Лабораторная работа №2

Людуховский В.В.

19 Декабрь 2020

# Лабораторная работа №2

Во второй лабораторной работе продолжим работу с примером данных, использовавшихся в предыдущей работе. Оценим параметры моделей, объясняющих поведение данного показателя несколькими факторами.

## Зависимая переменная модели:

* Y ORT.2014 - Оборот розничной торговли на душу населения (руб.)

## Независимые переменные модели:

* X1 KMP.2014 - Число малых предприятий на 10000 человек населения (шт.)
* X2 USEPC.2014 - Использование информационных и коммуникационных технологий в организациях: персональные компьютеры (% числа организаций, использующих эти технологии)
* X3 Budg.2013 - Расходы консолидированных бюджетов субъектов Российской Федерации: всего (млн. руб)
* X4 RDD.2014 - Реальные денежные доходы населения (% к пред. году)

## Исходная модель:

## Выполним оценку параметров модели линейной регрессии

#### Таблица № 1

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | t value | Pr(>|t|) |
| (Intercept) | 69137.1290 | 114925.7506 | 0.6016 | 0.5492 |
| KMP.2014 | 190.7609 | 66.6578 | 2.8618 | 0.0054 |
| USEPC.2014 | 1459.8081 | 825.2420 | 1.7689 | 0.0807 |
| Budg.2013 | 0.1481 | 0.0214 | 6.9173 | 0.0000 |
| RDD.2014 | -931.4891 | 949.2055 | -0.9813 | 0.3294 |

## Исколючаем из модели наименее значимый фактор RDD.2014

#### Таблица № 2

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | t value | Pr(>|t|) |
| (Intercept) | -13926.2453 | 77721.9329 | -0.1792 | 0.8582 |
| KMP.2014 | 191.6996 | 66.6358 | 2.8768 | 0.0051 |
| USEPC.2014 | 1332.9741 | 814.8721 | 1.6358 | 0.1058 |
| Budg.2013 | 0.1537 | 0.0206 | 7.4543 | 0.0000 |

## Исколючаем из модели второй наименее значимый фактор USEPC.2014

#### Таблица № 3

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | t value | Pr(>|t|) |
| (Intercept) | 112392.3844 | 8900.0230 | 12.6283 | 0.0000 |
| KMP.2014 | 181.9867 | 67.0455 | 2.7144 | 0.0081 |
| Budg.2013 | 0.1604 | 0.0204 | 7.8510 | 0.0000 |

## Включим в модель фиктивные переменные

#### Таблица № 4

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | t value | Pr(>|t|) |
| (Intercept) | 55599.1914 | 29190.0556 | 1.9047 | 0.0615 |
| KMP.2014 | 606.5100 | 209.2887 | 2.8980 | 0.0052 |
| FOПФО | 6215.0367 | 47835.7150 | 0.1299 | 0.8971 |
| FOСЗФО | 116476.9610 | 33717.6112 | 3.4545 | 0.0010 |
| FOСКФО | 9782.8237 | 47155.1762 | 0.2075 | 0.8363 |
| FOСФО | 19968.3710 | 33850.7881 | 0.5899 | 0.5574 |
| FOУФО | 48830.4488 | 42122.9758 | 1.1592 | 0.2509 |
| FOЦФО | 97327.9725 | 37841.3771 | 2.5720 | 0.0126 |
| FOЮФО | -21719.9925 | 47406.5982 | -0.4582 | 0.6485 |
| Budg.2013 | 0.2579 | 0.1837 | 1.4043 | 0.1653 |
| KMP.2014:FOПФО | -423.9569 | 345.1750 | -1.2282 | 0.2241 |
| KMP.2014:FOСЗФО | -794.3832 | 264.1497 | -3.0073 | 0.0038 |
| KMP.2014:FOСКФО | -637.8498 | 480.7266 | -1.3268 | 0.1895 |
| KMP.2014:FOСФО | -349.2996 | 255.1345 | -1.3691 | 0.1760 |
| KMP.2014:FOУФО | -109.5396 | 293.0449 | -0.3738 | 0.7099 |
| KMP.2014:FOЦФО | -692.7085 | 281.0034 | -2.4651 | 0.0165 |
| KMP.2014:FOЮФО | 505.6038 | 502.0892 | 1.0070 | 0.3179 |
| FOПФО:Budg.2013 | 0.3755 | 0.2157 | 1.7406 | 0.0868 |
| FOСЗФО:Budg.2013 | -0.0040 | 0.2330 | -0.0172 | 0.9863 |
| FOСКФО:Budg.2013 | 0.7300 | 0.3598 | 2.0290 | 0.0468 |
| FOСФО:Budg.2013 | -0.0575 | 0.2382 | -0.2416 | 0.8099 |
| FOУФО:Budg.2013 | -0.1847 | 0.1960 | -0.9422 | 0.3498 |
| FOЦФО:Budg.2013 | -0.1056 | 0.1846 | -0.5723 | 0.5692 |
| FOЮФО:Budg.2013 | -0.1856 | 0.2322 | -0.7991 | 0.4273 |

## Воспользуемся пользовательской функцией для исключения незначимых параметров

### Сравним методы применения данной функции

### Метод Бонферрони:

#### Таблица № 5

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | t value | Pr(>|t|) |
| (Intercept) | 84805.4967 | 13715.8828 | 6.1830 | 0.0000 |
| KMP.2014 | 535.1577 | 102.4939 | 5.2214 | 0.0000 |
| FOСЗФО | 70636.0907 | 27595.5324 | 2.5597 | 0.0123 |
| KMP.2014.FOСЗФО | -473.9858 | 169.0081 | -2.8045 | 0.0063 |

* Доля дисперсии
* *0.233*

### Исключение параметров без поправки:

#### Таблица № 6

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | t value | Pr(>|t|) |
| (Intercept) | 74529.0725 | 9964.3711 | 7.4796 | 0.0000 |
| KMP.2014 | 562.0497 | 75.2690 | 7.4672 | 0.0000 |
| FOСЗФО | 90814.1286 | 16728.3135 | 5.4288 | 0.0000 |
| FOЦФО | 77871.5018 | 24849.6818 | 3.1337 | 0.0025 |
| Budg.2013 | 0.1511 | 0.0161 | 9.4122 | 0.0000 |
| KMP.2014.FOПФО | -466.4121 | 94.3871 | -4.9415 | 0.0000 |
| KMP.2014.FOСЗФО | -649.1203 | 105.0922 | -6.1767 | 0.0000 |
| KMP.2014.FOСКФО | -688.3739 | 197.1470 | -3.4917 | 0.0008 |
| KMP.2014.FOСФО | -267.3866 | 55.9835 | -4.7762 | 0.0000 |
| KMP.2014.FOЦФО | -642.8408 | 185.2024 | -3.4710 | 0.0009 |
| FOПФО.Budg.2013 | 0.4676 | 0.1018 | 4.5936 | 0.0000 |
| FOСКФО.Budg.2013 | 0.7833 | 0.2201 | 3.5584 | 0.0007 |

* Доля дисперсии
* *0.7593*

# Сравнение нескольких моделей по качеству

## Analysis of Variance Table  
##   
## Model 1: ORT.2014 ~ KMP.2014 + Budg.2013  
## Model 2: ORT.2014 ~ KMP.2014 \* FO + Budg.2013 \* FO  
## Model 3: ORT.2014 ~ KMP.2014 + FOСЗФО + KMP.2014.FOСЗФО  
## Model 4: ORT.2014 ~ KMP.2014 + FOСЗФО + FOЦФО + Budg.2013 + KMP.2014.FOПФО +   
## KMP.2014.FOСЗФО + KMP.2014.FOСКФО + KMP.2014.FOСФО + KMP.2014.FOЦФО +   
## FOПФО.Budg.2013 + FOСКФО.Budg.2013  
## Res.Df RSS Df Sum of Sq F Pr(>F)   
## 1 82 8.2974e+10   
## 2 61 3.4191e+10 21 4.8783e+10 4.1445 7.149e-06 \*\*\*  
## 3 81 1.3244e+11 -20 -9.8249e+10 8.7643 2.100e-11 \*\*\*  
## 4 73 3.7453e+10 8 9.4987e+10 21.1833 5.814e-15 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

## Составим таблицу характеристик качества построенных моделей

#### Таблица № 7

|  |  |  |  |
| --- | --- | --- | --- |
| Модель | R.2.скорр | F.расч | Станд.Ошибка |
| fit.X1 | 0.525 | 47.49 | 31810.0 |
| fit.X1.fo | 0.737 | 11.24 | 23675.0 |
| fit.X1.fo1 | 0.233 | 9.51 | 40435.8 |
| fit.X1.fo2 | 0.759 | 25.10 | 22650.6 |

# Проделаем аналогичную работу над логарифмированными данными:

## Выполним оценку параметров модели линейной регрессии

#### Таблица № 8

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | t value | Pr(>|t|) |
| (Intercept) | 6.7432 | 3.6584 | 1.8432 | 0.0690 |
| KMP.2014 | 0.1805 | 0.0623 | 2.8950 | 0.0049 |
| USEPC.2014 | 0.9401 | 0.5217 | 1.8019 | 0.0753 |
| Budg.2013 | 0.2058 | 0.0318 | 6.4670 | 0.0000 |
| RDD.2014 | -0.4935 | 0.6565 | -0.7517 | 0.4544 |

## Исколючаем из модели наименее значимый фактор RDD.2014

#### Таблица № 9

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | t value | Pr(>|t|) |
| (Intercept) | 4.6404 | 2.3513 | 1.9736 | 0.0518 |
| KMP.2014 | 0.1826 | 0.0621 | 2.9401 | 0.0043 |
| USEPC.2014 | 0.8880 | 0.5157 | 1.7219 | 0.0889 |
| Budg.2013 | 0.2103 | 0.0311 | 6.7544 | 0.0000 |

## Исколючаем из модели второй наименее значимый фактор USEPC.2014

#### Таблица № 10

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | t value | Pr(>|t|) |
| (Intercept) | 8.6467 | 0.3436 | 25.1670 | 0.0000 |
| KMP.2014 | 0.1673 | 0.0622 | 2.6896 | 0.0087 |
| Budg.2013 | 0.2196 | 0.0310 | 7.0766 | 0.0000 |

## Включим в модель фиктивные переменные

#### Таблица № 11

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | t value | Pr(>|t|) |
| (Intercept) | 8.5121 | 1.0457 | 8.1401 | 0.0000 |
| KMP.2014 | 0.4539 | 0.2165 | 2.0966 | 0.0402 |
| FOПФО | -2.1804 | 1.9289 | -1.1303 | 0.2628 |
| FOСЗФО | 2.8069 | 1.3135 | 2.1369 | 0.0366 |
| FOСКФО | -1.4283 | 2.0988 | -0.6805 | 0.4988 |
| FOСФО | 0.2799 | 1.2853 | 0.2178 | 0.8283 |
| FOУФО | 0.1042 | 1.5410 | 0.0676 | 0.9463 |
| FOЦФО | 0.8035 | 1.3559 | 0.5926 | 0.5557 |
| FOЮФО | -0.6003 | 1.3424 | -0.4472 | 0.6563 |
| Budg.2013 | 0.1110 | 0.0913 | 1.2158 | 0.2287 |
| KMP.2014:FOПФО | -0.2232 | 0.3427 | -0.6513 | 0.5173 |
| KMP.2014:FOСЗФО | -0.5223 | 0.2659 | -1.9640 | 0.0541 |
| KMP.2014:FOСКФО | -0.5072 | 0.2786 | -1.8209 | 0.0735 |
| KMP.2014:FOСФО | -0.1287 | 0.2540 | -0.5065 | 0.6143 |
| KMP.2014:FOУФО | -0.1096 | 0.3790 | -0.2891 | 0.7735 |
| KMP.2014:FOЦФО | -0.4744 | 0.2825 | -1.6793 | 0.0982 |
| KMP.2014:FOЮФО | 0.5625 | 0.3806 | 1.4781 | 0.1445 |
| FOПФО:Budg.2013 | 0.2798 | 0.1207 | 2.3174 | 0.0239 |
| FOСЗФО:Budg.2013 | -0.0183 | 0.1251 | -0.1460 | 0.8844 |
| FOСКФО:Budg.2013 | 0.3289 | 0.1564 | 2.1028 | 0.0396 |
| FOСФО:Budg.2013 | 0.0124 | 0.1224 | 0.1015 | 0.9195 |
| FOУФО:Budg.2013 | 0.0412 | 0.1635 | 0.2520 | 0.8019 |
| FOЦФО:Budg.2013 | 0.1364 | 0.1033 | 1.3214 | 0.1913 |
| FOЮФО:Budg.2013 | -0.1699 | 0.1417 | -1.1985 | 0.2354 |

## Воспользуемся пользовательской функцией для исключения незначимых параметров

### Сравним методы применения данной функции

### Метод Бонферрони:

#### Таблица № 12

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | t value | Pr(>|t|) |
| (Intercept) | -77936.04 | 50150.25 | -1.5541 | 0.124 |
| KMP.2014 | 48479.15 | 10447.09 | 4.6404 | 0.000 |

* Доля дисперсии
* *0.1964*

### Исключение параметров без поправки:

#### Таблица № 13

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | t value | Pr(>|t|) |
| (Intercept) | -327568.158 | 44376.442 | -7.3816 | 0.0000 |
| KMP.2014 | 56106.637 | 11341.025 | 4.9472 | 0.0000 |
| FOСЗФО | 350583.172 | 91109.783 | 3.8479 | 0.0003 |
| Budg.2013 | 20008.805 | 4871.996 | 4.1069 | 0.0001 |
| KMP.2014.FOПФО | -70893.959 | 21763.230 | -3.2575 | 0.0017 |
| KMP.2014.FOСЗФО | -72040.324 | 18516.347 | -3.8906 | 0.0002 |
| KMP.2014.FOСКФО | -62420.052 | 17974.095 | -3.4728 | 0.0009 |
| KMP.2014.FOСФО | -7767.043 | 1753.142 | -4.4304 | 0.0000 |
| KMP.2014.FOЦФО | -72159.011 | 17697.160 | -4.0774 | 0.0001 |
| FOПФО.Budg.2013 | 28231.071 | 9276.603 | 3.0433 | 0.0033 |
| FOСКФО.Budg.2013 | 23862.473 | 7189.459 | 3.3191 | 0.0014 |
| FOЦФО.Budg.2013 | 31092.836 | 7606.903 | 4.0874 | 0.0001 |

* Доля дисперсии
* *0.7416*

# Сравним модели на логарифмированных данных и на обычных

## Analysis of Variance Table  
##   
## Model 1: ORT.2014 ~ KMP.2014 + Budg.2013  
## Model 2: ORT.2014 ~ KMP.2014 \* FO + Budg.2013 \* FO  
## Model 3: ORT.2014 ~ KMP.2014 + FOСЗФО + KMP.2014.FOСЗФО  
## Model 4: ORT.2014 ~ KMP.2014 + FOСЗФО + FOЦФО + Budg.2013 + KMP.2014.FOПФО +   
## KMP.2014.FOСЗФО + KMP.2014.FOСКФО + KMP.2014.FOСФО + KMP.2014.FOЦФО +   
## FOПФО.Budg.2013 + FOСКФО.Budg.2013  
## Model 5: ORT.2014 ~ KMP.2014 + Budg.2013  
## Model 6: ORT.2014 ~ KMP.2014 \* FO + Budg.2013 \* FO  
## Model 7: ORT.2014 ~ KMP.2014  
## Model 8: ORT.2014 ~ KMP.2014 + FOСЗФО + Budg.2013 + KMP.2014.FOПФО + KMP.2014.FOСЗФО +   
## KMP.2014.FOСКФО + KMP.2014.FOСФО + KMP.2014.FOЦФО + FOПФО.Budg.2013 +   
## FOСКФО.Budg.2013 + FOЦФО.Budg.2013  
## Res.Df RSS Df Sum of Sq F Pr(>F)   
## 1 82 8.2974e+10   
## 2 61 3.4191e+10 21 4.8783e+10 4.1445 7.149e-06 \*\*\*  
## 3 81 1.3244e+11 -20 -9.8249e+10 8.7643 2.100e-11 \*\*\*  
## 4 73 3.7453e+10 8 9.4987e+10 21.1833 5.814e-15 \*\*\*  
## 5 82 4.0000e+00 -9 3.7453e+10   
## 6 61 2.0000e+00 21 2.0000e+00 0.0000 1   
## 7 83 1.4219e+11 -22 -1.4219e+11 11.5309 2.266e-14 \*\*\*  
## 8 73 4.0208e+10 10 1.0198e+11 18.1945 8.013e-15 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

## Составим таблицу характеристик качества построенных моделей

#### Таблица № 14

|  |  |  |  |
| --- | --- | --- | --- |
| Модель | R.2.скорр | F.расч | Станд.Ошибка |
| fit.X1 | 0.525 | 47.49 | 31810.0 |
| fit.X1.fo | 0.737 | 11.24 | 23675.0 |
| fit.X1.fo1 | 0.233 | 9.51 | 40435.8 |
| fit.X1.fo2 | 0.759 | 25.10 | 22650.6 |
| fit.X1log | 0.517 | 46.02 | 0.2 |
| fit.X1log.fo | 0.680 | 8.78 | 0.2 |
| fit.X1log.fo1 | 0.196 | 21.53 | 41389.9 |
| fit.X1log.fo2 | 0.742 | 22.92 | 23468.9 |

### Результат:

По столбцу больше всего подходит четвертая модель , а именно без поправок по методу Бонферонни;

По столбцу F.расч - первая ;

Выберем лучшую модель по

Явный вид модели