Лабораторная работа №3

Полянцев И.С.

27 12 2020

# Раздел I.

Создать скрипт .Rmd, который генерирует отчёт в формате Word по регрессионному анализу данных варианта. Для 2-4 лучших моделей, полученных в лабораторной работе №2, провести тесты остатков.

В пятой модели обнаруживаются два влияющих наблюдения: регионы 83 и 87

## Label FO  
## 83 в том числе: Ненецкий автономный округ СЗФО  
## 87 Чукотский автономный округ ДФО

Проверим, нужно ли исключать эти наблюдения:

## .cooksd f.tabl p.val  
## 1 0.61 3.11 0.4546  
## 2 0.65 3.11 0.4756

Поскольку расстояния Кука не превосходят критическую границу (и, следовательно, соответствующие p-значения больше 0,05), принимаем нулевые гипотезы для регионов 83 и 87: их не стоит убирать из выборки для построения пятой модели.

Проверим гипотезу о равенстве среднего остатков нулю для модели 1

##   
## One Sample t-test  
##   
## data: models.list[[i]]$residuals  
## t = -1.9816e-16, df = 84, p-value = 1  
## alternative hypothesis: true mean is not equal to 0  
## 95 percent confidence interval:  
## -1175.709 1175.709  
## sample estimates:  
## mean of x   
## -1.17156e-13

##   
## Welch Two Sample t-test  
##   
## data: res.s1 and res.s2  
## t = 0.0051121, df = 81.155, p-value = 0.9959  
## alternative hypothesis: true difference in means is not equal to 0  
## 95 percent confidence interval:  
## -2146.127 2157.184  
## sample estimates:  
## mean of x mean of y   
## -264.2773 -269.8057

P-значение для модели 1 > 0,05, следовательно нулевая гипотеза не отклоняется

Проведем тесты на гетероскедастичность для 2 модели:

##   
## studentized Breusch-Pagan test  
##   
## data: models.list[[2]]  
## BP = 51.641, df = 9, p-value = 5.284e-08

Согласно первому тесту, P-значение < 0.05, следовательно, нулевая гипотеза отклоняется и в остатках модели 2 есть гетероскедастичность. Далее проведем тест Уайта

##   
## studentized Breusch-Pagan test  
##   
## data: models.list[[2]]  
## BP = 1.302, df = 2, p-value = 0.5215

Этот тест внезапно показывает, что гетероскедастичности нет. Далее тест Голдфельда-Квандта

##   
## Goldfeld-Quandt test  
##   
## data: models.list[[2]]  
## GQ = 7.0991, df1 = 24, df2 = 24, p-value = 4.509e-06  
## alternative hypothesis: variance increases from segment 1 to 2

Результат тест Г-К вновь показывает, что присутствует гетероскедастичность. Далее - тест Глейзера

## ! >>> Модель значима >>> beta = -1p-value = 0.001

## ! >>> Модель значима >>> beta = -0.95p-value = 8e-04

## ! >>> Модель значима >>> beta = -0.9p-value = 6e-04

## ! >>> Модель значима >>> beta = -0.85p-value = 5e-04

## ! >>> Модель значима >>> beta = -0.8p-value = 4e-04

## ! >>> Модель значима >>> beta = -0.75p-value = 3e-04

## ! >>> Модель значима >>> beta = -0.7p-value = 2e-04

## ! >>> Модель значима >>> beta = -0.65p-value = 2e-04

## ! >>> Модель значима >>> beta = -0.6p-value = 1e-04

## ! >>> Модель значима >>> beta = -0.55p-value = 1e-04

## ! >>> Модель значима >>> beta = -0.5p-value = 1e-04

## ! >>> Модель значима >>> beta = -0.45p-value = 1e-04

## ! >>> Модель значима >>> beta = -0.4p-value = 0

## ! >>> Модель значима >>> beta = -0.35p-value = 0

## ! >>> Модель значима >>> beta = -0.3p-value = 0

## ! >>> Модель значима >>> beta = -0.25p-value = 0

## ! >>> Модель значима >>> beta = -0.2p-value = 0

## ! >>> Модель значима >>> beta = -0.15p-value = 0

## ! >>> Модель значима >>> beta = -0.1p-value = 0

## ! >>> Модель значима >>> beta = -0.05p-value = 0

## ! >>> Модель значима >>> beta = 0.05p-value = 0

## ! >>> Модель значима >>> beta = 0.1p-value = 0

## ! >>> Модель значима >>> beta = 0.15p-value = 0

## ! >>> Модель значима >>> beta = 0.2p-value = 0

## ! >>> Модель значима >>> beta = 0.25p-value = 0

## ! >>> Модель значима >>> beta = 0.3p-value = 0

## ! >>> Модель значима >>> beta = 0.35p-value = 0

## ! >>> Модель значима >>> beta = 0.4p-value = 0

## ! >>> Модель значима >>> beta = 0.45p-value = 0

## ! >>> Модель значима >>> beta = 0.5p-value = 0

## ! >>> Модель значима >>> beta = 0.55p-value = 1e-04

## ! >>> Модель значима >>> beta = 0.6p-value = 1e-04

## ! >>> Модель значима >>> beta = 0.65p-value = 1e-04

## ! >>> Модель значима >>> beta = 0.7p-value = 1e-04

## ! >>> Модель значима >>> beta = 0.75p-value = 2e-04

## ! >>> Модель значима >>> beta = 0.8p-value = 2e-04

## ! >>> Модель значима >>> beta = 0.85p-value = 3e-04

## ! >>> Модель значима >>> beta = 0.9p-value = 4e-04

## ! >>> Модель значима >>> beta = 0.95p-value = 6e-04

## ! >>> Модель значима >>> beta = 1p-value = 7e-04

## ! >>> Модель значима >>> beta = 1.05p-value = 0.001

## ! >>> Модель значима >>> beta = 1.1p-value = 0.0013

## ! >>> Модель значима >>> beta = 1.15p-value = 0.0017

## ! >>> Модель значима >>> beta = 1.2p-value = 0.0022

## ! >>> Модель значима >>> beta = 1.25p-value = 0.0029

## ! >>> Модель значима >>> beta = 1.3p-value = 0.0037

## ! >>> Модель значима >>> beta = 1.35p-value = 0.0047

## ! >>> Модель значима >>> beta = 1.4p-value = 0.0059

## ! >>> Модель значима >>> beta = 1.45p-value = 0.0073

## ! >>> Модель значима >>> beta = 1.5p-value = 0.009

Поскольку ни одна из вспомогательных регрессий для теста Глейзера не оказалась значимой, нулевая гипотеза (остатки гомоскедастичны) не отвергается

Проверим модель 3 на автокорреляцию

##   
## Durbin-Watson test  
##   
## data: models.list[[i]]  
## DW = 1.6307, p-value = 0.08209  
## alternative hypothesis: true autocorrelation is not 0

##   
## Pearson's product-moment correlation  
##   
## data: models.list[[i]]$residuals[1:(n - 1)] and models.list[[i]]$residuals[2:n]  
## t = 0.58662, df = 82, p-value = 0.5591  
## alternative hypothesis: true correlation is not equal to 0  
## 95 percent confidence interval:  
## -0.1518545 0.2752260  
## sample estimates:  
## cor   
## 0.06464532

P-значение > 0.05 => нулевая гипотеза не отклоняется и в остатках 3 модели отсутствует автокорреляция первого порядка

Переоценка параметров модели 1 с учетом ошибок

##   
## t test of coefficients:  
##   
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 1.7641e+04 7.4434e+02 23.700 < 2.2e-16 \*\*\*  
## GRP 1.5264e-02 1.1137e-03 13.706 < 2.2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

##   
## t test of coefficients:  
##   
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 1.7641e+04 7.8660e+02 22.4266 < 2.2e-16 \*\*\*  
## GRP 1.5264e-02 2.0434e-03 7.4701 7.312e-11 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Тест на мультиколлинеарность факторов модели 6

## Social.Policy Social.Policy.FOЦФО   
## 1 1

Поскольку значения VIF-коэффициентов равны 1, в модели 6 нет мультиколлинеарности факторов.