

# Задачи по Эконометрике: Степень подгонки линейной регрессии

Н.В. Артамонов (МГИМО МИД России)

## Содержание

<b>1</b>	<b>Суммы квадратов, остатки, предсказанные значения, <math>R^2, R^2_{adj}</math> (Python)</b>	<b>1</b>
1.1	Sleep equation . . . . .	1
1.2	Sleep equation (smsa only) . . . . .	2
1.3	Sleep equation (men only) . . . . .	2
1.4	Sleep equation (women only) . . . . .	3
1.5	Labour equation (in levels) . . . . .	4
1.6	Labour equation (in log) . . . . .	4
1.7	Cost equation (in levels) . . . . .	5
1.8	Cost equation (in logs) . . . . .	6

## 1 Суммы квадратов, остатки, предсказанные значения, $R^2, R^2_{adj}$ (Python)

### 1.1 Sleep equation

Для датасета `sleep75` рассмотрим регрессию `sleep` на `totwrk`, `age`, `male`, `smsa`

Найдите коэффициенты подогнанной модели. **Ответ округлите до 2-х десятичных знаков.**

Ответ:

(Intercept)	totwrk	age	male	smsa
3494.22	-0.17	2.81	86.91	-75.29

Для наблюдений с индексом 0, 3, 5, 9, 507, 699 найдите фактическое значение зависимой переменной, предсказанное значение и остатки.

```
=====
i   Факт   Fitted   Остаток
-----
0   3113   3094.454   18.546
3   3083   2943.580   139.420
5   4063   3494.143   568.857
9   3018   3111.983   -93.983
507 3630   3476.914   153.086
699 2993   3295.398   -302.398
-----
```

Вычислите TSS, ESS, RSS

```
[1] "TSS=139239835.763456"
```

```
[1] "ESS=17188382.2336174"
```

```
[1] "RSS=122051453.529839"
```

Вычислите  $R^2$  и  $adj.R^2$

```
[1] 0.1234444
```

```
[1] 0.1184427
```

Вычислите стандартную ошибку регрессии

```
[1] 417.2655
```

## 1.2 Sleep equation (smsa only)

Для датасета `sleep75` рассмотрим регрессию `sleep` на `totwrk`, `age`, `male`, `south` только для жителей мегаполиса (`smsa=1`)

Найдите коэффициенты подогнанной модели. Ответ округлите до 2-х десятичных знаков.

Ответ:

(Intercept)	totwrk	age	male	south
3460.43	-0.18	2.28	92.65	168.85

Для наблюдений с индексом 0, 8, 55, 75, 99 найдите фактическое значение зависимой переменной, предсказанное значение и остатки.

```
=====
i  Факт  Fitted  Остаток
-----
0  2670  3132.491 -462.491
8  3090  3370.083 -280.083
55 2985  2835.360 149.640
75 3660  3248.540 411.460
99 2618  2863.070 -245.070
-----
```

Вычислите TSS, ESS, RSS

```
[1] "TSS=50055043.4042553"
```

```
[1] "ESS=7396259.39341296"
```

```
[1] "RSS=42658784.0108424"
```

Вычислите  $R^2$  и  $adj.R^2$

```
[1] 0.1477625
```

```
[1] 0.1354558
```

Вычислите стандартную ошибку регрессии

```
[1] 392.4319
```

## 1.3 Sleep equation (men only)

Для датасета `sleep75` рассмотрим регрессию `sleep` на `totwrk`, `age`, `smsa`, `south` только для мужчин (`male=1`)

Найдите коэффициенты подогнанной модели. Ответ округлите до 2-х десятичных знаков.

Ответ:

(Intercept)	totwrk	age	smsa	south
3600.96	-0.19	3.30	-77.69	60.11

Для наблюдений с индексом 0, 8, 44, 186, 279 найдите фактическое значение зависимой переменной, предсказанное значение и остатки.

```
=====
i   Факт   Fitted   Остаток
-----
0   3113   3061.481   51.519
8   3018   3087.341   -69.341
44  3363   3268.521   94.479
186 3390   3228.976   161.024
279 2955   3399.060   -444.060
-----
```

Вычислите TSS, ESS, RSS

```
[1] "TSS=75570140.5775"
[1] "ESS=12125135.5978652"
[1] "RSS=63445004.9796349"
```

Вычислите  $R^2$  и  $adj.R^2$

```
[1] 0.1604488
[1] 0.151947
```

Вычислите стандартную ошибку регрессии

```
[1] 400.7746
```

#### 1.4 Sleep equation (women only)

Для датасета `sleep75` рассмотрим регрессию **sleep** на **totwrk**, **age**, **smsa**, **south** *только для женщин (male=0)*

Найдите коэффициенты подогнанной модели. **Ответ округлите до 2-х десятичных знаков.**

Ответ:

(Intercept)	totwrk	age	smsa	south
3468.20	-0.15	1.65	-37.69	136.03

Для наблюдений с индексом 0, 8, 44, 186, 279 найдите фактическое значение зависимой переменной, предсказанное значение и остатки.

```
=====
i   Факт   Fitted   Остаток
-----
0   3083   2945.903   137.097
8   2858   2920.662   -62.662
44  3730   3328.629   401.371
186 3376   3503.722   -127.722
279 3405   3215.073   189.927
-----
```

Вычислите TSS, ESS, RSS

```
[1] "TSS=63490152.1176471"
```

```
[1] "ESS=6300990.56898593"
```

```
[1] "RSS=57189161.5486611"
```

Вычислите  $R^2$  и  $adj.R^2$

```
[1] 0.09924359
```

```
[1] 0.0872734
```

Вычислите стандартную ошибку регрессии

```
[1] 435.8867
```

## 1.5 Labour equation (in levels)

Для датасета Labour рассмотрим регрессию **output** на **capital**, **labour**, **wage**

Найдите коэффициенты подогнанной модели. **Ответ округлите до 2-х десятичных знаков.**

Ответ:

(Intercept)	capital	labour	wage
-16.75	0.40	0.05	0.44

Для наблюдений с индексом 1, 10, 23, 100, 348, 500 найдите фактическое значение зависимой переменной, предсказанное значение и остатки.

```
=====
i   Факт   Fitted Остаток
-----
1   3.664  0.457   3.207
10  3.246  5.666  -2.420
23  2.666 -0.625   3.292
100 2.181  6.952  -4.771
348 4.307  2.278   2.029
500 3.660 11.353 -7.694
-----
```

Вычислите TSS, ESS, RSS

```
[1] "TSS=2235073.2947312"
```

```
[1] "ESS=2190376.07672369"
```

```
[1] "RSS=44697.2180075161"
```

Вычислите  $R^2$  и  $adj.R^2$

```
[1] 0.9800019
```

```
[1] 0.9798957
```

Вычислите стандартную ошибку регрессии

```
[1] 8.894387
```

## 1.6 Labour equation (in log)

Для датасета Labour рассмотрим регрессию **log(output)** на **log(capital)**, **log(labour)**, **log(wage)**

Найдите коэффициенты подогнанной модели. **Ответ округлите до 2-х десятичных знаков.**

Ответ:

(Intercept)	log(capital)	log(labour)	log(wage)
-5.01	0.15	0.72	0.92

Для наблюдений с индексом 1, 10, 23, 100, 348, 500 найдите фактическое значение зависимой переменной, предсказанное значение и остатки.

```
=====
i   Факт   Fitted Остаток
-----
1   1.299  1.347  -0.049
10  1.177  1.427  -0.250
23  0.981  0.344   0.637
100 0.780  1.006  -0.226
348 1.460  1.177   0.283
500 1.297  1.730  -0.433
-----
```

Вычислите TSS, ESS, RSS

```
[1] "TSS=797.667404532845"
```

```
[1] "ESS=708.647370586093"
```

```
[1] "RSS=89.0200339467518"
```

Вычислите  $R^2$  и  $adj.R^2$

```
[1] 0.8883996
```

```
[1] 0.887807
```

Вычислите стандартную ошибку регрессии

```
[1] 0.3969352
```

## 1.7 Cost equation (in levels)

Для датасета `Electricity` рассмотрим регрессию `cost` на `q`, `q2`, `pk`, `pf`, `pl`

Найдите коэффициенты подогнанной модели. **Ответ округлите до 2-х десятичных знаков.**

Ответ:

(Intercept)	q	I(q <sup>2</sup> )	pk	pf	pl
-81.36	0.00	0.00	0.37	0.95	0.00

Для наблюдений с индексом 1, 16, 28, 47, 57, 93 найдите фактическое значение зависимой переменной, предсказанное значение и остатки.

```
=====
i   Факт   Fitted Остаток
-----
1   3.043  -2.512  5.555
16  5.597   2.747  2.850
28  7.439  -2.269  9.709
47 28.786 34.228 -5.442
57 41.902 43.941 -2.040
93  9.467  -4.009 13.477
-----
```

Вычислите TSS, ESS, RSS

```
[1] "TSS=1189954.20941659"
[1] "ESS=1143993.95587659"
[1] "RSS=45960.2535400048"
```

Вычислите  $R^2$  и  $adj.R^2$

```
[1] 0.9613765
[1] 0.9601059
```

Вычислите стандартную ошибку регрессии

```
[1] 17.38879
```

## 1.8 Cost equation (in logs)

Для датасета `Electricity` рассмотрим регрессию  $\log(\text{cost})$  на  $\log(q)$ ,  $\log^2(q)$ ,  $\log(pk)$ ,  $\log(pf)$ ,  $\log(pl)$

Найдите коэффициенты подогнанной модели. **Ответ округлите до 2-х десятичных знаков.**

Ответ:

(Intercept)	$\log(q)$	$I(\log(q)^2)$	$\log(pk)$	$\log(pf)$	$\log(pl)$
-6.74	0.40	0.03	0.16	0.68	0.15

Для наблюдений с индексом 1, 16, 28, 47, 57, 93 найдите фактическое значение зависимой переменной, предсказанное значение и остатки.

```
=====
i   Факт   Fitted Остаток
-----
1   1.113  1.452  -0.339
16  1.722  1.650   0.072
28  2.007  1.784   0.223
47  3.360  3.556  -0.196
57  3.735  3.846  -0.111
93  2.248  2.247   0.001
-----
```

Вычислите TSS, ESS, RSS

```
[1] "TSS=373.664560791614"
[1] "ESS=370.75994248779"
[1] "RSS=2.90461830382466"
```

Вычислите  $R^2$  и  $adj.R^2$

```
[1] 0.9922267
[1] 0.991971
```

Вычислите стандартную ошибку регрессии

```
[1] 0.1382365
```