Comparison of means using the t-test

motiende

3/16/2021

load required packages

```
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

One-sample t-test

Here we test whether the mean from a dataset is significantly different from a hypothesized mean

```
# load birthweight dataset
birthweight <- read.csv("bw.csv", header = T)

mean(birthweight$bweight)
```

```
## [1] 3129.137
```

```
t.test(birthweight$bweight, mu= 2500)
```

```
##
##  One Sample t-test
##
## data:  birthweight$bweight
## t = 24.401, df = 640, p-value < 2.2e-16
## alternative hypothesis: true mean is not equal to 2500
## 95 percent confidence interval:
##  3078.507 3179.767
## sample estimates:
## mean of x
##  3129.137
```

Two independent sample t-test

In this example test whether the birthweight in girsl is significantky different from the birthweight in boys

```
t.test(birthweight$bweight ~ birthweight$sex)
```

```
##
##  Welch Two Sample t-test
##
## data:  birthweight$bweight by birthweight$sex
## t = -3.2686, df = 638.65, p-value = 0.001139
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -267.57246  -66.73186
## sample estimates:
## mean in group Female   mean in group Male
##             3044.127             3211.279
```

Paired t-test

As an example of data, 20 newborns received a vitamin supplement X during 3 months. We want to know whether the supplement X has an impact on weight gain.

To answer to this question, the weight of the 20 children was measured before they were started on the supplements and after the 3 months. This gives us 20 sets of values before introduction of the supplements and 20 sets of values after the supplement was given. The 'pairdness' is resulting from measuring twice the weight of the same children - before and after giving the supplements.

In such situations, paired t-test can be used to compare the mean weights before and after treatment.

```
# Weight (in grams) of the children before giving supplements
before <-c(2000.1, 1090.9, 1920.7, 2130, 2401.4, 1960.9, 1702.2, 1085.5, 2005.2, 1930.7)

# Weight of the children after treatment
after <-c(3920.9, 3903.2, 3450.1, 3903, 4340, 4270.9, 4220, 3830.9, 3920.3, 3502.2)

# Create a data frame
my_data <- data.frame(
                group = rep(c("before", "after"), each = 10),
                weight = c(before,  after)
                )
```

```
## # A tibble: 2 x 4
##   group  count  mean    sd
##   <chr>  <int> <dbl> <dbl>
## 1 after     10 3926.  297.
## 2 before    10 1823.  425.
```

```
##
##  Paired t-test
##
## data:  before and after
```

```
## t = -14.302, df = 9, p-value = 1.705e-07
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -2436.083 -1770.697
## sample estimates:
## mean of the differences
##                -2103.39
```