

Прикладная статистика. Занятие 7. Регрессионный анализ.

5 апреля 2011 г.

Ericsson, Yount, Effect of aluminized covers on body temperature in patients having abdominal surgery (1991): в результате абдоминальной хирургии происходит снижение температуры пациента, для уменьшения этого эффекта могут быть использованы алюминизированные покрывала. В ходе исследования были собраны данные по 60 пациентам, известны возраст, индекс массы тела, длительность операции, измеренная в ухе температура до операции и по окончании анестезии, а также индикаторы использования алюминизированных покрывал (для головы, для тела) и согретых одеял.



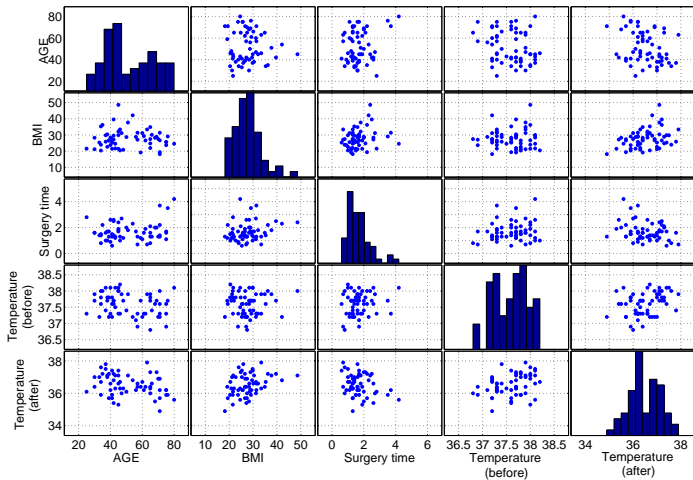
Построить функцию, предсказывающую температуру по окончании анестезии.

Температура при абдоминальной хирургии

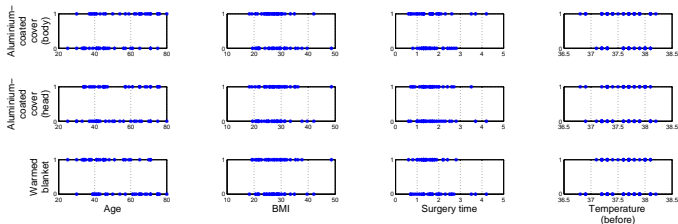
AGE	BMI	SURGTM	BODY	HEAD	BODYCOV	TTEMP1	TTEMP
59	19.2	1.2	1	1	1	37.7	36.4
39	26.6	1.3	0	0	0	37.2	35.7
75	23.7	1.7	1	0	0	36.9	35.9
34	24.0	0.8	0	1	1	38.0	37.6
71	18.2	1.3	1	1	0	37.2	34.9
65	22.0	1.3	0	1	1	37.7	36.3
41	25.3	0.6	1	0	1	37.6	37.4
46	20.5	1.0	1	0	0	38.2	36.7
56	28.8	1.7	0	0	1	37.1	36.2
42	27.2	2.6	0	1	0	37.6	35.4
51	37.7	1.8	0	0	1	37.9	37.1
38	22.7	1.0	1	0	1	37.8	37.0
68	28.3	2.0	1	1	0	37.6	35.5
37	29.8	1.0	0	0	1	38.1	37.5
35	36.2	2.2	0	1	1	38.0	37.2
65	34.9	1.6	1	1	0	37.9	36.4
71	31.4	3.7	1	0	0	37.3	36.2
65	27.5	0.8	1	1	0	36.8	36.0
60	31.2	1.1	0	0	1	37.2	36.7
48	20.9	1.2	0	0	1	37.7	36.3
37	25.9	1.6	1	1	1	37.4	37.8
66	30.1	1.3	1	0	0	37.4	36.6
71	26.7	1.4	0	1	1	38.0	36.9
30	21.2	1.6	1	0	0	37.9	37.0
69	28.9	2.0	1	1	0	37.7	37.3
47	31.2	2.7	0	1	0	37.9	36.0
30	28.3	1.6	0	0	1	37.7	36.4

...

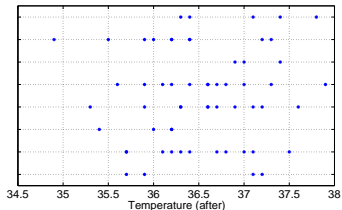
Температура при абдоминальной хирургии



Температура при абдоминальной хирургии



Способы согревания действительно распределены между пациентами случайным образом.



Влияние способов согревания на температуру по окончании анестезии.

Корреляции (Кендалла) числовых признаков:

τ	Age	BMI	Surgery time	Temperature before	Temperature after
Age	1.0000	0.0172	0.0743	-0.1481	-0.2306
BMI	0.0172	1.0000	0.1899	-0.0261	0.2323
Surgery time	0.0743	0.1899	1.0000	0.1199	-0.1375
Temperature before	-0.1481	-0.0261	0.1199	1.0000	0.2627
Temperature after	-0.2306	0.2323	-0.1375	0.2627	1.0000

p-value	Age	BMI	Surgery time	Temperature before	Temperature after
Age	0.0000	0.8531	0.4196	0.1124	0.0115
BMI	0.8531	0.0000	0.0369	0.7824	0.0104
Surgery time	0.4196	0.0369	0.0000	0.2044	0.1369
Temperature before	0.1124	0.7824	0.2044	0.0000	0.0051
Temperature after	0.0115	0.0104	0.1369	0.0051	0.0000

Частные корреляции (Спирмена) числовых признаков:

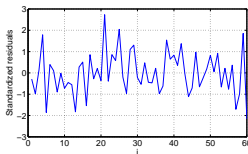
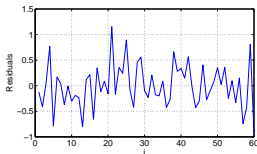
ρ	Age	BMI	Surgery time	Temperature before	Temperature after
Age	1.0000	0.1265	0.0125	-0.0762	-0.2651
BMI	0.1265	1.0000	0.4311	-0.2946	0.5017
Surgery time	0.0125	0.4311	1.0000	0.3498	-0.4210
Temperature before	-0.0762	-0.2946	0.3498	1.0000	0.4570
Temperature after	-0.2651	0.5017	-0.4210	0.4570	1.0000

p-value	Age	BMI	Surgery time	Temperature before	Temperature after
Age	0.0000	0.3485	0.9267	0.5733	0.0001
BMI	0.3485	0.0000	0.0008	0.0261	0.0104
Surgery time	0.9267	0.0008	0.0000	0.0076	0.0011
Temperature before	0.5733	0.0261	0.0076	0.0000	0.0004
Temperature after	0.0463	0.0001	0.0011	0.0004	0.0000

Стандартная линейная регрессия:

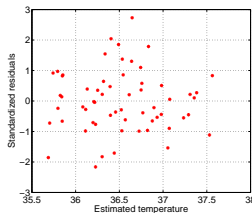
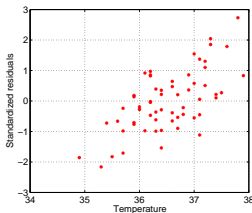
$$T_1 = 7.383 - 0.006 * Age + 0.005 * BMI - 0.367 * Time + 0.471 * Body - 0.248 * Head + 0.297 * Blanket + 0.752 * T_0.$$

$$MSE = 0.2066, R^2 = 0.5704, R_A^2 = 0.5125, F = 9.8618, p = 9 \times 10^{-8}.$$



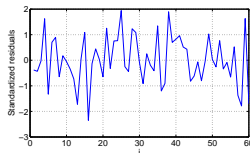
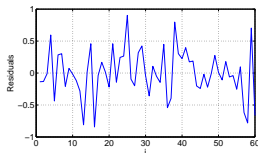
Критерий автокорреляции Дарбина-Уотсона: $p = 0.3045$.

Критерий нормальности Андерсона-Дарлинга: $p = 0.6952$.



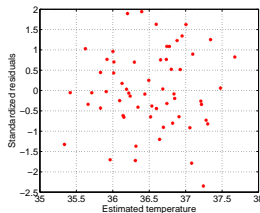
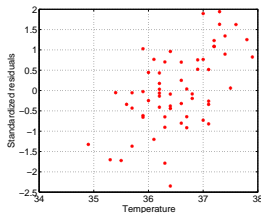
Линейная регрессия с добавлением взаимодействий:

$$MSE = 0.2629, R^2 = 0.6741, R_A^2 = 0.3797, F = 2.2899, p = 0.0132.$$



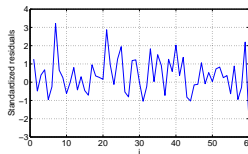
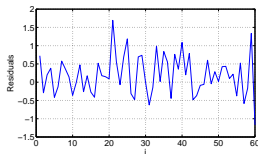
Критерий автокорреляции Дарбина-Уотсона: $p = 0.5743$.

Критерий нормальности Андерсона-Дарлинга: $p = 0.3833$.



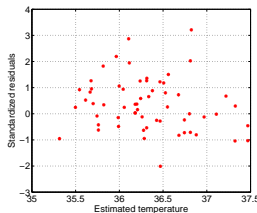
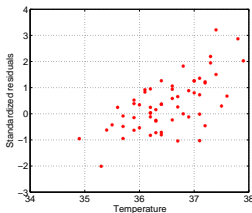
Линейная регрессия с добавлением квадратов предикторов:

$$MSE = 0.4273, R^2 = 0.2312, R_A^2 = -0.0080, F = 2.9333, p = 0.0031.$$



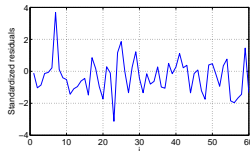
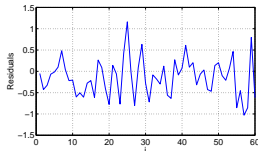
Критерий автокорреляции Дарбина-Уотсона: $p = 0.6775$.

Критерий нормальности Андерсона-Дарлинга: $p = 0.6966$.



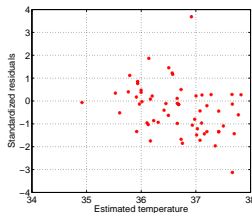
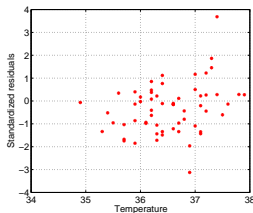
Линейная регрессия с добавлением взаимодействий и квадратов предикторов:

$$MSE = 0.5288, R^2 = 0.4925, R_A^2 = -0.2476, F = 1.4297, p = 0.1816.$$

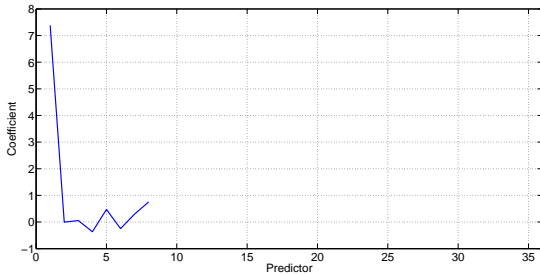


Критерий автокорреляции Дарбина-Уотсона: $p = 0.1433$.

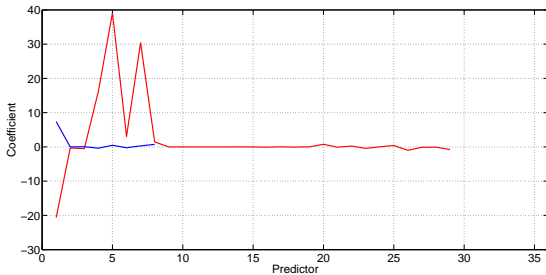
Критерий нормальности Андерсона-Дарлинга: $p = 0.6039$.



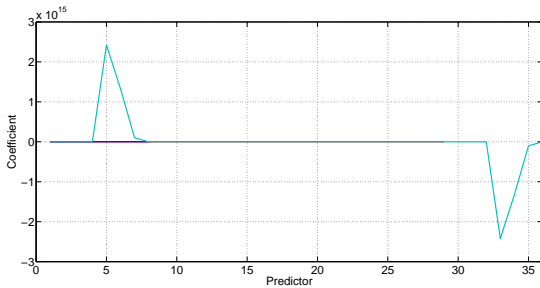
Температура при абдоминальной хирургии



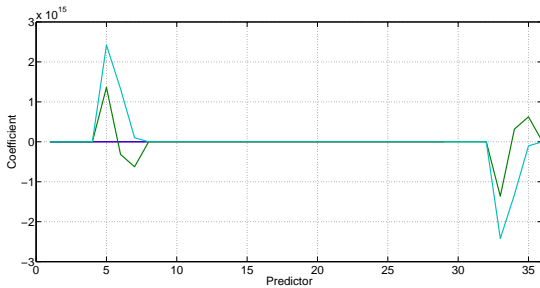
Температура при абдоминальной хирургии



Температура при абдоминальной хирургии



Температура при абдоминальной хирургии



Пошаговая регрессия:

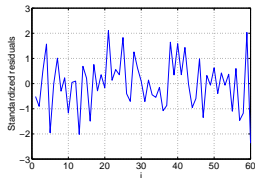
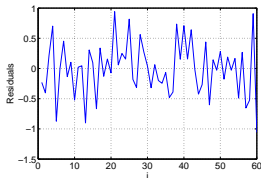
- Step 1, added $BMI * T_0$, $p = 0.00444823$;
- Step 2, added $Age * Time$, $p = 0.00110911$;
- Step 3, added T_0^2 , $p = 0.00153327$;
- Step 4, added $Time * Body$, $p = 0.00243268$;
- Step 5, added $Age * Blanket$, $p = 0.0452467$.

$$T_1 = -0.0082 * Age * Time + 0.0057 * Age * Blanket + 0.0013 * BMI * T_0 + 0.3220 * Time * Body + 0.0085 * T_0^2.$$

$$MSE: 0.6121 \rightarrow 0.5619 \rightarrow 0.5179 \rightarrow 0.4804 \rightarrow 0.4670.$$

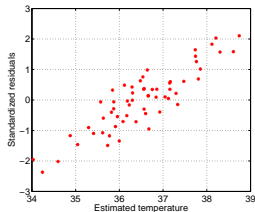
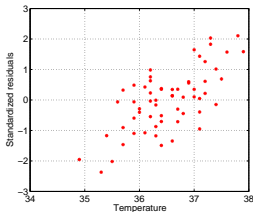
$$F = 12.1388, \quad p = 7 \times 10^{-8}.$$

Температура при абдоминальной хирургии

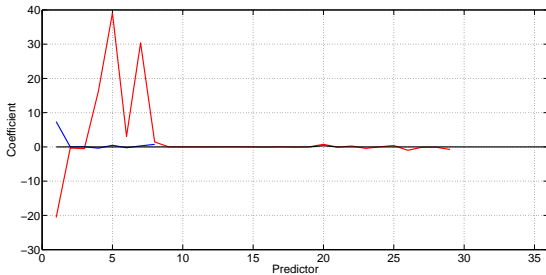


Критерий автокорреляции Дарбина-Уотсона: $p = 0.3265$.

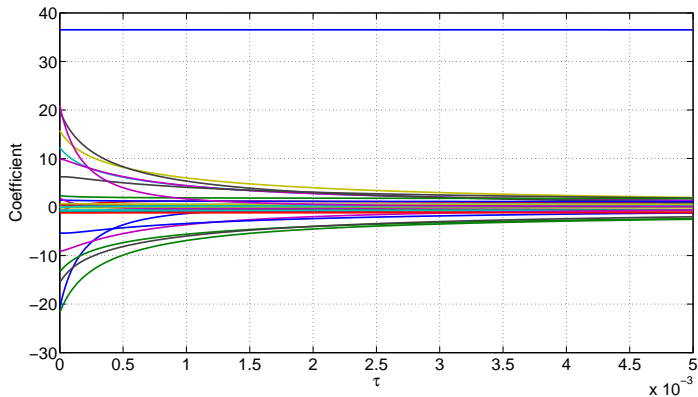
Критерий нормальности Андерсона-Дарлинга: $p = 0.7522$.



Температура при абдоминальной хирургии

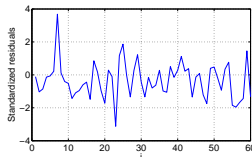
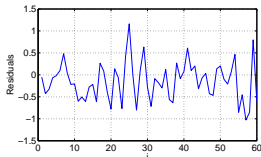


Гребневая регрессия, коэффициенты



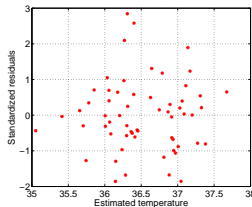
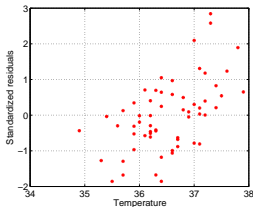
При $\tau = 0.004$:

$MSE = 0.1192$

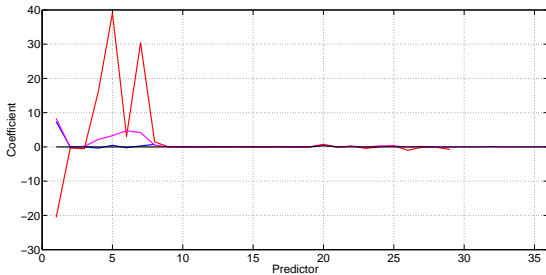


Критерий автокорреляции Дарбина-Уотсона: $p = 0.5442$.

Критерий нормальности Андерсона-Дарлинга: $p = 0.4094$.

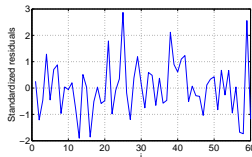
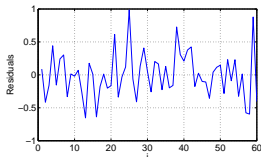


Температура при абдоминальной хирургии



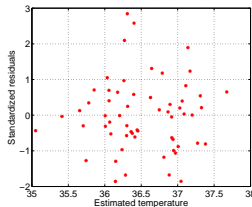
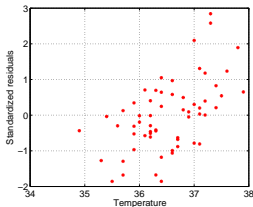
При $\tau = 0.002$:

$MSE = 0.1159$

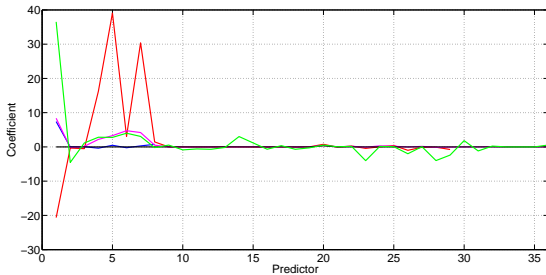


Критерий автокорреляции Дарбина-Уотсона: $p = 0.5746$.

Критерий нормальности Андерсона-Дарлинга: $p = 0.4318$.



Температура при абдоминальной хирургии



Что ещё можно сделать:

- применить лассо;
- разбить возраст на два признака;
- непараметрическая регрессия;
- скользящий контроль для выбора параметра гребневой регрессии;
- перебор и сравнение всех 8589934590 регрессионных моделей (появляется эффект множественных сравнений);
- построить эластичную сеть;
- ...

Прикладная статистика
Семинар 7. Регрессионный анализ.

Рябенко Евгений
riabenko.e@gmail.com