

# Lista 3

---

Mateusz Lewko

## Zadanie 1

Wariancje nieznane i różne

```
dane <- as.matrix(read.csv("w-03-01.csv", sep=';'))
(X <- dane[,1]) # M
(Y <- dane[,2]) # K
var(X)
var(Y)

# funkcja pomocnicza
printf <- function(...) cat(sprintf(...))

# test dwóch średnich, wariancje nieznane i różne
test_a4 <- function(X, Y, mi_diff) {
  nx <- length(X)
  ny <- length(Y)

  s2x <- var(X)
  s2y <- var(Y)

  # s <- ((nx - 1) * s2x + (ny - 1) * s2y) / (nx +
ny - 2)
  a <- s2x / nx + s2y / ny
  t <- (mean(X) - mean(Y) - mi_diff) / sqrt(a)
  n <- a^2 / (s2x^2 / (nx^2 * (nx - 1)) + s2y^2 /
(ny^2 * (ny - 1)))
  n <- as.integer(n)

  printf("t = %.4f; qt(0.975, df=%d) = %.4f;
pt(%.4f, df=%d) = %.4f\n",
```

```

t, n, qt(0.975, df=n), t, n, pt(t, df=n))

half <- mean(X) - mean(Y)
left <- half - qt(0.975, df=n) * sqrt(a)
right <- half + qt(0.975, df=n) * sqrt(a)

printf("half: %.4f, left: %.4f, right: %.4f\n",
half, left, right)
}

test_a4(X, Y, 0)

# t = 3.0594; qt(0.975, df=13) = 2.1604; pt(3.0594,
df=13) = 0.9954
# half: 5.2222, left: 1.5346, right: 8.9098

```

a) statystyka testowa = 3.0594,  $qt(0.975, df=13) = 2.1604$ , odrzucamy  $H_0$

b) przedział ufności dla różnicy wartości średnich: [1.5346, 8.9098]

## Zadanie 2

```

dane <- as.matrix(read.csv("w-03-02.csv", sep=';'))
(X <- dane[,1]) # M
(Y <- dane[,2]) # K
var(X) # 7.708333
var(Y) # 9.212585

# wariancje nieznane i różne

test_a4(X, Y, 2)

# t = 6.3901; qt(0.975, df=95) = 1.9853; pt(6.3901,
df=95) = 1.0000
# half: 5.7551, left: 4.5885, right: 6.9217

```

wartość\_p bliska 0, odrzucamy hipotezę  $H_0$

## Zadanie 3

Testujemy hipotezę  $H_0: m_{\text{przed}} = m_{\text{po}}$  (dieta nie ma wpływu) wobec  $H_1: m_{\text{przed}} > m_{\text{po}}$  (dieta ma wpływ)

```
dane <- as.matrix(read.csv("w-03-03.csv", sep=';'))
(X <- dane[,1]) # Przed
(Y <- dane[,2]) # Po
var(X) # 29.81029
var(Y) # 36.37981

# wariancje nieznane i różne

test_a4(X, Y, 0)
# t = 2.1581; qt(0.975, df=27) = 2.0518; pt(2.1581,
df=27) = 0.9800
# half: 4.5333, left: 0.2232, right: 8.8435

2 * (1 - 0.9800) # 0.04
```

wartość\_p: 0.04, można odrzucić  $H_0$  (dieta ma wpływ na wagę)

## Zadanie 4

```
n <- 200
p1 <- 159 / n # K
p2 <- 138 / n # M
p <- (n * p1 + n * p2) / (2 * n)

z <- (p1 - p2) / (sqrt(p * (1 - p) * (2/n)))
z # 2.401333
qnorm(0.975) # 1.959964
```

```
2 * (1 - pnorm(z))  
# 0.01633545
```

wartość\_p: 0.01633545, odrzucamy  $H_0 \Rightarrow$  kobiety częściej używają komunikatorów

## Zadanie 6

```
dane <- as.matrix(read.csv("w-03-06.csv", sep=';'))  
(X <- dane[,1])  
(Y <- dane[,2])  
  
var(X)  
# [1] 563.8673  
var(Y)  
# [1] 700.9337  
  
test(Y, X, 200)  
# t = 4.8122; qt(0.975, df=94) = 1.9855; pt(4.8122,  
df=94) = 1.0000  
# half: 224.4490, left: 214.3614, right: 234.5366
```

wartość\_p bliska 0, odrzucamy  $H_0 \Rightarrow$  wartość ubezpieczenia wzrosła o co najmniej 200zł

## Zadanie 7

```
dane <- as.matrix(read.csv("w-03-07.csv", sep=';'))  
(X <- dane[,1])  
(Y <- dane[,2])  
var(X)  
var(Y)  
  
# 58.99451
```

```
# 46.41758

test_a4(Y, X, 0)
# t = 0.8590; qt(0.975, df=25) = 2.0595; pt(0.8590,
df=25) = 0.8008
# half: 2.3571, left: -3.2942, right: 8.0085

2 * (1 - 0.8008)
```

wartość\_p: 0.3984, nie ma podstaw do odrzucenia H0 (nowy sposób nauczania daje podobne wyniki)

## Zadanie 8

```
dane <- as.matrix(read.csv("w-03-08.csv", sep=';'))
(X <- dane[,1])
(Y <- dane[,2])

var(X)
# [1] 254.8611
var(Y)
# [1] 195.6111

test_a4(X, Y, 0)
# t = 0.5968; qt(0.975, df=15) = 2.1314; pt(0.5968,
df=15) = 0.7202
2 * (1 - 0.7202)
# [1] 0.5596
```

wartość\_p: 0.5596, dodatkowy czynnik nie ma wpływu na jakość uprawy