

Часть 1. Тест.

Вопрос 1 ♣ В регрессии с константой сумма квадратов остатков равна 232, а число наблюдений равно 33. Точечная оценка дисперсии случайной составляющей равна

- | | | |
|---------------------------------------|--|--|
| <input type="checkbox"/> A 9 | <input checked="" type="checkbox"/> 8 | <input type="checkbox"/> G Нет верного ответа. |
| <input type="checkbox"/> B 10 | <input type="checkbox"/> E $\sqrt{10}$ | |
| <input type="checkbox"/> C $\sqrt{8}$ | <input type="checkbox"/> F $\sqrt{9}$ | |

Вопрос 2 ♣ Если для регрессора используется преобразование Бокса-Кокса с параметром $\theta = -1$, а для зависимой переменной — с параметром $\lambda = 1$, то регрессионное уравнение представимо в виде

- | | | |
|--|---|--|
| <input type="checkbox"/> A $\ln Y_i = \beta_1 + \beta_2 X_i + u_i$ | <input type="checkbox"/> D $\ln Y_i = \beta_1 - \beta_2 \ln X_i + u_i$ | <input type="checkbox"/> G Нет верного ответа. |
| <input type="checkbox"/> B $Y_i = \beta_1 + \beta_2 \ln X_i + u_i$ | <input type="checkbox"/> E $\ln Y_i = \beta_1 + \beta_2 \ln X_i + u_i$ | |
| <input type="checkbox"/> C $Y_i = \beta_1 + \beta_2 X_i + u_i$ | <input checked="" type="checkbox"/> $Y_i = \beta_1 + \beta_2 \frac{1}{X_i} + u_i$ | |

Вопрос 3 ♣ Известно, что регрессоры X и Z ортогональны, а истинная зависимость описывается уравнением $Y_i = \alpha_1 + \alpha_2 X_i + \alpha_3 Z_i + u_i$. Исследователь оценивает с помощью МНК две регрессии: $\hat{Y}_i = \hat{\beta}_1 + \hat{\beta}_2 X_i$ и $\hat{Y}_i = \hat{\gamma}_1 + \hat{\gamma}_2 Z_i$. При этом

- | |
|--|
| <input type="checkbox"/> A $\hat{\beta}_2$ — смещённая оценка для α_2 ; $\hat{\gamma}_2$ — смещённая оценка для α_3 |
| <input type="checkbox"/> B $\hat{\beta}_2$ — несмещённая оценка для α_2 ; $\hat{\gamma}_2$ — смещённая оценка для α_3 |
| <input type="checkbox"/> C $\hat{\beta}_2$ — смещённая оценка для α_2 ; $\hat{\gamma}_2$ — несмещённая оценка для α_3 |
| <input checked="" type="checkbox"/> $\hat{\beta}_2$ — несмещённая оценка для α_2 ; $\hat{\gamma}_2$ — несмещённая оценка для α_3 |
| <input type="checkbox"/> E $\hat{\beta}_2$ — эффективная оценка для α_2 ; $\hat{\gamma}_2$ — эффективная оценка для α_3 |
| <input type="checkbox"/> F Нет верного ответа. |

Вопрос 4 ♣ Гипотеза о том, что одновременно $\beta_1 + \beta_2 = 1$ и $\beta_3 = 0$ в множественной линейной регрессии построенной по n наблюдениям проверяется с помощью статистики, имеющей распределение

- | | | |
|--------------------------------------|---|--|
| <input type="checkbox"/> A t_{n-2} | <input type="checkbox"/> D Демешева-Мамонтова | <input type="checkbox"/> G Нет верного ответа. |
| <input type="checkbox"/> B t_{n-k} | <input checked="" type="checkbox"/> F | |
| <input type="checkbox"/> C t_n | <input type="checkbox"/> F $N(0; 1)$ | |

Вопрос 5 ♣ Элеонора исследует зависимость цены номера в отеле от звёздности отеля, $star$, (от 1 до 3 звёзд) и расстояния до моря, $dist$. Элеонора хочет оценить модель вида $price_i = \beta_1 + \beta_2 star_i + \beta_3 dist_i + u_i$. Чтобы считаться богиней эконометрики Элеоноре стоит

- ☐ A использовать МНК для оценки данной модели
- ☐ B добавить в модель переменную $z_i = star_i^2$, так как эффект звёздности наверняка нелинейный
- ☐ C добавить в модель переменную $z_i = star_i \cdot dist_i$
- ☐ D добавить дамми-переменные one_i, two_i и $three_i$, равные 1 для отелей с одной, двумя и тремя звёздами соответственно
- ☒ E заменить переменную $star_i$ на дамми-переменные one_i и two_i , равные 1 для отелей с одной и двумя звёздами соответственно
- ☐ F заменить переменную $star_i$ на дамми-переменные one_i, two_i и $three_i$, равные 1 для отелей с одной, двумя и тремя звёздами соответственно
- ☐ G Нет верного ответа.

Вопрос 6 ♣ Показатель R_{adj}^2 можно вычислить по формуле

- ☒ A $R_{adj}^2 = (-1) \cdot \frac{k-1}{n-k} + R^2 \cdot \frac{n-1}{n-k}$
- ☐ B $R_{adj}^2 = \frac{k-1}{n-k} + R^2 \cdot \frac{n-1}{n-k}$
- ☐ C $R_{adj}^2 = \frac{k-1}{n-k} - R^2 \cdot \frac{n-1}{n-k}$
- ☐ D $R_{adj}^2 = \frac{k}{n-k} + R^2 \cdot \frac{n-1}{n-k}$
- ☐ E $R_{adj}^2 = \frac{n-k}{k-1} + R^2 \cdot \frac{n-1}{n-k}$
- ☐ F $R_{adj}^2 = \frac{k-1}{n-k} + R^2 \cdot \frac{n-k}{n-1}$
- ☐ G Нет верного ответа.

Вопрос 7 ♣ Если гипотеза $\beta_2 + \beta_3 = 1$ верна, то модель $\ln Y_i = \beta_1 + \beta_2 \ln X_i + \beta_3 \ln Z_i + u_i$ совпадает с моделью

- ☐ A $\ln Y_i = \beta_1 + \beta_2 \ln(X_i/Z_i) + u_i$
- ☐ B $\ln(Y_i/Z_i) = \beta_1 + \beta_2 \ln(Y_i/Z_i) + u_i$
- ☐ C $\ln(Y_i/Z_i) = \beta_1 + \beta_2 \ln(Y_i/X_i) + u_i$
- ☒ D $\ln(Y_i/Z_i) = \beta_1 + \beta_2 \ln(X_i/Z_i) + u_i$
- ☐ E $\ln Y_i = \beta_1 + \beta_2 \ln(Z_i/Y_i) + u_i$
- ☐ F Нет верного ответа.

Вопрос 8 ♣ Гипотеза о неадекватности множественной регрессии проверяется с помощью статистики равной

- ☐ A $\frac{\hat{\beta} - \beta}{se(\hat{\beta})}$
- ☐ B $\frac{ESS}{TSS}$
- ☒ C $\frac{ESS/(k-1)}{RSS/(n-k)}$
- ☐ D $\frac{TSS/(n-1)}{RSS/(n-k)}$
- ☐ E $\frac{RSS}{TSS}$
- ☐ F $\frac{TSS/(n-1)}{ESS/(k-1)}$
- ☐ G Нет верного ответа.

Вопрос 9 ♣ Исследователь выполнил второй шаг в РЕ-тесте МакКиннона. В регрессии $\ln Y_i$ на исходные регрессоры и $Z_i = \hat{Y}_i - \exp(\ln \hat{Y}_i)$ коэффициент при Z_i оказался значимым. А в регрессии Y_i на исходные регрессоры и $W_i = \ln \hat{Y}_i - \ln Y_i$ коэффициент при W_i оказался незначимым. Из результатов следует сделать вывод, что

- | | | |
|--|---|---|
| <input type="checkbox"/> А тесты противоречат друг другу, ни одна из моделей не предпочитается | <input type="checkbox"/> С следует предпочесть полулогарифмическую модель | <input checked="" type="checkbox"/> следует предпочесть линейную модель |
| <input type="checkbox"/> В в исходной модели пропущен регрессор W_i | <input type="checkbox"/> D в исходной модели пропущен регрессор Z_i | <input type="checkbox"/> F следует предпочесть логарифмическую модель |
| | | <input type="checkbox"/> G Нет верного ответа. |

Вопрос 10 ♣ Истинной является модель $Y_i = \beta_1 + \beta_2 X_i + u_i$. Глафира оценивает две регрессии: $\hat{Y}_i = \hat{\beta}_1 + \hat{\beta}_2 X_i$ и $\hat{Y}_i = \hat{\gamma}_1 + \hat{\gamma}_2 X_i + \hat{\gamma}_3 Z_i$ с помощью МНК. Для коэффициента β_2

- | | |
|---|---|
| <input type="checkbox"/> А оценка $\hat{\beta}_2$ является несмещённой, а оценка $\hat{\gamma}_2$ — смещённой | <input type="checkbox"/> D оценка $\hat{\beta}_2$ является смещённой, а оценка $\hat{\gamma}_2$ — несмещённой |
| <input checked="" type="checkbox"/> оценки $\hat{\beta}_2$ и $\hat{\gamma}_2$ являются несмещёнными | <input type="checkbox"/> E оценки $\hat{\beta}_2$ и $\hat{\gamma}_2$ являются неэффективными |
| <input type="checkbox"/> C оценки $\hat{\beta}_2$ и $\hat{\gamma}_2$ являются эффективными | <input type="checkbox"/> F Нет верного ответа. |

Часть 2. Задачи.

1. На основании опроса 200 человек была оценена следующая модель:

$$\ln(wage_i) = \beta_1 + \beta_2 exper_i + \beta_3 exper_i^2 + \beta_4 sex_i + \varepsilon_i$$

где:

- $wage_i$ — величина заработной платы в долларах
- $exper_i$ — опыт работы в годах
- $exper_i^2$ — опыт работы в годах
- sex_i — пол (1 — мужской, 0 — женский)

Показатель	Значение
R^2	0.911
Скорректированный R^2	B7
Стандартная ошибка регрессии	B6
Количество наблюдений	B2

Результаты дисперсионного анализа:

	df	сумма квадратов	F	P-значение
Регрессия	3	B9	B5	0.000
Остаток	B1	830.1		
Итого	B3	B4		

	Оценка	Ст. ошибка	t-статистика	P-Значение
Константа	3.6869	1.1960	3.08	0.0023
$exper$	B8	0.3525	16.45	0.0000
$exper^2$	-0.1916	0.0254	-7.54	0.0000
sex	1.5745	0.2937	B10	0.0000

а) Найдите пропущенные числа **B1–B10**.

б) Как изменятся результаты оценки регрессии, если переменную sex_i переопределить так, чтобы 0 соответствовал мужчинам, 1 — женщинам?

Ответ округляйте до 2-х знаков после запятой. Кратко поясняйте, например, формулой, как были получены результаты.

2. Исследовательница Глафира изучает спрос на молоко. 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)	6.255 (4.448)	6.020 (5.151)	11.907 (7.384)
income	0.180*** (0.045)	0.136* (0.054)	0.225** (0.070)
price	-0.331* (0.139)	-0.117 (0.153)	-0.832** (0.251)
R-squared	0.2	0.1	0.4
adj. R-squared	0.2	0.1	0.3
sigma	4.5	4.2	4.5
F	11.9	3.8	11.8
P-value	0.0	0.0	0.0
RSS	1995.1	961.9	773.8
n observations	100	58	42

- Проверьте значимость в целом регрессии (All) на 5%-ом уровне значимости.
- На 5%-ом уровне значимости проверьте гипотезу, что зависимость спроса на молоко является единой для городской и сельской местности.

3. Исследовательница Глафира продолжает изучать спрос на молоко. В её распоряжении по-прежнему данные по трём переменным:

- *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

Имеются результаты оценивания модели $milk_i = \beta_1 + \beta_2 income_i + \beta_3 price_i + u_i$ по 100 наблюдениям:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	6.2554	4.4480	1.41	0.1628
income	0.1800	0.0448	4.01	0.0001
price	-0.3312	0.1389	-2.38	0.0191

Коэффициент детерминации R^2 оказался равен 0.197.

Глафира рассчитала оценку ковариационной матрицы исходных переменных:

	price	income	milk
price	10.86	-3.21	-4.18
income	-3.21	104.25	19.83
milk	-4.18	19.83	25.11

- Постройте точечный прогноз расходов на молоко семьи с доходом 100 тысяч рублей при цене на молоко 30 рублей за литр.
- Найдите выборочную корреляцию между фактическими расходами на молоко и их прогнозами.
- Разложите коэффициент детерминации R^2 в модели в сумму эффектов переменных *income* и *price*.

4. По квартальным данным 1958-1976 годов была оценена модель с тремя объясняющими факторами:

$$\hat{Y}_i = 2.2 + 0.104X_i - 3.48Z_i + 0.34W_i, ESS = 100, RSS = 20$$

- а) Какую модель необходимо оценить исследователю, если он считает, что в различные сезоны среднее значение зависимой переменной помимо зависимости от трёх регрессоров может отличаться на константу?
- б) При оценивании модели, допускающей сезонные эффекты, оказалось, что значение ESS увеличилось до 170. На уровне значимости 5% проверьте гипотезу о наличии сезонности.

5. По 24 наблюдениям была оценена модель:

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

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

$$(X'X)^{-1} = \begin{pmatrix} 0.223 & -0.132 & -0.065 \\ -0.132 & 0.144 & 0.014 \\ -0.065 & 0.014 & 0.046 \end{pmatrix}$$

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