

Поехали!!!

Короткие вопросы:

1. Методом наименьших квадратов оцените коэффициент β в модели $y_i = \beta \cdot i + \varepsilon_i$.
-

2. Априори известно, что парная регрессия должна проходить через точку $(0, 1)$. Найдите мнк оценки в регрессии $\hat{y}_i = \hat{\beta}_1 + \hat{\beta}_2 x_i$
-

3. Аня и Настя утверждают, что лектор опоздал на 10 минут. Таня считает, что лектор опоздал на 3 минуты. С помощью мнк оцените на сколько опоздал лектор.
-

4. Эконометрист Вовочка оценил линейную регрессионную модель, где y измерялся в тугриках. Затем он оценил ту же модель, но измерял y в мунгу (1 тугрик = 100 мунгу). Как изменятся оценки коэффициентов?
-

5. Какие из указанных моделей можно представить в линейном виде?

(a) $y_i = \beta_1 + \frac{\beta_2}{x_i} + \varepsilon_i$

(b) $y_i = \exp(\beta_1 + \beta_2 x_i + \varepsilon_i)$

(c) $y_i = 1 + \frac{1}{\exp(\beta_1 + \beta_2 x_i + \varepsilon_i)}$

(d) $y_i = \frac{1}{1 + \exp(\beta_1 + \beta_2 x_i + \varepsilon_i)}$

(e) $y_i = x_i^{\beta_2} e^{\beta_1 + \varepsilon_i}$

6. Сформулируйте предпосылки теоремы Гаусса-Маркова
-

7. Сформулируйте выводы теоремы Гаусса-Маркова
-

8. Неограниченная регрессионная модель имеет вид $y_i = \beta_1 + \beta_2 x_i + \beta_3 z_i + \beta_4 w_i + \varepsilon_i$. Какую регрессию следует оценить для проверки ограничений $\begin{cases} \beta_2 = \beta_3 \\ \beta_4 = 0 \end{cases}$?
-

9. Неограниченная регрессионная модель имеет вид $y_i = \beta_1 + \beta_2 x_i + \beta_3 z_i + \beta_4 w_i + \varepsilon_i$. Какую регрессию следует оценить для проверки ограничений $\begin{cases} \beta_2 + \beta_3 = 1 \\ \beta_3 + \beta_4 = 0 \end{cases}$?
-

10. Если соответствующее р-значение равно 0.03, то гипотеза $H_0 : \beta = 0$ будет отвергаться при уровнях значимости лежащих в интервале

11. Числом обусловленности матрицы называется отношение

12. Большие значения коэффициентов вздутия дисперсии, VIF, являются одним из признаков

Большие задачи:

1. Регрессионная модель задана в матричном виде при помощи уравнения $y = X\beta + \varepsilon$, где $\beta = (\beta_1, \beta_2, \beta_3)'$. Известно, что $\mathbb{E}(\varepsilon) = 0$ и $\text{Var}(\varepsilon) = \sigma^2 \cdot I$. Известно также, что

$$y = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{pmatrix}, X = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \end{pmatrix}.$$

Для удобства расчетов приведены матрицы

$$X'X = \begin{pmatrix} 5 & 2 & 1 \\ 2 & 2 & 0 \\ 1 & 0 & 1 \end{pmatrix} \text{ и } (X'X)^{-1} = \frac{1}{2} \begin{pmatrix} 1 & -1 & -1 \\ -1 & 2 & 1 \\ -1 & 1 & 3 \end{pmatrix}.$$

- (a) Укажите число наблюдений.
-

- (b) Укажите число регрессоров с учетом свободного члена.
-

- (c) Рассчитайте при помощи метода наименьших квадратов $\hat{\beta}$, оценку для вектора неизвестных коэффициентов.
-

- (d) Рассчитайте $TSS = \sum (y_i - \bar{y})^2$, $RSS = \sum (y_i - \hat{y}_i)^2$ и $ESS = \sum (\hat{y}_i - \bar{y})^2$.
-

- (e) Чему равен $\hat{\varepsilon}_5$, МНК-остаток регрессии, соответствующий 5-ому наблюдению?
-

- (f) Чему равен R^2 в модели? Прокомментируйте полученное значение с точки зрения качества оцененного уравнения регрессии.
-

- (g) Рассчитайте несмещенную оценку для неизвестного параметра σ^2 регрессионной модели.
-

- (h) Рассчитайте $\widehat{\text{Var}}(\hat{\beta})$, оценку для ковариационной матрицы вектора МНК-коэффициентов $\hat{\beta}$.
-

- (i) Найдите $\widehat{\text{Var}}(\hat{\beta}_1)$, несмещенную оценку дисперсии МНК-коэффициента $\hat{\beta}_1$.
-

(j) Найдите $\widehat{\text{Cov}}(\hat{\beta}_1, \hat{\beta}_2)$, несмещенную оценку ковариации МНК-коэффициентов $\hat{\beta}_1$ и $\hat{\beta}_2$.

(k) Найдите $\widehat{\text{Var}}(\hat{\beta}_1 + \hat{\beta}_2)$

(l) Найдите $\widehat{\text{Corr}}(\hat{\beta}_1, \hat{\beta}_2)$, оценку коэффициента корреляции МНК-коэффициентов $\hat{\beta}_1$ и $\hat{\beta}_2$.

(m) Найдите $s_{\hat{\beta}_1}$, стандартную ошибку МНК-коэффициента $\hat{\beta}_1$.

(n) Рассчитайте выборочную корреляцию y и \hat{y} .

2. Оценена модель зависимости продолжительности сна млекопитающих от массы мозга и массы тела:

```
lm(formula = sleep_total ~ brainwt + bodywt, data = msleep)
```

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	10.6722485	0.5898643	18.093	<2e-16 ***
brainwt	-2.3518943	1.6180072	-1.454	0.152
bodywt	0.0007953	0.0016703	0.476	0.636

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 4.193 on 53 degrees of freedom

(27 observations deleted due to missingness)

Multiple R-squared: 0.1337, Adjusted R-squared: 0.101

F-statistic: 4.088 on 2 and 53 DF, p-value: 0.02232

(a) Протестируйте на значимость регрессию «в целом» на уровне значимости 5%

i. Сформулируйте основную и альтернативную гипотезу, которые соответствуют тесту на значимость уравнения регрессии «в целом».

ii. Приведите формулу для тестовой статистики.

iii. Укажите распределение тестовой статистики.

iv. Вычислите наблюдаемое значение тестовой статистики.

v. Укажите границы области, где основная гипотеза не отвергается.

vi. Сделайте статистический вывод о значимости уравнения регрессии «в целом»

(b) Проверьте гипотезу $H_0 : \beta_{bodywt} = 1$ против альтернативной гипотезы $H_a : \beta_{bodywt} < 1$.
Уровень значимости 5%.

i. Приведите формулу для тестовой статистики.

ii. Укажите распределение тестовой статистики.

iii. Вычислите наблюдаемое значение тестовой статистики.

iv. Укажите границы области, где основная гипотеза не отвергается.

v. Сделайте статистический вывод.

(c) Укажите оценки коэффициентов, значимые на 5%-ом уровне значимости