

Многомерный статистический анализ в ПОЛИТОЛОГИИ

Вопросы для самопроверки после лекции 5

Вопрос

Что показывает внутриклассовый коэффициент корреляции (ICC)?

Вопрос

Что показывает внутриклассовый коэффициент корреляции (ICC)?

Ответ

$$ICC = \frac{Var(u_{0j})}{Var(u_{0j}) + Var(e_{ij})}$$

ICC показывает долю вариации зависимой переменной, объясняемой межгрупповыми различиями (страновой спецификой). В случае если наблюдаем очень маленькие или, наоборот, очень высокие значения ICC, это свидетельствует о том, что нет необходимости в моделях, учитывающих неоднородность данных (зависимость единиц внутри групп).

Вопрос

Обозначьте допущения модели со смешанными эффектами.

Вопрос

Обозначьте допущения модели со смешанными эффектами.

Ответ

- ❶ $E(u_j) = E(e_{ij}) = 0$
- ❷ нормальное распределение u_j, e_{ij}
- ❸ нет строгого допущения о гомоскедастичности e_{ij}, u_j .
Примечание: см. робастное оценивание моделей со смешанными эффектами
- ❹ $Cov(u_j, e_{ij}) = 0$
- ❺ $Cov(u_j, e_{ij}) = 0$
- ❻ случайные эффекты для константы и наклона могут быть скоррелированы

Вопрос

Что содержательно означает корреляция между случайными эффектами?

Вопрос

Что содержательно означает корреляция между случайными эффектами?

Ответ

Пример положительной корреляции: при более высоких значениях «стартовых» условий наблюдаем более выраженную положительную взаимосвязь зависимой переменной и ключевого предиктора.

Когда такое может быть? Те страны, у которых изначально менее благоприятные экономические условия, долго восстанавливаются после экономического кризиса.

Вопрос

Один из методов оценивания моделей со смешанными эффектами – метод максимального правдоподобия. Вспомните его ключевую идею.

Вопрос

Один из методов оценивания моделей со смешанными эффектами – метод максимального правдоподобия. Вспомните его ключевую идею.

Ответ

Правдоподобие выборки – вероятность получения указанной выборки.

$$f(x_1, x_2, \dots, x_n | \theta) = f(x_1 | \theta) * f(x_2 | \theta) * \dots * f(x_n | \theta)$$

Наша цель – найти такое значение параметра θ , при котором максимальна вероятность получения указанной выборки.

Интересующие нас параметры – фиксированные эффекты, вариации случайных эффектов.

Вопрос

В контексте моделирования со смешанными эффектами нас в ряде случаев интересует не то, что мы в среднем наблюдаем по выборке, а страновая специфика. В связи с этим возникает необходимость предсказания характера взаимосвязи ключевого предиктора и зависимой переменной в отдельно взятой стране (или «стартовых» условий в отдельно взятой стране). Что показывает BLUP-значение?

Вопрос

В контексте моделирования со смешанными эффектами нас в ряде случаев интересует не то, что мы в среднем наблюдаем по выборке, а страновая специфика. В связи с этим возникает необходимость предсказания характера взаимосвязи ключевого предиктора и зависимой переменной в отдельно взятой стране (или «стартовых» условий в отдельно взятой стране). Что показывает BLUP-значение?

Ответ

BLUP – предсказанное значение отклонения в характере взаимосвязи (или отклонения в «стартовых условиях») в j -ой стране в отличие от того, что мы наблюдаем в среднем по всей выборке. (\hat{u}_j)

Вопрос

Каков механизм получения оценки «уникального» значения характера взаимосвязи (или «стартовых» условий) для j -ой страны?

Вопрос

Каков механизм получения оценки «уникального» значения характера взаимосвязи (или «стартовых» условий) для j -ой страны?

Ответ (продолжение на след. слайде)

Представьте, что мы разделили массив на N подвыборок (всего N стран) и на каждой подвыборке оценили взаимосвязь ключевого предиктора и зависимой переменной. Полученные оценки коэффициентов при предикторе обозначим \hat{k}_j . Ту же взаимосвязь мы оценили на всей выборке, соответствующая оценка (средний характер взаимосвязи по всей выборке = $\hat{\gamma}$).

Ответ (продолжение)

$$\hat{b}_j = \alpha_j * \hat{k}_j + (1 - \alpha_j) * \hat{\gamma},$$

где \hat{b}_j – уникальный характер взаимосвязи предиктора и зависимой переменной в j-ой стране;

α_j – «вес» для оценки по отдельной подвыборке;

$(1 - \alpha_j)$ – «вес» для среднего по всей выборке, насколько средний характер взаимосвязи по всей выборке значим в формировании оценки \hat{b}_j

Вопрос

В каком случае оценка взаимосвязи по отдельной подвыборке имеет больший вес в формировании оценки \hat{b}_j ?

Вопрос

В каком случае оценка взаимосвязи по отдельной подвыборке имеет больший вес в формировании оценки \hat{b}_j ?

Ответ

$$\alpha_j = \frac{Var(u_{0j})}{Var(u_{0j}) + \frac{Var(e_{ij})}{n_j}}$$

- ❶ при более высокой межгрупповой вариации
- ❷ при большем размере подвыборки (количества наблюдений, доступных для j-ой страны)

при прочих равных условиях оценка взаимосвязи по отдельной подвыборке становится более значимой при формировании оценки \hat{b}_j .

Вопрос

Как осуществляется тестирование альтернативных спецификаций моделей со смешанными эффектами?

Вопрос

Как осуществляется тестирование альтернативных спецификаций моделей со смешанными эффектами?

Ответ

- 1 тест отношения правдоподобий (для вложенных моделей)
- 2 информационные критерии