# СТАТИСТИКА И ЕМПИРИЧНИ МЕТОДИ

# Домашно 3

Антоан Венциславов Стефанов

Ф.Н.: 61797

гр.5

Софтуерно инженерство

*22.01.2017*

# Задача 1.

a) Формулирайте хипотеза:

Нека :

r\_e: “Резултатите от измерванията на експерта“

r\_m: “Резултатите от измерванията на машината“

H\_0 : r\_e == r\_m

H\_a : r\_e != r\_m

б) К.О. при α = 0.1

> qt(0.05, length(blood) - 1, lower.tail = FALSE)

[1] 6.313752

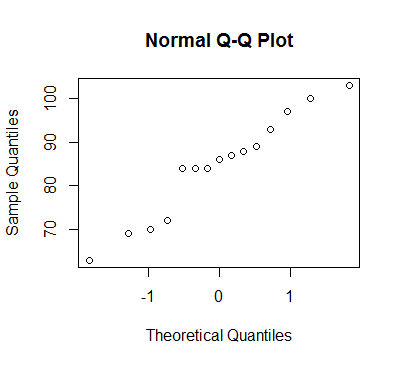
К.О. при α = 0.1 е (-безкр. , -6.313752) u (6.313752, +безкр.)

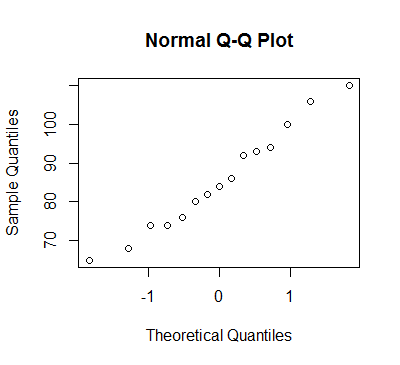
в) Тествайте хипотезата

> expert = blood$Expert

> machine = blood$Machine

>

> qqnorm(expert)

> qqnorm(machine)

>

> t.test(expert, machine, paired = TRUE)

Paired t-test

data: expert and machine

t = -0.68162, df = 14, p-value = 0.5066

alternative hypothesis: true difference in means is not equal to 0

95 percent confidence interval:

-4.146615 2.146615

sample estimates:

mean of the differences

-1

От теста не можем да отхвърлим нулевата хипотеза, защото p-value е много голямо (0.5).

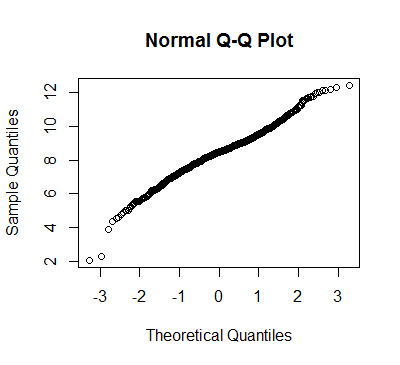
Това означава, че приликата между данните имат 50% шанс да бъдат произволно породени.

Тест статистиката (t) показва връзката между двете извадки от данни.

95% доверителен интервал има съответния процентов шанс да съдържа истинската стойност на случайната величина.

# Задача 2.

а) Удачно е да се използва тест за средната стойност, защото можем да приемем, че разпределението на случайната величина „Стойност на яркост“ е нормално.

> qqnorm(brightness)

б) Постройте 93% доверителен интервал за средната стойност:

> t.test(brightness, conf.level = 0.93)

One Sample t-test

data: brightness

t = 202.15, df = 965, p-value < 2.2e-16

alternative hypothesis: true mean is not equal to 0

93 percent confidence interval:

8.342209 8.493277

sample estimates:

mean of x

8.417743

# Задача 3.

От 1337 души, 1000 отговарят, че използват интернет всеки ден. Използвайте това, за да тествате твърдението, че повече от 70% от хората използват интернет всеки ден,

при ниво на съгласие α=0.05.

> prop.test(1000, 1337, alt = "greater", p = 0.7, conf.level = 0.95)

1-sample proportions test with continuity correction

data: 1000 out of 1337, null probability 0.7

X-squared = 14.407, df = 1, p-value = 7.364e-05

alternative hypothesis: true p is greater than 0.7

95 percent confidence interval:

0.727541 1.000000

sample estimates:

p

0.7479432

От това, че p-value-то е много малко и, че стойността, за която проверяваме (0.7) е по-малка от 0.727541 (от доверителния интервал), можем да приемем, че повече от 70% от хората използват интернет всеки ден.