

ДОМАШНЕЕ ЗАДАНИЕ 1

Deadline: 23.59 14 октября 2022

Задание выполняется на базе данных [RAPDC_hw1.dta](#).
Описание переменных представлено ниже:

county	номер округа штата Северная Каролина
year	год
lncrime	натуральный логарим числа преступлений на человека
lnpolice	натуральный логарифм числа полицейских на душу населения
lndensity	натуральный логарифм плотности населения
nonwhite	небелое население – натуральный логарифм

Задание 1

Оцените регрессионную модель (без учета панельной структуры данных), в которой откликом выступает логарифм числа преступлений на человека, предикторами – логарифм числа полицейских на душу населения, логарифм плотности населения, логарифм показателя небелого населения. При необходимости или для удобства интерпретации Вы можете преобразовать исходные переменные.

1.1 Проинтерпретируйте полученные результаты. В частности, сделайте вывод о характере взаимосвязи между зависимой переменной и объясняющими переменными.

1.2 К чему может привести оценивание модели объединенной регрессии (pooled model) на массиве панельных данных?

Задание 2

Оцените модель посредством МНК с фиктивными переменными (LSDV-модель с набором дамми-переменных на константы). Сравните полученные результаты с результатами модели, построенной без учета панельной структуры данных. Проинтерпретируйте оценки коэффициентов при двух любых статистически-значимых дамми-переменных. Каковы недостатки указанной спецификации модели?

Задание 3

Оцените модель, используя внутригрупповое преобразование (within-group transformation). Коэффициент при каком предикторе не позволяет оценить модель с внутригрупповым преобразованием? Объясните, почему. Можно ли сказать, что это недостаток модели с фиксированными эффектами? Если нет, объясните, почему.

Задание 4

Проверьте гипотезу о том, что все индивидуальные эффекты равны нулю. Сделайте вывод о том, какая модель более адекватна данным: модель с фиксированными эффектами или модель без учета панельной структуры.

Задание 5

Оцените модель со случайными эффектами.

5.1 Проверьте гипотезу о том, что дисперсия случайного индивидуального эффекта равна 0. В пользу какой модели Вы сделаете выбор: модели со случайными эффектами или модели без учета панельной структуры данных?

5.2 Сравните результаты с результатами модели с фиксированными эффектами. Какое допущение главным образом отличает RE-model от FE-model? Поразмышляйте, правдоподобно ли это допущение применительно к нашим данным (поясните, почему, проиллюстрируйте конкретным примером)?

5.3 Протестируйте посредством теста Хаусмана отсутствие корреляции между индивидуальными эффектами и предикторами. Сделайте вывод. Назовите ограничения теста Хаусмана.

Задание 6

Протестируйте альтернативные спецификации модели: с включением временных эффектов и *twoway model* (с включением и пространственных, и временных эффектов). Объясните различия в интерпретации оценок коэффициентов при предикторах в разных моделях, сравните с соответствующей интерпретацией в модели с FE на округа. Какая модель, на Ваш взгляд, более адекватна в данном случае? Свой ответ аргументируйте. Выбранную модель переоцените с поправкой на гетероскедастичность.

Задание 7

Протестируйте, устойчива ли выбранная модель. Переоцените модель на усеченной выборке: без единиц анализа, в которых корреляция предсказанного отклика и наблюдаемого мала. Изменились ли значительно результаты?

Задание 8 (на 9-10 баллов)

8.1 Оцените регрессионную модель с фиксированными эффектами на округа, в которой откликом является по-прежнему логарифм числа преступлений на человека, а предиктор только один – логарифм числа полицейских на душу населения. Покажите, как получить оценку коэффициента при данном предикторе на основе оценок коэффициентов регрессионных моделей, оцененных на отдельных подгруппах (округа формируют подгруппы). В каком случае подгруппа (в данном случае – округ) не учитывается (то есть, не имеет веса) при расчете оценки коэффициента модели с фиксированными эффектами?

8.2 Алгоритм реализации процедуры взвешивания в случае множественной регрессии следующий:

1. Нужно очистить y_{it} от эффекта контрольных переменных z_{it} . Для этого нужно сохранить остатки регрессии y_{it} на z_{it} .
2. По такому же принципу очищаем x_{it} от эффекта z_{it}
3. Далее повторяем уже знакомую процедуру взвешивания, однако вместо y_{it} и x_{it} используем сохраненные остатки (очищенный эффект y_{it} и x_{it})

Добавьте в модель контрольную переменную «натуральный логарифм плотности населения». Покажите, как в случае множественной регрессии получить оценку коэффициента при предикторе «логарифм числа полицейских на душу населения» на основе оценок коэффициентов регрессионных моделей, оцененных на отдельных подгруппах (округа формируют подгруппы). Как формируются веса для коэффициентов по отдельным подгруппам (округам) в случае, если мы оцениваем множественную регрессию (включаем больше одного предиктора)? Критически прокомментируйте такую процедуру взвешивания.