepancy

How unlikely is the observed discrepancy, resp. $t \approx 3.1$, if H_0 holds true, i.e., if there is no difference in the population means?

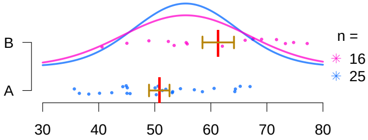


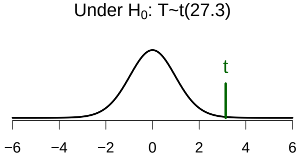
'Business as usual'

- Under $H_0 : \mu_x = \mu_y$ it is $T \sim t(27.3)$

Judgement of the discrepancy

How unlikely is the observed discrepancy, resp. $t \approx 3.1$, if H_0 holds true, i.e., if there is no difference in the population means?



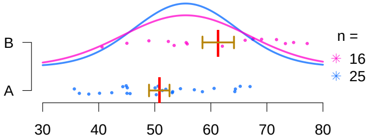


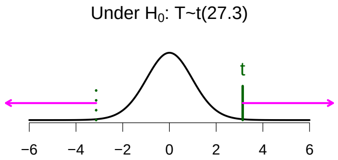
'Business as usual'

- Under $H_0 : \mu_x = \mu_y$ it is $T \sim t(27.3)$
- From the data we compute $t \approx 3.1$

Judgement of the discrepancy

How unlikely is the observed discrepancy, resp. $t \approx 3.1$, if H_0 holds true, i.e., if there is no difference in the population means?



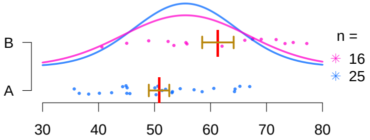


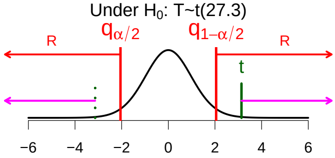
'Business as usual'

- Under $H_0 : \mu_x = \mu_y$ it is $T \sim t(27.3)$
- From the data we compute $t \approx 3.1$
- p-value: $p = \mathbb{P}_{H_0}(|T| \geqslant |t|) \approx 0.004$

Judgement of the discrepancy

How unlikely is the observed discrepancy, resp. $t \approx 3.1$, if H_0 holds true, i.e., if there is no difference in the population means?



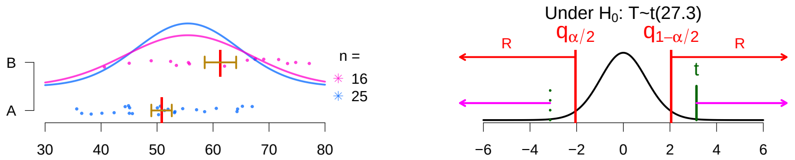


'Business as usual'

- Under $H_0 : \mu_x = \mu_y$ it is $T \sim t(27.3)$
- From the data we compute $t \approx 3.1$
- p-value: $p = \mathbb{P}_{H_0}(|T| \geqslant |t|) \approx 0.004$
- For $\alpha = 5\%$ (previously set) it is $t \in R$ ($\Leftrightarrow p \leqslant \alpha$), i.e., reject H_0 on the 5%-level

Judgement of the discrepancy

How unlikely is the observed discrepancy, resp. $t \approx 3.1$, if H_0 holds true, i.e., if there is no difference in the population means?



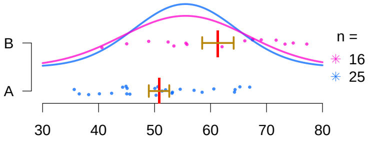
'Business as usual'

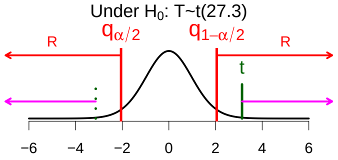
- Under $H_0 : \mu_x = \mu_y$ it is $T \sim t(27.3)$
- From the data we compute $t \approx 3.1$
- p-value: $p = \mathbb{P}_{H_0}(|T| \geqslant |t|) \approx 0.004$
- For $\alpha = 5\%$ (previously set) it is $t \in R$ ($\Leftrightarrow p \leqslant \alpha$), i.e., reject H_0 on the 5%-level
- Interpretation: A discrepancy as extreme as observed in the data appears only in about 4 of 1000 cases, if there is no difference between the Unis. In this sense the data speak against the null hypothesis.

Remark

Uni B: $y_1, \ldots, y_{n_2}$
Uni A: $x_1, \ldots, x_{n_1}$

$$t = \frac{\bar{y} - \bar{x}}{\sqrt{sem_y^2 + sem_x^2}}$$



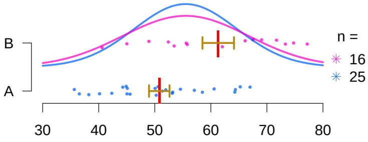


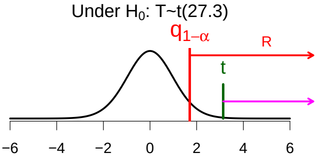
- We perfomed a two-sided test (null hypothesis: $H_0 : \mu_1 = \mu_2$)
  Alternative $H_A : \mu_1 \neq \mu_2$
  Extreme values of t speak against H_0 (here: H_0 rejected)

Remark

Uni B: $y_1, \ldots, y_{n_2}$
Uni A: $x_1, \ldots, x_{n_1}$

$$t = \frac{\bar{y} - \bar{x}}{\sqrt{sem_y^2 + sem_x^2}}$$



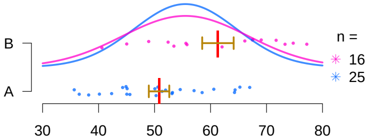


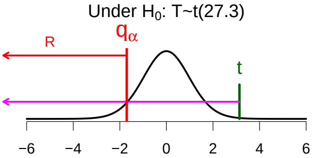
- We perfomed a two-sided test (null hypothesis: $H_0 : \mu_1 = \mu_2$)
  Alternative $H_A : \mu_1 \neq \mu_2$
  Extreme values of t speak against H_0 (here: H_0 rejected)
- One-sided tests:
  - right-sided: $H_A : \mu_1 < \mu_2$
    Large values of t speak against H_0 (here: H_0 rejected)

Remark

Uni B: $y_1, \ldots, y_{n_2}$
Uni A: $x_1, \ldots, x_{n_1}$

$$t = \frac{\bar{y} - \bar{x}}{\sqrt{sem_y^2 + sem_x^2}}$$





- We perfomed a two-sided test (null hypothesis: $H_0 : \mu_1 = \mu_2$)
  Alternative $H_A : \mu_1 \neq \mu_2$
  Extreme values of t speak against H_0 (here: H_0 rejected)
- One-sided tests:
  - right-sided: $H_A : \mu_1 < \mu_2$
    Large values of t speak against H_0 (here: H_0 rejected)
  - left-sided: $H_A : \mu_1 > \mu_2$
    Small values of t speak against H_0 (here: H_0 not rejected)

t-test in R (according to 'Welch')

```
# Enter data
x <- c(...)
y <- c(...)
# Perform t-test
t.test(y,x,...)
# Output

        Welch Two Sample t-test

data:  y and x
t = 3.1365, df = 27.311, p-value = 0.004067
alternative hypothesis:
true difference in means is not equal to 0
95 percent confidence interval:
  3.641362 17.396356
sample estimates:
mean of x mean of y
 61.31892  50.80006
```

- The t-test presented is also known as the *Welch*-test
- The degrees of freedom 27.3 can be read from here

Generalization

Let $X_1, \dots, X_{n_1}, Y_1, \dots, Y_{n_2}$ be independent RVs with
$X_i \sim N(\mu_1, \sigma_1^2)$ for $i = 1, \dots, n_1$ and
$Y_j \sim N(\mu_2, \sigma_2^2)$ for $j = 1, \dots, n_2$ with $(\mu_1, \sigma_1^2, \mu_2, \sigma_2^2) \in \mathbb{R} \times \mathbb{R}^+ \times \mathbb{R} \times \mathbb{R}^+$

Under $H_0 : \mu_1 = \mu_2$ it holds (approx) that

$$T = \frac{\bar{Y} - \bar{X}}{\sqrt{SEM_y^2 + SEM_x^2}} \sim t(\nu)$$

Generalization

Let $X_1, \dots, X_{n_1}, Y_1, \dots, Y_{n_2}$ be independent RVs with
$X_i \sim N(\mu_1, \sigma_1^2)$ for $i = 1, \dots, n_1$ and
$Y_j \sim N(\mu_2, \sigma_2^2)$ for $j = 1, \dots, n_2$ with $(\mu_1, \sigma_1^2, \mu_2, \sigma_2^2) \in \mathbb{R} \times \mathbb{R}^+ \times \mathbb{R} \times \mathbb{R}^+$

Under $H_0 : \mu_1 = \mu_2$ it holds (approx) that

$$T = \frac{\bar{Y} - \bar{X}}{\sqrt{SEM_y^2 + SEM_x^2}} \sim t(\nu)$$

H_0 reformulated: $d := \mu_2 - \mu_1 = 0$ (no difference)

Generalization

Let $X_1, \dots, X_{n_1}, Y_1, \dots, Y_{n_2}$ be independent RVs with
$X_i \sim N(\mu_1, \sigma_1^2)$ for $i = 1, \dots, n_1$ and
$Y_j \sim N(\mu_2, \sigma_2^2)$ for $j = 1, \dots, n_2$ with $(\mu_1, \sigma_1^2, \mu_2, \sigma_2^2) \in \mathbb{R} \times \mathbb{R}^+ \times \mathbb{R} \times \mathbb{R}^+$

Under $H_0 : \mu_1 = \mu_2$ it holds (approx) that

$$T = \frac{\bar{Y} - \bar{X}}{\sqrt{SEM_y^2 + SEM_x^2}} \sim t(\nu)$$

H_0 reformulated: $d := \mu_2 - \mu_1 = 0$ (no difference)

Generalization of the null hypothesis: $H_0 : d = d_0$ (difference is $d_0 \in \mathbb{R}$)

Generalization

Let $X_1, \dots, X_{n_1}, Y_1, \dots, Y_{n_2}$ be independent RVs with
$X_i \sim N(\mu_1, \sigma_1^2)$ for $i = 1, \dots, n_1$ and
$Y_j \sim N(\mu_2, \sigma_2^2)$ for $j = 1, \dots, n_2$ with $(\mu_1, \sigma_1^2, \mu_2, \sigma_2^2) \in \mathbb{R} \times \mathbb{R}^+ \times \mathbb{R} \times \mathbb{R}^+$

Under $H_0 : \mu_1 = \mu_2$ it holds (approx) that

$$T = \frac{\bar{Y} - \bar{X}}{\sqrt{SEM_y^2 + SEM_x^2}} \sim t(\nu)$$

H_0 reformulated: $d := \mu_2 - \mu_1 = 0$ (no difference)

Generalization of the null hypothesis: $H_0 : d = d_0$ (difference is $d_0 \in \mathbb{R}$)

Under $H_0 : d = d_0$ it holds (approx) that

$$T := \frac{(\bar{Y} - \bar{X}) - d_0}{\sqrt{SEM_y^2 + SEM_x^2}} \sim t(\nu)$$

Generalization

Let $X_1, \ldots, X_{n_1}, Y_1, \ldots, Y_{n_2}$ be independent RVs with
$X_i \sim N(\mu_1, \sigma_1^2)$ for $i = 1, \ldots, n_1$ and
$Y_j \sim N(\mu_2, \sigma_2^2)$ for $j = 1, \ldots, n_2$ with $(\mu_1, \sigma_1^2, \mu_2, \sigma_2^2) \in \mathbb{R} \times \mathbb{R}^+ \times \mathbb{R} \times \mathbb{R}^+$

Under $H_0 : \mu_1 = \mu_2$ it holds (approx) that

$$T = \frac{\bar{Y} - \bar{X}}{\sqrt{SEM_y^2 + SEM_x^2}} \sim t(\nu)$$

H_0 reformulated: $d := \mu_2 - \mu_1 = 0$ (no difference)

Generalization of the null hypothesis: $H_0 : d = d_0$ (difference is $d_0 \in \mathbb{R}$)

Under $H_0 : d = d_0$ it holds (approx) that

$$T := \frac{(\bar{Y} - \bar{X}) - d_0}{\sqrt{SEM_y^2 + SEM_x^2}} \sim t(\nu)$$

Analogous procedure, the only difference is that d_0 has to be subtracted in the numerator of t

Generalization

Let $X_1, \dots, X_{n_1}, Y_1, \dots, Y_{n_2}$ be independent RVs with
$X_i \sim N(\mu_1, \sigma_1^2)$ for $i = 1, \dots, n_1$ and
$Y_j \sim N(\mu_2, \sigma_2^2)$ for $j = 1, \dots, n_2$ with $(\mu_1, \sigma_1^2, \mu_2, \sigma_2^2) \in \mathbb{R} \times \mathbb{R}^+ \times \mathbb{R} \times \mathbb{R}^+$

Under $H_0 : \mu_1 = \mu_2$ it holds (approx) that

$$T = \frac{\bar{Y} - \bar{X}}{\sqrt{SEM_y^2 + SEM_x^2}} \sim t(\nu)$$

H_0 reformulated: $d := \mu_2 - \mu_1 = 0$ (no difference)

Generalization of the null hypothesis: $H_0 : d = d_0$ (difference is $d_0 \in \mathbb{R}$)

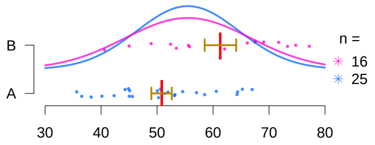
Under $H_0 : d = d_0$ it holds (approx) that

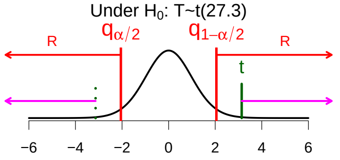
$$T := \frac{(\bar{Y} - \bar{X}) - d_0}{\sqrt{SEM_y^2 + SEM_x^2}} \sim t(\nu)$$

Analogous procedure, the only difference is that d_0 has to be subtracted in the numerator of t (Statistic has the same structure as in the one-sample case: $T = [\spadesuit - \spadesuit]/\spadesuit$)

Example

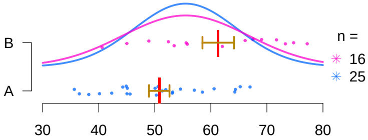
Under $H_0 : d = {\color{blue}0}$ ($\Leftrightarrow \mu_1 = \mu_2$) $\rightarrow$ reject H_0

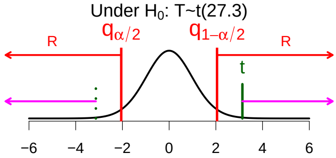




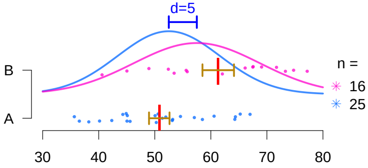
Example

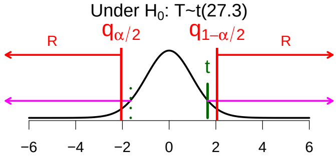
Under $H_0 : d = \color{blue}{0}$ $(\Leftrightarrow \mu_1 = \mu_2) \rightarrow$ reject H_0





Unter $H_0 : d = \color{blue}{5}$ $(\Leftrightarrow \mu_2 = \mu_1 + 5) \rightarrow$ do <u>not</u> reject H_0





t-test in R

```
# Enter data
x <- c(...)
y <- c(...)
# perform t-test
t.test(y,x,mu=5,...)
# Output

        Welch Two Sample t-test

        data:  y and x
t = 1.6456, df = 27.311, p-value = 0.1113
alternative hypothesis:
true difference in means is not equal to 5
95 percent confidence interval:
  3.641362 17.396356
sample estimates:
mean of x mean of y
 61.31892  50.80006
```

- Where does the confidence interval come from?

Two-sample t-test and confidence interval

- Let $X_1, \ldots, X_{n_1}, Y_1, \ldots, Y_{n_2}$ be independent RVs with
  $X_i \sim N(\mu_1, \sigma_1^2)$ for $i = 1, \ldots, n_1$ and
  $Y_j \sim N(\mu_2, \sigma_2^2)$ for $j = 1, \ldots, n_2$ with $(\mu_1, \sigma_1^2, \mu_2, \sigma_2^2) \in \mathbb{R} \times \mathbb{R}^+ \times \mathbb{R} \times \mathbb{R}^+$
- $d = \mu_2 - \mu_1$
- Let $q_{1-\alpha/2}$ be the $(1 - \alpha/2)$-quantile of the $t(\nu)$-distribution (R knows ν)

Under $H_0 : d = d_0$ it holds (approx)

$$T := \frac{(\bar{Y} - \bar{X}) - d_0}{\sqrt{SEM_y^2 + SEM_x^2}} \sim t(\nu)$$

Two-sample t-test and confidence interval

- Let $X_1, \ldots, X_{n_1}, Y_1, \ldots, Y_{n_2}$ be independent RVs with
  $X_i \sim N(\mu_1, \sigma_1^2)$ for $i = 1, \ldots, n_1$ and
  $Y_j \sim N(\mu_2, \sigma_2^2)$ for $j = 1, \ldots, n_2$ with $(\mu_1, \sigma_1^2, \mu_2, \sigma_2^2) \in \mathbb{R} \times \mathbb{R}^+ \times \mathbb{R} \times \mathbb{R}^+$
- $d = \mu_2 - \mu_1$
- Let $q_{1-\alpha/2}$ be the $(1-\alpha/2)$-quantile of the $t(\nu)$-distribution (R knows ν)

Under $H_0 : d = d_0$ it holds (approx)

$$T := \frac{(\bar{Y} - \bar{X}) - d_0}{\sqrt{SEM_y^2 + SEM_x^2}} \sim t(\nu)$$

and equivalently: The confidence interval

$$I := \left((\bar{Y} - \bar{X}) - q_{1-\alpha/2} \cdot \sqrt{SEM_y^2 + SEM_x^2},\ (\bar{Y} - \bar{X}) + q_{1-\alpha/2} \cdot \sqrt{SEM_y^2 + SEM_x^2} \right)$$

overlaps the parameter d_0 with probability (approx) $1 - \alpha$

- C.I. has 'structure' as in the one-sample case: $I = (\spadesuit - q \cdot \spadesuit,\ \spadesuit + q \cdot \spadesuit)$

Two-sample t-test and confidence interval

- Let $X_1, \dots, X_{n_1}, Y_1, \dots, Y_{n_2}$ be independent RVs with
  $X_i \sim N(\mu_1, \sigma_1^2)$ for $i = 1, \dots, n_1$ and
  $Y_j \sim N(\mu_2, \sigma_2^2)$ for $j = 1, \dots, n_2$ with $(\mu_1, \sigma_1^2, \mu_2, \sigma_2^2) \in \mathbb{R} \times \mathbb{R}^+ \times \mathbb{R} \times \mathbb{R}^+$
- $d = \mu_2 - \mu_1$
- Let $q_{1-\alpha/2}$ be the $(1-\alpha/2)$-quantile of the $t(\nu)$-distribution (R knows ν)

Under $H_0 : d = d_0$ it holds (approx)

$$T := \frac{(\bar{Y} - \bar{X}) - d_0}{\sqrt{SEM_y^2 + SEM_x^2}} \sim t(\nu)$$

and equivalently: The confidence interval

$$I := \left((\bar{Y} - \bar{X}) - q_{1-\alpha/2} \cdot \sqrt{SEM_y^2 + SEM_x^2},\ (\bar{Y} - \bar{X}) + q_{1-\alpha/2} \cdot \sqrt{SEM_y^2 + SEM_x^2} \right)$$

overlaps the parameter d_0 with probability (approx) $1 - \alpha$

- C.I. has 'structure' as in the one-sample case: $I = (\spadesuit - q \cdot \spadesuit,\ \spadesuit + q \cdot \spadesuit)$
- Derivation analogously: $\alpha = \mathbb{P}_{H_0}(T \in R) = \cdots = \mathbb{P}_{H_0}(I \not\ni d_0)$

Two-sample t-test and confidence interval

- Let $X_1, \dots, X_{n_1}, Y_1, \dots, Y_{n_2}$ be independent RVs with
  $X_i \sim N(\mu_1, \sigma_1^2)$ for $i = 1, \dots, n_1$ and
  $Y_j \sim N(\mu_2, \sigma_2^2)$ for $j = 1, \dots, n_2$ with $(\mu_1, \sigma_1^2, \mu_2, \sigma_2^2) \in \mathbb{R} \times \mathbb{R}^+ \times \mathbb{R} \times \mathbb{R}^+$
- $d = \mu_2 - \mu_1$
- Let $q_{1-\alpha/2}$ be the $(1-\alpha/2)$-quantile of the $t(\nu)$-distribution (R knows ν)

Under $H_0 : d = d_0$ it holds (approx)

$$T := \frac{(\bar{Y} - \bar{X}) - d_0}{\sqrt{SEM_y^2 + SEM_x^2}} \sim t(\nu)$$

and equivalently: The confidence interval

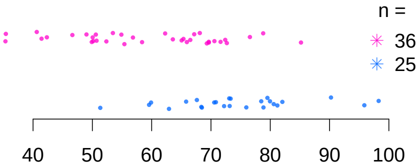
$$I := \left((\bar{Y} - \bar{X}) - q_{1-\alpha/2} \cdot \sqrt{SEM_y^2 + SEM_x^2},\ (\bar{Y} - \bar{X}) + q_{1-\alpha/2} \cdot \sqrt{SEM_y^2 + SEM_x^2} \right)$$

overlaps the parameter d_0 with probability (approx) $1 - \alpha$

- C.I. has 'structure' as in the one-sample case: $I = (\spadesuit - q \cdot \spadesuit,\ \spadesuit + q \cdot \spadesuit)$
- Derivation analogously: $\alpha = \mathbb{P}_{H_0}(T \in R) = \cdots = \mathbb{P}_{H_0}(I \not\ni d_0)$
- For n_1, n_2 large: possibly use normal approximation, i.e., replace $t(\nu)$ with $N(0,1)$. In this case the normality assumption of the RVs can be dropped

Question

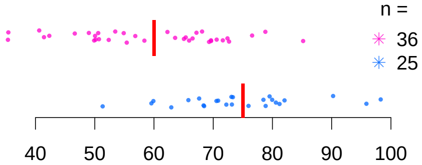
Can $H_0 : \mu_2 = \mu_1$ be rejected on the 5%-level?



Here *t*-test naively:

Question

Can $H_0 : \mu_2 = \mu_1$ be rejected on the 5%-level?

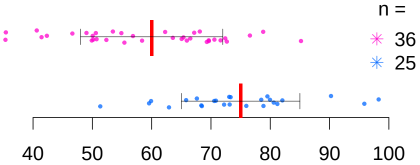


Here t-test naively:

- $\bar{x} \approx 75$ and $\bar{y} \approx 60$ (balances in equilibrium)

Question

Can $H_0 : \mu_2 = \mu_1$ be rejected on the 5%-level?

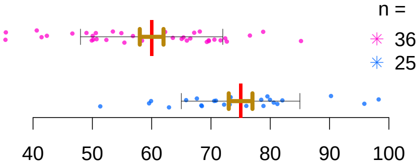


Here t-test naively:

- $\bar{x} \approx 75$ and $\bar{y} \approx 60$ (balances in equilibrium)
- $s_x \approx 10$ and $s_y \approx 12$ (approx bell-shaped distributions, $\approx 2/3$ of the data captured)

Question

Can $H_0 : \mu_2 = \mu_1$ be rejected on the 5%-level?

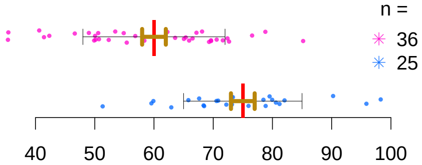


Here *t*-test naively:

- $\bar{x} \approx 75$ and $\bar{y} \approx 60$ (balances in equilibrium)
- $s_x \approx 10$ and $s_y \approx 12$ (approx bell-shaped distributions, $\approx 2/3$ of the data captured)
- $sem_x \approx 10/5 = 2$ and $sem_y \approx 12/6 = 2 \rightarrow$ means are far apart

Question

Can $H_0 : \mu_2 = \mu_1$ be rejected on the 5%-level?

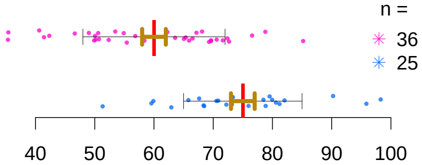


Here t-test naively:

- $\bar{x} \approx 75$ and $\bar{y} \approx 60$ (balances in equilibrium)
- $s_x \approx 10$ and $s_y \approx 12$ (approx bell-shaped distributions, $\approx 2/3$ of the data captured)
- $sem_x \approx 10/5 = 2$ and $sem_y \approx 12/6 = 2 \rightarrow$ means are far apart
- $$t \approx \frac{60 - 75}{\sqrt{2^2 + 2^2}} = \frac{-15}{\sqrt{8}} \approx \frac{-15}{\sqrt{9}} = -5 \qquad \text{(extreme value!)}$$
- $R \approx (-\infty, -2] \cup [2, \infty)$ (as 97.5%-quantile of $N(0,1)$ is $q \approx 1.96 \approx 2$)

Question

Can $H_0 : \mu_2 = \mu_1$ be rejected on the 5%-level?

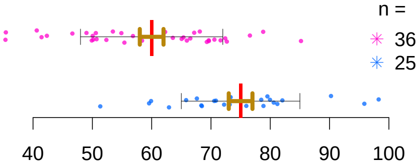


Here t-test naively:

- $\bar{x} \approx 75$ and $\bar{y} \approx 60$ (balances in equilibrium)
- $s_x \approx 10$ and $s_y \approx 12$ (approx bell-shaped distributions, $\approx 2/3$ of the data captured)
- $sem_x \approx 10/5 = 2$ and $sem_y \approx 12/6 = 2 \to$ means are far apart
- $$t \approx \frac{60 - 75}{\sqrt{2^2 + 2^2}} = \frac{-15}{\sqrt{8}} \approx \frac{-15}{\sqrt{9}} = -5 \qquad \text{(extreme value!)}$$
- $R \approx (-\infty, -2] \cup [2, \infty)$ (as 97.5%-quantile of $N(0, 1)$ is $q \approx 1.96 \approx 2$)
- $t \in R$, thus reject H_0

Question

Can $H_0 : \mu_2 = \mu_1$ be rejected on the 5%-level?

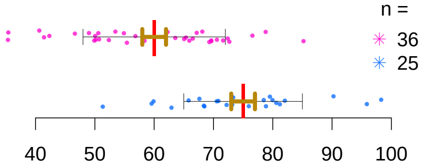


Here t-test naively:

- $\bar{x} \approx 75$ and $\bar{y} \approx 60$ (balances in equilibrium)
- $s_x \approx 10$ and $s_y \approx 12$ (approx bell-shaped distributions, $\approx 2/3$ of the data captured)
- $sem_x \approx 10/5 = 2$ and $sem_y \approx 12/6 = 2 \rightarrow$ means are far apart
- $$t \approx \frac{60 - 75}{\sqrt{2^2 + 2^2}} = \frac{-15}{\sqrt{8}} \approx \frac{-15}{\sqrt{9}} = -5 \qquad \text{(extreme value!)}$$
- $R \approx (-\infty, -2] \cup [2, \infty)$ (as 97.5%-quantile of $N(0,1)$ is $q \approx 1.96 \approx 2$)
- $t \in R$, thus reject H_0 … or equivalently
- C.I. $\approx (-15 - 2 \cdot 3, -15 + 2 \cdot 3) = (-21, -9)$ does not overlap $d = 0$

Question

Can $H_0 : \mu_2 = \mu_1$ be rejected on the 5%-level?



Here t-test naively:

- $\bar{x} \approx 75$ and $\bar{y} \approx 60$ (balances in equilibrium)
- $s_x \approx 10$ and $s_y \approx 12$ (approx bell-shaped distributions, $\approx 2/3$ of the data captured)
- $sem_x \approx 10/5 = 2$ and $sem_y \approx 12/6 = 2 \rightarrow$ means are far apart
- $$t \approx \frac{60 - 75}{\sqrt{2^2 + 2^2}} = \frac{-15}{\sqrt{8}} \approx \frac{-15}{\sqrt{9}} = -5 \qquad \text{(extreme value!)}$$
- $R \approx (-\infty, -2] \cup [2, \infty)$ (as 97.5%-quantile of $N(0, 1)$ is $q \approx 1.96 \approx 2$)
- $t \in R$, thus reject H_0 ... or equivalently
- C.I. $\approx (-15 - 2 \cdot 3, -15 + 2 \cdot 3) = (-21, -9)$ does not overlap $d = 0$

Again: Clear, only naive estimations. But $t \approx -5$ is extreme, regardless of our rough estimation. What does R say?

t-test in R

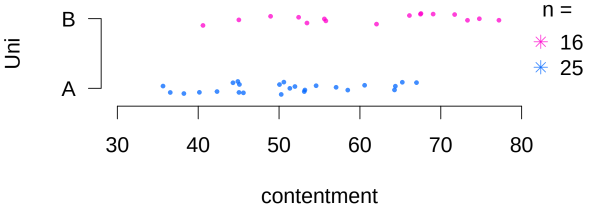
```
# Enter data
x <- c(...)
y <- c(...)
# Perform t-Test
t.test(y,x,...)
# Output

        Welch Two Sample t-test

data:  y and x
t = -4.9829, df = 56.052, p-value = 6.358e-06
alternative hypothesis:
true difference in means is not equal to 0
95 percent confidence interval:
 -21.097942  -8.998583
sample estimates:
mean of x mean of y
 59.27828  74.32654
```

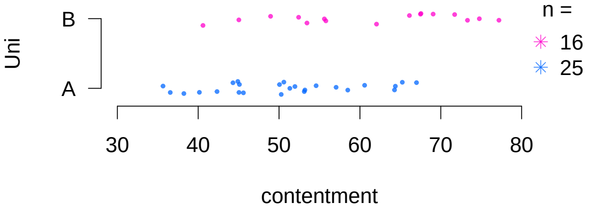
Our naive estimations were very precise (absurd how well it fits). But again: Even if the estimations in the picture were biased and we had obtained e.g., $t \approx -4$, then still this would have been an extreme value!

Welch vs. Student



Model 'Welch': Let $X_1, \ldots, X_{n_1}, Y_1, \ldots, Y_{n_2}$ be independent RVs with
$X_i \sim N(\mu_1, \sigma_1^2)$ for $i = 1, \ldots, n_1$ and
$Y_j \sim N(\mu_2, \sigma_2^2)$ for $j = 1, \ldots, n_2$ with $(\mu_1, \sigma_1^2, \mu_2, \sigma_2^2) \in \mathbb{R} \times \mathbb{R}^+ \times \mathbb{R} \times \mathbb{R}^+$

Welch vs. Student

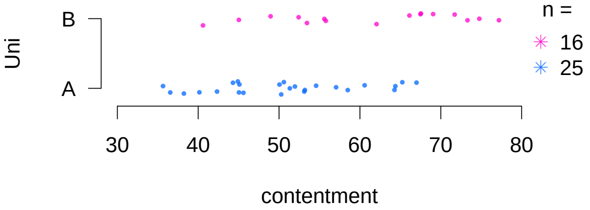


Model 'Welch': Let $X_1, \ldots, X_{n_1}, Y_1, \ldots, Y_{n_2}$ be independent RVs with
$X_i \sim N(\mu_1, \sigma_1^2)$ for $i = 1, \ldots, n_1$ and
$Y_j \sim N(\mu_2, \sigma_2^2)$ for $j = 1, \ldots, n_2$ with $(\mu_1, \sigma_1^2, \mu_2, \sigma_2^2) \in \mathbb{R} \times \mathbb{R}^+ \times \mathbb{R} \times \mathbb{R}^+$

A different approach, also known as Student's t-test, assumes the variances to be equal in both groups

Model 'Student': Let $X_1, \ldots, X_{n_1}, Y_1, \ldots, Y_{n_2}$ be independent RVs with
$X_i \sim N(\mu_1, \sigma^2)$ for $i = 1, \ldots, n_1$ and
$Y_j \sim N(\mu_2, \sigma^2)$ for $j = 1, \ldots, n_2$ with $(\mu_1, \mu_2, \sigma^2) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R}^+$

Welch vs. Student



Model 'Welch': Let $X_1, \dots, X_{n_1}, Y_1, \dots, Y_{n_2}$ be independent RVs with
$X_i \sim N(\mu_1, \sigma_1^2)$ for $i = 1, \dots, n_1$ and
$Y_j \sim N(\mu_2, \sigma_2^2)$ for $j = 1, \dots, n_2$ with $(\mu_1, \sigma_1^2, \mu_2, \sigma_2^2) \in \mathbb{R} \times \mathbb{R}^+ \times \mathbb{R} \times \mathbb{R}^+$

A different approach, also known as Student's t-test, assumes the variances to be equal in both groups

Model 'Student': Let $X_1, \dots, X_{n_1}, Y_1, \dots, Y_{n_2}$ be independent RVs with
$X_i \sim N(\mu_1, \sigma^2)$ for $i = 1, \dots, n_1$ and
$Y_j \sim N(\mu_2, \sigma^2)$ for $j = 1, \dots, n_2$ with $(\mu_1, \mu_2, \sigma^2) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R}^+$

In our data, the spread in both groups is similar, so Student's additional assumption is plausible. (In general, Welch's version is applicable to a wider range of data, as it allows for different variances in both groups)

Welch vs. Student

Model 'Welch': Let $X_1, \dots, X_{n_1}, Y_1, \dots, Y_{n_2}$ be independent RVs with
$X_i \sim N(\mu_1, \sigma_1^2)$ for $i = 1, \dots, n_1$ and
$Y_j \sim N(\mu_2, \sigma_2^2)$ for $j = 1, \dots, n_2$ with $(\mu_1, \sigma_1^2, \mu_2, \sigma_2^2) \in \mathbb{R} \times \mathbb{R}^+ \times \mathbb{R} \times \mathbb{R}^+$

Under $H_0 : \mu_1 = \mu_2$ it holds (approx), that

$$T := \frac{\bar{Y} - \bar{X}}{\sqrt{SEM_y^2 + SEM_x^2}} \sim t(\nu)$$

Welch vs. Student

Model 'Welch': Let $X_1, \ldots, X_{n_1}, Y_1, \ldots, Y_{n_2}$ be independent RVs with
$X_i \sim N(\mu_1, \sigma_1^2)$ for $i = 1, \ldots, n_1$ and
$Y_j \sim N(\mu_2, \sigma_2^2)$ for $j = 1, \ldots, n_2$ with $(\mu_1, \sigma_1^2, \mu_2, \sigma_2^2) \in \mathbb{R} \times \mathbb{R}^+ \times \mathbb{R} \times \mathbb{R}^+$

Under $H_0 : \mu_1 = \mu_2$ it holds (approx), that

$$T := \frac{\bar{Y} - \bar{X}}{\sqrt{SEM_y^2 + SEM_x^2}} \sim t(\nu)$$

Model 'Student': Let $X_1, \ldots, X_{n_1}, Y_1, \ldots, Y_{n_2}$ be independent RVs with
$X_i \sim N(\mu_1, \sigma^2)$ for $i = 1, \ldots, n_1$ and
$Y_j \sim N(\mu_2, \sigma^2)$ for $j = 1, \ldots, n_2$ with $(\mu_1, \mu_2, \sigma^2) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R}^+$

Under $H_0 : \mu_1 = \mu_2$ it holds (exactly), that

$$T := \frac{\bar{Y} - \bar{X}}{\sqrt{n_2^{-1} + n_1^{-1}} \cdot S_p} \sim t(n_2 + n_1 - 2)$$

It is $S_p^2 := \frac{(n_2-1)S_y^2 + (n_1-1)S_x^2}{n_2+n_1-2}$, which is often called the *pooled empirical variance*

Welch vs. Student

Model 'Welch': Let $X_1, \dots, X_{n_1}, Y_1, \dots, Y_{n_2}$ be independent RVs with
$X_i \sim N(\mu_1, \sigma_1^2)$ for $i = 1, \dots, n_1$ and
$Y_j \sim N(\mu_2, \sigma_2^2)$ for $j = 1, \dots, n_2$ with $(\mu_1, \sigma_1^2, \mu_2, \sigma_2^2) \in \mathbb{R} \times \mathbb{R}^+ \times \mathbb{R} \times \mathbb{R}^+$

Under $H_0 : \mu_1 = \mu_2$ it holds (approx), that

$$T := \frac{\bar{Y} - \bar{X}}{\sqrt{SEM_y^2 + SEM_x^2}} \sim t(\nu)$$

Model 'Student': Let $X_1, \dots, X_{n_1}, Y_1, \dots, Y_{n_2}$ be independent RVs with
$X_i \sim N(\mu_1, \sigma^2)$ for $i = 1, \dots, n_1$ and
$Y_j \sim N(\mu_2, \sigma^2)$ for $j = 1, \dots, n_2$ with $(\mu_1, \mu_2, \sigma^2) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R}^+$

Under $H_0 : \mu_1 = \mu_2$ it holds (exactly), that

$$T := \frac{\bar{Y} - \bar{X}}{\sqrt{n_2^{-1} + n_1^{-1}} \cdot S_p} \sim t(n_2 + n_1 - 2)$$

It is $S_p^2 := \frac{(n_2-1)S_y^2 + (n_1-1)S_x^2}{n_2+n_1-2}$, which is often called the *pooled empirical variance*

- downside Student: need spread to be similar in both groups
- upside Student: the test is exact and the degrees of freedom have a simple form

Welch vs. Student

Model 'Welch': Let $X_1, \ldots, X_{n_1}, Y_1, \ldots, Y_{n_2}$ be independent RVs with
$X_i \sim N(\mu_1, \sigma_1^2)$ for $i = 1, \ldots, n_1$ and
$Y_j \sim N(\mu_2, \sigma_2^2)$ for $j = 1, \ldots, n_2$ with $(\mu_1, \sigma_1^2, \mu_2, \sigma_2^2) \in \mathbb{R} \times \mathbb{R}^+ \times \mathbb{R} \times \mathbb{R}^+$

Under $H_0 : \mu_1 = \mu_2$ it holds (approx), that

$$T := \frac{\bar{Y} - \bar{X}}{\sqrt{SEM_y^2 + SEM_x^2}} \sim t(\nu)$$

Model 'Student': Let $X_1, \ldots, X_{n_1}, Y_1, \ldots, Y_{n_2}$ be independent RVs with
$X_i \sim N(\mu_1, \sigma^2)$ for $i = 1, \ldots, n_1$ and
$Y_j \sim N(\mu_2, \sigma^2)$ for $j = 1, \ldots, n_2$ with $(\mu_1, \mu_2, \sigma^2) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R}^+$

Under $H_0 : \mu_1 = \mu_2$ it holds (exactly), that

$$T := \frac{\bar{Y} - \bar{X}}{\sqrt{n_2^{-1} + n_1^{-1}} \cdot S_p} \sim t(n_2 + n_1 - 2)$$

It is $S_p^2 := \frac{(n_2-1)S_y^2 + (n_1-1)S_x^2}{n_2+n_1-2}$, which is often called the *pooled empirical variance*

- downside Student: need spread to be similar in both groups
- upside Student: the test is exact and the degrees of freedom have a simple form

For a derivation of Student's t-test see e.g., Messer, M. and Schneider, G. *Statistik: Theorie und Praxis im Dialog*,

t-test in R (according to 'Student')

```
# Enter data
x <- c(...)
y <- c(...)
# Perform t-test
t.test(y,x,var.equal=TRUE,...)
# Output

        Two Sample t-test

data:  y and x
t = 3.2846, df = 39, p-value = 0.002163
alternative hypothesis:
true difference in means is not equal to 0
95 percent confidence interval:
  4.041178 16.996539
sample estimates:
mean of x mean of y
 61.31892  50.80006
```

- The argument `var.equal` allows to assume equal variances ('Student')
- The degrees of freedom are $n_2 + n_1 - 2 = 16 + 25 - 2 = 39$

Thank you!