

Obvezna domača naloga 4

Linearni modeli

Alen Kahteran

9. 12. 2020

3.1 model.spol

```
# read data
tlak <- read.table(file="SKT.txt", header=TRUE, stringsAsFactors=TRUE)
# create linear model based on spol
model.spol <- lm(data=tlak, SKT~spol)

# create matrix of coefficients to be tested
test_matr <- matrix(c(1,0,1,1), nrow=2, byrow=TRUE)
# rename for easier viewing
colnames(test_matr) <- c("b0", "b1")
rownames(test_matr) <- c("M", "F")

# use glht with test_matr for multiple testing
model.spol_test <- glht(model.spol, linfct=test_matr)
# get values
summary(model.spol_test)

##
## Simultaneous Tests for General Linear Hypotheses
##
## Fit: lm(formula = SKT ~ spol, data = tlak)
##
## Linear Hypotheses:
##      Estimate Std. Error t value Pr(>|t|)
## M == 0  155.150      2.682   57.84  <1e-10 ***
## F == 0  139.862      3.150   44.40  <1e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## (Adjusted p values reported -- single-step method)

confint(model.spol_test)

##
## Simultaneous Confidence Intervals
##
## Fit: lm(formula = SKT ~ spol, data = tlak)
##
## Quantile = 2.2862
## 95% family-wise confidence level
##
##
```

```
## Linear Hypotheses:
##      Estimate lwr      upr
## M == 0 155.1500 149.0176 161.2824
## F == 0 139.8621 132.6600 147.0642
```

Povprečna vrednost SKT pri moških je bila 155.15mm z intervalom zaupanja [149.02mm, 161.28mm] medtem ko je bila pri ženskah 139.86mm z intervalom zaupanja [132.66mm, 147.06mm] pri stopnji tveganja $\alpha = 0.05$.

3.2 model.razlicne

```
# get correct data
tlak$skupina <- factor(rep(c("A", "B", "C"),
                          times=c(20, 20, 29)))
tlak$starost.50 <- tlak$starost - 50
model.razlicne <- lm(data=tlak, SKT~skupina*starost.50)
```

```
# create matrix of coefficients to be tested
test_matr <- matrix(c(0, 1, 0, 0, 0, 0,
                      0, 0, 1, 0, 0, 0,
                      0, -1, 1, 0, 0, 0,
                      0, 0, 0, 1, 0, 0,
                      0, 0, 0, 0, 1, 0,
                      0, 0, 0, 0, 0, 1,
                      0, 0, 0, 0, -1, 1),
                    nrow=7,
                    byrow=TRUE)
```

```
# rename for easier viewing
rownames(test_matr) <- c("mu|B-A", "mu|C-A", "mu|C-B",
                        "k|A", "k|B-A", "k|C-A", "k|C-B")

colnames(test_matr) <- c("b0", "b1", "b2", "b3", "b4", "b5")
test_matr
```

```
##      b0 b1 b2 b3 b4 b5
## mu|B-A  0  1  0  0  0  0
## mu|C-A  0  0  1  0  0  0
## mu|C-B  0 -1  1  0  0  0
## k|A      0  0  0  1  0  0
## k|B-A    0  0  0  0  1  0
## k|C-A    0  0  0  0  0  1
## k|C-B    0  0  0  0 -1  1
```

```
# use glht with test_matr for multiple testing
test.razlicne <- glht(model.razlicne, linfct=test_matr)
# get values
summary(test.razlicne)
```

```
##
## Simultaneous Tests for General Linear Hypotheses
##
## Fit: lm(formula = SKT ~ skupina * starost.50, data = tlak)
##
## Linear Hypotheses:
##      Estimate Std. Error t value Pr(>|t|)
```

```
## mu|B-A == 0  2.24975    2.91292    0.772 0.935081
## mu|C-A == 0 -12.37731    2.68078   -4.617 0.000105 ***
## mu|C-B == 0 -14.62706    2.72917   -5.360 < 1e-04 ***
## k|A == 0      1.03526    0.13512    7.662 < 1e-04 ***
## k|B-A == 0   -0.13881    0.19416   -0.715 0.951601
## k|C-A == 0   -0.08594    0.17370   -0.495 0.989400
## k|C-B == 0    0.05287    0.17707    0.299 0.998943
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## (Adjusted p values reported -- single-step method)
```

```
confint(test.razlicne)
```

```
##
## Simultaneous Confidence Intervals
##
## Fit: lm(formula = SKT ~ skupina * starost.50, data = tlak)
##
## Quantile = 2.7076
## 95% family-wise confidence level
##
## Linear Hypotheses:
##      Estimate   lwr      upr
## mu|B-A == 0    2.24975 -5.63737 10.13686
## mu|C-A == 0   -12.37731 -19.63588 -5.11875
## mu|C-B == 0   -14.62706 -22.01665 -7.23747
## k|A == 0       1.03526  0.66940  1.40112
## k|B-A == 0    -0.13881 -0.66454  0.38691
## k|C-A == 0    -0.08594 -0.55625  0.38437
## k|C-B == 0     0.05287 -0.42658  0.53232
```

1. Povprečni SKT med skupino A in skupino B ni statistično značilno drugačen ($p = 0.94$).
2. Povprečni SKT med skupino A in skupino C je statistično značilno drugačen ($p = 0.0001$). Povprečni SKT skupine A je višji za $12.38mm$ s pripadajočim intervalom zaupanja, pri $\alpha = 0.05$, $[5.12mm, 19.63mm]$.
3. Podobno kot pri prejšnji točki je povprečni SKT med skupino B in C statistično značilno drugačen ($p < 10^{-4}$). Povprečni SKT skupine B je višji za $14.63mm$ s pripadajočim intervalom zaupanja, pri $\alpha = 0.05$, $[7.24mm, 22.01mm]$.
4. Ko gledamo odvisnost SKT skupine A od starosti, vidimo da je statistično značilno ($p < 10^{-4}$) drugačna od 0. V povprečju se vsako leto SKT poveča za $1.04mm$. Interval zaupanja ($\alpha = 0.05$) je $[0.67mm, 1.40mm]$.
5. Razlike odvisnosti SKT od starosti po skupinah (BA , CA in CB), niso statistično značilno drugačne od 0.