

Задачи по Эконометрике: Доверительные интервалы

Н.В. Артамонов (МГИМО МИД России)

Содержание

1 Построение доверительных интервалов	1
1.1 sleep equation	1
1.2 output equation	2
1.3 cost equation	2
2 Доверительные интервалы и t-тест	2
2.1 sleep equation	2
2.2 cost equation	4
3 Доверительные интервалы и сравнение моделей	6

1 Построение доверительных интервалов

1.1 sleep equation

Для набора данных sleep75 рассмотрим линейную регрессию $\text{sleep} \sim 1 + \text{totwrk} + \text{age} + \text{south} + \text{male} + \text{smsa} + \text{yngkid} + \text{marr}$.

Постройте 90%-доверительные интервалы (неробастные и робастные) для каждого коэффициента. **Ответ округлите до 2-х десятичных знаков.**

Ответ (неробастные)

	5 %	95 %
(Intercept)	3317.954	3583.872
totwrk	-0.199	-0.139
age	0.269	5.109
south	32.662	170.475
male	29.851	145.487
smsa	-109.302	-0.193
yngkid	-96.876	68.952
marr	-38.349	100.770

Ответ (робастные)

	5 %	95 %
(Intercept)	3301.820	3600.005
totwrk	-0.203	-0.136
age	0.318	5.061
south	31.548	171.588
male	28.619	146.719
smsa	-109.835	0.340
yngkid	-103.344	75.419
marr	-44.867	107.289

1.2 output equation

Для набора данных `Labour` рассмотрим линейную регрессию $\log(\text{output}) \sim 1 + \log(\text{capital}) + \log(\text{labour})$.

Постройте 95%-доверительные интервалы (неробастные и робастные) для каждого коэффициента. **Ответ округлите до 2-х десятичных знаков.**

Ответ (неробастные)

	2.5 %	97.5 %
(Intercept)	-1.901	-1.522
$\log(\text{capital})$	0.174	0.241
$\log(\text{labour})$	0.669	0.760

Ответ (робастные)

	2.5 %	97.5 %
(Intercept)	-2.073	-1.350
$\log(\text{capital})$	0.146	0.270
$\log(\text{labour})$	0.629	0.801

1.3 cost equation

Для набора данных `Electricity` рассмотрим линейную регрессию $\log(\text{cost}) \sim 1 + \log(q) + I(\log(q)^2) + \log(pl) + \log(pk) + \log(pf)$.

Постройте 99%-доверительные интервалы (неробастные и робастные) для каждого коэффициента. **Ответ округлите до 2-х десятичных знаков.**

Ответ (неробастные)

	0.5 %	99.5 %
(Intercept)	-8.581	-4.896
$\log(q)$	0.320	0.486
$I(\log(q)^2)$	0.025	0.036
$\log(pl)$	-0.038	0.330
$\log(pk)$	0.007	0.308
$\log(pf)$	0.573	0.796

Ответ (робастные)

	0.5 %	99.5 %
(Intercept)	-8.949	-4.529
$\log(q)$	0.230	0.576
$I(\log(q)^2)$	0.020	0.041
$\log(pl)$	-0.077	0.369
$\log(pk)$	-0.005	0.320
$\log(pf)$	0.549	0.820

2 Доверительные интервалы и t-тест

2.1 sleep equation

Для набора данных `sleep75` рассмотрим линейную регрессию `sleep` на `totwrk`, `age`, `age^2`, `south`, `male`, `smsa`, `ynghkid`, `marr`, `union`.

Ниже приведены (неробастные) 90%-доверительные интервалы для коэффициентов.

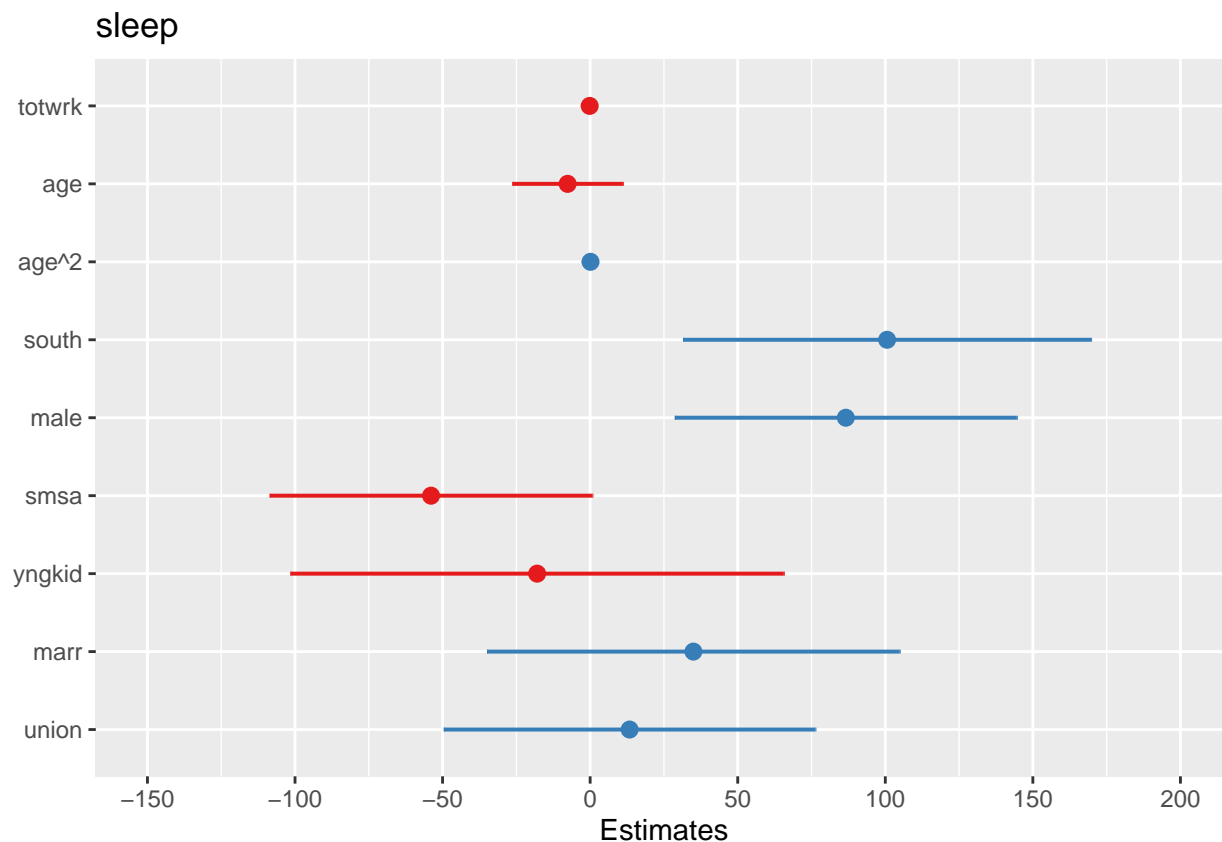
	5 %	95 %
(Intercept)	3269.356	4007.222

totwrk	-0.197	-0.137
age	-26.347	11.084
I (age^2)	-0.098	0.347
south	31.483	169.733
male	28.681	144.571
smsa	-108.584	0.775
yngkid	-101.543	65.622
marr	-34.936	104.904
union	-49.590	76.328

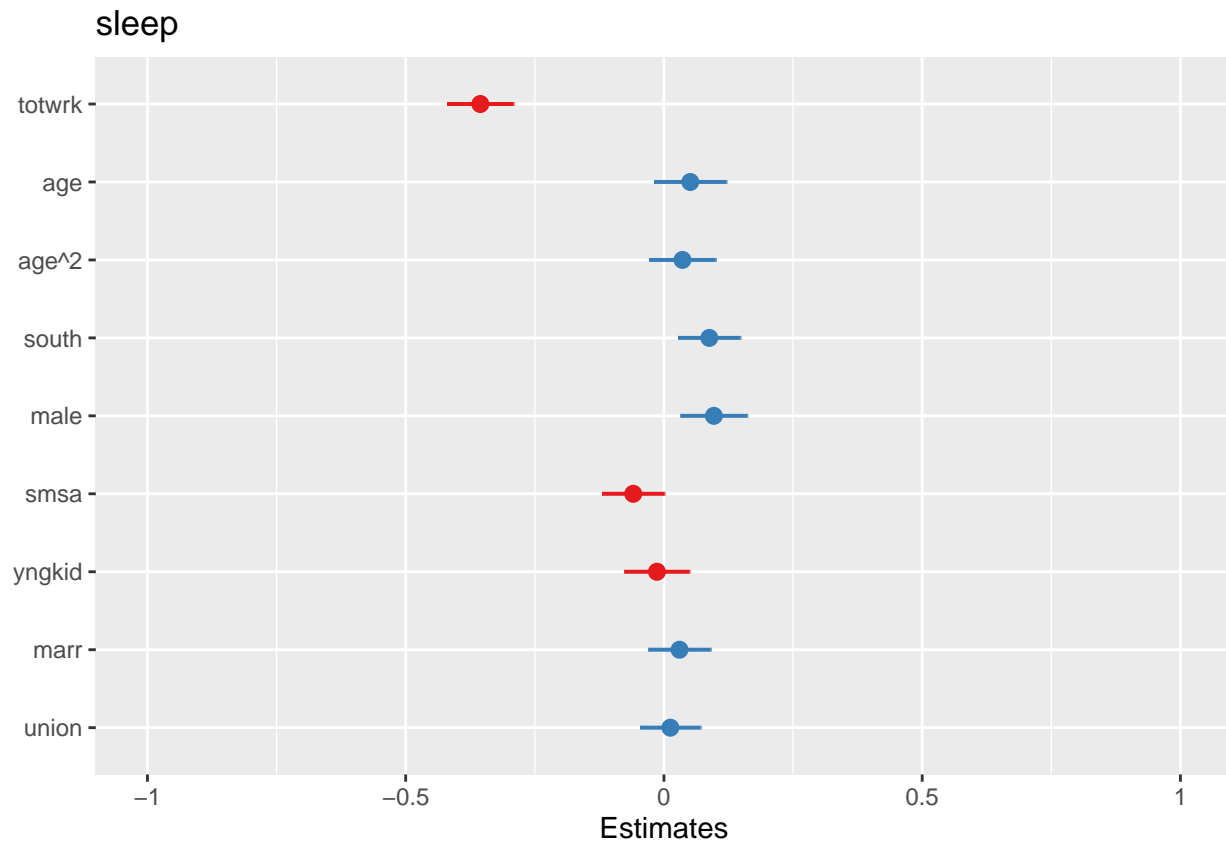
Какие коэффициенты значимы при уровне значимости 10%? Ответ

```
=====
Коэффициент  Значимость
-----
(Intercept)   Значим
totwrk         Значим
age            Незначим
I (age2)       Незначим
south          Значим
male           Значим
smsa           Незначим
yngkid         Незначим
marr           Незначим
union          Незначим
-----
```

Визуализация доверительных интервалов как есть (в одном масштабе)



Шкалированные доверительные интервалы



2.2 cost equation

Для набора данных *Electricity* рассмотрим линейную регрессию $\log(\text{cost})$ на $\log(q)$, $\log^2(q)$, $\log(pl)$, $\log^2(pl)$, $\log(pk)$, $\log^2(pk)$, $\log(pf)$, $\log^2(pf)$

Ниже приведены (робастные) 99%-доверительные интервалы для коэффициентов.

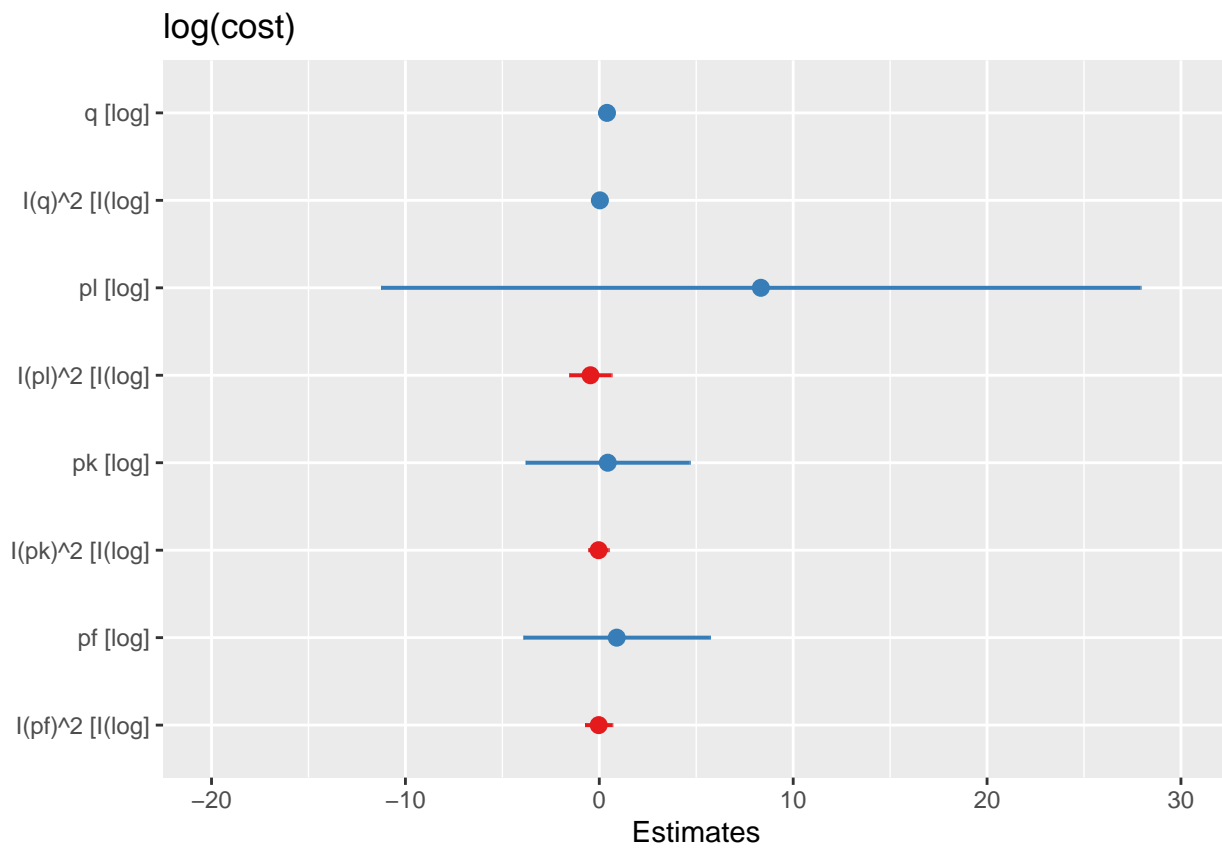
	0.5 %	99.5 %
(Intercept)	-129.940	41.133
$\log(q)$	0.223	0.570
$I(\log(q)^2)$	0.020	0.042
$\log(pl)$	-11.254	27.921
$I(\log(pl)^2)$	-1.544	0.633
$\log(pk)$	-3.795	4.668
$I(\log(pk)^2)$	-0.559	0.487
$\log(pf)$	-3.909	5.707
$I(\log(pf)^2)$	-0.734	0.673

Какие коэффициенты значимы при уровне значимости 1%? Ответ

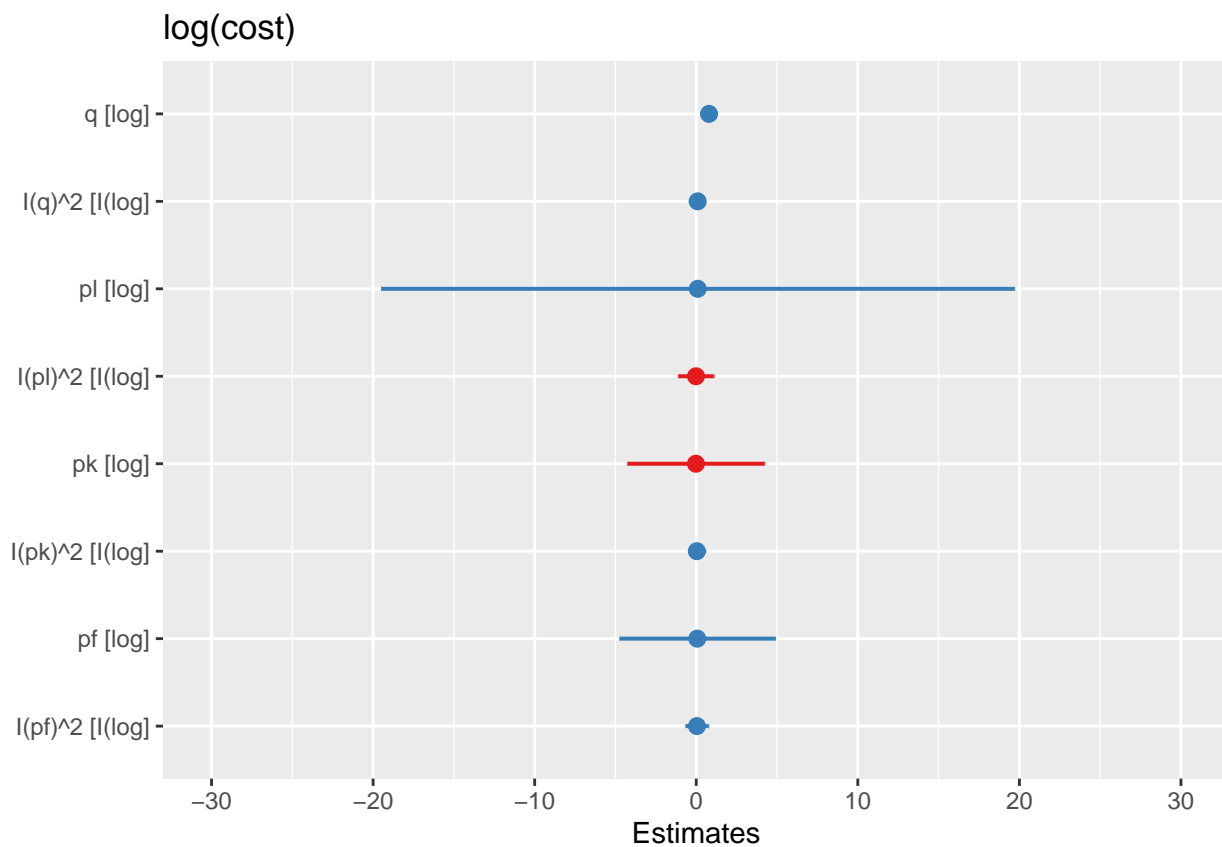
```
=====
Коэффициент  Значимость
-----
(Intercept)  Незначим
log(q)       Значим
I(log(q)^2)  Значим
```

$\log(p_l)$	Незначим
$I(\log(p_l)^2)$	Незначим
$\log(p_k)$	Незначим
$I(\log(p_k)^2)$	Незначим
$\log(p_f)$	Незначим
$I(\log(p_f)^2)$	Незначим

Визуализация доверительных интервалов как есть (в одном масштабе)



Шкалированные доверительные интервалы



3 Доверительные интервалы и сравнение моделей

Результаты подгонки:

Dependent variable:			
	sleep		
	Север (1)	Юг (2)	общая (3)
totwrk	-0.159*** (0.018)	-0.074 (0.045)	-0.147*** (0.017)
age	-5.507 (12.557)	-12.882 (25.039)	-8.675 (11.271)
$I(\text{age}^2)$	0.084 (0.150)	0.273 (0.299)	0.140 (0.135)
smsa	-59.812* (35.379)	16.055 (96.789)	-69.924** (32.330)
marr	28.034 (46.960)	152.296* (86.153)	47.598 (41.450)

Constant	3664.188*** (244.380)	3435.410*** (494.979)	3675.413*** (219.803)
----------	--------------------------	--------------------------	--------------------------

Observations	576	130	706
R2	0.129	0.140	0.118
Adjusted R2	0.122	0.106	0.112
Residual Std. Error	421.078	393.300	418.809
F Statistic	16.928***	4.043***	18.768***

Note: *p<0.1; **p<0.05; ***p<0.01

