

Зачетная работа. Вариант 3

Кремер Алексей. 431 группа.

30 05 2020

Номер 1.

```
library(Ecdat)

## Loading required package: Ecfun

##
## Attaching package: 'Ecfun'

## The following object is masked from 'package:base':
##
##   sign

##
## Attaching package: 'Ecdat'

## The following object is masked from 'package:datasets':
##
##   Orange

library(MASS)

##
## Attaching package: 'MASS'

## The following object is masked from 'package:Ecdat':
##
##   SP500

library(lmtest)

## Loading required package: zoo

##
## Attaching package: 'zoo'

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

library(orcutt)
library(sandwich)
library(AER)

## Loading required package: car

## Loading required package: carData
```

```
##
## Attaching package: 'carData'

## The following object is masked from 'package:Ecdat':
##
##      Mroz

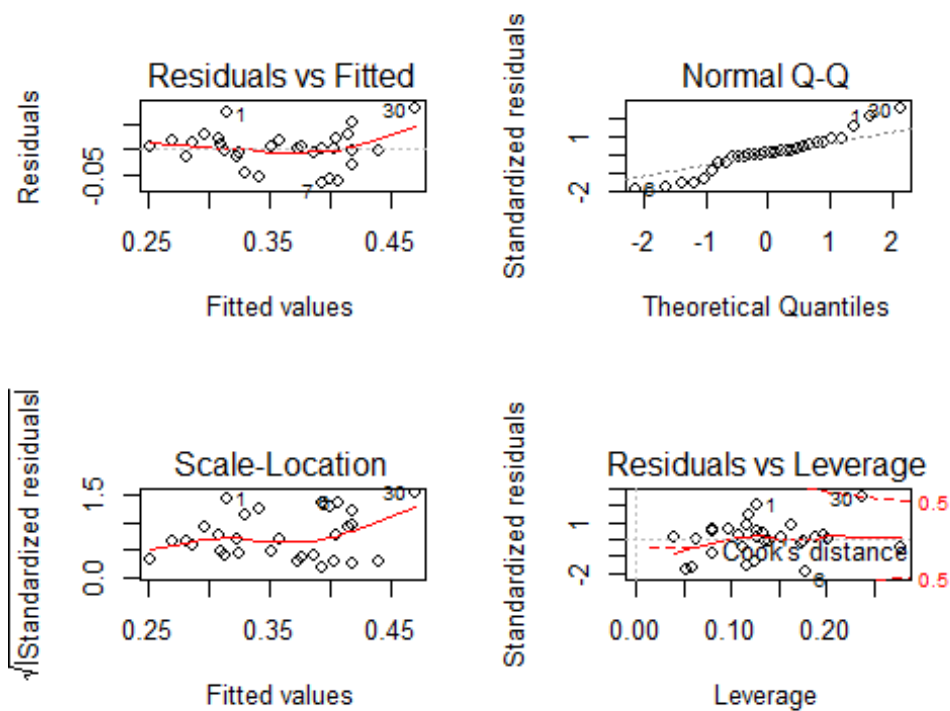
## Loading required package: survival
```

Пункт 1)

```
data(Icecream)
model <- lm(cons ~ ., data = Icecream)
summary(model)

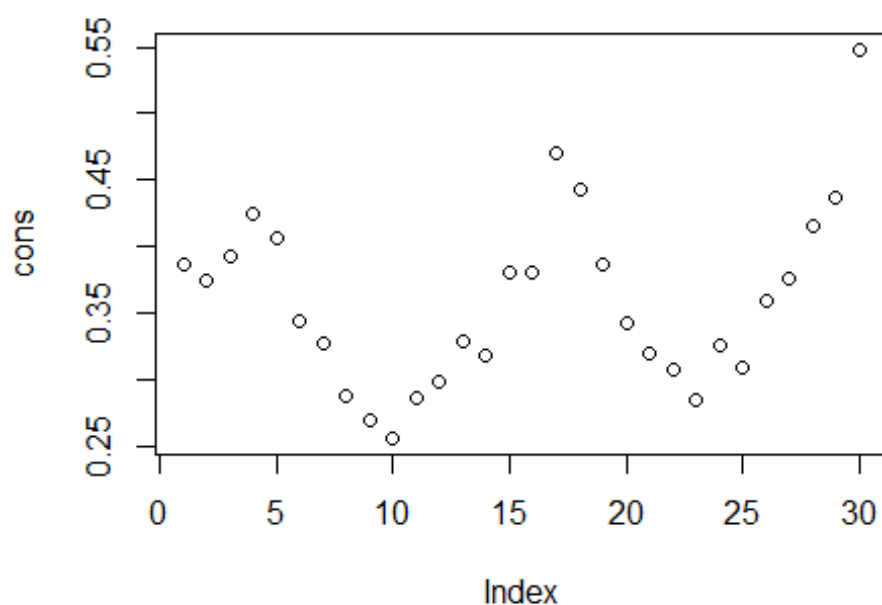
##
## Call:
## lm(formula = cons ~ ., data = Icecream)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.065302 -0.011873  0.002737  0.015953  0.078986
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.1973151  0.2702162   0.730  0.47179
## income       0.0033078  0.0011714   2.824  0.00899 **
## price      -1.0444140  0.8343573  -1.252  0.22180
## temp        0.0034584  0.0004455   7.762  3.1e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.03683 on 26 degrees of freedom
## Multiple R-squared:  0.719, Adjusted R-squared:  0.6866
## F-statistic: 22.17 on 3 and 26 DF, p-value: 2.451e-07

par(mfrow = c(2,2))
plot(model)
```

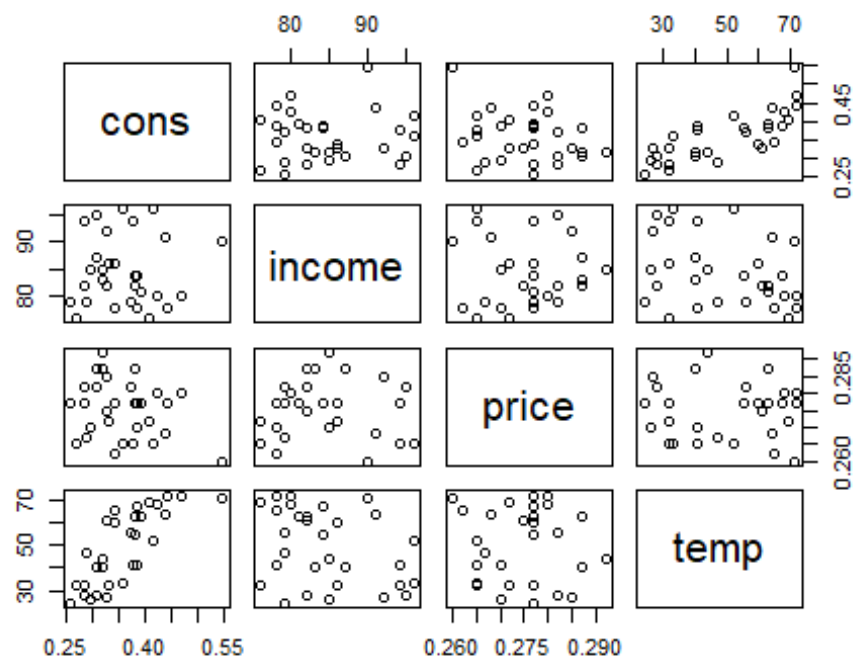


Пункт 2)

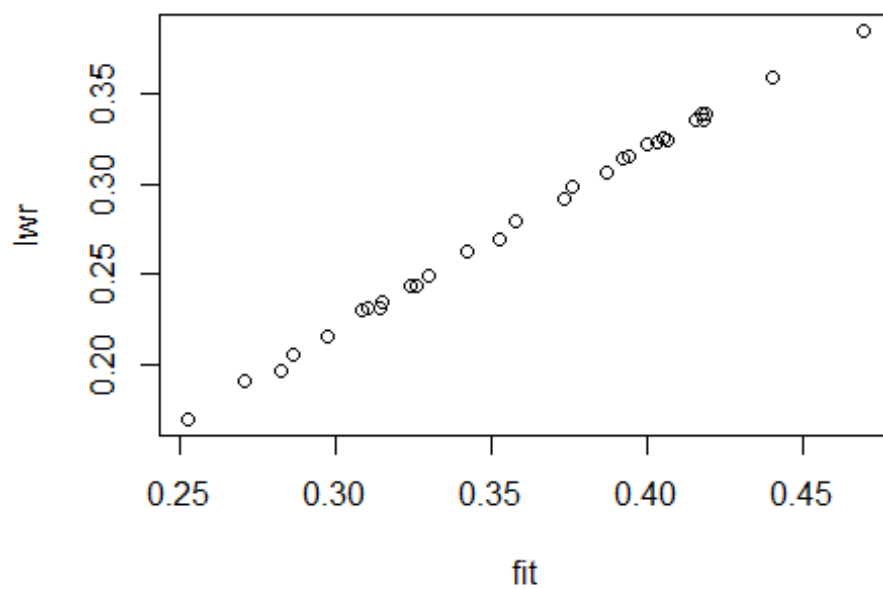
```
dataIce <- data.frame(Icecream)
confs <- predict(model, dataIce, interval = "confidence")
preds <- predict(model, dataIce, interval = "prediction")
plot(confs ~ 1, data = Icecream)
```



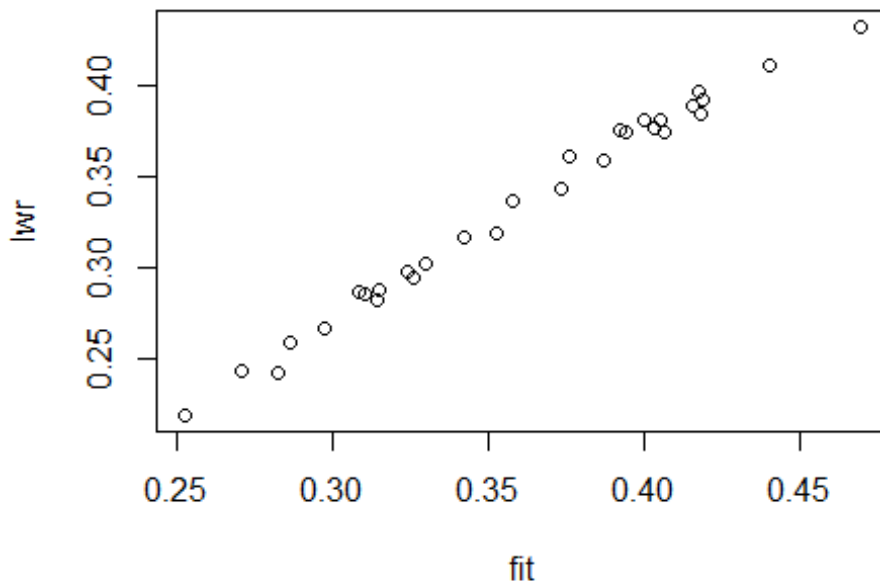
```
plot(dataIce)
```



```
plot(preds)
```



```
plot(confs)
```



Пункт 3)

```
bgtest(model, order = 2)
```

```
##
## Breusch-Godfrey test for serial correlation of order up to 2
##
## data: model
## LM test = 4.4872, df = 2, p-value = 0.1061
```

Пункт 4)

```
model_orcutt = cochrane.orcutt(model, max.iter = 1000)
model_orcutt
```

```
## Cochrane-orcutt estimation for first order autocorrelation
##
## Call:
## lm(formula = cons ~ ., data = Icecream)
##
## number of interaction: 11
## rho 0.400926
##
## Durbin-Watson statistic
## (original): 1.02117 , p-value: 3.024e-04
## (transformed): 1.54884 , p-value: 5.061e-02
##
## coefficients:
## (Intercept) income price temp
## 0.157148 0.003203 -0.892396 0.003558
```

```
coefficients(model)
```

```
## (Intercept)      income      price      temp  
## 0.19731507 0.00330776 -1.04441399 0.00345843
```

Пункт 5)

```
NeweyWest(model)
```

```
##          (Intercept)          income          price          temp  
## (Intercept) 1.030769e-01 -2.496457e-04 -0.3015182861 7.269565e-05  
## income      -2.496457e-04 8.769416e-07 0.0006296449 -6.511207e-08  
## price       -3.015183e-01 6.296449e-04 0.9230633344 -2.800484e-04  
## temp        7.269565e-05 -6.511207e-08 -0.0002800484 2.931894e-07
```

```
coefficients(model)
```

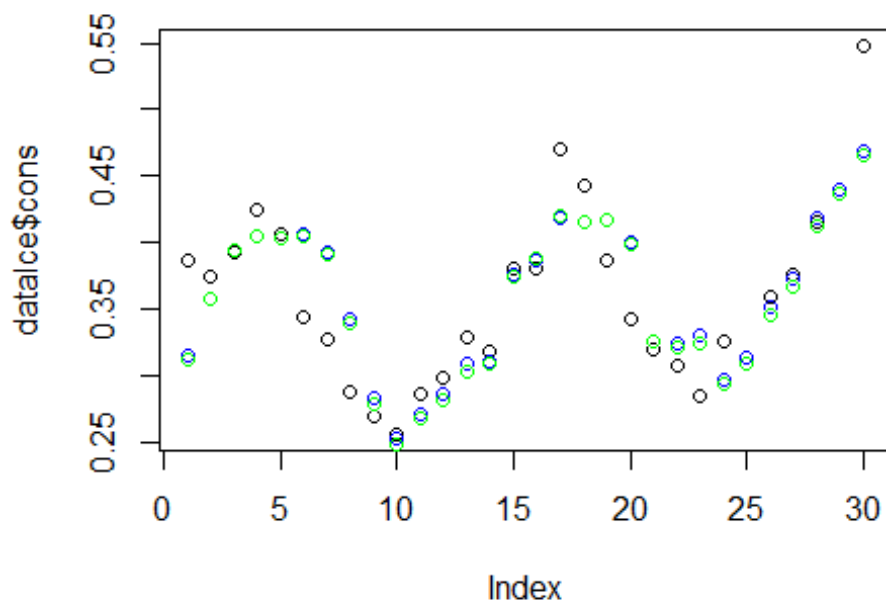
```
## (Intercept)      income      price      temp  
## 0.19731507 0.00330776 -1.04441399 0.00345843
```

```
pt(-1.0444140 / sqrt(0.9230633344), 30-3-1)*2
```

```
## [1] 0.286982
```

Пункт 6)

```
preds <- predict(model, dataIce)  
preds0rc <- predict(model_orcutt, dataIce)  
plot(dataIce$cons)  
points(preds, col = "blue")  
points(preds0rc, col = "green")
```



Номер 2

Пункт 1)

```
data(HousePrices)
dataHouse <- data.frame(HousePrices)
dataHouse$price = log(dataHouse$price)
dataHouse$lotsize = log(dataHouse$lotsize)
min.model = lm(price ~ 1, data=HousePrices)
biggest <- formula(lm(price~.,HousePrices))
summary(HousePrices)
```

	price	lotsize	bedrooms	bathrooms		
## Min.	: 25000	Min. : 1650	Min. :1.000	Min. :1.000		
## 1st Qu.:	49125	1st Qu.: 3600	1st Qu.:2.000	1st Qu.:1.000		
## Median :	62000	Median : 4600	Median :3.000	Median :1.000		
## Mean :	68122	Mean : 5150	Mean :2.965	Mean :1.286		
## 3rd Qu.:	82000	3rd Qu.: 6360	3rd Qu.:3.000	3rd Qu.:2.000		
## Max.	:190000	Max. :16200	Max. :6.000	Max. :4.000		
	stories	driveway	recreation	fullbase	gasheat	aircon
## Min.	:1.000	no : 77	no :449	no :355	no :521	no :373
## 1st Qu.:	1.000	yes:469	yes: 97	yes:191	yes: 25	yes:173
## Median :	2.000					
## Mean :	1.808					
## 3rd Qu.:	2.000					
## Max.	:4.000					
	garage	prefer				
## Min.	:0.0000	no :418				
## 1st Qu.:	0.0000	yes:128				
## Median :	0.0000					
## Mean :	0.6923					
## 3rd Qu.:	1.0000					
## Max.	:3.0000					

Акаике:

```
lm_a = stepAIC(min.model, direction = "forward", k=2, scope=biggest, trace = FALSE)
lm_a
```

```
##
## Call:
## lm(formula = price ~ lotsize + bathrooms + aircon + stories +
##     prefer + garage + fullbase + gasheat + driveway + recreation +
##     bedrooms, data = HousePrices)
##
## Coefficients:
## (Intercept)      lotsize      bathrooms      airconyes      stories
## -4038.350         3.546       14335.558       12632.890       6556.946
## preferyes      garage      fullbaseyes      gasheatyes      drivewayyes
##  9369.513       4244.829        5452.386       12831.406       6687.779
## recreationyes      bedrooms
##  4511.284        1832.003
```

Шварц:

```
lm_b = stepAIC(min.model, direction = "forward", k=log(nrow(HousePrices)), scope
=biggest, trace = FALSE)
lm_b

##
## Call:
## lm(formula = price ~ lotsize + bathrooms + aircon + stories +
##     prefer + garage + fullbase + gasheat + driveway, data = HousePrices)
##
## Coefficients:
## (Intercept)      lotsize    bathrooms    airconyes      stories    preferyes
##   -1123.144         3.666     15072.868     12875.657      7241.264      9595.888
##      garage fullbaseyes    gasheatyes    drivewayyes
##    4265.862     7134.099     12954.080      6428.566
```

Пункт 2)

Акаике:

```
summary(lm_a)$r.squared

## [1] 0.6731236
```

Шварц:

```
summary(lm_b)$r.squared

## [1] 0.6679045
```

Пункт 3)

Акаике:

```
resettest(lm_a)

##
## RESET test
##
## data:  lm_a
## RESET = 13.481, df1 = 2, df2 = 532, p-value = 1.944e-06
```

Шварц:

```
resettest(lm_b)

##
## RESET test
##
## data:  lm_b
## RESET = 11.2, df1 = 2, df2 = 534, p-value = 1.719e-05
```

В обоих случаях гипотеза о том, что в модели нет отсутствующих переменных или модель линейная, отвергается.