

Часть 1. Тест.

Вопрос 1 ♣ Предпосылка об отсутствии систематической ошибки в модели означает, что для всех наблюдений

☐ A $\text{Var}(\varepsilon_i) \neq 0$

☐ C Нет верного ответа.

☐ B $\text{Var}(\varepsilon_i) = 0$

Вопрос 2 ♣ После применения МНК к модели $y_i = \beta x_i + \varepsilon_i$ сумма $ESS + RSS$

☐ A обязательно равна TSS

☐ B может быть не равна TSS

Вопрос 3 ♣ При наличии ошибок измерения зависимой переменной МНК-оценки коэффициентов модели

☐ A состоятельны

☐ B несостоятельны

Вопрос 4 ♣ Индексы вздутия дисперсии (VIF) в случае отсутствия мультиколлинеарности лежат в интервале

☐ A $[0; 1]$

☐ B $[1; +\infty)$

Вопрос 5 ♣ Незначимость всех коэффициентов регрессии

☐ A может быть не связана с мультиколлинеарностью

☐ B обязательно свидетельствует о наличии мультиколлинеарности

Вопрос 6 ♣ Если выполнены все предпосылки теоремы Гаусса-Маркова, но остатки модели не подчиняются нормальному закону распределения, то МНК-оценки коэффициентов регрессии являются

☐ A несмещёнными

☐ B смещёнными

Вопрос 7 ♣ Если в модель добавили незначимый фактор, то коэффициент детерминации R^2

☐ A не изменится

☐ B упадёт

☐ C вырастет

Вопрос 8 ♣ Нулевая гипотеза в тесте Дарбина-Уотсона состоит в

☐ A отсутствии автокорреляции

☐ B наличии автокорреляции

Вопрос 9 ♣ При диагностике автокорреляции третьего порядка тест Бройша-Годфри

☐ A применим

☐ B неприменим

Вопрос 10 ♣ Стандартные ошибки в форме Уайта в случае гетероскедастичности помогают устранить несостоятельность оценок коэффициентов

☐ A верно

☐ B неверно

Часть 2. Задачи.

1. На основании опроса была оценена следующая модель:

$$\ln(wage_i) = \beta_1 + \beta_2 exper_i + \beta_3 exper_i^2 + \beta_4 married_i + \beta_5 educ_i + \beta_6 black_i + \varepsilon_i$$

где:

- $wage_i$ — величина заработной платы в долларах
- $exper_i$ — опыт работы в годах
- $educ_i$ — количество лет обучения
- $married_i$ — наличие супруга/супруги (1 — есть, 0 — нет)
- $black_i$ — принадлежность к негроидной расе (1 — да, 0 — нет)

Показатель	Значение
R^2	B7
Скорректированный R^2	0.219
Стандартная ошибка регрессии	B6
Количество наблюдений	B2

Результаты дисперсионного анализа:

	df	SS	MS	F	Р-значение
Регрессия	B1	5.993	1.199	B5	0.000
Остаток	134	18.240	0.136		
Итого	B3	B4			

Коэффициент	Оценка	$se(\hat{\beta})$	t-статистика	Р-Значение	Нижние 95%	Верхние 95%
Константа	4.529	0.331	13.688	0.000	3.874	5.183
$exper$	0.090	0.037	2.419	0.017	0.016	0.164
$exper^2$	-0.003	0.002	-1.790	0.076	-0.006	0.000
$married$	0.240	0.079	3.045	0.003	B8	B9
$educ$	0.078	0.017	B10	0.000	0.045	0.111
$black$	0.073	0.171	0.424	0.672	-0.266	0.411

Найдите пропущенные числа **B1–B10**.

Ответ округляйте до 3-х знаков после запятой. Кратко поясняйте, например, формулой, как были получены результаты.

2. На основании данных по ценам на квартиры в Москве были построена модель

$$\ln(price_i) = \beta_1 + \beta_2 totsp_i + \beta_3 metrdist_i + \beta_4 dist_i + \beta_5 floor_i + \varepsilon_i,$$

где:

- $\ln(price_i)$ — логарифм цены квартиры в тысячах долларов
- $totsp_i$ — общая площадь квартиры в кв. м.
- $metrdist_i$ — расстояние до метро в минутах
- $dist_i$ — расстояние до центра города в км
- $floor_i$ — дамми-переменная (1 — если квартира не на первом и последнем этажах, 0 — иначе)

Модели были оценены на пяти разных выборках, результаты представлены в таблице:

Коэффициент	Выборка А	Выборка В	Выборка С	Выборка D	Выборка Е
Константа	3.980***	3.926***	3.929***	3.719***	4.224***
$totsp$	0.0155***	0.0148***	0.0163***	0.0179***	0.0139***
$metrdist$	-0.00858***	-0.0169***	-0.00566**	-0.0108***	-0.0077
$dist$	-0.0267***	-0.0186***	-0.0253***	-0.0150***	-0.0350***
$floor$	0.0419**	0.0633*	0.0224	0.0225	0.0228
Наблюдений	460	145	315	150	150
R^2	0.693	0.684	0.723	0.328	0.520
RSS	15.120	4.503	9.408	2.163	8.545

* — значимость на 10%, ** — значимость на 5%, *** — значимость на 1%.

- Для всей выборки (выборка А) проинтерпретируйте коэффициент при переменной $dist_i$.
- Определите на 5%-ом уровне значимости, можно ли использовать одну модель для квартир, находящихся в пешей доступности от метро (выборка С), и квартир, находящихся в транспортной доступности (выборка В).
- Исследователь предположил, что дисперсия ошибок модели возрастает с увеличением площади квартиры. Проверьте, есть ли в модели гетероскедастичность на 10% уровне значимости на основании соответствующего теста. В выборку D включены 150 квартир с наименьшей общей площадью, в выборку Е — 150 квартир с наибольшей общей площадью.

При проверке гипотез: выпишите H_0 , H_a , найдите значение тестовой статистики, укажите её распределение, найдите критическое значение, сделайте выводы

3. Эконометресса Эвридика хочет оценить модель $y_i = \beta_1 + \beta_2 x_i + \beta_3 z_i + \varepsilon_i$. К сожалению, она измеряет зависимую переменную с ошибкой. Т.е. вместо y_i она знает значение $y_i^* = y_i + u_i$ и использует его в качестве зависимой переменной при оценке регрессии. Ошибки измерения u_i некоррелированы между собой и с ε_i .
- а) Будут ли оценки Эвридики несмещенными?
 - б) Могут ли дисперсии оценок Эвридики быть ниже чем дисперсии МНК оценок при использовании настоящего y_i ?
 - в) Могут ли оценки дисперсий оценок Эвридики быть ниже оценок дисперсий МНК оценок при использовании настоящего y_i ?

Аргументируйте ответы.

4. По наблюдениям $x = (1, 2, 3)'$, $y = (2, -1, 3)'$ оценивается модель $y = x + \varepsilon$. Ошибки ε_i гетероскедастичны и известно, что $\text{Var}(\varepsilon_i) = \sigma^2 \cdot x_i^2$.
- а) Найдите оценку $\hat{\beta}_{ols}$ с помощью МНК и её дисперсию
 - б) Найдите оценку $\hat{\beta}_{gls}$ с помощью обобщенного МНК и её дисперсию

Часть 3. Теоретические вопросы

5. Дана модель $y_t = \beta_1 + \beta_2 x_t + \varepsilon_t$, в которой ошибки модели подчиняются авторегрессионной схеме первого порядка, $\varepsilon_t = \rho \varepsilon_{t-1} + u_t$, где $u_t \sim WN(0, \sigma^2)$ и ρ известно. Здесь WN означает белый шум. Опишите процедуру получения эффективных оценок коэффициентов для такой модели.
6. В модели $y_i = \beta_1 + \beta_2 x_i + \varepsilon_i$ переменная x_i эндогенна. Для нее был найден инструмент z_i . Опишите процедуру получения состоятельных оценок коэффициентов регрессии.
7. Дайте определение гомоскедастичности и гетероскедастичности. Укажите последствия гетероскедастичности. Какие действия нужно сделать для получения корректных доверительных интервалов для коэффициентов при гетероскедастичности неизвестной формы?