Ukol 2

Moznost práce az ve skupine 4 lidí. Deadline je 21.5. Cím dríve to vsak poslete tím lepe. Práci poslete na lukas.fryd@gmail.com.

Problem 1

- *violent* crime rate (incidents per 100,000 members of the population).
- *murder* murder rate (incidents per 100,000).
- *robbery* robbery rate (incidents per 100,000).
- $afam$ percent of state population that is African-American, ages 10 to 64.
- *income* real per capita personal income in the state (US dollars).
- *density* population per square mile of land area, divided by 1,000.

První regresní funkce:

```
regrese1=lm(log(murder)~violent+robbery+afam+income+density,data = Guns)
summary(regrese1)
```

```
##
## Call:
## lm(formula = log(murder) ~ violent + robbery + afam + income +
##     density, data = Guns)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.61436 -0.21320  0.03374  0.25911  1.17988
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  1.690e+00  6.652e-02  25.411   <2e-16 ***
## violent      1.629e-03  8.374e-05  19.458   <2e-16 ***
## robbery      4.960e-04  1.941e-04   2.555   0.0107 *
## afam         4.032e-02  2.956e-03  13.641   <2e-16 ***
## income      -7.085e-05  4.937e-06 -14.352   <2e-16 ***
## density     -1.434e-01  1.401e-02 -10.236   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.391 on 1167 degrees of freedom
## Multiple R-squared:  0.6923, Adjusted R-squared:  0.691
## F-statistic: 525.1 on 5 and 1167 DF,  p-value: < 2.2e-16
```

Druhá regresní funkce:

```
regrese2=lm(murder~violent+robbery+afam+income+density,data = Guns)
summary(regrese2)
```

```
##
## Call:
```

```
## lm(formula = murder ~ violent + robbery + afam + income + density,
##     data = Guns)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -20.3920  -1.3722  -0.1531   1.2278  28.7689
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  5.982e+00  5.891e-01  10.155  < 2e-16 ***
## violent      1.527e-02  7.415e-04  20.592  < 2e-16 ***
## robbery     -5.015e-03  1.719e-03  -2.917   0.0036 **
## afam         1.844e-01  2.617e-02   7.043  3.2e-12 ***
## income      -5.037e-04  4.372e-05 -11.522  < 2e-16 ***
## density      2.110e+00  1.241e-01  17.001  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.463 on 1167 degrees of freedom
## Multiple R-squared:  0.789,  Adjusted R-squared:  0.7881
## F-statistic: 872.8 on 5 and 1167 DF,  p-value: < 2.2e-16
```

Testy

```
residua=regrese1$residuals
regrese3=lm(residua^2~violent+robbery+afam+income+density,data = Guns)
summary(regrese3)
```

```
##
## Call:
## lm(formula = residua^2 ~ violent + robbery + afam + income +
##     density, data = Guns)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.4973 -0.1303 -0.0694  0.0163  6.6171
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  1.850e-01  5.383e-02   3.436 0.000611 ***
## violent     -2.286e-04  6.776e-05  -3.373 0.000767 ***
## robbery     -9.385e-05  1.571e-04  -0.598 0.550269
## afam         7.533e-03  2.392e-03   3.150 0.001675 **
## income       2.854e-06  3.995e-06   0.714 0.475148
## density      5.104e-02  1.134e-02   4.501 7.44e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3164 on 1167 degrees of freedom
## Multiple R-squared:  0.04374,    Adjusted R-squared:  0.03965
## F-statistic: 10.68 on 5 and 1167 DF,  p-value: 4.755e-10
```

```r
bptest(murder~violent+robbery+afam+income+density,data = Guns)
```

```
##
##  studentized Breusch-Pagan test
##
## data:  murder ~ violent + robbery + afam + income + density
## BP = 807.03, df = 5, p-value < 2.2e-16
```

- Napiste populacní regresní funkci pro odhad $regrese1$ (1)

- Jaký je porusen GM a proc? Jaký to má dopad? Jakou metodu byste pouzili pro nápravu. $regrese1$ (3)

- Jaká je interpretace odhadu parametru pro income? $regrese1$ (1)

- Který model byste vybrali? $regrese1$ vs $regrese2$ a proc? (1)

- Spoctete pro regresi 2, $E(murder|violent = 420, robbery = 100, afam = 8, income = 0, density = 0)$ (2)

Problem 2

- *age* of the borrower
- *school* Years of schooling for the borrower.
- *interest* Fixed interest rate.
- *networth* - Net worth of the borrower.
- *married* Factor. Is the borrower married?
- *liability* Dluhy
- *assets* Aktiva

```
v=rnorm(N,1,1)

age=runif(N,20,60)
age2=age^2

school=runif(N,8,15)
school2=school^2

interest=rlnorm(N, meanlog = 4, sdlog = 2)+v

networth=rlnorm(N, meanlog = 10, sdlog = 3)


sigma=1+exp(age)
e=rnorm(N,0,sigma)+v
```

```
regrese=lm(liabilities~age+age2+school+school2+interest+log(networth)+married+assets)
summary(regrese)
```

```
##
## Call:
## lm(formula = liabilities ~ age + age2 + school + school2 + interest +
##     log(networth) + married + assets)
##
## Residuals:
##        Min         1Q     Median         3Q        Max
## -1.402e+26 -1.187e+24  1.307e+23  1.423e+24  1.595e+26
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)   -3.753e+25  1.583e+25  -2.372   0.0179 *
## age            5.519e+23  2.763e+23   1.997   0.0461 *
## age2          -7.673e+21  3.421e+21  -2.243   0.0251 *
## school         4.146e+24  2.634e+24   1.574   0.1159
## school2       -1.875e+23  1.148e+23  -1.633   0.1028
## interest      -1.956e+20  3.763e+20  -0.520   0.6034
## log(networth)  3.000e+23  1.381e+23   2.172   0.0301 *
## married        4.773e+23  8.335e+23   0.573   0.5670
## assets         3.024e+21  1.411e+21   2.143   0.0323 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.314e+25 on 991 degrees of freedom
```

```
## Multiple R-squared:  0.02056,    Adjusted R-squared:  0.01265
## F-statistic:   2.6 on 8 and 991 DF,  p-value: 0.008154
```

- Které GM jsou poruseny? Jsou celkem 2. Napsat které a jak jste to poznali. Jaké to má dopady? (4)

- Jaké metody byste pouzili k náprave. (3)

- Interpretujte odhad parametru pro log(networth) a married. (2)

- Interpretujte dopad veku na dluhy. (1)

- Jaký bude rozdíl mezi osobou A a B. Kdy osoba A bude zenatá/vdaná a bude navíc vlastnit akcie v hodnote 1000 USD. (2)