

Многомерный статистический анализ

Вопросы для самопроверки после лекции 3

Вопрос

Давайте рассмотрим разные спецификации FE-моделей:

- ❶ классическая модель (с FE с дамми-переменными на пространственные единицы)
- ❷ модель с дамми-переменными на временные периоды
- ❸ twoway model (эффекты и на пространственные единицы, и на временные периоды)

Объясните различия: что содержательно отражают оценки коэффициентов при предикторах в этих спецификациях?

Смотря как «пирожок» разрежешь



Смотря как «пирожок» разрежешь



Hint

Вспомните, как мы в процессе условно делим массив данных при оценивании той или иной модели

Вопрос

Начнем с классической FE-модели. Вспомните ее исходную спецификацию. Что происходит при внутригрупповом преобразовании? Что отражает оценка коэффициента при предикторах?

Вопрос

Начнем с классической FE-модели. Вспомните ее исходную спецификацию. Что происходит при внутригрупповом преобразовании? Что отражает оценка коэффициента при предикторах?

Ответ (продолжение следует)

- $$y_{it} = b_0 + \gamma_1 * D_{1i} + ... \gamma_{n-1} * D_{(n-1)i} + b_1 * x_{it} + e_{it}$$

Вопрос

Начнем с классической FE-модели. Вспомните ее исходную спецификацию. Что происходит при внутригрупповом преобразовании? Что отражает оценка коэффициента при предикторах?

Ответ (продолжение следует)

- $y_{it} = b_0 + \gamma_1 * D_{1i} + ... \gamma_{n-1} * D_{(n-1)i} + b_1 * x_{it} + e_{it}$
- При внутригрупповом преобразовании мы вычитаем среднее, рассчитанное по каждой пространственной единице. Допустим, у нас в массиве N стран. После внутригруппового преобразования получим N подмассивов со значениями переменных в терминах отклонения от среднего по стране.

Продолжение

В каждом кусочке массива (а таких кусочков у нас N) выявляем взаимосвязь отклика и предикторов. На выходе получаем, как в среднем при увеличении предиктора на 1 во временной перспективе(!) изменяется значение отклика при прочих равных.

Вопрос

А что происходит в модели с фиксированными эффектами на временные периоды? Зачем она нам нужна?

Вопрос

А что происходит в модели с фиксированными эффектами на временные периоды? Зачем она нам нужна?

Ответ (продолжение следует)

$$y_{it} = b_0 + \gamma_1 * T_{1t} + ... \gamma_{T-1} * T_{(T-1)t} + b_1 * x_{it} + e_{it}$$

Вопрос

А что происходит в модели с фиксированными эффектами на временные периоды? Зачем она нам нужна?

Ответ (продолжение следует)

$$y_{it} = b_0 + \gamma_1 * T_{1t} + \dots \gamma_{T-1} * T_{(T-1)t} + b_1 * x_{it} + e_{it}$$

- \hat{b}_0 – чему в среднем равно значение зависимой переменной во временном периоде – базовой категории – при равенстве предикторов 0

Вопрос

А что происходит в модели с фиксированными эффектами на временные периоды? Зачем она нам нужна?

Ответ (продолжение следует)

$$y_{it} = b_0 + \gamma_1 * T_{1t} + ... \gamma_{T-1} * T_{(T-1)t} + b_1 * x_{it} + e_{it}$$

- \hat{b}_0 – чему в среднем равно значение зависимой переменной во временном периоде – базовой категории – при равенстве предикторов 0
- $\hat{\gamma}_t$ – на сколько в среднем отклоняется значение зависимой переменной в t-ой временном периоде в отличие от временного периода – базовой категории – при прочих равных

Продолжение

Такая модель контролирует характеристики, изменяющиеся во времени, но при этом неизменные в межстрановой перспективе.

Что означает оценить модель с FE на временные периоды?

Продолжение

Такая модель контролирует характеристики, изменяющиеся во времени, но при этом неизменные в межстрановой перспективе.

Что означает оценить модель с ФЕ на временные периоды? Мы оцениваем модель, предварительно вычитая из исходных значений переменных среднее по каждому временному периоду. То есть, к примеру, зафиксировали первый временной период, рассчитали среднее по выборке стран за этот временной период, и так далее.

Продолжение

В итоге получаем массив, разрезанный на T подвыборок. После внутригруппового преобразования получим T подмассивов со значениями переменных в терминах отклонения от среднего по временному периоду. Внутри каждого из этих T кусочков выявляем взаимосвязь отклика и предикторов. То есть, коэффициент при предикторе теперь показывает, как в среднем при увеличении предиктора на 1 в межстрановой перспективе(!) изменяется значение отклика при прочих равных.

А теперь та самая tricky model! Twoway FE-model

Ответ

Выглядит эта красotka вот так:

$$y_{it} = b_0 + \gamma_1 * D_{1i} + \dots \gamma_{n-1} * D_{(n-1)i} + c_1 * T_{1t} + \dots c_{T-1} * T_{(T-1)t} + b_1 * x_{it} + e_{it}$$

Равносильно тому, что центрировать по стране и временному периоду, а потом оценить модель. Берем исходный массив, совершаем внутригрупповое преобразование как в классической FE-модели, а потом берем срез по каждому временному периоду.

Продолжение

В итоге получаем массив, разрезанный на T подвыборки. Внутри каждой такой подвыборки уже значения переменных в терминах отклонений от среднего по стране. Внутри каждого из этих T кусочков выявляем взаимосвязь отклика и предикторов, а затем взвешиваем все эти результаты и получаем итоговую оценку коэффициента при предикторе. Теперь коэффициент при предикторе показывает, как в среднем при увеличении предиктора на 1 как в межстрановой, так и временной перспективе(!) изменяется значение отклика при прочих равных.

Вопрос

Вспомним, как работать с полученными результатами FE-модели. Какие дополнительные диагностики применять?

Вопрос

Вспомним, как работать с полученными результатами FE-модели. Какие дополнительные диагностики применять?

Ответ

Для начала можно оценить модель, которая учитывает только разные стартовые условия. Далее проверим, одинаково ли хорошо для каждой пространственной единицы модель объясняет вариацию отклика. Для этого рассчитаем корреляцию между предсказанным и наблюдаемым значением отклика по подгруппам. Далее перепроверим, сохраняются ли результаты при исключении стран с маленькой корреляцией (со значимостью аккуратнее, если маленькая подвыборка).

Вопрос

Что если все «сломалось»: получили другие результаты?

Вопрос

Что если все «сломалось»: получили другие результаты?

Ответ

Тогда следует уточнить модель. Подумайте, а действительно ли правдоподобно, что взаимосвязь предиктора и отклика одинакова во всех странах? Протестируйте:

$$y_{it} = b_0 + \gamma_1 * D_{1i} + \dots \gamma_{n-1} * D_{(n-1)i} + b_1 * x_{it} + \mu_1 * D_{1i} * x_{it} + \dots \mu_{n-1} * D_{(n-1)i} * x_{it} + e_{it}$$

Продолжение: интерпретация модели

Продолжение: интерпретация модели

- 1 b_0 – то, чему в среднем равен y_{it} в стране – базовой категории – при всех предикторах равных 0
- 2 γ_i – отклонение y_{it} в среднем в i -ой стране от y_{it} в базовой категории при всех предикторах равных 0
- 3 b_1 – насколько в среднем при прочих равных при увеличении x_{it} на 1 изменяется отклик в стране – базовой категории
- 4 μ_i – насколько в среднем отличается взаимосвязь x_{it} и y_{it} в i -ой стране в отличие от базовой категории при прочих равных