

Проверочная работа №1

Имя и фамилия студента:

Задание 1. Ниже приведены аргументы, использующиеся некоторыми авторами для обоснования оценивания «усеченной» модели с переменными взаимодействия, то есть, такой модели, которая включает не все составные части переменной взаимодействия в качестве предикторов.

«First, some claim that they do not believe that Z has any effect on Y on average and that, as a result, they do not need to include it as a separate term in the model. Second, others claim that they do not believe that Z has an effect on Y when X is zero and this means they can exclude it as a separate variable from their model.»

Объясните, почему данные аргументы являются некорректными или слабыми. (2 балла)

Задание 2. Ниже представлены результаты анализа разложения вариации по результатам оценивания линейной регрессионной модели:

ANOVA

	sum_sq	mean_sq	df	f	PR(>F)
x	2.5771				
Residual	11.5366		490		
Total			495		

Используя информацию из данной таблицы,

1. рассчитайте коэффициент детерминации и проинтерпретируйте его (1.5 балла)

2. проверьте гипотезу о том, что регрессия на константу (то есть, модель без объясняющих переменных) не хуже модели с предикторами, на фиксированном уровне значимости 0.05. Запишите нулевую и альтернативную гипотезы на статистическом языке, рассчитайте значение статистики, а также выберите необходимую критическую точку – квантиль – из списка ниже. Сделайте вывод (1.5 балла)

- (a) квантиль распределения Фишера, 0.95, $df1 = 490$; $df2 = 5$: 4.373
- (b) квантиль распределения Фишера, 0.975, $df1 = 490$; $df2 = 5$: 6.029
- (c) квантиль распределения Фишера, 0.95, $df1 = 6$; $df2 = 490$: 2.117
- (d) квантиль распределения Фишера, 0.975, $df1 = 6$; $df2 = 490$: 2.434
- (e) квантиль распределения Фишера, 0.95, $df1 = 5$; $df2 = 490$: 2.232
- (f) квантиль распределения Фишера, 0.975, $df1 = 5$; $df2 = 490$: 2.592

Задание 3. Начинаящий исследователь пытается проверить, страдает ли его линейная регрессионная модель, оцененная посредством метода наименьших квадратов, от эндогенности. Для этого он сохранил остатки и рассчитал корреляцию между остатками и X (предиктором) в модели. Чему будет равна данная корреляция и почему? Критически оцените действия начинающего исследователя, свой ответ математически обоснуйте. (2 балла)

Задание 4. Ниже представлены данные об индексе протестной готовности и проценте безработного населения в 5 регионах. Большее значение отклика соответствует более высокой протестной активности.

y — индекс протестной готовности	13	7	4	8	8
x — процент безработного населения	3	2	4	6	10

1. Выполнив необходимые предварительные расчеты при помощи общей векторно-матричной формулы получения оценок коэффициентов, запишите спецификацию модели, подставив полученные оценки коэффициентов. (2 балла)
2. Проинтерпретируйте полученную оценку коэффициента при предикторе. (1 балл)
3. Если переоценить модель с использованием вместо исходного предиктора переменную – результат деления x на 100, каким образом изменится оценка коэффициента при предикторе в модели? Свой ответ поясните. (1 балл)

Задание 5. Ответьте на нижеприведенные вопросы по следующей спецификации регрессионной модели: $y_i = 8 - 1.8x_i + 0.1x_i^2 + \epsilon_i$

- Отметьте верное утверждение: (1 балл)
 1. взаимосвязь отклика и x_i задается графически прямой с положительным угловым коэффициентом
 2. взаимосвязь отклика и x_i задается графически прямой с отрицательным угловым коэффициентом
 3. взаимосвязь отклика и x_i задается графиком U-формы
 4. взаимосвязь отклика и x_i задается графиком перевернутой U-формы
- Проинтерпретируйте оценку коэффициента при x_i^2 (1 балл)
- Рассчитайте «переломную» точку (значение x_i), после которой эффект x_i на y_i меняет знак. (1 балл)

Задание 6. Изучается эффект введения новых социальных выплат на доверие политическим институтам. Рассматривается пять лет до и после реализации специальной социальной программы. Те регионы, в которых социальная программа не была введена, рассматривается как контрольная группа. Ниже представлены результаты оценивания средних значений доверия политическим институтам (шкала уровня доверия – 5-балльная, более высокое значение соответствует более высокому уровню доверия).

Группа/Период	До введения соц. выплат	После введения соц. выплат
Регионы: Группа воздействия	3.6	4.2
Регионы: Контрольная группа	3.4	3.3

1. Вычислите все оценки модели difference-in-differences без включения контрольных переменных и запишите соответствующую спецификацию модели (2 балла)
2. Проинтерпретируйте оценку коэффициента при дамми-переменной для периода (1 балл)

Задание 7. Выберите ВСЕ верные утверждения из списка ниже. Если верные утверждения отсутствуют, обязательно в ответе напишите "НЕТ" (1 балл)

1. Коэффициент детерминации это квадрат косинуса угла между векторами y и \hat{y}
2. Параллельность трендов в рамках DiD предполагает отсутствие статистически значимых различий между средними значениями рассматриваемой зависимой переменной в группе воздействия и контрольной группе в период, предшествующий воздействию
3. Согласно теореме Фриша-Во-Ловелла, МНК-оценка коэффициента при предикторе x_i в множественной линейной регрессии идентична МНК-оценке коэффициента наклона в линейной регрессии очищенного от влияния третьих факторов отклика (y_i) на исходный x_i
4. Оценена следующая регрессионная модель: $\hat{y}_i = 10 - 4.7x_i + 0.05D_i + 0.3x_iD_i$, где x_i – непрерывная переменная в шкале от 0 до 10, D_i – дамми-переменная, то есть, принимающая значение 1 или 0. Справедливо следующее: если оценить регрессию y_i на x_i на подвыборке, где $D_i = 1$, то оценка коэффициента при x_i будет равна -4.4

Задание 8. Дана следующая спецификация модели и соответствующая ковариационная матрица для оценок коэффициентов:

$$\hat{y}_i = -13.788 - 8.122x_i + 3.595z_i + 1.149x_iz_i$$

	Intercept	X	Z	X:Z
Intercept	25.6279	3.1163	-2.9567	-0.0114
X	3.1163	18.0029	-0.1440	-2.1206
Z	-2.9567	-0.1440	0.3505	-0.0256
X:Z	-0.0114	-2.1206	-0.0256	0.2629

Рассчитайте предельный эффект z_i на отклик при условии, что $x_i = 2$, и проверьте значимость этого предельного эффекта (2 балла)