

Метрика. Экзамен 2016

Задача 1. С помощью МНК оцените модель $y_i = \beta x_i + \varepsilon_i$ по трём наблюдениям: $x_1 = 1$, $x_2 = 1$, $x_3 = -2$, $y_1 = 0$, $y_2 = 1$, $y_3 = 4$. Найдите $\hat{\beta}$, RSS , TSS , ESS , R^2 .

Задача 2.

Исследователь Тимофей следующую функцию спроса с сезонными переменными *SPRING* (весна), *SUMMER* (лето), *FALL* (осень):

$$\widehat{\ln Q} = \hat{\beta}_1 + \hat{\beta}_2 \cdot \ln P + \hat{\beta}_3 \cdot \text{SPRING} + \hat{\beta}_4 \cdot \text{SUMMER} + \hat{\beta}_5 \cdot \text{FALL}, \quad R^2 = 0.24, n = 24.$$

Тимофей желает протестировать гипотезу $H_0 : \begin{cases} \beta_3 = 0, \\ \beta_4 = \beta_5 \end{cases}$.

1. Дайте интерпретацию проверяемой гипотезе.
2. Какую ограниченную регрессию надо оценить Тимофею для проверки данной гипотезы?
3. Пусть для регрессии с ограничениями $R^2 = 0.13$. На уровне значимости 5% проверьте нулевую гипотезу.

Задача 3.

Винни-Пух построил логит-модель для предсказания качества мёда: $y_i = 0$ означает неправильный мёд, а $y_i = 1$ — правильный. Правильность мёда зависит от разных характеристик дупла, в частности от правильности пчёл (x_i) и их количества (z_i). Оценка модели выглядит так:

$$\Lambda(P[y_i = 1]) = -0.2 + 2x_i + 0.03z_i$$

Здесь $\Lambda()$ — логистическая функция, $\Lambda(t) = e^t / (1 + e^t)$.

1. Спрогнозируйте вероятность правильного мёда для дупла с сотней правильных пчёл.
2. Рассмотрим дупло с сотней правильных пчёл. Рассчитайте предельный эффект увеличения от количества пчёл на вероятность правильности мёда.
3. При каком количестве правильных пчёл предельный эффект от количества пчёл на вероятность правильности мёда будет максимальным?

	Model 1
(Intercept)	8.45*** (0.00)
log(carat)	1.68*** (0.00)
R ²	0.93
Adj. R ²	0.93
Num. obs.	53940
RMSE	0.26
*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$	

Таблица 1: Statistical models

Задача 4.

Исследовательница Мишель оценила зависимость логарифма цены бриллиантов от логарифма массы. Результаты приведены в таблице 1.

```
model <- lm(data = diamonds, log(price) ~ log(carat))
texreg(model)
```

А затем провела тест Голдфелда-Квандта (таблица 2):

```
gqtest(model, fraction = 0.2,
       order.by = diamonds$carat) %>% pander
```

Таблица 2: Goldfeld-Quandt test: model

Test statistic	df1	df2	P value
1.363	21574	21574	1.501e-114 * * *

1. Аккуратно сформулируйте H_0 и H_a данного теста
2. К каким выводам пришла Мишель по результатам теста?
3. Что можно посоветовать Мишель, если она желает строить предиктивные доверительные интервалы для цены бриллиантов, зная их массу?

Задача 5. Рассмотрим стационарный случайный процесс удовлетворяющий уравнению

$$y_t = 3 + 0.5y_{t-1} - 0.06y_{t-2} + \varepsilon_t, \quad \varepsilon_t \sim WN(0; \sigma^2).$$

1. Найдите $\mathbb{E}(y_t)$.
2. Найдите первые два значения корреляционной и автокорреляционной функций.
3. Дополнительно известно, что остатки имеют нормальное распределение, $\varepsilon_t \sim N(0; 1)$, $y_{99} = 6$, $y_{100} = 5$. Постройте точечный прогноз y_{101} и 95%-ый предиктивный интервал для y_{101} .