

Часть 1. Тест.

Вопрос 1 ♣ Мультиколлинеарность приводит к смещению оценок коэффициентов регрессии.

☐ A верно

☐ не верно

Вопрос 2 ♣ Для устранения мультиколлинеарности применяется обобщенный метод наименьших квадратов.

☐ A верно

☐ не верно

Вопрос 3 ♣ Признаком мультиколлинеарности является значимость модели в целом при незначимости отдельных коэффициентов.

☐ верно

☐ B не верно

Вопрос 4 ♣ Тест Дарбина-Уотсона применим только в случае автокорреляции первого порядка.

☐ верно

☐ B не верно

Вопрос 5 ♣ Мультиколлинеарность приводит к высокой дисперсии оценок коэффициентов.

☐ верно

☐ B не верно

Вопрос 6 ♣ В случае гетероскедастичности применение стандартных ошибок в форме Уайта помогает сделать оценки коэффициентов эффективными.

☐ A верно

☐ не верно

Вопрос 7 ♣ Если регрессор коррелирован с ошибкой модели, то оценки коэффициентов становятся несостоятельными.

☐ A верно

☐ не верно

Вопрос 8 ♣ Нулевая гипотеза в тесте Дарбина-Уотсона – наличие автокорреляции.

☐ A верно

☐ не верно

Вопрос 9 ♣ В случае автокорреляции оценки дисперсий коэффициентов оказываются смещенными.

☐ A верно

☐ не верно

Вопрос 10 ♣ Мультиколлинеарность приводит к смещению оценок дисперсий коэффициентов регрессии.

☐ верно

☐ не верно

Часть 2. Задачи.

1. Ниже представлена оценка следующей модели для выборки, состоящей из женщин 18–60 лет:

$$\ln(wage_i) = \beta_1 + \beta_2 educ_i + \beta_3 marst_i + \beta_4 age_i + \beta_5 child_i + \epsilon_i,$$

где $wage_i$ – величина заработной платы, руб.

$educ_i$ – бинарная переменная (1 – в случае наличия высшего образования, 0 – иначе);

$marst_i$ – бинарная переменная (1 - в случае наличия постоянного партнера, 0 - иначе);

age_i – возраст женщины, лет;

$child_i$ – количество детей младше 18 лет, чел.

Регрессионная статистика	
Множественный R	0.318
R-квадрат	0.101
Нормированный R-квадрат	B10
Стандартная ошибка	B7
Наблюдения	B4

	df	SS	MS	F	Значимость F
Регрессия	B5	23.073	5.768	B9	0.000
Остаток	B6	204.907	0.414		
Итого	499	B8			

	Коэф-ты	Стандартная ошибка	t- статистика	P- значение	Нижние 95%	Верхние 95%
Y-пересечение	9.296	0.205	45.333	0.000	8.893	9.699
educ	B1	B2	B3	0.000	0.283	0.524
marst	-0.227	0.065	-3.498	0.001	-0.355	-0.100
age	0.003	0.005	0.760	0.448	-0.005	0.012
nchild18	-0.017	0.037	-0.470	0.638	-0.089	0.055

Вставьте пропущенные числа на месте пропусков (округляйте ответ до 3 знака после запятой):

B1=

B5=

B9=

B2=

B6=

B10=

B3=

B7=

B4=

B8=

2. Для 500 квартир Москвы была оценена зависимость их цены ($price$, тыс. \$) от общей площади ($totsp$, кв. м), расстояния до метро ($metrdist$, мин), возможности добраться до метро пешком ($walk$, 1 — если такая возможность есть, 0 — иначе):

$$\widehat{price}_i = -28.87 + 2.26totsp_i - 2.57metrdist_i - 33.94walk_i + 1.70walk_i * metrdisst_i + 0.51walk_i * totsp_i$$

$$R^2 = 0.608$$

- Кроме того, была получена следующая модель:

$$\widehat{price}_i = -49.30 + 2.56totsp_i - 1.37metrdist_i$$

$$R^2 = 0.576$$

- Проверьте гипотезу об адекватности регрессии для модели из пункта а) (выпишите H_0 , H_1 , вычислите расчетную статистику, укажите её распределение, найдите критическую статистику, сделайте выводы)
 - Проверьте, можно ли использовать объединенную модель для квартир, находящихся в пешей доступности от метро, и остальных квартир.
 - Проинтерпретируйте для модели а) коэффициент при переменной $walk_i * metrdisst_i$ (учитывая, что все коэффициенты модели значимы на 10% уровне значимости).
3. На основании опроса 100 студентов ВШЭ была оценена зависимость выпитого кофе в день ($coffee_i$, чашек) от прорешанных задач по эконометрике ($metrics_i$):

$$\widehat{coffee}_i = 1 + 0.1metrics_i.$$

Оценка ковариационной матрицы коэффициентов выглядит следующим образом:

$$Var(\hat{\beta}) = \begin{pmatrix} 19 & 0.1 \\ 0.1 & 1 \end{pmatrix}$$

- Сколько чашек кофе выпьет студент Петя, если решит 10 задач по эконометрике?
 - Постройте 95%-ый доверительный интервал для $E(coffee_i | metrics_i = 10)$ ожидаемой величины выпитого кофе при 10 прорешанных задачах.
4. По ежемесячным данным с января 2003 по февраль 2014 года была оценена зависимость сбережений домашних хозяйств РФ ($save_t$) от располагаемого дохода (inc_t) и индекса потребительских цен (ipc_t):

$$\widehat{save}_i = 211.75 + 0.067inc_t - 10.19ipc_t, R^2 = 0.54$$

Кроме того, были оценены следующие модели:

$$\hat{e}_t = -1.64 + 0.000069inc_t + 0.11ipc_t + 0.12e_{t-1} + 0.12e_{t-2}, R^2 = 0.03$$

$$\hat{e}_t^2 = 95030.38 - 38.36inc_t + 0.0035inc_t^2 - 11810.24ipc_t + 364.98ipc_t^2 + 2.75ipc_tinc_t, R^2 = 0.068$$

- а) Проведите соответствующий тест на гетероскедастичность (укажите H_o , H_a , вычислите расчетную статистику, укажите её распределение, найдите критическую статистику и сделайте выводы).
- б) Проведите соответствующий тест на автокорреляцию (укажите H_o , H_a , вычислите расчетную статистику, укажите её распределение, найдите критическую статистику и сделайте выводы).

Часть 3. Теоретические вопросы.

1. Опишите тест Дарбина-Уотсона (сформулируйте нулевую и альтернативную гипотезы, условия применимости, методику расчета статистики и её распределение).
2. Сформулируйте теорему Гаусса-Маркова.
3. Опишите тест Чоу (сформулируйте нулевую и альтернативную гипотезы, методику расчета статистики и её распределение).