R skripta LRM

Tihana Škrinjarić

Last edited on 23. veljača, 2023

# POGLAVLJE 1: TEMELJNI POJMOVI REGRESIJSKOG MODELA

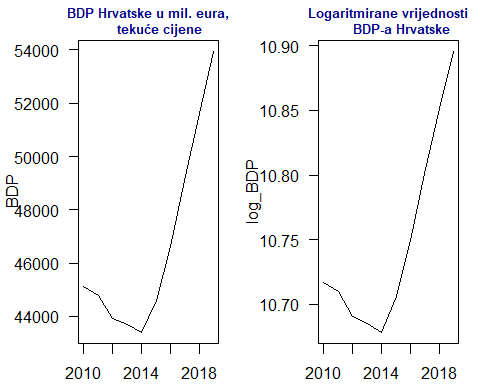
## 1.4. Priprema podataka za ekonometrijsku analizu i grafičko predočavanje

#### Slika 1.1. Usporedba BDP-a Hrvatske u razinama, te logaritmirane vrijednosti

library(readxl)  
BDP\_hrvatske <- read\_excel("BDP hrvatske.xlsx")  
BDP<-ts(BDP\_hrvatske$`BDP (tekuce cijene, mil E)`,start=c(2010),frequency = 1)  
log\_BDP<-log(BDP)

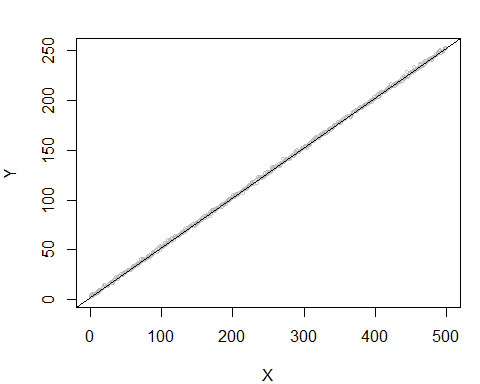
Crtanje dvije slike, u jednom retku i dva stupca:

par(mfrow=c(1,2),mar=c(2,4,2,1)+0.1)  
par(las=1)  
plot(BDP)  
title("BDP Hrvatske u mil. eura,  
 tekuće cijene", cex.main = 0.8, col.main = "darkblue", font.main = 2)  
plot(log\_BDP)  
title ("Logaritmirane vrijednosti  
 BDP-a Hrvatske",cex.main = 0.8, col.main ="darkblue", font.main=2)



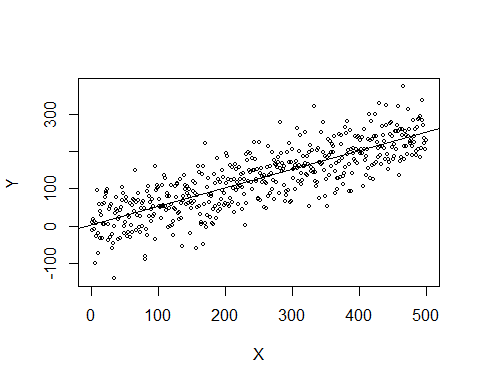
#### Slika 1.2. Dijagram rasipanja između varijabli x i y, jaka pozitivna linearna veza

x<-seq(1,500,1)  
epsilon<-rnorm(500,0,1)  
y<-2+0.5\*x+epsilon  
par(mfrow = c(1, 1),mar=c(4,4,2,1))  
  
plot(x,y,cex=0.5,col="grey",ylab="Y",xlab="X")  
abline(2,0.5)



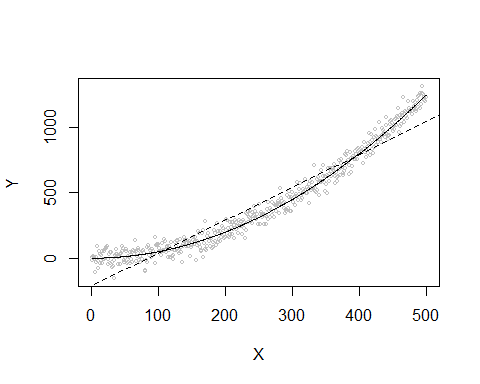
#### Slika 1.3. Dijagram rasipanja između varijabli x i y, umjerena pozitivna linearna veza

x<-seq(1,500,1)  
epsilon2<-rnorm(500,0,50)  
y2<-2+0.5\*x+epsilon2  
  
plot(x,y2,cex=0.5,ylab="Y",xlab="X")  
abline(2,0.5)



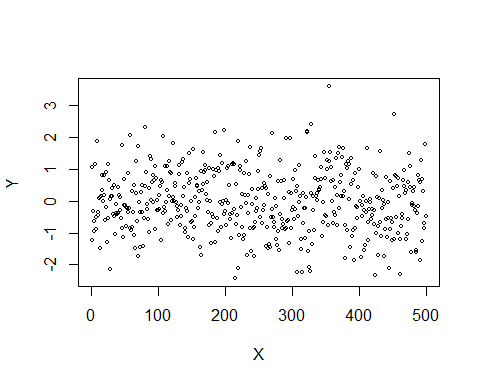
#### Slika 1.4. Dijagram rasipanja između varijabli x i y, parabolična veza

x<-seq(1,500,1)  
y3<-(0.005)\*x^2+epsilon2  
  
plot(x,y3,cex=0.5,col="grey",ylab="Y",xlab="X")  
abline(lm(y3~x),lty="dashed")  
lines(fitted(lm(y3~0+I(x^2))))



#### Slika 1.5. Dijagram rasipanja između varijabli x i y, ne postoji veza

x<-seq(1,500,1)  
y4<-rnorm(500,0,1)  
  
plot(x,y4,cex=0.5,ylab="Y",xlab="X")



#### Primjer 1.5.

getwd()

## [1] "C:/Users/njerak/Desktop/R datoteke LRM"

podaci<-read.table("BDP\_i\_HICP.txt",header=T, sep="\t")  
  
y<-podaci$GDP  
x<-podaci$HICP

#### Slika 1.6. Ispis izvršene naredbe i radnog direktorija

getwd()

## [1] "C:/Users/njerak/Desktop/R datoteke LRM"

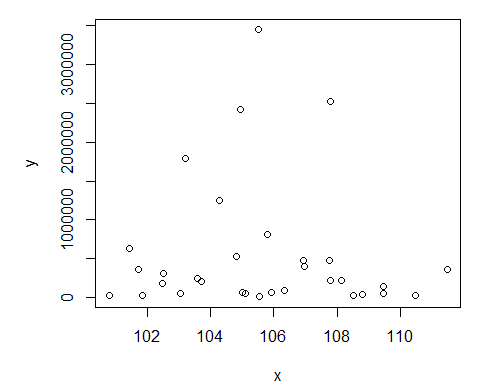
podaci<-read.table("BDP\_i\_HICP.txt",header=T, sep="\t")

#### Slika 1.7. Učitavanje datoteke BDP\_i\_HICP.txt

y<-podaci$GDP  
x<-podaci$HICP

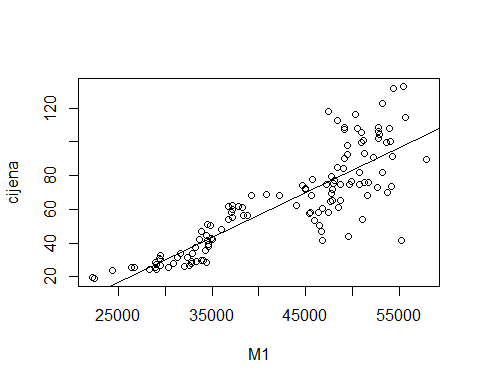
#### Slika 1.10. Dijagram rasipanja između HICP-a (x) i BDP-a (y)

par(mfrow=c(1,1),mar=c(4,5,1,1))  
plot(x,y)

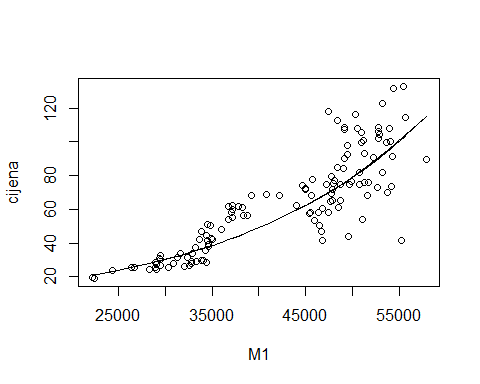


#### 1.5. Pitanja za ponavljanje, 8 zadatak

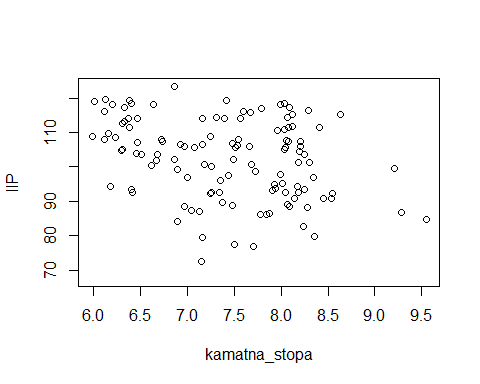
library(readxl)  
makro\_varijable <- read\_excel("makro varijable.xls")  
  
cijena<-makro\_varijable$NAFTA  
M1<-makro\_varijable$M1  
  
#slika lijevi panel  
plot(M1,cijena)  
abline(lm(cijena~M1))



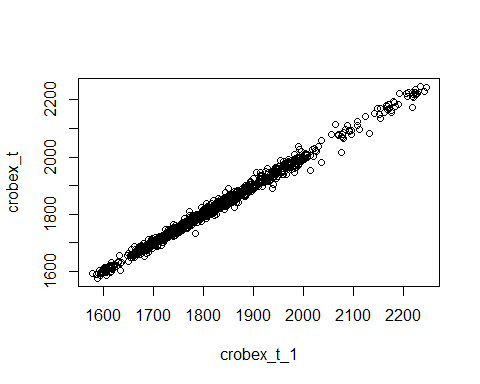
#slika desni panel  
plot(M1,cijena)  
kvadrat<-lm(log(cijena)~M1)  
b<-as.vector(rep(NA,length(M1)))  
for(i in 1:length(cijena))  
{b[i]<-2.7^(kvadrat[["coefficients"]][1]+kvadrat[["coefficients"]][2]\*M1[i])  
}  
points(M1,b, type="l")



#slika treci panel (IIP i K stopa)  
IIP<-makro\_varijable$IIP  
kamatna\_stopa<-makro\_varijable$kamatna\_stopa  
plot(kamatna\_stopa,IIP)

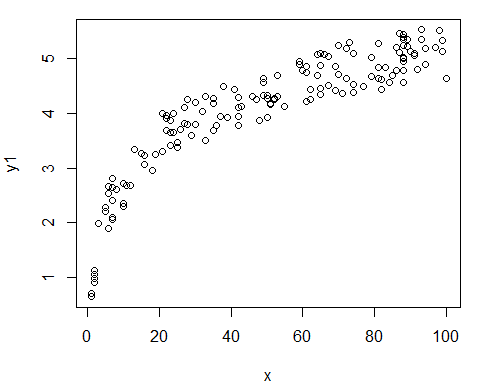


#Crobex slika  
library(readxl)  
crobex <- read\_excel("crobex.xlsx")  
crobex\_t<-crobex$crobex  
crobex\_t\_1<-c(NA,crobex\_t[-1225])  
plot(crobex\_t\_1,crobex\_t)

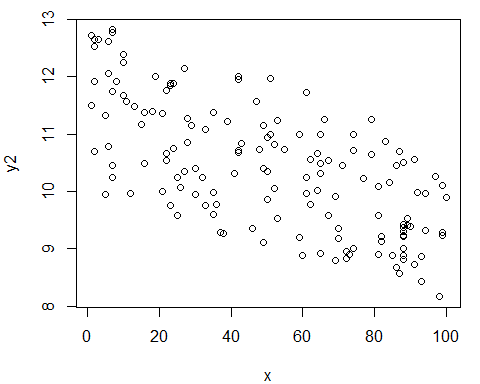


#### 1.5. Pitanja za ponavljanje, zadatak 10

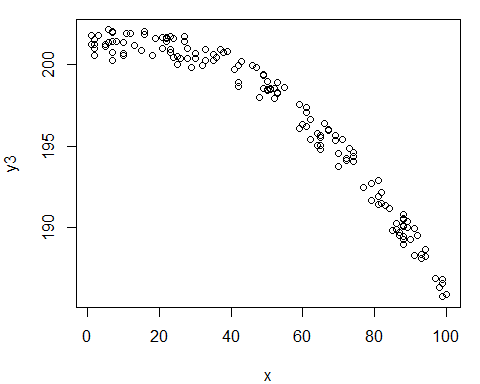
podaci<-read.table("dijagram.txt",header=T,sep="\t")  
y1<-podaci$y1  
y2<-podaci$y2  
y3<-podaci$y3  
x<-podaci$x  
  
par(mfrow = c(1, 1),mar=c(4,4,1,1))  
plot(x,y1)



plot(x,y2)



plot(x,y3)



# POGLAVLJE 2: LINEARNI REGRESIJSKI MODEL

## 2.1. Model jednostavne linearne regresije

##### 2.1.3.1. Metoda najmanjih kvadrata

#### Slika 2.5. Unos potrebnih naredbi kako bi se procijenili parametri regresijskog modela u primjeru 2.1.

x<-c(10,15,12,7,4,14,22,1)  
y<-c(18,29,21,11,7,25,44,1)  
  
x\_potez<-mean(x)  
y\_potez<-mean(y)  
  
umnozak<-t(x)%\*%(y)  
x\_potez\_2<-mean(x)^2  
x\_2<-sum(x^2)  
  
beta\_1<-(umnozak-8\*x\_potez\*y\_potez)/(x\_2-8\*x\_potez\_2)  
beta\_0<-y\_potez-beta\_1\*x\_potez

#### Slika 2.6. Ispis procijenjenih parametara modela u primjeru 2.1.

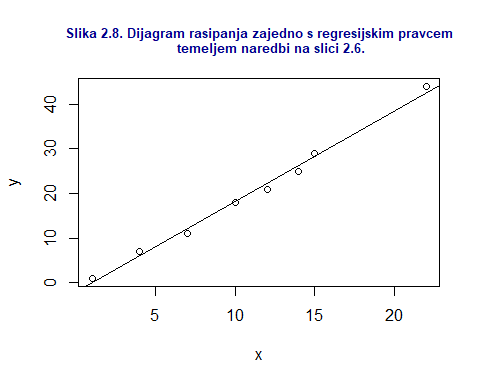
beta\_1;beta\_0

## [,1]  
## [1,] 2.031263

## [,1]  
## [1,] -2.082164

#### Slika 2.7. Naredbe potrebne za prikaz dijagrama rasipanja i regresijskog pravca

plot(x,y)  
title("Slika 2.8. Dijagram rasipanja zajedno s regresijskim pravcem  
 temeljem naredbi na slici 2.6.",cex.main = 0.8, col.main ="darkblue", font.main=2)  
abline(a=beta\_0,b=beta\_1)



#### Slika 2.9. Naredbe potrebne za izračun procijenjenih vrijednosti i rezidualnih odstupanja

procjena<-beta\_0+beta\_1\*x  
rezidual<-y-procjena  
procjena

## [1] 18.2304609 28.3867735 22.2929860 12.1366733 6.0428858 26.3555110 42.6056112  
## [8] -0.0509018

rezidual

## [1] -0.2304609 0.6132265 -1.2929860 -1.1366733 0.9571142 -1.3555110 1.3943888  
## [8] 1.0509018

#### Slika 2.10. Ispis temeljem naredbi na slici 2.9

procjena<-beta\_0+beta\_1\*x  
rezidual<-y-procjena  
procjena

## [1] 18.2304609 28.3867735 22.2929860 12.1366733 6.0428858 26.3555110 42.6056112  
## [8] -0.0509018

rezidual

## [1] -0.2304609 0.6132265 -1.2929860 -1.1366733 0.9571142 -1.3555110 1.3943888  
## [8] 1.0509018

##### 2.1.3.2. Metoda najmanjih kvadrata u matričnom zapisu

#### Slika 2.11. Naredbe za procjenu parametara regresijskog modela u primjeru 2.2.

x<-c(10,15,12,7,4,14,22,1)  
y<-c(18,29,21,11,7,25,44,1)  
  
y<-as.matrix(y)  
jed<-c(rep(1,each=8))  
jed<-as.matrix(jed)  
  
x<-as.matrix(x)  
X<-cbind(jed,x)  
  
a<-t(X)%\*%X  
b<-solve(a)  
c<-t(X)%\*%y  
beta<-b%\*%c

#### Slika 2.12. Ispis procijenjenih parametara modela u primjeru 2.2.

beta

## [,1]  
## [1,] -2.082164  
## [2,] 2.031263

#### Slika 2.13. Naredbe potrebne za izračun procijenjenih vrijednosti i rezidualnih odstupanja

procjena<-X%\*%beta  
rezidual<-y-procjena  
procjena

## [,1]  
## [1,] 18.2304609  
## [2,] 28.3867735  
## [3,] 22.2929860  
## [4,] 12.1366733  
## [5,] 6.0428858  
## [6,] 26.3555110  
## [7,] 42.6056112  
## [8,] -0.0509018

rezidual

## [,1]  
## [1,] -0.2304609  
## [2,] 0.6132265  
## [3,] -1.2929860  
## [4,] -1.1366733  
## [5,] 0.9571142  
## [6,] -1.3555110  
## [7,] 1.3943888  
## [8,] 1.0509018

#### Slika 2.14. Ispis temeljem naredbi na slici 2.13.

procjena<-X%\*%beta  
rezidual<-y-procjena  
procjena

## [,1]  
## [1,] 18.2304609  
## [2,] 28.3867735  
## [3,] 22.2929860  
## [4,] 12.1366733  
## [5,] 6.0428858  
## [6,] 26.3555110  
## [7,] 42.6056112  
## [8,] -0.0509018

rezidual

## [,1]  
## [1,] -0.2304609  
## [2,] 0.6132265  
## [3,] -1.2929860  
## [4,] -1.1366733  
## [5,] 0.9571142  
## [6,] -1.3555110  
## [7,] 1.3943888  
## [8,] 1.0509018

##### 2.1.3.4. Algebarska svojstva metode najmanjih kvadrata

#### Slika 2.16. Naredbe za provjeru algebarskih svojstava metode najmanjih kvadrata

#svojstvo 1:  
round(sum(rezidual),6)

## [1] 0

#svojstvo 2:  
mean(y)==beta\_0+beta\_1\*mean(x)

## [,1]  
## [1,] TRUE

#svojstvo 3:  
mean(y)==mean(procjena)

## [1] FALSE

#### Slika 2.17. Rezultat temeljem naredbi zadanih na slici 2.16.

#svojstvo 1:  
round(sum(rezidual),6)

## [1] 0

#svojstvo 2:  
mean(y)==beta\_0+beta\_1\*mean(x)

## [,1]  
## [1,] TRUE

#svojstvo 3:  
mean(y)==mean(procjena)

## [1] FALSE

#### Slika 2.18. Naredbe za provjeru algebarskih svojstava metode najmanjih kvadrata

round(sum(rezidual),6)

## [1] 0

mean(y); beta[1,]+beta[2,]\*mean(x); mean(procjena)

## [1] 19.5

## [1] 19.5

## [1] 19.5

#### Slika 2.19. Rezultat temeljem naredbi zadanih na slici 2.18

round(sum(rezidual),6)

## [1] 0

mean(y); beta[1,]+beta[2,]\*mean(x); mean(procjena)

## [1] 19.5

## [1] 19.5

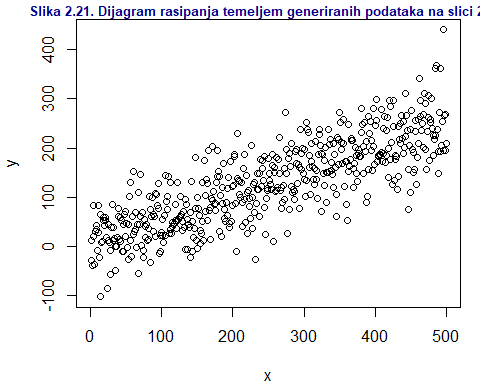
## [1] 19.5

#### Slika 2.20. Generiranje podataka za varijable x i y

set.seed(1)  
x<-seq(1,500,1)  
epsilon<-rnorm(500,0,50)  
y<-(2+0.5\*x+epsilon)

#### Slika 2.21. Dijagram rasipanja temeljem generiranih podataka na slici 2.19.

par(mfrow = c(1, 1),mar=c(4,4,1,1))  
plot(x,y)  
title("Slika 2.21. Dijagram rasipanja temeljem generiranih podataka na slici 2.19.", cex.main = 0.8, col.main ="darkblue", font.main=2)



#### Slika 2.22. Ispis temeljem naredbe lm(y~x)

lm(y~x)

##   
## Call:  
## lm(formula = y ~ x)  
##   
## Coefficients:  
## (Intercept) x   
## 5.637 0.490

#### Slika 2.23. Naredbe za izračun procijenjenih vrijednosti i tablice usporedbe vrijednosti

procjena<-fitted(lm(y~x))  
head(cbind(y,procjena,y-procjena))

## y procjena   
## 1 -28.82269 6.126688 -34.949379  
## 2 12.18217 6.616690 5.565476  
## 3 -38.28143 7.106693 -45.388123  
## 4 83.76404 7.596695 76.167346  
## 5 20.97539 8.086697 12.888692  
## 6 -36.02342 8.576699 -44.600118

#### Slika 2.24. Ispis stvarnih, procijenjenih vrijednosti zavisne varijable te rezidualnih odstupanja

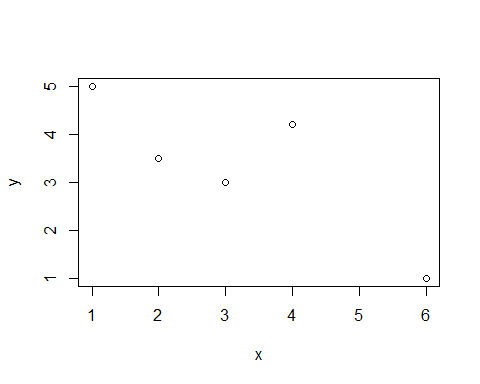
procjena<-fitted(lm(y~x))  
head(cbind(y,procjena,y-procjena))

## y procjena   
## 1 -28.82269 6.126688 -34.949379  
## 2 12.18217 6.616690 5.565476  
## 3 -38.28143 7.106693 -45.388123  
## 4 83.76404 7.596695 76.167346  
## 5 20.97539 8.086697 12.888692  
## 6 -36.02342 8.576699 -44.600118

#### 2.1.5. Pitanja za ponavljanje

#### Zadatak 15:

y<-c(5,1,3,3.5,4.2)  
x<-c(1,6,3,2,4)  
  
plot(x,y)



lm(y~x)

##   
## Call:  
## lm(formula = y ~ x)  
##   
## Coefficients:  
## (Intercept) x   
## 5.4243 -0.6514

### 2.1.6. Interpretacija parametara u modelu jednostavne linearne regresije

#### Slika 2.25. Naredbe potrebne za procjene 4 modela

podaci<-read.table("BDP\_i\_HICP.txt",header=T, sep="\t")  
  
m1<-lm(GDP~HICP,data=podaci)  
m2<-lm(log(GDP)~log(HICP),data=podaci)  
m3<-lm(log(GDP)~HICP,data=podaci)  
m4<-lm(GDP~log(HICP),data=podaci)  
  
m1$coefficients

## (Intercept) HICP   
## 2216385.89 -15790.89

m2$coefficients

## (Intercept) log(HICP)   
## 28.929835 -3.590287

m3$coefficients

## (Intercept) HICP   
## 15.98955011 -0.03587012

m4$coefficients

## (Intercept) log(HICP)   
## 7769301 -1549691

#### Slika 2.26. Ispis koeficijenata sva 4 modela

m1$coefficients

## (Intercept) HICP   
## 2216385.89 -15790.89

m2$coefficients

## (Intercept) log(HICP)   
## 28.929835 -3.590287

m3$coefficients

## (Intercept) HICP   
## 15.98955011 -0.03587012

m4$coefficients

## (Intercept) log(HICP)   
## 7769301 -1549691

#### Slika 2.28. Rezultat procjene standardiziranog modela

lm(scale(GDP)~0+scale(HICP),data=podaci)

##   
## Call:  
## lm(formula = scale(GDP) ~ 0 + scale(HICP), data = podaci)  
##   
## Coefficients:  
## scale(HICP)   
## -0.05293

### 2.1.7. Intervalna procjena parametara jednostavne linearne regresije

#### Slika 2.30. Naredbe potrebne za intervalnu procjenu parametara u sva četiri modela

confint(m1,level=0.95)

## 2.5 % 97.5 %  
## (Intercept) -9524310.3 13957082.02  
## HICP -126871.4 95289.59

confint(m2,level=0.95)

## 2.5 % 97.5 %  
## (Intercept) -70.33952 128.19919  
## log(HICP) -24.89296 17.71239

confint(m3,level=0.95)

## 2.5 % 97.5 %  
## (Intercept) -5.2878607 37.2669610  
## HICP -0.2371789 0.1654386

confint(m4,level=0.95)

## 2.5 % 97.5 %  
## (Intercept) -47004746 62543347  
## log(HICP) -13303910 10204528

#### Slika 2.31. Intervalne procjene parametara u sva četiri modela, razina pouzdanosti 95%

confint(m1,level=0.95)

## 2.5 % 97.5 %  
## (Intercept) -9524310.3 13957082.02  
## HICP -126871.4 95289.59

confint(m2,level=0.95)

## 2.5 % 97.5 %  
## (Intercept) -70.33952 128.19919  
## log(HICP) -24.89296 17.71239

confint(m3,level=0.95)

## 2.5 % 97.5 %  
## (Intercept) -5.2878607 37.2669610  
## HICP -0.2371789 0.1654386

confint(m4,level=0.95)

## 2.5 % 97.5 %  
## (Intercept) -47004746 62543347  
## log(HICP) -13303910 10204528

#### Slika 2.32. Intervalne procjene parametara u sva četiri modela, razina pouzdanosti 90% (lijevi panel) i 99% (desni panel)

confint(m1,level=0.9)

## 5 % 95 %  
## (Intercept) -7540894 11973666.0  
## HICP -108106 76524.2

confint(m2,level=0.9)

## 5 % 95 %  
## (Intercept) -53.56944 111.42911  
## log(HICP) -21.29419 14.11362

confint(m3,level=0.9)

## 5 % 95 %  
## (Intercept) -1.6933585 33.6724587  
## HICP -0.2031708 0.1314305

confint(m4,level=0.9)

## 5 % 95 %  
## (Intercept) -37751485 53290087  
## log(HICP) -11318210 8218827

confint(m1,level=0.99)

## 0.5 % 99.5 %  
## (Intercept) -13592897.2 18025668.9  
## HICP -165364.9 133783.1

confint(m2,level=0.99)

## 0.5 % 99.5 %  
## (Intercept) -104.74004 162.59971  
## log(HICP) -32.27513 25.09456

confint(m3,level=0.99)

## 0.5 % 99.5 %  
## (Intercept) -12.6612732 44.6403735  
## HICP -0.3069398 0.2351996

confint(m4,level=0.99)

## 0.5 % 99.5 %  
## (Intercept) -65985986 81524587  
## log(HICP) -17377183 14277801

### 2.1.8. Analiza varijance u modelu jednostavne linearne regresije

#### Slika 2.34. Ispis procijenjenog linearnog regresijskog modela

podaci<-read.table("BDP\_i\_HICP.txt",header=T, sep="\t")  
m1<-lm(GDP~HICP,data=podaci)  
summary(m1)

##   
## Call:  
## lm(formula = GDP ~ HICP, data = podaci)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -603036 -457350 -315196 -39255 2898603   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)  
## (Intercept) 2216386 5748839 0.386 0.703  
## HICP -15791 54391 -0.290 0.774  
##   
## Residual standard error: 852700 on 30 degrees of freedom  
## Multiple R-squared: 0.002802, Adjusted R-squared: -0.03044   
## F-statistic: 0.08429 on 1 and 30 DF, p-value: 0.7736

#### Slika 2.35. Procjena standardne devijacije i koeficijenta varijacije regresije

sazetak<-summary(m1)  
  
sazetak$sigma

## [1] 852701.5

GDP<-podaci$GDP  
koef\_v<-(sazetak$sigma/mean(GDP))\*100  
koef\_v

## [1] 155.6208

#### Slika 2.36. Tablica ANOVA za razmatrani primjer

anova(m1)

## Analysis of Variance Table  
##   
## Response: GDP  
## Df Sum Sq Mean Sq F value Pr(>F)  
## HICP 1 6.1286e+10 6.1286e+10 0.0843 0.7736  
## Residuals 30 2.1813e+13 7.2710e+11

#### Slika 2.37. Koeficijent korelacije između varijabli BDP i HICP

HICP<-podaci$HICP  
cor(HICP, GDP)

## [1] -0.05293131

### 2.1.9. Testiranje hipoteza u modelu jednostavne linearne regresije

#### Slika 2.40. Ispis modela jednostavne linearne regresije

podaci<-read.table("BDP\_i\_HICP.txt",header=T, sep="\t")  
m1<-lm(GDP~HICP,data=podaci)  
summary(m1)

##   
## Call:  
## lm(formula = GDP ~ HICP, data = podaci)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -603036 -457350 -315196 -39255 2898603   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)  
## (Intercept) 2216386 5748839 0.386 0.703  
## HICP -15791 54391 -0.290 0.774  
##   
## Residual standard error: 852700 on 30 degrees of freedom  
## Multiple R-squared: 0.002802, Adjusted R-squared: -0.03044   
## F-statistic: 0.08429 on 1 and 30 DF, p-value: 0.7736

#### Slika 2.41. Naredbe potrebne za izračun teorijskih veličina (teorijskih t-omjera)

#kritična granica za dvosmjerni test  
kriticna<-abs(qt(0.05/2,32-2))  
kriticna

## [1] 2.042272

#za jednosmjerni (na gornju granicu):  
kriticna2<-qt(1-0.05,32-2)  
kriticna2

## [1] 1.697261

#za jednosmjerni (na donju granicu):  
kriticna2<-qt(0.05,32-2)  
kriticna2

## [1] -1.697261

#### Slika 2.42. Naredbe potrebne za izračun teorijskih t-omjera

#lijevi panel  
#kritična granica za dvosmjerni test  
kriticna<-abs(qt(0.01/2,32-2))  
kriticna

## [1] 2.749996

#za jednosmjerni (na gornju granicu):  
kriticna2<-qt(0.01,32-2)  
kriticna2

## [1] -2.457262

#desni panel:  
#kritična granica za dvosmjerni test  
kriticna<-abs(qt(0.1/2,32-2))  
kriticna

## [1] 1.697261

#za jednosmjerni (na gornju granicu):  
kriticna2<-qt(0.1,32-2)  
kriticna2

## [1] -1.310415

#### Slika 2.44. Ispis modela jednostavne linearne regresije

podaci<-read.table("BDP\_i\_HICP.txt",header=T, sep="\t")  
m1<-lm(GDP~HICP,data=podaci)  
summary(m1)

##   
## Call:  
## lm(formula = GDP ~ HICP, data = podaci)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -603036 -457350 -315196 -39255 2898603   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)  
## (Intercept) 2216386 5748839 0.386 0.703  
## HICP -15791 54391 -0.290 0.774  
##   
## Residual standard error: 852700 on 30 degrees of freedom  
## Multiple R-squared: 0.002802, Adjusted R-squared: -0.03044   
## F-statistic: 0.08429 on 1 and 30 DF, p-value: 0.7736

#### Slika 2.45. Tablica ANOVA

anova(m1)

## Analysis of Variance Table  
##   
## Response: GDP  
## Df Sum Sq Mean Sq F value Pr(>F)  
## HICP 1 6.1286e+10 6.1286e+10 0.0843 0.7736  
## Residuals 30 2.1813e+13 7.2710e+11

#### Slika 2.46. Naredbe potrebne za izračun teorijskog F-omjera

#za F test:  
qf(1-0.05,1,32-2)

## [1] 4.170877

#### Slika 2.47. Naredbe potrebne za provedbu Waldova testa, zajedno s rezultatom

library(car)

## Loading required package: carData

ogranicenje<-"HICP=0"  
linearHypothesis(m1,ogranicenje,test="Chisq")

## Linear hypothesis test  
##   
## Hypothesis:  
## HICP = 0  
##   
## Model 1: restricted model  
## Model 2: GDP ~ HICP  
##   
## Res.Df RSS Df Sum of Sq Chisq Pr(>Chisq)  
## 1 31 2.1874e+13   
## 2 30 2.1813e+13 1 6.1286e+10 0.0843 0.7716

qchisq(0.95,1)

## [1] 3.841459

#### Slika 2.48. Naredbe potrebne za provedbu LR testa, zajedno s rezultatom

m1<-lm(GDP~HICP,data=podaci)  
library(lmtest)

## Loading required package: zoo

##   
## Attaching package: 'zoo'

## The following objects are masked from 'package:base':  
##   
## as.Date, as.Date.numeric

m2<-lm(GDP~1,data=podaci)  
lrtest(m1,m2)

## Likelihood ratio test  
##   
## Model 1: GDP ~ HICP  
## Model 2: GDP ~ 1  
## #Df LogLik Df Chisq Pr(>Chisq)  
## 1 3 -481.37   
## 2 2 -481.42 -1 0.0898 0.7645

### 2.1.10. Predviđanje modelom jednostavne linearne regresije

#### Slika 2.50. Naredbe potrebne za predviđanje vrijednosti zavisne varijable i interval predviđanja

novo <- data.frame(HICP=120)  
predict(m1, newdata = novo, interval = 'confidence',level = 0.95)

## fit lwr upr  
## 1 321479.1 -1300992 1943951

#### Slika 2.51. Ispis procijenjenog modela u primjeru

podaci<-read.table("udio\_65.txt",header=T, sep="\t")  
model<-lm(health~Udio\_65,data=podaci)  
summary(model)

##   
## Call:  
## lm(formula = health ~ Udio\_65, data = podaci)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -73662 -58529 -16401 11478 288591   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)  
## (Intercept) -146700 134536 -1.09 0.285  
## Udio\_65 1052420 701823 1.50 0.145  
##   
## Residual standard error: 86530 on 29 degrees of freedom  
## Multiple R-squared: 0.07196, Adjusted R-squared: 0.03996   
## F-statistic: 2.249 on 1 and 29 DF, p-value: 0.1445

#### Slika 2.52. Kritične granice za dvosmjerni i jednosmjerni t-test

abs(qt(0.05/2,29))

## [1] 2.04523

qt(1-0.05,29)

## [1] 1.699127

#### Slika 2.53. Naredbe za izračun R

sqrt(summary(model)$r.squared)

## [1] 0.2682537

summary(model)$sigma

## [1] 86527.65

potrosnja<-podaci$health  
(summary(model)$sigma/mean(potrosnja))\*100

## [1] 161.1532

#### Slika 2.54. Procjena modela sa standardiziranim varijablama

lm(scale(health)~0+scale(Udio\_65),data=podaci)

##   
## Call:  
## lm(formula = scale(health) ~ 0 + scale(Udio\_65), data = podaci)  
##   
## Coefficients:  
## scale(Udio\_65)   
## 0.2683

#### Slika 2.55. Tablica ANOVA i teorijski F-omjer

qf(1-0.05,1,29)

## [1] 4.182964

anova(model)

## Analysis of Variance Table  
##   
## Response: health  
## Df Sum Sq Mean Sq F value Pr(>F)  
## Udio\_65 1 1.6836e+10 1.6836e+10 2.2487 0.1445  
## Residuals 29 2.1712e+11 7.4870e+09

#### Slika 2.56. Intervalna procjena parametara modela

confint(model,level=0.90)

## 5 % 95 %  
## (Intercept) -375293.4 81893.38  
## Udio\_65 -140066.8 2244905.84

#### Slika 2.57. Ispis Waldova testa

library(car)  
ogranicenje<-"Udio\_65=0"  
linearHypothesis(model,ogranicenje,test="Chisq")

## Linear hypothesis test  
##   
## Hypothesis:  
## Udio\_65 = 0  
##   
## Model 1: restricted model  
## Model 2: health ~ Udio\_65  
##   
## Res.Df RSS Df Sum of Sq Chisq Pr(>Chisq)  
## 1 30 2.3396e+11   
## 2 29 2.1712e+11 1 1.6836e+10 2.2487 0.1337

#### Slika 2.58. Ispis LR testa

library(lmtest)  
m2<-lm(health~1,data=podaci)  
lrtest(model,m2)

## Likelihood ratio test  
##   
## Model 1: health ~ Udio\_65  
## Model 2: health ~ 1  
## #Df LogLik Df Chisq Pr(>Chisq)  
## 1 3 -395.37   
## 2 2 -396.53 -1 2.3151 0.1281

#### Slika 2.59. Usporedba lin-lin i log-lin modela

model2<-lm(log(health)~Udio\_65,data=podaci)  
library(stargazer)

##   
## Please cite as:

## Hlavac, Marek (2022). stargazer: Well-Formatted Regression and Summary Statistics Tables.

## R package version 5.2.3. https://CRAN.R-project.org/package=stargazer

stargazer(list(model,model2),type="text")

##   
## ==========================================================  
## Dependent variable:   
## ----------------------------  
## health log(health)   
## (1) (2)   
## ----------------------------------------------------------  
## Udio\_65 1,052,420.000 21.082   
## (701,822.900) (13.387)   
##   
## Constant -146,700.000 5.670\*\*   
## (134,535.800) (2.566)   
##   
## ----------------------------------------------------------  
## Observations 31 31   
## R2 0.072 0.079   
## Adjusted R2 0.040 0.047   
## Residual Std. Error (df = 29) 86,527.650 1.650   
## F Statistic (df = 1; 29) 2.249 2.480   
## ==========================================================  
## Note: \*p<0.1; \*\*p<0.05; \*\*\*p<0.01

#### Slika 2.60. Naredbe potrebne za predviđanje vrijednosti zavisne varijable i interval predviđanja

novo <- data.frame(Udio\_65=0.2)  
predict(model, newdata = novo, interval = 'confidence',level = 0.95)

## fit lwr upr  
## 1 63783.91 29147.46 98420.36

#### 2.1.12. Pitanja za ponavljanje, Zadatak 24:

podaci<-read.table("phillips.txt",header=T, sep="\t")  
model<-lm(s\_inflacije~s\_nezaposlenosti,data=podaci)  
summary(model)

##   
## Call:  
## lm(formula = s\_inflacije ~ s\_nezaposlenosti, data = podaci)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -5.150 -3.199 -1.385 1.936 28.257   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 17.419 2.087 8.345 1.56e-09 \*\*\*  
## s\_nezaposlenosti -23.286 5.202 -4.476 9.04e-05 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 5.952 on 32 degrees of freedom  
## Multiple R-squared: 0.385, Adjusted R-squared: 0.3658   
## F-statistic: 20.04 on 1 and 32 DF, p-value: 9.04e-05

abs(qt(0.05/2,34-2))

## [1] 2.036933

qt(1-0.05,34-2)

## [1] 1.693889

sqrt(summary(model)$r.squared)

## [1] 0.6205206

summary(model)$sigma

## [1] 5.951779

s\_inflacije<-podaci$s\_inflacije  
(summary(model)$sigma/mean(s\_inflacije))\*100

## [1] 64.21288

lm(scale(s\_inflacije)~0+scale(s\_nezaposlenosti),data=podaci)

##   
## Call:  
## lm(formula = scale(s\_inflacije) ~ 0 + scale(s\_nezaposlenosti),   
## data = podaci)  
##   
## Coefficients:  
## scale(s\_nezaposlenosti)   
## -0.6205

qf(1-0.05,1,34-2)

## [1] 4.149097

anova(model)

## Analysis of Variance Table  
##   
## Response: s\_inflacije  
## Df Sum Sq Mean Sq F value Pr(>F)   
## s\_nezaposlenosti 1 709.76 709.76 20.036 9.04e-05 \*\*\*  
## Residuals 32 1133.56 35.42   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

confint(model,level=0.90)

## 5 % 95 %  
## (Intercept) 13.88308 20.95446  
## s\_nezaposlenosti -32.09730 -14.47382

library(car)  
ogranicenje<-"s\_nezaposlenosti=0"  
linearHypothesis(model,ogranicenje,test="Chisq")

## Linear hypothesis test  
##   
## Hypothesis:  
## s\_nezaposlenosti = 0  
##   
## Model 1: restricted model  
## Model 2: s\_inflacije ~ s\_nezaposlenosti  
##   
## Res.Df RSS Df Sum of Sq Chisq Pr(>Chisq)   
## 1 33 1843.3   
## 2 32 1133.6 1 709.76 20.036 7.598e-06 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

library(lmtest)  
m2<-lm(s\_inflacije~1,data=podaci)  
lrtest(model,m2)

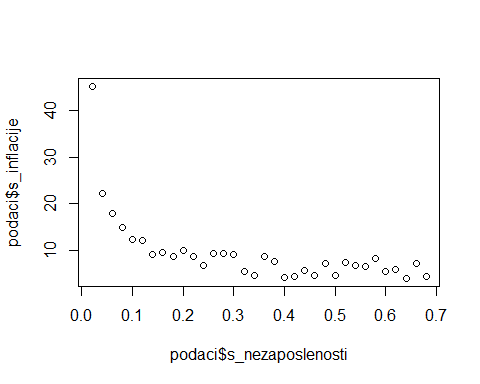
## Likelihood ratio test  
##   
## Model 1: s\_inflacije ~ s\_nezaposlenosti  
## Model 2: s\_inflacije ~ 1  
## #Df LogLik Df Chisq Pr(>Chisq)   
## 1 3 -107.86   
## 2 2 -116.12 -1 16.531 4.786e-05 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

#predviđanje:  
novo <- data.frame(s\_nezaposlenosti=0.15)  
predict(model, newdata = novo, interval = 'confidence',level = 0.95)

## fit lwr upr  
## 1 13.92594 10.95708 16.89479

#### 2.1.12. Pitanja za ponavljanje, Zadatak 25:

plot(podaci$s\_nezaposlenosti,podaci$s\_inflacije)



model<-lm(s\_inflacije~I(1/s\_nezaposlenosti),data=podaci)  
summary(model)

##   
## Call:  
## lm(formula = s\_inflacije ~ I(1/s\_nezaposlenosti), data = podaci)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -2.26387 -1.23447 0.08129 1.06495 2.44700   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 4.43759 0.29624 14.98 5.17e-16 \*\*\*  
## I(1/s\_nezaposlenosti) 0.79773 0.02718 29.35 < 2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 1.436 on 32 degrees of freedom  
## Multiple R-squared: 0.9642, Adjusted R-squared: 0.9631   
## F-statistic: 861.6 on 1 and 32 DF, p-value: < 2.2e-16

abs(qt(0.05/2,34-2))

## [1] 2.036933

qt(1-0.05,34-2)

## [1] 1.693889

sqrt(summary(model)$r.squared)

## [1] 0.9819316

summary(model)$sigma

## [1] 1.436249

s\_inflacije<-podaci$s\_inflacije  
(summary(model)$sigma/mean(s\_inflacije))\*100

## [1] 15.49548

lm(scale(s\_inflacije)~0+scale(I(1/s\_nezaposlenosti)),data=podaci)

##   
## Call:  
## lm(formula = scale(s\_inflacije) ~ 0 + scale(I(1/s\_nezaposlenosti)),   
## data = podaci)  
##   
## Coefficients:  
## scale(I(1/s\_nezaposlenosti))   
## 0.9819

qf(1-0.05,1,34-2)

## [1] 4.149097

anova(model)

## Analysis of Variance Table  
##   
## Response: s\_inflacije  
## Df Sum Sq Mean Sq F value Pr(>F)   
## I(1/s\_nezaposlenosti) 1 1777.31 1777.31 861.6 < 2.2e-16 \*\*\*  
## Residuals 32 66.01 2.06   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

confint(model,level=0.90)

## 5 % 95 %  
## (Intercept) 3.9357854 4.9393975  
## I(1/s\_nezaposlenosti) 0.7516991 0.8437697

library(car)  
ogranicenje<-"I(1/s\_nezaposlenosti)=0"  
linearHypothesis(model,ogranicenje,test="Chisq")

## Linear hypothesis test  
##   
## Hypothesis:  
## I(1/s\_nezaposlenosti) = 0  
##   
## Model 1: restricted model  
## Model 2: s\_inflacije ~ I(1/s\_nezaposlenosti)  
##   
## Res.Df RSS Df Sum of Sq Chisq Pr(>Chisq)   
## 1 33 1843.32   
## 2 32 66.01 1 1777.3 861.6 < 2.2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

library(lmtest)  
m2<-lm(s\_inflacije~1,data=podaci)  
lrtest(model,m2)

## Likelihood ratio test  
##   
## Model 1: s\_inflacije ~ I(1/s\_nezaposlenosti)  
## Model 2: s\_inflacije ~ 1  
## #Df LogLik Df Chisq Pr(>Chisq)   
## 1 3 -59.522   
## 2 2 -116.124 -1 113.2 < 2.2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

#predviđanje:  
novo <- data.frame(s\_nezaposlenosti=0.15)  
predict(model, newdata = novo, interval = 'confidence',level = 0.95)

## fit lwr upr  
## 1 9.755821 9.252957 10.25868

#### 2.1.12. Pitanja za ponavljanje, Zadatak 26:

m1<-lm(s\_inflacije~s\_nezaposlenosti,data=podaci)  
m2<-lm(s\_inflacije~I(1/s\_nezaposlenosti),data=podaci)  
library(stargazer)  
stargazer(list(m1,m2),type="text")

##   
## ==========================================================  
## Dependent variable:   
## ----------------------------  
## s\_inflacije   
## (1) (2)   
## ----------------------------------------------------------  
## s\_nezaposlenosti -23.286\*\*\*   
## (5.202)   
##   
## I(1/s\_nezaposlenosti) 0.798\*\*\*   
## (0.027)   
##   
## Constant 17.419\*\*\* 4.438\*\*\*   
## (2.087) (0.296)   
##   
## ----------------------------------------------------------  
## Observations 34 34   
## R2 0.385 0.964   
## Adjusted R2 0.366 0.963   
## Residual Std. Error (df = 32) 5.952 1.436   
## F Statistic (df = 1; 32) 20.036\*\*\* 861.596\*\*\*   
## ==========================================================  
## Note: \*p<0.1; \*\*p<0.05; \*\*\*p<0.01

## 2.2. Model višestruke linearne regresije

### 2.2.3. Metoda najmanjih kvadrata i procjenitelj za slučaj višestruke linearne regresije

#### Slika 2.61. i Slika 2.62. Unos potrebnih naredbi i Ispis procijenjenih parametara

y<-c(18,29,21,11,7,25,44,1)  
x1<-c(10,15,12,7,4,14,22,1)  
x2<-c(13,12,16,8,10,11,25,5)  
x3<-c(9,3,7,14,18,2,1,22)  
y<-as.matrix(y)  
jed<-c(rep(1,each=8))  
jed<-as.matrix(jed)  
  
x<-as.matrix(cbind(x1,x2,x3))  
X<-cbind(jed,x)  
  
a<-t(X)%\*%X  
b<-solve(a)  
c<-t(X)%\*%y  
beta<-b%\*%c  
beta

## [,1]  
## -10.4810160  
## x1 2.5542314  
## x2 -0.0848511  
## x3 0.4108364

#### Slika 2.63. Procjena modela višestruke linearne regresije

summary(lm(y~x1+x2+x3))

##   
## Call:  
## lm(formula = y ~ x1 + x2 + x3)  
##   
## Residuals:  
## 1 2 3 4 5 6 7 8   
## 0.344238 0.953249 -0.687998 -1.471505 0.717546 -0.166534 -0.001634 0.312639   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -10.48102 4.41491 -2.374 0.07649 .   
## x1 2.55423 0.37535 6.805 0.00244 \*\*  
## x2 -0.08485 0.18530 -0.458 0.67078   
## x3 0.41084 0.22836 1.799 0.14640   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 1.038 on 4 degrees of freedom  
## Multiple R-squared: 0.9967, Adjusted R-squared: 0.9942   
## F-statistic: 399.9 on 3 and 4 DF, p-value: 2.068e-05

### 2.2.4. Interpretacija parametara u modelu višestruke linearne regresije

#### Slika 2.64. Rezultat procjene standardiziranog modela

lm(scale(y)~0+scale(x1)+scale(x2)+scale(x3))

##   
## Call:  
## lm(formula = scale(y) ~ 0 + scale(x1) + scale(x2) + scale(x3))  
##   
## Coefficients:  
## scale(x1) scale(x2) scale(x3)   
## 1.25299 -0.03756 0.23554

### 2.2.5. Intervalna procjena parametara višestruke linearne regresije

#### Slika 2.65. Intervalne procjene parametara, razina pouzdanosti 92%

model<-lm(y~x1+x2+x3)  
confint(model,level=0.92)

## 4 % 96 %  
## (Intercept) -20.7804435 -0.1815886  
## x1 1.6785836 3.4298792  
## x2 -0.5171283 0.3474261  
## x3 -0.1219095 0.9435824

### 2.2.6. Analiza varijance u modelu višestruke linearne regresije

#### Slika 2.66. Ispis procijenjenog linearnog regresijskog modela

summary(lm(y~x1+x2+x3))

##   
## Call:  
## lm(formula = y ~ x1 + x2 + x3)  
##   
## Residuals:  
## 1 2 3 4 5 6 7 8   
## 0.344238 0.953249 -0.687998 -1.471505 0.717546 -0.166534 -0.001634 0.312639   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -10.48102 4.41491 -2.374 0.07649 .   
## x1 2.55423 0.37535 6.805 0.00244 \*\*  
## x2 -0.08485 0.18530 -0.458 0.67078   
## x3 0.41084 0.22836 1.799 0.14640   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 1.038 on 4 degrees of freedom  
## Multiple R-squared: 0.9967, Adjusted R-squared: 0.9942   
## F-statistic: 399.9 on 3 and 4 DF, p-value: 2.068e-05

#### Slika 2.67. Naredbe za procjenu greške regresije i koeficijenta varijacije regresije

model<-lm(y~x1+x2+x3)  
sazetak<-summary(model)  
  
sazetak$sigma

## [1] 1.03757

sazetak$sigma/mean(y)

## [1] 0.0532087

#### Slika 2.68. Tablica ANOVA za razmatrani primjer

model2<-lm(y~1)  
anova(model,model2)

## Analysis of Variance Table  
##   
## Model 1: y ~ x1 + x2 + x3  
## Model 2: y ~ 1  
## Res.Df RSS Df Sum of Sq F Pr(>F)   
## 1 4 4.31   
## 2 7 1296.00 -3 -1291.7 399.95 2.068e-05 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

### 2.2.7. Testiranje hipoteza u modelu višestruke linearne regresije

#### Slika 2.69. Naredbe i ispis Waldova testa

library(car)  
ogranicenje<-"x2=0"  
linearHypothesis(model,ogranicenje,test="Chisq")

## Linear hypothesis test  
##   
## Hypothesis:  
## x2 = 0  
##   
## Model 1: restricted model  
## Model 2: y ~ x1 + x2 + x3  
##   
## Res.Df RSS Df Sum of Sq Chisq Pr(>Chisq)  
## 1 5 4.5319   
## 2 4 4.3062 1 0.22574 0.2097 0.647

#### Slika 2.70. Naredbe i ispis Waldova testa

library(car)  
ogranicenje<-c("x1=0","x3=0")  
linearHypothesis(model,ogranicenje,test="Chisq")

## Linear hypothesis test  
##   
## Hypothesis:  
## x1 = 0  
## x3 = 0  
##   
## Model 1: restricted model  
## Model 2: y ~ x1 + x2 + x3  
##   
## Res.Df RSS Df Sum of Sq Chisq Pr(>Chisq)   
## 1 6 287.984   
## 2 4 4.306 2 283.68 263.51 < 2.2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

#### Slika 2.71. Naredbe i ispis Waldova testa

library(car)  
ogranicenje<-c("x3=0.5")  
linearHypothesis(model,ogranicenje,test="Chisq")

## Linear hypothesis test  
##   
## Hypothesis:  
## x3 = 0.5  
##   
## Model 1: restricted model  
## Model 2: y ~ x1 + x2 + x3  
##   
## Res.Df RSS Df Sum of Sq Chisq Pr(>Chisq)  
## 1 5 4.4703   
## 2 4 4.3062 1 0.16412 0.1524 0.6962

#### Slika 2.72. Naredbe i ispis F-testa

ogranicenje<-"x2=0"  
linearHypothesis(model,ogranicenje,test="F")

## Linear hypothesis test  
##   
## Hypothesis:  
## x2 = 0  
##   
## Model 1: restricted model  
## Model 2: y ~ x1 + x2 + x3  
##   
## Res.Df RSS Df Sum of Sq F Pr(>F)  
## 1 5 4.5319   
## 2 4 4.3062 1 0.22574 0.2097 0.6708

#### Slika 2.73. Naredbe i ispis F-testa

ogranicenje<-c("x1=0","x3=0")  
linearHypothesis(model,ogranicenje,test="F")

## Linear hypothesis test  
##   
## Hypothesis:  
## x1 = 0  
## x3 = 0  
##   
## Model 1: restricted model  
## Model 2: y ~ x1 + x2 + x3  
##   
## Res.Df RSS Df Sum of Sq F Pr(>F)   
## 1 6 287.984   
## 2 4 4.306 2 283.68 131.75 0.0002236 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

#### Slika 2.74. Naredbe i ispis F-testa

ogranicenje<-c("x3=0.5")  
linearHypothesis(model,ogranicenje,test="F")

## Linear hypothesis test  
##   
## Hypothesis:  
## x3 = 0.5  
##   
## Model 1: restricted model  
## Model 2: y ~ x1 + x2 + x3  
##   
## Res.Df RSS Df Sum of Sq F Pr(>F)  
## 1 5 4.4703   
## 2 4 4.3062 1 0.16412 0.1524 0.7161

#### Slika 2.75. Parcijalni F-test o izostavljenoj značajnoj regresijskoj varijabli

model1<-lm(y~x1+x2+x3)  
ssr\_1<-sum(resid(model1)^2)  
model2<-lm(y~x2+x3)  
ssr\_2<-sum(resid(model2)^2)  
ssr\_1;ssr\_2

## [1] 4.306203

## [1] 54.15769

emp<-((ssr\_2-ssr\_1)/1)/(ssr\_1/(8-2-1-1))  
emp

## [1] 46.30667

#### Slika 2.76. LR test o izostavljenoj značajnoj regresijskoj varijabli

library(lmtest)  
model2<-lm(y~x2+x3)  
lrtest(model1,model2)

## Likelihood ratio test  
##   
## Model 1: y ~ x1 + x2 + x3  
## Model 2: y ~ x2 + x3  
## #Df LogLik Df Chisq Pr(>Chisq)   
## 1 5 -8.874   
## 2 4 -19.001 -1 20.255 6.779e-06 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

#### Slika 2.77. Usporedba procijenjenih parametara modela iz primjera 2.38

model

##   
## Call:  
## lm(formula = y ~ x1 + x2 + x3)  
##   
## Coefficients:  
## (Intercept) x1 x2 x3   
## -10.48102 2.55423 -0.08485 0.41084

model2

##   
## Call:  
## lm(formula = y ~ x2 + x3)  
##   
## Coefficients:  
## (Intercept) x2 x3   
## 17.054 1.004 -1.064

#### Slika 2.78. Parcijalni F-test o uključenoj nepotrebnoj regresijskoj varijabli

model<-lm(y~x1+x2+x3)  
library(car)  
ogranicenje<-"x1=0"  
linearHypothesis(model,ogranicenje,test="F")

## Linear hypothesis test  
##   
## Hypothesis:  
## x1 = 0  
##   
## Model 1: restricted model  
## Model 2: y ~ x1 + x2 + x3  
##   
## Res.Df RSS Df Sum of Sq F Pr(>F)   
## 1 5 54.158   
## 2 4 4.306 1 49.851 46.307 0.002437 \*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

#### Slika 2.79. Test o stabilnosti parametara

y1<-y[1:4];x11<-x1[1:4]  
y2<-y[5:8];x12<-x1[5:8]  
library(gap)

## Loading required package: gap.datasets

## gap version 1.5-1

chow.test(y1,x11,y2,x12)

## F value d.f.1 d.f.2 P value   
## 1.0683583 2.0000000 4.0000000 0.4248619

#### Slika 2.80. Ispis modela za oba poduzorka

#lijevi panel  
summary(lm(y1~x11))

##   
## Call:  
## lm(formula = y1 ~ x11)  
##   
## Residuals:  
## 1 2 3 4   
## 0.45588 0.42647 -0.95588 0.07353   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -4.5147 1.5788 -2.86 0.10363   
## x11 2.2059 0.1387 15.90 0.00393 \*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.809 on 2 degrees of freedom  
## Multiple R-squared: 0.9922, Adjusted R-squared: 0.9882   
## F-statistic: 252.8 on 1 and 2 DF, p-value: 0.003932

#desni panel  
summary(lm(y1~x11))

##   
## Call:  
## lm(formula = y1 ~ x11)  
##   
## Residuals:  
## 1 2 3 4   
## 0.45588 0.42647 -0.95588 0.07353   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -4.5147 1.5788 -2.86 0.10363   
## x11 2.2059 0.1387 15.90 0.00393 \*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.809 on 2 degrees of freedom  
## Multiple R-squared: 0.9922, Adjusted R-squared: 0.9882   
## F-statistic: 252.8 on 1 and 2 DF, p-value: 0.003932

#### Slika 2.81. Sume kvadrata rezidualnih odstupanja potrebnih za izračun empirijskog F-omjera

sum(resid(lm(y1~x11))^2)

## [1] 1.308824

sum(resid(lm(y2~x12))^2)

## [1] 4.684734

sum(resid(lm(y~x1))^2)

## [1] 9.19519

#### Slika 2.83. RESET test

model<-lm(y~x1+x2+x3)  
library(lmtest)  
resettest(model,power=2:3,type="fitted")

##   
## RESET test  
##   
## data: model  
## RESET = 0.44191, df1 = 2, df2 = 2, p-value = 0.6935

#### Slika 2.84. RESET test pomoću naredbi o linearnim ograničenjima na parametre

model<-lm(y~x1+x2+x3)  
y1<-fitted(model)  
model\_novo<-lm(y~x1+x2+x3+I(y1^2)+I(y1^3))  
library(car)  
ogranicenje<-c("I(y1^2)=0","I(y1^3)=0")  
linearHypothesis(model\_novo,ogranicenje,test="F")

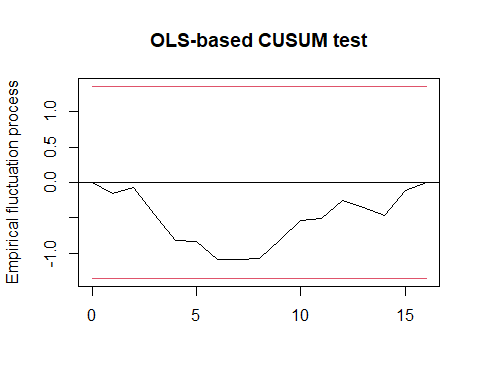
## Linear hypothesis test  
##   
## Hypothesis:  
## I(y1^2) = 0  
## I(y1^3) = 0  
##   
## Model 1: restricted model  
## Model 2: y ~ x1 + x2 + x3 + I(y1^2) + I(y1^3)  
##   
## Res.Df RSS Df Sum of Sq F Pr(>F)  
## 1 4 4.3062   
## 2 2 2.9864 2 1.3197 0.4419 0.6935

#### Slika 2.85. i 2.86. Naredbe potrebne za CUSUM test i CUSUM dijagram

y<-c(18,29,21,11,7,25,44,1,0,43,24,6,10,20,28,17)  
x1<-c(10,15,12,7,4,14,22,1,0,21,13,3,6,11,14,9)  
x2<-c(13,12,16,8,10,11,25,5,4,24,10,9,7,15,11,12)  
x3<-c(9,3,7,14,18,2,1,22,21,0,1,17,13,6,2,8)  
  
library(strucchange)

## Loading required package: sandwich

y<-ts(y,start=1,frequency = 1)  
x1<-ts(x1,start=1,frequency = 1)  
x2<-ts(x2,start=1,frequency = 1)  
x3<-ts(x3,start=1,frequency = 1)  
cusum <- efp(y~x1+x2+x3, type = "OLS-CUSUM")  
plot(cusum,xlab=NA)



fs <- Fstats(y~x1+x2+x3, from = 10, to = 11)  
sctest(fs, type="aveF")

##   
## aveF test  
##   
## data: fs  
## ave.F = 1.8958, p-value = 0.773

#### Slika 2.87. Naredbe potrebne za predviđanje vrijednosti zavisne varijable i interval predviđanja

novo<-data.frame(x1=10,x2=5,x3=10)  
predict(model,newdata = novo,interval = "confidence",level=.95)

## fit lwr upr  
## 1 18.74541 15.07897 22.41184

#### Slika 2.88. Naredbe potrebne za ispis modela m1, m2 i m3

potrosnja<-read.table("potrosnja.txt",header=T,sep="\t")  
m1<-lm(potrosnja~bdp+cijene,data=potrosnja)  
m2<-lm(log(potrosnja)~log(bdp)+log(cijene),data=potrosnja)  
m3<-lm(scale(potrosnja)~0+scale(bdp)+scale(cijene),data=potrosnja)  
library(stargazer)  
stargazer(list(m1,m2,m3),type="text")

##   
## ======================================================================================================  
## Dependent variable:   
## ----------------------------------------------------------------------------------  
## potrosnja log(potrosnja) scale(potrosnja)   
## (1) (2) (3)   
## ------------------------------------------------------------------------------------------------------  
## bdp 0.645\*\*\*   
## (0.003)   
##   
## cijene 4.202   
## (38.637)   
##   
## log(bdp) 0.956\*\*\*   
## (0.016)   
##   
## log(cijene) 0.116   
## (0.166)   
##   
## scale(bdp) 0.998\*\*\*   
## (0.005)   
##   
## scale(cijene) 0.001   
## (0.005)   
##   
## Constant 3,951.826 -0.458   
## (6,268.220) (0.816)   
##   
## ------------------------------------------------------------------------------------------------------  
## Observations 140 140 140   
## R2 0.997 0.965 0.997   
## Adjusted R2 0.997 0.965 0.997   
## Residual Std. Error 34,192.400 (df = 137) 0.556 (df = 137) 0.056 (df = 138)   
## F Statistic 21,855.880\*\*\* (df = 2; 137) 1,906.151\*\*\* (df = 2; 137) 22,015.410\*\*\* (df = 2; 138)  
## ======================================================================================================  
## Note: \*p<0.1; \*\*p<0.05; \*\*\*p<0.01

#### Slika 2.90. Ispis modela m1

summary(m1)

##   
## Call:  
## lm(formula = potrosnja ~ bdp + cijene, data = potrosnja)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -126036 -4467 -4373 -3738 329885   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 3.952e+03 6.268e+03 0.630 0.529   
## bdp 6.450e-01 3.085e-03 209.073 <2e-16 \*\*\*  
## cijene 4.202e+00 3.864e+01 0.109 0.914   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 34190 on 137 degrees of freedom  
## Multiple R-squared: 0.9969, Adjusted R-squared: 0.9968   
## F-statistic: 2.186e+04 on 2 and 137 DF, p-value: < 2.2e-16

#### Slika 2.91. Teorijske vrijednosti t-omjera i F-omjera

cbind(qt(0.95,137), qf(0.95,2,137))

## [,1] [,2]  
## [1,] 1.656052 3.062204

#### Slika 2.92. Ispis modela m1, m2, m4 i m5

m4<-lm(log(potrosnja)~bdp+cijene,data=potrosnja)  
m5<-lm(potrosnja~log(bdp)+log(cijene),data=potrosnja)  
stargazer(list(m1,m2,m4,m5),type="text")

##   
## ==================================================================================  
## Dependent variable:   
## ---------------------------------------------------  
## potrosnja log(potrosnja) potrosnja   
## (1) (2) (3) (4)   
## ----------------------------------------------------------------------------------  
## bdp 0.645\*\*\* 0.00000\*\*\*   
## (0.003) (0.00000)   
##   
## cijene 4.202 -0.0005   
## (38.637) (0.003)   
##   
## log(bdp) 0.956\*\*\* 78,786.460\*\*\*  
## (0.016) (15,707.190)   
##   
## log(cijene) 0.116 2,633.681   
## (0.166) (167,918.800)  
##   
## Constant 3,951.826 -0.458 6.758\*\*\* -477,697.900   
## (6,268.220) (0.816) (0.506) (825,124.000)  
##   
## ----------------------------------------------------------------------------------  
## Observations 140 140 140 140   
## R2 0.997 0.965 0.146 0.156   
## Adjusted R2 0.997 0.965 0.134 0.144   
## Residual Std. Error (df = 137) 34,192.400 0.556 2.758 561,989.900   
## F Statistic (df = 2; 137) 21,855.880\*\*\* 1,906.151\*\*\* 11.717\*\*\* 12.658\*\*\*   
## ==================================================================================  
## Note: \*p<0.1; \*\*p<0.05; \*\*\*p<0.01

#### Slika 2.93. Intervalna procjena parametara za model m1

confint(m1,level=.9)

## 5 % 95 %  
## (Intercept) -6428.6717521 1.433232e+04  
## bdp 0.6399042 6.501225e-01  
## cijene -59.7832864 6.818696e+01

#### Slika 2.94. Procjena standardne devijacije regresije i koeficijenta varijacije regresije za m1

y<-potrosnja$potrosnja  
sazetak<-summary(m1)  
cbind(sazetak$sigma,sazetak$sigma/mean(y))

## [,1] [,2]  
## [1,] 34192.4 0.3766832

#### Slika 2.95. Waldov test

ogranicenje<-"bdp=2\*cijene"  
linearHypothesis(m1,ogranicenje,test="Chisq")

## Linear hypothesis test  
##   
## Hypothesis:  
## bdp - 2 cijene = 0  
##   
## Model 1: restricted model  
## Model 2: potrosnja ~ bdp + cijene  
##   
## Res.Df RSS Df Sum of Sq Chisq Pr(>Chisq)  
## 1 138 1.6018e+11   
## 2 137 1.6017e+11 1 11785882 0.0101 0.92

#### Slika 2.96. Waldov test

ogranicenje<-c("bdp=.7","cijene=0")  
linearHypothesis(m1,ogranicenje,test="Chisq")

## Linear hypothesis test  
##   
## Hypothesis:  
## bdp = 0.7  
## cijene = 0  
##   
## Model 1: restricted model  
## Model 2: potrosnja ~ bdp + cijene  
##   
## Res.Df RSS Df Sum of Sq Chisq Pr(>Chisq)   
## 1 139 5.3157e+11   
## 2 137 1.6017e+11 2 3.714e+11 317.67 < 2.2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

#### Slika 2.97. F-test za postupak h)

ogranicenje<-"bdp=2\*cijene"  
linearHypothesis(m1,ogranicenje,test="F")

## Linear hypothesis test  
##   
## Hypothesis:  
## bdp - 2 cijene = 0  
##   
## Model 1: restricted model  
## Model 2: potrosnja ~ bdp + cijene  
##   
## Res.Df RSS Df Sum of Sq F Pr(>F)  
## 1 138 1.6018e+11   
## 2 137 1.6017e+11 1 11785882 0.0101 0.9202

#### Slika 2.98. F-test za postupak i)

ogranicenje<-c("bdp=.7","cijene=0")  
linearHypothesis(m1,ogranicenje,test="F")

## Linear hypothesis test  
##   
## Hypothesis:  
## bdp = 0.7  
## cijene = 0  
##   
## Model 1: restricted model  
## Model 2: potrosnja ~ bdp + cijene  
##   
## Res.Df RSS Df Sum of Sq F Pr(>F)   
## 1 139 5.3157e+11   
## 2 137 1.6017e+11 2 3.714e+11 158.84 < 2.2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

#### Slika 2.99. Test o uključenoj nepotrebnoj varijabli

ogranicenje<-c("bdp=0")  
linearHypothesis(m1,ogranicenje,test="F")

## Linear hypothesis test  
##   
## Hypothesis:  
## bdp = 0  
##   
## Model 1: restricted model  
## Model 2: potrosnja ~ bdp + cijene  
##   
## Res.Df RSS Df Sum of Sq F Pr(>F)   
## 1 138 5.1264e+13   
## 2 137 1.6017e+11 1 5.1104e+13 43711 < 2.2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

#### Slika 2.100. Test o izostavljenoj značajnoj varijabli

library(lmtest)  
model2<-lm(potrosnja~bdp,data=potrosnja)  
lrtest(m1,model2)

## Likelihood ratio test  
##   
## Model 1: potrosnja ~ bdp + cijene  
## Model 2: potrosnja ~ bdp  
## #Df LogLik Df Chisq Pr(>Chisq)  
## 1 4 -1658.7   
## 2 3 -1658.7 -1 0.0121 0.9125

#### Slika 2.101. Test o stabilnosti parametara

bdp<-potrosnja$bdp;cijene<-potrosnja$cijene  
y1<-y[1:70];bdp1<-bdp[1:70];cijene1<-cijene[1:70]  
y2<-y[71:140];bdp2<-bdp[71:140];cijene2<-cijene[71:140]  
library(gap)  
chow.test(y1,c(bdp1,cijene1),y2,c(bdp2,cijene2))

## F value d.f.1 d.f.2 P value   
## 0.3498094 2.0000000 276.0000000 0.7051344

#### Slika 2.102. RESET test o adekvatnosti linearnog modela

library(lmtest)  
resettest(m1,power=2:3,type="fitted")

##   
## RESET test  
##   
## data: m1  
## RESET = 218.19, df1 = 2, df2 = 135, p-value < 2.2e-16

#### Slika 2.103. Ispis modela s uključenim kvadratom i kubom procijenjene vrijednosti varijable potrošnja

summary(lm(potrosnja~bdp+cijene+I(fitted(m1)^2)+I(fitted(m1)^3),data=potrosnja))

##   
## Call:  
## lm(formula = potrosnja ~ bdp + cijene + I(fitted(m1)^2) + I(fitted(m1)^3),   
## data = potrosnja)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -133688 -1404 -1305 -920 125675   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 8.265e+02 3.100e+03 0.267 0.79015   
## bdp 6.788e-01 1.645e-02 41.263 < 2e-16 \*\*\*  
## cijene 4.492e+00 1.893e+01 0.237 0.81281   
## I(fitted(m1)^2) 5.908e-08 1.692e-08 3.491 0.00065 \*\*\*  
## I(fitted(m1)^3) -1.034e-14 1.983e-15 -5.212 6.82e-07 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 16740 on 135 degrees of freedom  
## Multiple R-squared: 0.9993, Adjusted R-squared: 0.9992   
## F-statistic: 4.569e+04 on 4 and 135 DF, p-value: < 2.2e-16

#### Slika 2.104. Predviđena vrijednost potrošnje u modelu m1

novo<-data.frame(bdp=1000,cijene=134)  
predict(m1,newdata = novo,interval = "confidence",level=.90)

## fit lwr upr  
## 1 5159.886 287.8734 10031.9

### 2.2.10. Pitanja za ponavljanje, Rješenja

#### 2.2.10. Pitanja za ponavljanje, Rješenja, Zadatak 27

place<-read.table("placa.txt",header=T,sep="\t")  
  
m1<-lm(placa~staz+obrazovanje,data=place)  
m2<-lm(log(placa)~log(staz)+log(obrazovanje),data=place)  
m3<-lm(scale(placa)~0+scale(staz)+scale(obrazovanje),data=place)  
  
library(stargazer)  
stargazer(list(m1,m2,m3),type="text")

##   
## ==============================================================================================  
## Dependent variable:   
## --------------------------------------------------------------------------  
## placa log(placa) scale(placa)   
## (1) (2) (3)   
## ----------------------------------------------------------------------------------------------  
## staz 1.390\*\*\*   
## (0.233)   
##   
## obrazovanje 0.015   
## (0.239)   
##   
## log(staz) 0.002\*\*\*   
## (0.0004)   
##   
## log(obrazovanje) 0.002\*\*\*   
## (0.0004)   
##   
## scale(staz) 0.938\*\*\*   
## (0.157)   
##   
## scale(obrazovanje) 0.010   
## (0.157)   
##   
## Constant 4,008.723\*\*\* 8.293\*\*\*   
## (0.995) (0.0005)   
##   
## ----------------------------------------------------------------------------------------------  
## Observations 150 150 150   
## R2 0.899 0.791 0.899   
## Adjusted R2 0.898 0.789 0.898   
## Residual Std. Error 5.706 (df = 147) 0.002 (df = 147) 0.318 (df = 148)   
## F Statistic 657.821\*\*\* (df = 2; 147) 278.873\*\*\* (df = 2; 147) 662.296\*\*\* (df = 2; 148)  
## ==============================================================================================  
## Note: \*p<0.1; \*\*p<0.05; \*\*\*p<0.01

m4<-lm(log(placa)~staz+obrazovanje,data=place)  
m5<-lm(placa~log(staz)+log(obrazovanje),data=place)  
stargazer(list(m1,m2,m4,m5),type="text")

##   
## ==============================================================================  
## Dependent variable:   
## -----------------------------------------------  
## placa log(placa) placa   
## (1) (2) (3) (4)   
## ------------------------------------------------------------------------------  
## staz 1.390\*\*\* 0.0003\*\*\*   
## (0.233) (0.0001)   
##   
## obrazovanje 0.015 0.00000   
## (0.239) (0.0001)   
##   
## log(staz) 0.002\*\*\* 7.780\*\*\*   
## (0.0004) (1.636)   
##   
## log(obrazovanje) 0.002\*\*\* 8.353\*\*\*   
## (0.0004) (1.614)   
##   
## Constant 4,008.723\*\*\* 8.293\*\*\* 8.296\*\*\* 3,995.443\*\*\*  
## (0.995) (0.0005) (0.0002) (1.954)   
##   
## ------------------------------------------------------------------------------  
## Observations 150 150 150 150   
## R2 0.899 0.791 0.899 0.791   
## Adjusted R2 0.898 0.789 0.898 0.788   
## Residual Std. Error (df = 147) 5.706 0.002 0.001 8.237   
## F Statistic (df = 2; 147) 657.821\*\*\* 278.873\*\*\* 657.624\*\*\* 277.371\*\*\*   
## ==============================================================================  
## Note: \*p<0.1; \*\*p<0.05; \*\*\*p<0.01

confint(m1,level=.9)

## 5 % 95 %  
## (Intercept) 4007.0755655 4010.3710507  
## staz 1.0046636 1.7760848  
## obrazovanje -0.3809056 0.4116561

summary(m1)

##   
## Call:  
## lm(formula = placa ~ staz + obrazovanje, data = place)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -11.6656 -4.0054 0.0725 4.0248 10.7120   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 4.009e+03 9.954e-01 4027.074 < 2e-16 \*\*\*  
## staz 1.390e+00 2.330e-01 5.967 1.73e-08 \*\*\*  
## obrazovanje 1.538e-02 2.394e-01 0.064 0.949   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 5.706 on 147 degrees of freedom  
## Multiple R-squared: 0.8995, Adjusted R-squared: 0.8981   
## F-statistic: 657.8 on 2 and 147 DF, p-value: < 2.2e-16

cbind(qt(0.95,147), qf(0.95,2,147))

## [,1] [,2]  
## [1,] 1.655285 3.057621

y<-place$placa  
sazetak<-summary(m1)  
cbind(sazetak$sigma,sazetak$sigma/mean(y))

## [,1] [,2]  
## [1,] 5.705662 0.001412991

ogranicenje<-"staz=500+obrazovanje"  
linearHypothesis(m1,ogranicenje,test="Chisq")

## Linear hypothesis test  
##   
## Hypothesis:  
## staz - obrazovanje = 500  
##   
## Model 1: restricted model  
## Model 2: placa ~ staz + obrazovanje  
##   
## Res.Df RSS Df Sum of Sq Chisq Pr(>Chisq)   
## 1 148 36525046   
## 2 147 4786 1 36520260 1121816 < 2.2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

ogranicenje<-"staz=obrazovanje"  
linearHypothesis(m1,ogranicenje,test="Chisq")

## Linear hypothesis test  
##   
## Hypothesis:  
## staz - obrazovanje = 0  
##   
## Model 1: restricted model  
## Model 2: placa ~ staz + obrazovanje  
##   
## Res.Df RSS Df Sum of Sq Chisq Pr(>Chisq)   
## 1 148 5063.2   
## 2 147 4785.5 1 277.71 8.5306 0.003492 \*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

ogranicenje<-"staz=500+obrazovanje"  
linearHypothesis(m1,ogranicenje,test="F")

## Linear hypothesis test  
##   
## Hypothesis:  
## staz - obrazovanje = 500  
##   
## Model 1: restricted model  
## Model 2: placa ~ staz + obrazovanje  
##   
## Res.Df RSS Df Sum of Sq F Pr(>F)   
## 1 148 36525046   
## 2 147 4786 1 36520260 1121816 < 2.2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

ogranicenje<-"staz=obrazovanje"  
linearHypothesis(m1,ogranicenje,test="F")

## Linear hypothesis test  
##   
## Hypothesis:  
## staz - obrazovanje = 0  
##   
## Model 1: restricted model  
## Model 2: placa ~ staz + obrazovanje  
##   
## Res.Df RSS Df Sum of Sq F Pr(>F)   
## 1 148 5063.2   
## 2 147 4785.5 1 277.71 8.5306 0.004043 \*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

ogranicenje<-c("obrazovanje=0")  
linearHypothesis(m1,ogranicenje,test="F")

## Linear hypothesis test  
##   
## Hypothesis:  
## obrazovanje = 0  
##   
## Model 1: restricted model  
## Model 2: placa ~ staz + obrazovanje  
##   
## Res.Df RSS Df Sum of Sq F Pr(>F)  
## 1 148 4785.7   
## 2 147 4785.5 1 0.13428 0.0041 0.9489

library(lmtest)  
model2<-lm(placa~obrazovanje,data=place)  
lrtest(m1,model2)

## Likelihood ratio test  
##   
## Model 1: placa ~ staz + obrazovanje  
## Model 2: placa ~ obrazovanje  
## #Df LogLik Df Chisq Pr(>Chisq)   
## 1 4 -472.54   
## 2 3 -488.81 -1 32.532 1.172e-08 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

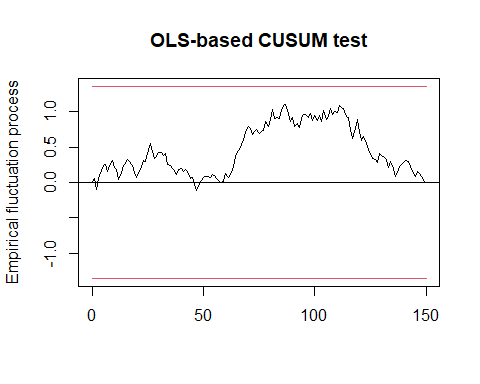
placa<-place$placa;staz<-place$staz;obraz<-place$obrazovanje  
placa1<-y[1:75];staz1<-staz[1:75];obraz1<-obraz[1:75]  
placa2<-y[76:150];staz2<-staz[76:150];obraz2<-obraz[76:150]  
library(gap)  
chow.test(placa,c(obraz1,staz1),placa2,c(obraz2,staz2))

## F value d.f.1 d.f.2 P value   
## 2.252555e+01 2.000000e+00 2.960000e+02 7.832376e-10

library(lmtest)  
resettest(m1,power=2:4,type="fitted")

##   
## RESET test  
##   
## data: m1  
## RESET = 0.60545, df1 = 3, df2 = 144, p-value = 0.6125

library(strucchange)  
placa<-ts(place$placa,start=1,frequency = 1)  
staz<-ts(place$staz,start=1,frequency = 1)  
obrazovanje<-ts(place$obrazovanje,start=1,frequency = 1)  
  
cusum <- efp(placa~staz+obrazovanje, type = "OLS-CUSUM")  
plot(cusum,xlab=NA)



fs <- Fstats(placa~staz+obrazovanje, from = 80, to = 120)  
sctest(fs, type="aveF")

##   
## aveF test  
##   
## data: fs  
## ave.F = 5.9081, p-value = 0.08881

novo<-data.frame(staz=19,obrazovanje=15)  
predict(m1,newdata = novo,interval = "confidence",level=.99)

## fit lwr upr  
## 1 4035.371 4033.712 4037.03

# 3. DALJNJA ANALIZA REGRESIJSKOG MODELA

## 3.1. Kvalitativne regresorske varijable

#### Slika 3.2. Učitavanje podataka i definiranje binarne varijable te procjena modela

bin<-read.table("binarne.txt",header=T,sep="\t")  
binarna<-ifelse(bin$spol=="m",1,0)  
summary(lm(placa~binarna,data=bin))

##   
## Call:  
## lm(formula = placa ~ binarna, data = bin)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -24.8617 -7.2568 0.2717 7.5248 16.7607   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 4023.544 1.087 3700.8 <2e-16 \*\*\*  
## binarna 30.122 1.569 19.2 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 9.602 on 148 degrees of freedom  
## Multiple R-squared: 0.7134, Adjusted R-squared: 0.7115   
## F-statistic: 368.4 on 1 and 148 DF, p-value: < 2.2e-16

#### Slika 3.3. Procijenjeni model

summary(lm(placa~binarna+staz+obrazovanje,data=bin))

##   
## Call:  
## lm(formula = placa ~ binarna + staz + obrazovanje, data = bin)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -11.9445 -4.0947 0.3051 4.0821 11.9770   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 4009.8377 1.0471 3829.499 < 2e-16 \*\*\*  
## binarna 5.0384 1.7609 2.861 0.00484 \*\*   
## staz 1.3311 0.2285 5.826 3.48e-08 \*\*\*  
## obrazovanje -0.1084 0.2377 -0.456 0.64918   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 5.571 on 146 degrees of freedom  
## Multiple R-squared: 0.9048, Adjusted R-squared: 0.9029   
## F-statistic: 462.7 on 3 and 146 DF, p-value: < 2.2e-16

#### Slika 3.4. Procijenjeni model

summary(lm(placa~staz+obrazovanje+I(staz\*binarna)+I(obrazovanje\*binarna),data=bin))

##   
## Call:  
## lm(formula = placa ~ staz + obrazovanje + I(staz \* binarna) +   
## I(obrazovanje \* binarna), data = bin)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -11.9881 -4.2011 0.2059 3.9496 12.1290   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 4009.9901 1.1464 3497.869 < 2e-16 \*\*\*  
## staz 1.4493 0.2915 4.972 1.85e-06 \*\*\*  
## obrazovanje -0.2500 0.3195 -0.782 0.435   
## I(staz \* binarna) -0.1302 0.4386 -0.297 0.767   
## I(obrazovanje \* binarna) 0.3163 0.4879 0.648 0.518   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 5.638 on 145 degrees of freedom  
## Multiple R-squared: 0.9032, Adjusted R-squared: 0.9005   
## F-statistic: 338.3 on 4 and 145 DF, p-value: < 2.2e-16

#### Slika 3.5. Generiranje binarnih varijabli i procijenjeni model

bin1<-ifelse(bin$staz>=15,1,0)  
bin2<-ifelse(bin$obrazovanje>=12,1,0)  
summary(lm(placa~bin1\*bin2,data=bin))

##   
## Call:  
## lm(formula = placa ~ bin1 \* bin2, data = bin)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -26.1027 -7.3073 0.0007 6.7974 21.9165   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 4017.410 1.490 2697.007 <2e-16 \*\*\*  
## bin1 13.659 10.533 1.297 0.1967   
## bin2 13.711 7.522 1.823 0.0704 .   
## bin1:bin2 3.730 12.900 0.289 0.7729   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 10.43 on 146 degrees of freedom  
## Multiple R-squared: 0.6666, Adjusted R-squared: 0.6598   
## F-statistic: 97.32 on 3 and 146 DF, p-value: < 2.2e-16

#### Slika 3.6. Procijenjen model

cd<-read.table("cobb-douglas.txt",sep="\t",header=T)  
summary(lm(log(proizvodnja)~log(rad)+log(kapital),data=cd))

##   
## Call:  
## lm(formula = log(proizvodnja) ~ log(rad) + log(kapital), data = cd)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.0034333 -0.0015728 -0.0000916 0.0013707 0.0055202   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -0.047580 0.014192 -3.353 0.00186 \*\*   
## log(rad) 0.371616 0.003244 114.546 < 2e-16 \*\*\*  
## log(kapital) 0.624127 0.003436 181.667 < 2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.002326 on 37 degrees of freedom  
## Multiple R-squared: 0.9998, Adjusted R-squared: 0.9998   
## F-statistic: 1.15e+05 on 2 and 37 DF, p-value: < 2.2e-16

#### Slika 3.7. Waldov test

model<-lm(log(proizvodnja)~log(rad)+log(kapital),data=cd)  
library(car)  
ogranicenje<-"log(rad)+log(kapital)=1"  
linearHypothesis(model,ogranicenje,test="Chisq")

## Linear hypothesis test  
##   
## Hypothesis:  
## log(rad) + log(kapital) = 1  
##   
## Model 1: restricted model  
## Model 2: log(proizvodnja) ~ log(rad) + log(kapital)  
##   
## Res.Df RSS Df Sum of Sq Chisq Pr(>Chisq)   
## 1 38 0.00022289   
## 2 37 0.00020014 1 2.2741e-05 4.2041 0.04033 \*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

## 3.3. Pitanja za ponavljanje, Rješenja

#### 3.3. Pitanja za ponavljanje, Rješenja, Zadatak 6

potrosnja<-read.table("potrosnja.txt",header=T,sep="\t")  
  
bin<-ifelse(potrosnja$cijene>=125,1,0)  
  
m1<-lm(potrosnja~bdp+bin,data=potrosnja)  
m2<-lm(potrosnja~bdp+I(bin\*bdp),data=potrosnja)  
m3<-lm(potrosnja~bin,data=potrosnja)  
  
library(stargazer)  
stargazer(list(m1,m2,m3),type="text")

##   
## ==================================================================================================  
## Dependent variable:   
## ------------------------------------------------------------------------------  
## potrosnja   
## (1) (2) (3)   
## --------------------------------------------------------------------------------------------------  
## bdp 0.645\*\*\* 0.647\*\*\*   
## (0.003) (0.018)   
##   
## bin 7,458.418 135,676.000   
## (5,780.006) (102,416.000)   
##   
## I(bin \* bdp) -0.002   
## (0.018)   
##   
## Constant 988.872 4,520.722 24,872.530   
## (4,007.148) (2,933.061) (71,377.000)   
##   
## --------------------------------------------------------------------------------------------------  
## Observations 140 140 140   
## R2 0.997 0.997 0.013   
## Adjusted R2 0.997 0.997 0.005   
## Residual Std. Error 33,987.950 (df = 137) 34,192.080 (df = 137) 605,654.000 (df = 138)  
## F Statistic 22,120.430\*\*\* (df = 2; 137) 21,856.290\*\*\* (df = 2; 137) 1.755 (df = 1; 138)   
## ==================================================================================================  
## Note: \*p<0.1; \*\*p<0.05; \*\*\*p<0.01

bin2<-ifelse(potrosnja$bdp>=1500,1,0)  
  
m4<-lm(potrosnja~bin+bin2,data=potrosnja)  
m5<-lm(potrosnja~bin\*bin2,data=potrosnja)  
  
stargazer(list(m4,m5),type="text")

##   
## =================================================================  
## Dependent variable:   
## ---------------------------------------------  
## potrosnja   
## (1) (2)   
## -----------------------------------------------------------------  
## bin 116,490.000 34.242   
## (102,469.500) (145,249.100)   
##   
## bin2 167,740.600 55,444.060   
## (102,427.700) (142,631.800)   
##   
## bin:bin2 231,379.100   
## (204,736.500)   
##   
## Constant -49,678.860 230.725   
## (84,295.390) (95,087.870)   
##   
## -----------------------------------------------------------------  
## Observations 140 140   
## R2 0.032 0.041   
## Adjusted R2 0.017 0.019   
## Residual Std. Error 601,996.600 (df = 137) 601,388.500 (df = 136)  
## F Statistic 2.229 (df = 2; 137) 1.915 (df = 3; 136)   
## =================================================================  
## Note: \*p<0.1; \*\*p<0.05; \*\*\*p<0.01

# 4. NARUŠAVANJE PRETPOSTAVKI REGRESIJSKOG MODELA

#### Slika 4.1. Ispitivanje multikolinearnosti u modelu

stanovi<-read.table("stanovi.txt",sep="\t",header=T)  
model<-lm(cijena~kvadrat+sobe+godine+udaljenost,data=stanovi)  
library(car)  
vif(model)

## kvadrat sobe godine udaljenost   
## 26.697603 26.456214 1.088931 1.103061

1/vif(model)

## kvadrat sobe godine udaljenost   
## 0.03745655 0.03779830 0.91833149 0.90656791

#### Slika 4.2. Koeficijenti determinacije Rj2

m1<-lm(kvadrat~sobe+godine+udaljenost,data=stanovi)  
m2<-lm(sobe~kvadrat+godine+udaljenost,data=stanovi)  
summary(m1)$r.squared

## [1] 0.9625435

summary(m2)$r.squared

## [1] 0.9622017

#### Slika 4.3. Korelacijska matrica i koeficijent korelacije regresije

cor(stanovi)

## cijena kvadrat sobe godine udaljenost  
## cijena 1.0000000 0.78350150 0.772192289 -0.52535133 -0.417120020  
## kvadrat 0.7835015 1.00000000 0.978988399 -0.09359588 0.043776437  
## sobe 0.7721923 0.97898840 1.000000000 -0.05742016 -0.001503659  
## godine -0.5253513 -0.09359588 -0.057420159 1.00000000 0.165709301  
## udaljenost -0.4171200 0.04377644 -0.001503659 0.16570930 1.000000000

sqrt(summary(model)$r.squared)

## [1] 0.9820341

#### Slika 4.4. Ispis procijenjenog modela

summary(model)

##   
## Call:  
## lm(formula = cijena ~ kvadrat + sobe + godine + udaljenost, data = stanovi)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -12.2746 -5.0720 -0.2008 4.8410 11.9819   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 2009.7457 2.9255 686.966 < 2e-16 \*\*\*  
## kvadrat 0.6240 0.1246 5.010 8.9e-06 \*\*\*  
## sobe 0.6457 2.5767 0.251 0.803   
## godine -1.4330 0.1075 -13.336 < 2e-16 \*\*\*  
## udaljenost -2.6140 0.2011 -13.000 < 2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 6.186 on 45 degrees of freedom  
## Multiple R-squared: 0.9644, Adjusted R-squared: 0.9612   
## F-statistic: 304.7 on 4 and 45 DF, p-value: < 2.2e-16

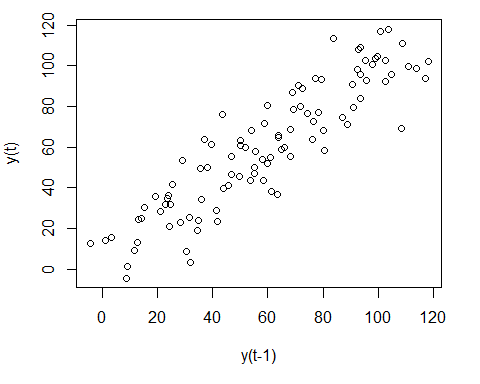
#### Slika 4.5. Procijenjeni modeli uz isključivanje neke od nezavisnih varijabli

m2<-lm(cijena~sobe+godine+udaljenost,data=stanovi)  
m3<-lm(cijena~kvadrat+godine+udaljenost,data=stanovi)  
library(stargazer)  
stargazer(list(m2,m3),type="text")

##   
## ==========================================================  
## Dependent variable:   
## ----------------------------  
## cijena   
## (1) (2)   
## ----------------------------------------------------------  
## sobe 13.308\*\*\*   
## (0.619)   
##   
## kvadrat 0.655\*\*\*   
## (0.024)   
##   
## godine -1.556\*\*\* -1.427\*\*\*   
## (0.129) (0.104)   
##   
## udaljenost -2.352\*\*\* -2.627\*\*\*   
## (0.240) (0.193)   
##   
## Constant 2,008.432\*\*\* 2,009.915\*\*\*   
## (3.597) (2.818)   
##   
## ----------------------------------------------------------  
## Observations 50 50   
## R2 0.945 0.964   
## Adjusted R2 0.941 0.962   
## Residual Std. Error (df = 46) 7.637 6.123   
## F Statistic (df = 3; 46) 261.092\*\*\* 414.669\*\*\*   
## ==========================================================  
## Note: \*p<0.1; \*\*p<0.05; \*\*\*p<0.01

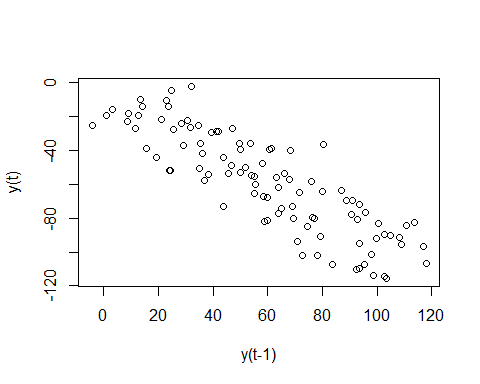
#### Slika 4.6. Pozitivna autokorelacija prvog reda

eps<-seq(1:100)  
eps<-eps+10\*rnorm(100,1,1)  
eps\_1<-c(NA,eps[1:99])  
  
par(mfrow = c(1,1),oma=c(1,0,0,1),mar=c(4,4,1,1))  
plot(eps\_1,eps,xlab="y(t-1)",ylab="y(t)")



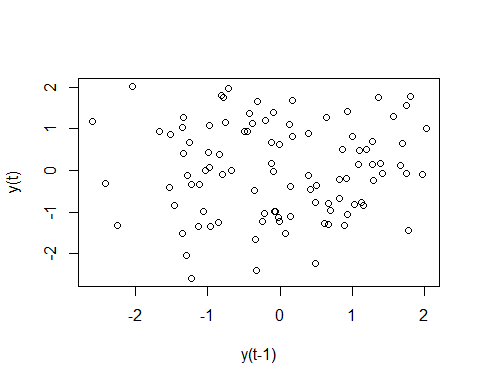
#### Slika 4.7. Negativna autokorelacija prvog reda

eps2<-seq(1:100)  
eps2<-(-eps2-10\*rnorm(100,1,1))  
eps2\_1<-c(NA,eps[1:99])  
plot(eps2\_1,eps2,xlab="y(t-1)",ylab="y(t)")



#### Slika 4.8. Nepostojanje autokorelacija prvog reda

eps3<-rnorm(100,0,1)  
eps3\_1<-c(NA,eps3[1:99])  
plot(eps3\_1,eps3,xlab="y(t-1)",ylab="y(t)")



#### Slika 4.10. Durbin-Watson test

library(car)  
durbinWatsonTest(model)

## lag Autocorrelation D-W Statistic p-value  
## 1 0.006259401 1.970795 0.936  
## Alternative hypothesis: rho != 0

durbinWatsonTest(model,alternative = "positive")

## lag Autocorrelation D-W Statistic p-value  
## 1 0.006259401 1.970795 0.424  
## Alternative hypothesis: rho > 0

durbinWatsonTest(model,alternative = "negative")

## lag Autocorrelation D-W Statistic p-value  
## 1 0.006259401 1.970795 0.558  
## Alternative hypothesis: rho < 0

#### Slika 4.11. Provedba Breusch-Godfrey testa

library(quantmod)

## Loading required package: xts

## Loading required package: TTR

## Registered S3 method overwritten by 'quantmod':  
## method from  
## as.zoo.data.frame zoo

reziduali<-residuals(model)  
rez1<-Lag(reziduali,1);rez2<-Lag(reziduali,2);rez3<-Lag(reziduali,3)  
summary(pomocna<-lm(reziduali~kvadrat+sobe+godine+udaljenost+rez1+rez2+rez3,data=stanovi))

##   
## Call:  
## lm(formula = reziduali ~ kvadrat + sobe + godine + udaljenost +   
## rez1 + rez2 + rez3, data = stanovi)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -12.2809 -4.9404 0.4371 4.2712 12.3717   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)  
## (Intercept) -0.232455 3.156806 -0.074 0.942  
## kvadrat -0.031593 0.133236 -0.237 0.814  
## sobe 0.587398 2.748807 0.214 0.832  
## godine 0.010039 0.121023 0.083 0.934  
## udaljenost -0.003460 0.217382 -0.016 0.987  
## rez1 -0.001735 0.162590 -0.011 0.992  
## rez2 0.086990 0.166745 0.522 0.605  
## rez3 -0.024371 0.164186 -0.148 0.883  
##   
## Residual standard error: 6.449 on 39 degrees of freedom  
## (3 observations deleted due to missingness)  
## Multiple R-squared: 0.008903, Adjusted R-squared: -0.169   
## F-statistic: 0.05005 on 7 and 39 DF, p-value: 0.9998

#### Slika 4.12. Izračun empirijske veličine Breusch-Godfrey testa i pripadajuće p-vrijednosti

test\_vel<-nobs(pomocna)\*summary(pomocna)$r.squared  
test\_vel

## [1] 0.4184306

p\_v<-1-pchisq(test\_vel,3)  
p\_v

## [1] 0.9364094

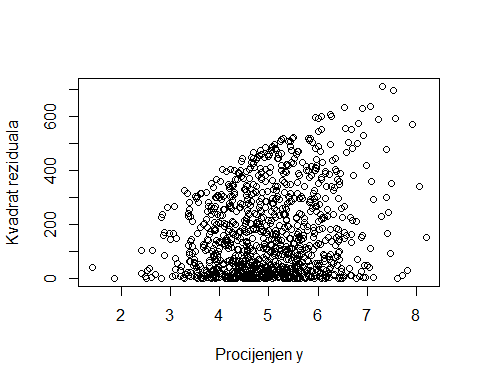
#### Slika 4.13. Ljung-Box test autokorelacije

Box.test(reziduali,lag=3,type="Ljung-Box")

##   
## Box-Ljung test  
##   
## data: reziduali  
## X-squared = 0.31157, df = 3, p-value = 0.9578

#### Slika 4.15. Grafičko utvrđivanje postojanja heteroskedastičnosti

procijenjen\_y<-rnorm(1000,5,1)  
res2<-procijenjen\_y  
  
for (i in 1:1000)  
{  
 res2[i]<-procijenjen\_y[i]\*(i^2)/10000+2  
}  
  
plot(procijenjen\_y,res2,xlab="Procijenjen y",ylab="Kvadrat reziduala")



##### 4.3.3.2. Newey-West korekcija standardnih pogrešaka procjenitelja

#### Slika 4.16. Breusch-Paganov test heteroskedastičnosti

summary(bptest<-lm(residuals(model)^2~kvadrat+sobe+godine+udaljenost,data=stanovi))

##   
## Call:  
## lm(formula = residuals(model)^2 ~ kvadrat + sobe + godine + udaljenost,   
## data = stanovi)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -48.832 -23.290 -8.723 14.187 121.484   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)  
## (Intercept) 12.3898 18.1963 0.681 0.499  
## kvadrat -0.7793 0.7747 -1.006 0.320  
## sobe 19.0070 16.0267 1.186 0.242  
## godine 0.2025 0.6684 0.303 0.763  
## udaljenost 0.6638 1.2507 0.531 0.598  
##   
## Residual standard error: 38.48 on 45 degrees of freedom  
## Multiple R-squared: 0.0514, Adjusted R-squared: -0.03292   
## F-statistic: 0.6096 on 4 and 45 DF, p-value: 0.6578

nobs(bptest);(summary(bptest))$r.squared

## [1] 50

## [1] 0.05140362

nobs(bptest)\*(summary(bptest))$r.squared

## [1] 2.570181

p\_vrijednost<-1-pchisq(nobs(bptest)\*(summary(bptest))$r.squared,4)  
p\_vrijednost

## [1] 0.6321144

#### Slika 4.17. Whiteov test heteroskedastičnosti

summary(white<-lm(residuals(model)^2~  
 kvadrat+sobe+godine+udaljenost  
 +I(kvadrat^2)+I(sobe^2)+I(godine^2)+I(udaljenost^2)  
 +I(kvadrat\*sobe)+I(kvadrat\*godine)+I(kvadrat\*udaljenost)  
 +I(sobe\*godine)+I(sobe\*udaljenost)+I(godine\*udaljenost)  
 ,data=stanovi))

##   
## Call:  
## lm(formula = residuals(model)^2 ~ kvadrat + sobe + godine + udaljenost +   
## I(kvadrat^2) + I(sobe^2) + I(godine^2) + I(udaljenost^2) +   
## I(kvadrat \* sobe) + I(kvadrat \* godine) + I(kvadrat \* udaljenost) +   
## I(sobe \* godine) + I(sobe \* udaljenost) + I(godine \* udaljenost),   
## data = stanovi)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -52.691 -21.373 -3.651 15.247 85.288   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 30.60218 55.66675 0.550 0.5860   
## kvadrat 2.17481 2.63676 0.825 0.4151   
## sobe -58.03190 58.33160 -0.995 0.3266   
## godine 5.44016 3.69522 1.472 0.1499   
## udaljenost -2.28104 6.50207 -0.351 0.7278   
## I(kvadrat^2) 0.02174 0.11802 0.184 0.8549   
## I(sobe^2) 10.06745 51.40239 0.196 0.8459   
## I(godine^2) -0.21858 0.11503 -1.900 0.0657 .  
## I(udaljenost^2) 0.14422 0.36655 0.393 0.6964   
## I(kvadrat \* sobe) -0.80112 4.90070 -0.163 0.8711   
## I(kvadrat \* godine) -0.19547 0.11578 -1.688 0.1002   
## I(kvadrat \* udaljenost) -0.15968 0.21289 -0.750 0.4582   
## I(sobe \* godine) 3.64219 2.43552 1.495 0.1438   
## I(sobe \* udaljenost) 2.56008 4.47629 0.572 0.5710   
## I(godine \* udaljenost) 0.26275 0.17227 1.525 0.1362   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 38.86 on 35 degrees of freedom  
## Multiple R-squared: 0.2473, Adjusted R-squared: -0.05374   
## F-statistic: 0.8215 on 14 and 35 DF, p-value: 0.6421

nobs(white);(summary(white))$r.squared

## [1] 50

## [1] 0.2473302

nobs(white)\*(summary(white))$r.squared

## [1] 12.36651

p\_vrijednost<-1-pchisq(nobs(white)\*(summary(white))$r.squared,14)  
p\_vrijednost

## [1] 0.5768947

#### Slika 4.18. Whiteova korekcija standardnih pogrešaka procjenitelja

#White korekcija:  
library(car)  
mat<-hccm(model,type="hc0")  
mat #ovo je matrica var-kovar procjenitelja uz White korekciju

## (Intercept) kvadrat sobe godine udaljenost  
## (Intercept) 7.2228921 0.1026964675 -3.128843180 -0.1022419712 -0.151749376  
## kvadrat 0.1026965 0.0142705101 -0.298465323 -0.0001960663 -0.007060468  
## sobe -3.1288432 -0.2984653227 6.518833756 0.0026099914 0.134398624  
## godine -0.1022420 -0.0001960663 0.002609991 0.0079664313 -0.001587154  
## udaljenost -0.1517494 -0.0070604678 0.134398624 -0.0015871545 0.036076413

#t-test :  
library(lmtest)  
coeftest(model,vcov=mat)

##   
## t test of coefficients:  
##   
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 2009.745687 2.687544 747.8001 < 2.2e-16 \*\*\*  
## kvadrat 0.624027 0.119459 5.2238 4.352e-06 \*\*\*  
## sobe 0.645730 2.553201 0.2529 0.8015   
## godine -1.433028 0.089255 -16.0555 < 2.2e-16 \*\*\*  
## udaljenost -2.614038 0.189938 -13.7626 < 2.2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

#### Slika 4.19. Newey-West korekcija standardnih pogrešaka procjenitelja

#Newey-West korekcija:  
library(sandwich)  
mat2<-NeweyWest(model,lag=2)  
mat2 #ovo je matrica var-kovar procjenitelja uz Newey-West korekciju

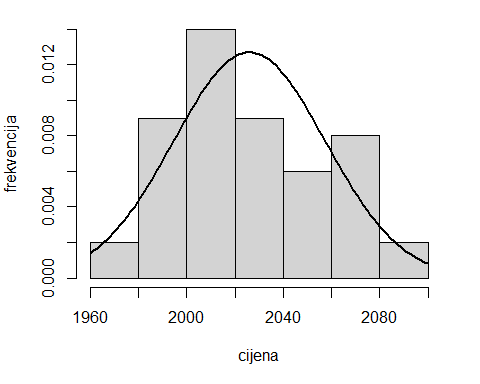
## (Intercept) kvadrat sobe godine udaljenost  
## (Intercept) 9.5825437 0.130658529 -3.79284882 -0.210233655 -0.128148233  
## kvadrat 0.1306585 0.010619590 -0.22636983 -0.002575908 -0.007324410  
## sobe -3.7928488 -0.226369830 5.09602924 0.065306161 0.118564355  
## godine -0.2102337 -0.002575908 0.06530616 0.007715074 0.003935095  
## udaljenost -0.1281482 -0.007324410 0.11856435 0.003935095 0.028147677

#t-test :  
library(lmtest)  
coeftest(model,vcov=mat2)

##   
## t test of coefficients:  
##   
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 2009.745687 3.095568 649.2332 < 2.2e-16 \*\*\*  
## kvadrat 0.624027 0.103051 6.0555 2.583e-07 \*\*\*  
## sobe 0.645730 2.257439 0.2860 0.7762   
## godine -1.433028 0.087835 -16.3149 < 2.2e-16 \*\*\*  
## udaljenost -2.614038 0.167773 -15.5808 < 2.2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

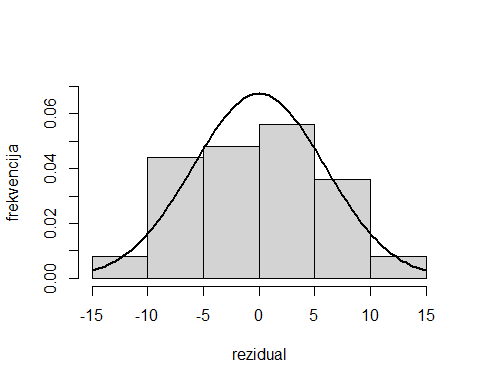
#### Slika 4.20. Usporedba histograma s normalnom distribucijom , cijena m2 stana

stanovi<-read.table("stanovi.txt",sep="\t",header=T)  
  
par(mfrow = c(1,1),oma=c(1,0,0,1),mar=c(4,4,1,1))  
hist(stanovi$cijena,main=NA,xlab="cijena",ylab="frekvencija",probability = T)  
curve(dnorm(x,mean(stanovi$cijena),sd(stanovi$cijena)),  
 lwd=2, add=TRUE, yaxt="n")



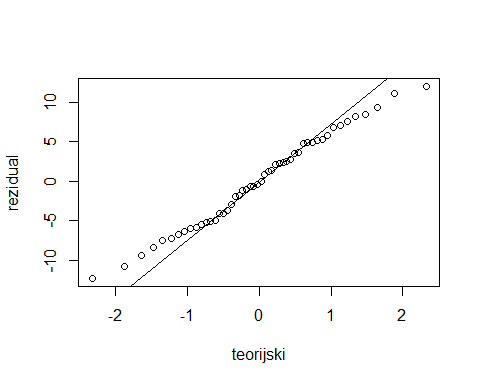
#### Slika 4.21. Usporedba histograma s normalnom distribucijom, rezidualna odstupanja regresijskog modela

model<-lm(cijena~kvadrat+sobe+godine+udaljenost,data=stanovi)  
rezidual<-resid(model)  
hist(rezidual,main=NA,xlab="rezidual",ylab="frekvencija",probability = T,ylim=c(0,.07))  
curve(dnorm(x,mean(rezidual),sd(rezidual)),  
 lwd=2, add=TRUE, yaxt="n")



#### Slika 4.22. Dijagram vjerojatnosti rezidualnih odstupanja regresijskog modela

qqnorm(rezidual, pch = 1, frame = T,main=NA,xlab="teorijski",ylab="rezidual")  
qqline(rezidual, col = "black", lwd = 1)



### 4.4.3. Ublažavanje/uklanjanje problema nenormalnosti distribucije grešaka relacije

#### Slika 4.23. Jarque-Bera test normalnosti

#Jarque-Bera test:  
reziduali<-resid(model)  
library("moments")  
jarque.test(reziduali)

##   
## Jarque-Bera Normality Test  
##   
## data: reziduali  
## JB = 1.3873, p-value = 0.4998  
## alternative hypothesis: greater

#N, koef. asimetrije i zaobljenosti:  
n<-length(reziduali); s<-skewness(reziduali); k<-kurtosis(reziduali)  
n;s;k

## [1] 50

## [1] 0.002649004

## [1] 2.183994

JB<-n\*(s^2/6+(k-3)^2/24)  
JB

## [1] 1.387279

#### Slika 4.24. Anderson-Darlingov, Cramér-von-Misesov i Lilliefors testovi normalnosti

library(nortest)  
ad.test(reziduali)

##   
## Anderson-Darling normality test  
##   
## data: reziduali  
## A = 0.25149, p-value = 0.7265

cvm.test(reziduali)

##   
## Cramer-von Mises normality test  
##   
## data: reziduali  
## W = 0.041388, p-value = 0.6485

lillie.test(reziduali)

##   
## Lilliefors (Kolmogorov-Smirnov) normality test  
##   
## data: reziduali  
## D = 0.078498, p-value = 0.6177

## 4.5. Alternativne metode procjene parametara

### 4.5.1. Generalizirana metoda najmanjih kvadrata

#### Slika 4.25. Procijenjeni modeli metodom najmanjih kvadrata (OLS) i generaliziranom metodom najmanjih kvadrata (GLS)

OLS<-lm(cijena~kvadrat+sobe+godine+udaljenost,data=stanovi)  
library(nlme)  
GLS<-gls(cijena~kvadrat+sobe+godine+udaljenost,  
 data=stanovi,weights=varExp())  
#usporedba procjena OLS i GLS metodom:  
summary(OLS)$coefficients

## Estimate Std. Error t value Pr(>|t|)  
## (Intercept) 2009.7456865 2.9255378 686.9662462 4.047604e-92  
## kvadrat 0.6240274 0.1245594 5.0098793 8.901349e-06  
## sobe 0.6457300 2.5767277 0.2506008 8.032635e-01  
## godine -1.4330282 0.1074557 -13.3359863 3.066576e-17  
## udaljenost -2.6140377 0.2010817 -12.9998786 7.671376e-17

summary(GLS)$tTable

## Value Std.Error t-value p-value  
## (Intercept) 2009.6828429 2.9202770 688.182259 3.737988e-92  
## kvadrat 0.6324484 0.1243709 5.085182 6.923700e-06  
## sobe 0.4781664 2.5740122 0.185767 8.534624e-01  
## godine -1.4270009 0.1073192 -13.296785 3.410254e-17  
## udaljenost -2.6174081 0.2004430 -13.058115 6.537902e-17

#### Slika 4.26. Procijenjeni modeli metodom najmanjih kvadrata (OLS) i ponderiranom metodom najmanjih kvadrata (WLS)

OLS<-lm(cijena~kvadrat+sobe+godine+udaljenost,data=stanovi)  
#WLS:  
WLS<-lm(cijena~kvadrat+sobe+godine+udaljenost,data=stanovi,weights=1/sobe)  
summary(OLS)$coefficients

## Estimate Std. Error t value Pr(>|t|)  
## (Intercept) 2009.7456865 2.9255378 686.9662462 4.047604e-92  
## kvadrat 0.6240274 0.1245594 5.0098793 8.901349e-06  
## sobe 0.6457300 2.5767277 0.2506008 8.032635e-01  
## godine -1.4330282 0.1074557 -13.3359863 3.066576e-17  
## udaljenost -2.6140377 0.2010817 -12.9998786 7.671376e-17

summary(WLS)$coefficients

## Estimate Std. Error t value Pr(>|t|)  
## (Intercept) 2011.9939690 2.6863136 748.979545 8.284906e-94  
## kvadrat 0.8166275 0.1220448 6.691213 2.928920e-08  
## sobe -3.7453205 2.5374195 -1.476035 1.468982e-01  
## godine -1.3830705 0.1044945 -13.235816 4.024384e-17  
## udaljenost -2.7022614 0.2003145 -13.490094 2.023451e-17

#### Slika 4.27. Ispitivanje multikolinearnosti varijabli u modelu

cd<-read.table("cobb-douglas.txt",sep="\t",header=T)  
  
#multikolinearnost  
model<-lm(log(proizvodnja)~log(rad)+log(kapital),data=cd)  
library(car)  
vif(model)

## log(rad) log(kapital)   
## 2.884414 2.884414

1/vif(model)

## log(rad) log(kapital)   
## 0.3466909 0.3466909

m2<-lm(log(rad)~log(kapital),data=cd)  
summary(m2)$r.squared

## [1] 0.6533091

cor(cd)

## proizvodnja rad kapital  
## proizvodnja 1.0000000 0.9244237 0.9706802  
## rad 0.9244237 1.0000000 0.8057383  
## kapital 0.9706802 0.8057383 1.0000000

sqrt(summary(model)$r.squared)

## [1] 0.9999196

#### Slika 4.28. Durbin-Watsonov test autokorelacije

library(car)  
durbinWatsonTest(model)

## lag Autocorrelation D-W Statistic p-value  
## 1 0.236226 1.481046 0.108  
## Alternative hypothesis: rho != 0

durbinWatsonTest(model,alternative = "positive")

## lag Autocorrelation D-W Statistic p-value  
## 1 0.236226 1.481046 0.049  
## Alternative hypothesis: rho > 0

durbinWatsonTest(model,alternative = "negative")

## lag Autocorrelation D-W Statistic p-value  
## 1 0.236226 1.481046 0.933  
## Alternative hypothesis: rho < 0

#### Slika 4.29. Breusch-Godfrey test autokorelacije

reziduali<-residuals(model)  
rez1<-Lag(reziduali,1);rez2<-Lag(reziduali,2)  
summary(pomocna<-lm(reziduali~log(rad)+log(kapital)+rez1+rez2,data=cd))

##   
## Call:  
## lm(formula = reziduali ~ log(rad) + log(kapital) + rez1 + rez2,   
## data = cd)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.0031862 -0.0014355 0.0000960 0.0009503 0.0057225   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -0.0003672 0.0142200 -0.026 0.9796   
## log(rad) 0.0030470 0.0033999 0.896 0.3766   
## log(kapital) -0.0034138 0.0036401 -0.938 0.3551   
## rez1 0.3456974 0.1790120 1.931 0.0621 .  
## rez2 -0.1107569 0.1672579 -0.662 0.5124   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.002261 on 33 degrees of freedom  
## (2 observations deleted due to missingness)  
## Multiple R-squared: 0.1074, Adjusted R-squared: -0.0007681   
## F-statistic: 0.9929 on 4 and 33 DF, p-value: 0.4251

test\_vel<-nobs(pomocna)\*summary(pomocna)$r.squared  
test\_vel

## [1] 4.082075

p\_v<-1-pchisq(test\_vel,2)  
p\_v

## [1] 0.1298939

#### Slika 4.30. Ljung-Box test autokorelacije

Box.test(reziduali,lag=2,type="Ljung-Box")

##   
## Box-Ljung test  
##   
## data: reziduali  
## X-squared = 2.4395, df = 2, p-value = 0.2953

#### Slika 4.31. Breusch-Paganov test heteroskedastičnosti

summary(bptest<-lm(residuals(model)^2~log(rad)+log(kapital),data=cd))

##   
## Call:  
## lm(formula = residuals(model)^2 ~ log(rad) + log(kapital), data = cd)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -8.994e-06 -4.282e-06 -1.471e-06 1.695e-06 2.271e-05   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 8.527e-05 4.176e-05 2.042 0.0483 \*  
## log(rad) -1.903e-05 9.546e-06 -1.994 0.0536 .  
## log(kapital) 9.093e-06 1.011e-05 0.899 0.3742   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 6.844e-06 on 37 degrees of freedom  
## Multiple R-squared: 0.1281, Adjusted R-squared: 0.08096   
## F-statistic: 2.718 on 2 and 37 DF, p-value: 0.0792

nobs(bptest);(summary(bptest))$r.squared

## [1] 40

## [1] 0.1280905

nobs(bptest)\*(summary(bptest))$r.squared

## [1] 5.123619

p\_vrijednost<-1-pchisq(nobs(bptest)\*(summary(bptest))$r.squared,2)  
p\_vrijednost

## [1] 0.07716498

#### Slika 4.32. Whiteov test heteroskedastičnosti

summary(white<-lm(residuals(model)^2~log(rad)+log(kapital)  
 +I(log(rad)\*log(kapital))  
 +I(log(rad)^2)+I(log(kapital)^2),data=cd))

##   
## Call:  
## lm(formula = residuals(model)^2 ~ log(rad) + log(kapital) + I(log(rad) \*   
## log(kapital)) + I(log(rad)^2) + I(log(kapital)^2), data = cd)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -8.799e-06 -3.395e-06 -5.574e-07 3.149e-06 1.396e-05   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -2.088e-04 1.292e-03 -0.162 0.872582   
## log(rad) -9.164e-04 3.770e-04 -2.431 0.020494 \*   
## log(kapital) 1.117e-03 3.069e-04 3.641 0.000894 \*\*\*  
## I(log(rad) \* log(kapital)) -5.738e-04 1.281e-04 -4.481 8.01e-05 \*\*\*  
## I(log(rad)^2) 3.141e-04 6.083e-05 5.163 1.06e-05 \*\*\*  
## I(log(kapital)^2) 2.397e-04 7.946e-05 3.017 0.004812 \*\*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 4.899e-06 on 34 degrees of freedom  
## Multiple R-squared: 0.5895, Adjusted R-squared: 0.5291   
## F-statistic: 9.765 on 5 and 34 DF, p-value: 7.505e-06

nobs(white);(summary(white))$r.squared

## [1] 40

## [1] 0.5894874

nobs(white)\*(summary(white))$r.squared

## [1] 23.5795

p\_vrijednost<-1-pchisq(nobs(white)\*(summary(white))$r.squared,5)  
p\_vrijednost

## [1] 0.0002614601

#### Slika 4.33. Whiteova i Newey-Westova korekcija standardnih pogrešaka

OLS<-lm(log(proizvodnja)~log(rad)+log(kapital),data=cd)  
summary(OLS)$coefficients

## Estimate Std. Error t value Pr(>|t|)  
## (Intercept) -0.04758001 0.014192190 -3.352549 1.856875e-03  
## log(rad) 0.37161629 0.003244253 114.546040 8.362369e-49  
## log(kapital) 0.62412677 0.003435547 181.667369 3.346718e-56

#White korekcija:  
library(car)  
mat<-hccm(model,type="hc0")  
#t-test :  
library(lmtest)  
coeftest(model,vcov=mat)

##   
## t test of coefficients:  
##   
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -0.0475800 0.0134602 -3.5349 0.001116 \*\*   
## log(rad) 0.3716163 0.0045908 80.9485 < 2.2e-16 \*\*\*  
## log(kapital) 0.6241268 0.0039069 159.7493 < 2.2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

#Newey-West korekcija:  
library(sandwich)  
mat2<-NeweyWest(model,lag=2)  
#t-test :  
library(lmtest)  
coeftest(model,vcov=mat2)

##   
## t test of coefficients:  
##   
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -0.0475800 0.0144590 -3.2907 0.002202 \*\*   
## log(rad) 0.3716163 0.0042342 87.7656 < 2.2e-16 \*\*\*  
## log(kapital) 0.6241268 0.0033445 186.6146 < 2.2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

#### Slika 4.34. Testiranje normalnosti distribucije rezidualnih odstupanja

reziduali<-resid(model)  
library("moments")  
jarque.test(reziduali)

##   
## Jarque-Bera Normality Test  
##   
## data: reziduali  
## JB = 3.9295, p-value = 0.1402  
## alternative hypothesis: greater

#N, koef. asimetrije i zaobljenosti:  
n<-length(reziduali); s<-skewness(reziduali); k<-kurtosis(reziduali)  
n;s;k

## [1] 40

## [1] 0.7676979

## [1] 2.984667

JB<-n\*(s^2/6+(k-3)^2/24)  
JB

## [1] 3.929459

#  
library(nortest)  
ad.test(reziduali)

##   
## Anderson-Darling normality test  
##   
## data: reziduali  
## A = 0.82998, p-value = 0.0294

cvm.test(reziduali)

##   
## Cramer-von Mises normality test  
##   
## data: reziduali  
## W = 0.13803, p-value = 0.03259

lillie.test(reziduali)

##   
## Lilliefors (Kolmogorov-Smirnov) normality test  
##   
## data: reziduali  
## D = 0.16657, p-value = 0.006836

#### Slika 4.35. Rezultati procjena OLS i GLS metodom

GLS<-gls(log(proizvodnja)~log(rad)+log(kapital),data=cd,weights=varExp())  
summary(OLS)$coefficients

## Estimate Std. Error t value Pr(>|t|)  
## (Intercept) -0.04758001 0.014192190 -3.352549 1.856875e-03  
## log(rad) 0.37161629 0.003244253 114.546040 8.362369e-49  
## log(kapital) 0.62412677 0.003435547 181.667369 3.346718e-56

summary(GLS)$tTable

## Value Std.Error t-value p-value  
## (Intercept) -0.02643646 0.014688625 -1.799791 8.005225e-02  
## log(rad) 0.37617018 0.002700396 139.301860 6.099844e-52  
## log(kapital) 0.61568310 0.002680711 229.671581 5.756774e-60

#### Slika 4.36. Rezultati procjena OLS i WLS metodom

WLS<-lm(log(proizvodnja)~log(rad)+log(kapital),data=cd,weights=1/log(rad))  
summary(OLS)$coefficients

## Estimate Std. Error t value Pr(>|t|)  
## (Intercept) -0.04758001 0.014192190 -3.352549 1.856875e-03  
## log(rad) 0.37161629 0.003244253 114.546040 8.362369e-49  
## log(kapital) 0.62412677 0.003435547 181.667369 3.346718e-56

summary(WLS)$coefficients

## Estimate Std. Error t value Pr(>|t|)  
## (Intercept) -0.04762223 0.014208005 -3.351789 1.860777e-03  
## log(rad) 0.37130233 0.003254334 114.094726 9.673678e-49  
## log(kapital) 0.62449061 0.003449353 181.045712 3.798745e-56

## 4.7. Pitanja za ponavljanje, Rješenja

#### 4.7. Pitanja za ponavljanje, Rješenja, Zadatak 37

place<-read.table("placa.txt",header=T,sep="\t")  
  
#multikolinearnost  
model<-lm(placa~staz+obrazovanje,data=place)  
library(car)  
vif(model)

## staz obrazovanje   
## 36.18098 36.18098

1/vif(model)

## staz obrazovanje   
## 0.02763883 0.02763883

m1<-lm(staz~obrazovanje,data=place)  
summary(m1)$r.squared

## [1] 0.9723612

cor(place)

## pojedinac placa staz obrazovanje  
## pojedinac 1.00000000 -0.07015636 -0.03102125 -0.01702086  
## placa -0.07015636 1.00000000 0.94841667 0.93549745  
## staz -0.03102125 0.94841667 1.00000000 0.98608375  
## obrazovanje -0.01702086 0.93549745 0.98608375 1.00000000

sqrt(summary(model)$r.squared)

## [1] 0.9484182

#durbin watson  
library(car)  
durbinWatsonTest(model)

## lag Autocorrelation D-W Statistic p-value  
## 1 -0.01298824 2.023118 0.84  
## Alternative hypothesis: rho != 0

durbinWatsonTest(model,alternative = "positive")

## lag Autocorrelation D-W Statistic p-value  
## 1 -0.01298824 2.023118 0.575  
## Alternative hypothesis: rho > 0

durbinWatsonTest(model,alternative = "negative")

## lag Autocorrelation D-W Statistic p-value  
## 1 -0.01298824 2.023118 0.428  
## Alternative hypothesis: rho < 0

#breusch-godfrey  
library(quantmod)  
reziduali<-residuals(model)  
rez1<-Lag(reziduali,1);rez2<-Lag(reziduali,2);rez3<-Lag(reziduali,3)  
summary(pomocna<-lm(reziduali~staz+obrazovanje+rez1+rez2+rez3,data=place))

##   
## Call:  
## lm(formula = reziduali ~ staz + obrazovanje + rez1 + rez2 + rez3,   
## data = place)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -11.8499 -3.8661 0.0219 4.1305 10.6759   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)  
## (Intercept) 0.008743 1.014943 0.009 0.993  
## staz 0.039127 0.233730 0.167 0.867  
## obrazovanje -0.045741 0.239974 -0.191 0.849  
## rez1 0.014716 0.083610 0.176 0.861  
## rez2 -0.024190 0.084354 -0.287 0.775  
## rez3 -0.016795 0.082939 -0.203 0.840  
##   
## Residual standard error: 5.687 on 141 degrees of freedom  
## (3 observations deleted due to missingness)  
## Multiple R-squared: 0.001624, Adjusted R-squared: -0.03378   
## F-statistic: 0.04588 on 5 and 141 DF, p-value: 0.9987

test\_vel<-nobs(pomocna)\*summary(pomocna)$r.squared  
test\_vel

## [1] 0.2387691

p\_v<-1-pchisq(test\_vel,3)  
p\_v

## [1] 0.9711005

#ako ne radi Lag:  
length(reziduali)

## [1] 150

rez1<-c(NA,reziduali[1:149])  
rez2<-c(NA,rez1[1:149])  
rez3<-c(NA,rez2[1:149])  
  
#ljung-box  
Box.test(reziduali,lag=3,type="Ljung-Box")

##   
## Box-Ljung test  
##   
## data: reziduali  
## X-squared = 0.11924, df = 3, p-value = 0.9894

#heteroskedastičnost:  
  
#BG test:  
summary(bptest<-lm(residuals(model)^2~staz+obrazovanje,data=place))

##   
## Call:  
## lm(formula = residuals(model)^2 ~ staz + obrazovanje, data = place)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -36.54 -24.65 -13.46 19.14 110.78   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 37.504 5.760 6.511 1.11e-09 \*\*\*  
## staz -2.440 1.348 -1.809 0.0724 .   
## obrazovanje 2.440 1.385 1.761 0.0803 .   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 33.02 on 147 degrees of freedom  
## Multiple R-squared: 0.02191, Adjusted R-squared: 0.008607   
## F-statistic: 1.647 on 2 and 147 DF, p-value: 0.1962

nobs(bptest);(summary(bptest))$r.squared

## [1] 150

## [1] 0.02191474

nobs(bptest)\*(summary(bptest))$r.squared

## [1] 3.287211

p\_vrijednost<-1-pchisq(nobs(bptest)\*(summary(bptest))$r.squared,2)  
p\_vrijednost

## [1] 0.1932819

#White test:  
summary(white<-lm(residuals(model)^2~staz+obrazovanje  
 +I(staz\*obrazovanje)  
 +I(staz^2)+I(obrazovanje^2),data=place))

##   
## Call:  
## lm(formula = residuals(model)^2 ~ staz + obrazovanje + I(staz \*   
## obrazovanje) + I(staz^2) + I(obrazovanje^2), data = place)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -44.88 -24.77 -12.94 20.14 100.81   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 26.7511 8.4884 3.151 0.00198 \*\*  
## staz -6.1412 2.7691 -2.218 0.02814 \*   
## obrazovanje 7.1512 3.0573 2.339 0.02071 \*   
## I(staz \* obrazovanje) -3.3170 1.2807 -2.590 0.01059 \*   
## I(staz^2) 1.5660 0.6136 2.552 0.01175 \*   
## I(obrazovanje^2) 1.7388 0.6697 2.596 0.01040 \*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 32.59 on 144 degrees of freedom  
## Multiple R-squared: 0.06631, Adjusted R-squared: 0.03389   
## F-statistic: 2.045 on 5 and 144 DF, p-value: 0.07575

nobs(white);(summary(white))$r.squared

## [1] 150

## [1] 0.06630995

nobs(white)\*(summary(white))$r.squared

## [1] 9.946492

p\_vrijednost<-1-pchisq(nobs(white)\*(summary(white))$r.squared,5)  
p\_vrijednost

## [1] 0.07676564

#korekcije  
  
#White korekcija:  
library(car)  
mat<-hccm(model,type="hc0")  
mat #ovo je matrica var-kovar procjenitelja uz White korekciju

## (Intercept) staz obrazovanje  
## (Intercept) 1.1171687 -0.15555411 0.12805398  
## staz -0.1555541 0.06693509 -0.06770103  
## obrazovanje 0.1280540 -0.06770103 0.07004318

#t-test :  
library(lmtest)  
coeftest(model,vcov=mat)

##   
## t test of coefficients:  
##   
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 4.0087e+03 1.0570e+00 3792.6843 < 2.2e-16 \*\*\*  
## staz 1.3904e+00 2.5872e-01 5.3741 2.954e-07 \*\*\*  
## obrazovanje 1.5375e-02 2.6466e-01 0.0581 0.9538   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

#Newey-West korekcija:  
library(sandwich)  
mat2<-NeweyWest(model,lag=3)  
mat2 #ovo je matrica var-kovar procjenitelja uz Newey-West korekciju

## (Intercept) staz obrazovanje  
## (Intercept) 0.7735279 -0.13957166 0.12058891  
## staz -0.1395717 0.05786712 -0.05704447  
## obrazovanje 0.1205889 -0.05704447 0.05737971

#t-test :  
library(lmtest)  
coeftest(model,vcov=mat2)

##   
## t test of coefficients:  
##   
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 4.0087e+03 8.7950e-01 4557.9345 < 2.2e-16 \*\*\*  
## staz 1.3904e+00 2.4056e-01 5.7798 4.315e-08 \*\*\*  
## obrazovanje 1.5375e-02 2.3954e-01 0.0642 0.9489   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

#Normalnost:  
  
#Jarque-Bera test:  
reziduali<-resid(model)  
library("moments")  
jarque.test(reziduali)

##   
## Jarque-Bera Normality Test  
##   
## data: reziduali  
## JB = 5.5857, p-value = 0.06125  
## alternative hypothesis: greater

#N, koef. asimetrije i zaobljenosti:  
n<-length(reziduali); s<-skewness(reziduali); k<-kurtosis(reziduali)  
n;s;k

## [1] 150

## [1] -0.09295569

## [1] 2.073095

JB<-n\*(s^2/6+(k-3)^2/24)  
JB

## [1] 5.585722

#  
library(nortest)  
ad.test(reziduali)

##   
## Anderson-Darling normality test  
##   
## data: reziduali  
## A = 0.82084, p-value = 0.03325

cvm.test(reziduali)

##   
## Cramer-von Mises normality test  
##   
## data: reziduali  
## W = 0.12102, p-value = 0.0574

lillie.test(reziduali)

##   
## Lilliefors (Kolmogorov-Smirnov) normality test  
##   
## data: reziduali  
## D = 0.07755, p-value = 0.02776

OLS<-lm(placa~staz+obrazovanje,data=place)  
library(nlme)  
GLS<-gls(placa~staz+obrazovanje,data=place,weights=varExp())  
summary(OLS)$coefficients

## Estimate Std. Error t value Pr(>|t|)  
## (Intercept) 4.008723e+03 0.9954432 4.027074e+03 0.000000e+00  
## staz 1.390374e+00 0.2330176 5.966821e+00 1.730807e-08  
## obrazovanje 1.537528e-02 0.2394033 6.422334e-02 9.488796e-01

summary(GLS)$tTable

## Value Std.Error t-value p-value  
## (Intercept) 4008.7624912 1.0155900 3947.2251424 0.000000e+00  
## staz 1.3741493 0.2346451 5.8562886 2.976270e-08  
## obrazovanje 0.0315033 0.2407076 0.1308779 8.960508e-01

WLS<-lm(placa~staz+obrazovanje,data=place,weights=1/obrazovanje)  
summary(OLS)$coefficients

## Estimate Std. Error t value Pr(>|t|)  
## (Intercept) 4.008723e+03 0.9954432 4.027074e+03 0.000000e+00  
## staz 1.390374e+00 0.2330176 5.966821e+00 1.730807e-08  
## obrazovanje 1.537528e-02 0.2394033 6.422334e-02 9.488796e-01

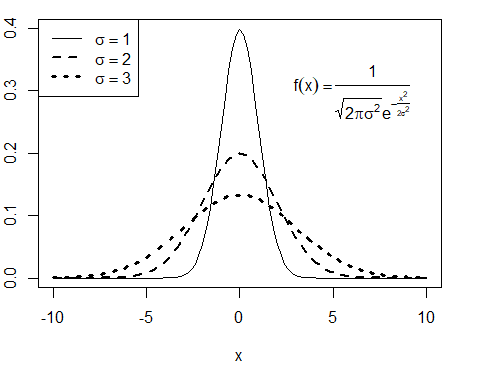
summary(WLS)$coefficients

## Estimate Std. Error t value Pr(>|t|)  
## (Intercept) 4009.3466372 0.6337507 6326.378163 0.000000e+00  
## staz 1.6478526 0.1929244 8.541441 1.540287e-14  
## obrazovanje -0.3075025 0.2050179 -1.499881 1.357899e-01

# 5. DODACI

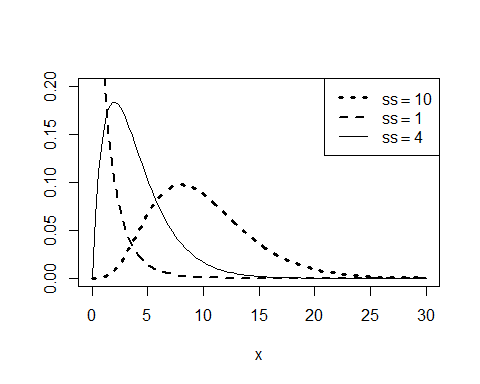
#### Slika A1. Tri normalne distribucije

par(mfrow = c(1,1),oma=c(1,0,0,1),mar=c(4,2,1,1))  
  
curve(dnorm(x,0,1),-10,10,lwd=1,lty=1,ylab=NA)  
curve(dnorm(x,0,2),add=T,lwd=2,lty=2)  
curve(dnorm(x,0,3),add=T,lwd=3,lty=3)  
legend("topleft",expression(sigma==1,sigma==2, sigma==3),  
 lwd=1:3,lty=1:3)  
text(6,.3,expression(f(x)==frac(1,sqrt(2\*pi\*sigma^2)\*  
 e^{ -frac(x^2,2\*sigma^2)})))



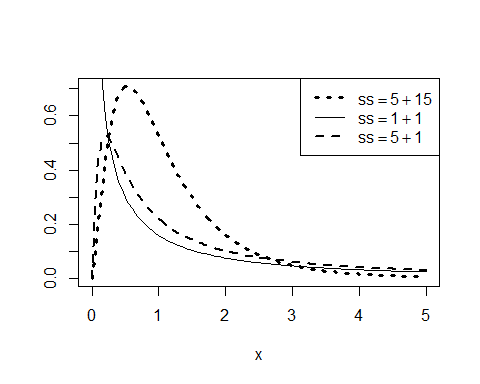
#### Slika A2. Tri hi-kvadrat distribucije

curve(dchisq(x,10),0,30,lwd=3,lty=3,ylab=NA,ylim=c(0,.2))  
curve(dchisq(x,1),add=T,lwd=2,lty=2)  
curve(dchisq(x,4),add=T,lwd=1,lty=1)  
legend("topright",expression(ss==10,ss==1, ss==4),  
 lwd=c(3,2,1),lty=c(3,2,1))



#### Slika A3. Tri F-kvadrat distribucije

curve(df(x,5,15),0,5,lwd=3,lty=3,ylab=NA)  
curve(df(x,1,1),add=T,lwd=1,lty=1)  
curve(df(x,5,1),add=T,lwd=2,lty=2)  
legend("topright",expression(ss==5+15,ss==1+1, ss==5+1),  
 lwd=c(3,1,2),lty=c(3,1,2))



#### Slika A4. Tri Studentove distribucije

curve(dt(x,1),-5,5,lwd=3,lty=3,ylab=NA,ylim=c(0,.4))  
curve(dt(x,5),add=T,lwd=1,lty=1)  
curve(dt(x,30),add=T,lwd=2,lty=2)  
legend("topleft",expression(ss==1,ss==5, ss==30),  
 lwd=c(3,1,2),lty=c(3,1,2))

