

Тест	1	2	3	4	Итого

← для проверяющего!

Фамилия, имя, номер группы:

.....

Ответы на тест:

1	2	3	4	5	6	7	8	9	10

Тест

Вопрос 1. При проверке модели $Y_i = \beta_0 + \beta_1 X_{i1} + \dots + \beta_k X_{ik} + \varepsilon_i$ на адекватность нулевая гипотеза имеет вид:

- ☐ $\beta_0 = \beta_1 = \dots = \beta_k = 0$
☐ $\beta_0 = \beta_1 = \dots = \beta_k$
☐ $X_{i1} = \dots = X_{ik} = 0$
☐ $\beta_1 = \dots = \beta_k = 0$
☐ $\beta_1 = \dots = \beta_k$
☐ нет верного ответа

Вопрос 2. При проверке модели множественной регрессии $Y_i = \beta_0 + \beta_1 X_i + \beta_2 Z_i + \varepsilon_i$ на мультиколлинеарность оказалось, что $VIF_X = VIF_Z = 10.26$. Из этого можно сделать вывод, что

- ☐ МНК-оценки коэффициентов будут несмещённые, но не состоятельные
 ☐ МНК-оценки коэффициентов не существуют
☐ МНК-оценки коэффициентов будут несмещённые и состоятельные
 ☐ оценок коэффициентов $\hat{\beta}_1$ и $\hat{\beta}_2$ будут незначимы
☐ МНК-оценки коэффициентов будут смещённые
 ☐ нет верного ответа

Вопрос 3. МНК-оценка уравнения регрессии в отклонениях имеет вид $\hat{y}_i = 5x_i$, оценка остаточной дисперсии равна $\hat{\sigma}^2 = 2$. Вектор регрессоров имеет вид $x = (-2, 1, 1)^T$.

Оценка дисперсии ошибки прогноза индивидуального значения y_{n+1} при $x_{n+1} = 3$ равна

- ☐ 1
 ☐ 1.5
 ☐ 0.5
☐ 5
 ☐ 2.25
 ☐ нет верного ответа

Вопрос 4. Для переменных X , Z и W известны выборочные корреляции, $\widehat{\text{Corr}}(Z, W) = 0$, $\widehat{\text{Corr}}(Z, W) = 0$ и $\widehat{\text{Corr}}(X, Z) = -0.7$. Наибольшее собственное число выборочной корреляционной матрицы равно 1.7.

Первая главная компонента, выраженная через стандартизованные переменные, имеет вид

- ☐ $(x + z)/\sqrt{2}$
☐ $(x + z - \sqrt{2}w)/2$
☐ $(x - \sqrt{2}z + w)/2$
☐ $(x - z)/\sqrt{2}$
☐ $(\sqrt{2}x + z + w)/2$
☐ нет верного ответа

Вопрос 5. По данным 28 фирм была оценена зависимость выпуска Y от труда L и капитала K с помощью двух моделей: $\ln Y_i = \beta_0 + \beta_1 \ln L_i + \beta_2 \ln K_i + \varepsilon_i$ и $\ln Y_i = \beta_0 + \beta_1 \ln(L_i K_i) + \varepsilon_i$. Коэффициенты детерминации равны 0.9 и 0.8.

Значение F -статистики для проверки гипотезы о равенстве эластичностей по труду и по капиталу равно

- | | | |
|---------------------------------|-------------------------------|---|
| <input type="checkbox"/> A 25 | <input type="checkbox"/> C 20 | <input type="checkbox"/> E 0.04 |
| <input type="checkbox"/> B 12.5 | <input type="checkbox"/> D 12 | <input type="checkbox"/> F нет верного ответа |

Вопрос 6. Зависимость величины спроса Y в штуках от цены в тысячах рублей имеет вид $\ln \hat{Y}_i = 30 - 0.03P_i$. Все коэффициенты регрессии значимы. Спрос снизится на 3% при увеличении цены примерно на

- | | | |
|---|--|---|
| <input type="checkbox"/> A 100 тысяч рублей | <input type="checkbox"/> C 1 тысячу рублей | <input type="checkbox"/> E 10% |
| <input type="checkbox"/> B 10 тысяч рублей | <input type="checkbox"/> D 1% | <input type="checkbox"/> F нет верного ответа |

Вопрос 7. Исследователь интересуется зависимостью среднегодового прироста работающих E от прироста валового национального продукта X . Обе величины измеряются в процентах. Исследователь оценил три парных регрессии E на X : по выборке для 30 развитых стран, по выборке для 24 развивающихся стран и по общей выборке.

В этих регрессиях суммы квадратов остатков оказались равны 25, 35 и 120.

Значение F -статистики для проверки гипотезы о том, что изучаемая зависимость одинакова для развитых и развивающихся стран, равно

- | | | |
|---------------------------------|---------------------------------|---|
| <input type="checkbox"/> A 25 | <input type="checkbox"/> C 26 | <input type="checkbox"/> E 27 |
| <input type="checkbox"/> B 25.5 | <input type="checkbox"/> D 26.5 | <input type="checkbox"/> F нет верного ответа |

Вопрос 8. Известно, что выборочная корреляция между переменными Z и W равна 0.5. Величина VIF_X в регрессии $Y_i = \beta_1 + \beta_2 X_i + \beta_3 Z_i + \beta_4 W_i + u_i$

- | | | |
|---|---|--|
| <input type="checkbox"/> A не менее 2 | <input type="checkbox"/> C не более 2 | <input type="checkbox"/> E не может быть оценена ни сверху, ни снизу |
| <input type="checkbox"/> B не менее 4/3 | <input type="checkbox"/> D не более 4/3 | <input type="checkbox"/> F нет верного ответа |

Вопрос 9. Исследователь исключил из регрессии со свободным членом переменную, t -статистика коэффициента при которой меньше 1 по модулю. Скорректированный коэффициент множественной детерминации при этом

- | | |
|---|---|
| <input type="checkbox"/> A не увеличится | <input type="checkbox"/> D станет равным нулю |
| <input type="checkbox"/> B не уменьшится | <input type="checkbox"/> E станет равным $1/n$, где n — число наблюдений |
| <input type="checkbox"/> C может измениться в любую сторону | <input type="checkbox"/> F нет верного ответа |

Вопрос 10. С помощью метода максимального правдоподобия оценили зависимость веса индивида W_i от его роста H_i , $W_i^{(\theta)} = \beta_0 + \beta_1 H_i^{(\lambda)} + \varepsilon_i$. Здесь $W_i^{(\theta)}$ и $H_i^{(\lambda)}$ – вес и рост после преобразований Бокса-Кокса с параметрами θ и λ . Были проверены три гипотезы:

H_0	$\theta = \lambda = -1$	$\theta = \lambda = 0$	$\theta = \lambda = 1$
P -значение	0.00	0.53	0.00

На основании имеющейся информации исследователю следует предпочесть модель

☐ A $W_i = \beta_0 + \beta_1 H_i + \varepsilon_i$

☐ C $\ln W_i = \beta_0 + \beta_1 \ln H_i + \varepsilon_i$

☐ E $W_i = \beta_0 + \beta_1 H_i + \beta_2 H_i^2 + \varepsilon_i$

☐ B $\ln W_i = \beta_0 + \beta_1 H_i + \varepsilon_i$

☐ D $W_i = \beta_0 + \beta_1 \ln H_i + \varepsilon_i$

☐ F нет верного ответа

Фамилия, имя, номер группы:

.....

Задачи

1. Исследователь рассматривает уравнение зависимости расходов на питание (W) от доходов (Income), с учетом сезона. Переменная сезон (S) принимает следующие значения: 1 – зима, 2 – весна, 3 – лето и 4 – осень. Исследователь предполагает, что в каждый сезон может выполняться своя линейная зависимость.
 - а) (2 балла) Выпишите уравнение оцениваемой модели. Укажите смысл всех включенных в модель переменных.
 - б) (2 балла) Как проверить гипотезу о единой линейной зависимости расходов на питание для всех сезонов? Выпишите аккуратно основную и альтернативную гипотезы, формулу расчета статистики и способ проверки.
2. Рассмотрим модель $y_i = \beta_1 x_i + \beta_2 z_i + \varepsilon_i$ в стандартизированных переменных, оцениваемую по n наблюдениям с помощью гребневой (ridge) регрессии с параметром регуляризации λ .
 - а) (2 балла) Выпишите условия первого порядка для задачи гребневой регрессии.
 - б) (3 балла) Выведите оценки гребневой регрессии $\hat{\beta}_1$ и $\hat{\beta}_2$.
 - в) (1 балл) Что произойдет с оценками при $\lambda = 0$?
 - г) (1 балл) Что произойдет с оценками при $\lambda \rightarrow +\infty$?
3. По 24 наблюдениям была оценена модель:

$$\hat{Y}_i = 15 - 4Z_i + 3W_i$$

Известно, что случайные ошибки нормально распределены, $RSS = 180$, и

$$(X'X)^{-1} = \begin{pmatrix} 0.216 & -0.112 & -0.075 \\ -0.112 & 0.119 & 0.021 \\ -0.075 & 0.021 & 0.047 \end{pmatrix}$$

- а) (1 балл) Проверьте гипотезу $H_0 : \beta_Z = 0$ против $H_a : \beta_Z \neq 0$ на уровне значимости 5%.
- б) (3 балла) Проверьте гипотезу $H_0 : \beta_Z + \beta_W = 0$ против $H_a : \beta_Z + \beta_W \neq 0$ на уровне значимости 5%.
- в) (2 балла) Выпишите использованные при проверке гипотез предположения о случайных ошибках модели.

4. Исследовательница Глафира изучает зависимость спроса на молоко от цены молока и дохода семьи. В её распоряжении есть следующие переменные:

- *price* — цена молока в рублях за литр
- *income* — ежемесячный доход семьи в тысячах рублей
- *milk* — расходы семьи на молоко за последние семь дней в рублях

В данных указано, проживает ли семья в сельской или городской местности. Поэтому Глафира оценила три регрессии: (All) — по всем данным, (Urban) — по городским семьям, (Rural) — по сельским семьям.

	(All)	(Urban)	(Rural)
(Intercept)	−1.765 (4.943)	−4.059 (6.601)	−0.155 (7.812)
income	0.308*** (0.052)	0.341*** (0.072)	0.281*** (0.079)
price	−0.383* (0.161)	−0.352 (0.253)	−0.391 (0.221)
R-squared	0.304	0.356	0.273
adj. R-squared	0.290	0.325	0.245
sigma	4.912	4.857	5.036
F	21.216	11.593	9.744
P-value	0.000	0.000	0.000
RSS	2340.080	990.839	1318.741
n observations	100	45	55

Выборочная ковариационная матрица регрессоров по полной выборке имеет вид:

	price	income	milk
price	9.45	−1.73	−4.15
income	−1.73	90.19	28.43
milk	−4.15	28.43	33.98

- (1 балл) Проверьте значимость в целом регрессии (All) на 5%-ом уровне значимости.
- (2 балла) На 5%-ом уровне значимости проверьте гипотезу, что зависимость спроса на молоко является единой для городской и сельской местности.
- (3 балла) Разложите коэффициент детерминации R^2 в модели (All) в сумму эффектов переменных *income* и *price*.