

ДОМАШНЕЕ ЗАДАНИЕ 1

Deadline: 23.59 8 октября 2020

Задание выполняется на базе данных MSA_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	процент небелого населения

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, процент небелого населения.

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 трех любых статистически-значимых дамми-переменных.

Задание 3

Оцените модель, используя внутригрупповое преобразование (within-group transformation).

3.1 Проинтерпретируйте полученные результаты. Сравните коэффициенты при предикторах с коэффициентами при соответствующих предикторах в модели с набором дамми-переменных на константы.

3.2 Коэффициент при каком предикторе не позволяет оценить модель с внутригрупповым преобразованием? 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. Какая модель более адекватна в данном случае? Выбранную модель переоцените с поправкой на гетероскедастичность.

Задание 7

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

Бонусное задание

8.1 Оцените регрессионную модель с фиксированными эффектами на константу, в которой откликом является по-прежнему логарифм числа преступлений на человека, а предиктор только один – логарифм плотности 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 Покажите, как в случае множественной регрессии получить оценку коэффициента при предикторе «логарифм плотности 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.