

Семинар 12.

Выбор функциональной зависимости.

- Для 400 голландских магазинов модной одежды с помощью трёх моделей оценили зависимость продаж в расчете на квадратный метр в гульденах, $Sales$, от:

- общей площади магазина, $Size$, в m^2 ;
- количества сотрудников, работающих целый день, $Nfull$;
- количества временных рабочих, $Ntemp$;
- дамми-переменной $Owner$, равной единице, если собственник один, и нулю иначе.

$$\widehat{Sales}_i = 6083 - 15.25Size_i + 1452.8Nfull_i + 420.15Ntemp_i - 1464.1Owner_i$$

(718) (1.59) (171) (423) (361)

$$\ln \widehat{Sales}_i = 8.59 - 0.0024Size_i + 0.183Nfull_i + 0.102Ntemp_i - 0.209Owner_i$$

(0.11) (0.00024) (0.026) (0.066) (0.056)

$$\ln \widehat{Sales}_i = 10.08 - 0.31 \ln Size_i + 0.22 \ln Nfull_i + 0.066 \ln Ntemp_i - 0.19 \ln Owner_i$$

(0.21) (0.043) (0.061) (0.118) (0.059)

В скобках приведены стандартные ошибки.

- Дайте интерпретацию коэффициента при переменной $Size$ в каждой из трёх моделей;
- Подробно опишите, как выбрать наилучшую из этих моделей.

- С помощью теста Бокса-Кокса оценили зависимость веса индивида (в килограммах) от его роста (в сантиметрах):

Log likelihood = -2659.5656 Number of obs = 540
 LR chi2(2) = 230.68
 Prob > chi2 = 0.000

	w	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
/lambda		1.055498	1.892654	0.56	0.577	-2.654035 4.76503
/theta		-.0263371	.1471576	-0.18	0.858	-.3147607 .2620865

Estimates of scale-variant parameters

	Coef.
Notrans	
_cons	2.936809
Trans	
h	.0237224
/sigma	.1660251

Test H0:	Restricted log likelihood	chi2	Prob > chi2
theta=lambda = -1	-2680.8693	42.61	0.000
theta=lambda = 0	-2659.7618	0.39	0.531
theta=lambda = 1	-2685.5201	51.91	0.000

Какую спецификацию модели (линейную, линейную в логарифмах, полулогарифмическую) следует предпочесть и почему?

3. Для выбора между линейной и полулогарифмической моделями (где *EARNINGS* — почасовая заработная плата в \$, *S* — длительность обучения, *ASVABC* — результаты тестов, характеризующие успеваемость) был проведен тест Дэвидсона, Уайта и МакКиннона и получены следующие результаты:

	Зависимая: Y	Зависимая: $\ln Y$
(Intercept)	-26.148 (4.17)	-1.941 (3.2499)
S	2.008 (0.276)	0.087 (0.035)
ASVABC	0.393 (0.079)	0.017 (0.007)
lin_add	-15.373 (5.984)	
semilog_add		-0.029 (0.065)
R^2	0.2071	0.2212
F	46.59	50.74
Adj. R^2	0.2027	0.2168
Num. obs.	540	540
RSS	90975.57	148.1
$\hat{\sigma}$	13.04	0.5256

Где $\text{lin_add} = \ln Y - \ln Y$, $\text{semilog_add} = Y - \exp(\ln Y)$ и в скобках указаны стандартные ошибки. С помощью РЕ-теста на уровне значимости 5% сделайте выбор между двумя оцененными моделями.