

Метода анализа неоднородных данных и паттерн-анализ

О том, что нас ожидает,
на что / кого надеяться и зачем все это нужно

13 января 2021

В каком направлении мы движемся

- приобрести уверенность в обращении с классической линейной регрессионной моделью
- научиться выявлять неоднородный характер взаимосвязи
- выйти хорошо вооруженными (подходящими инструментами для работы с гетерогенной выборкой)
- практиковать разборчивость (понимать, когда и какой инструмент использовать)
- изучить методы анализа паттернов

Наш план:

- повторить основы статистики, в частности, модель классической линейной регрессии

Наш план:

- повторить основы статистики, в частности, модель классической линейной регрессии
- изучить и научиться применять регрессионные модели с переменными взаимодействия

Наш план:

- повторить основы статистики, в частности, модель классической линейной регрессии
- изучить и научиться применять регрессионные модели с переменными взаимодействия
- изучить и научиться применять модели с фиксированными эффектами и модели со случайными эффектами

Наш план:

- повторить основы статистики, в частности, модель классической линейной регрессии
- изучить и научиться применять регрессионные модели с переменными взаимодействия
- изучить и научиться применять модели с фиксированными эффектами и модели со случайными эффектами
- изучить и научиться применять модели со смешанными эффектами

Наш план:

- повторить основы статистики, в частности, модель классической линейной регрессии
- изучить и научиться применять регрессионные модели с переменными взаимодействия
- изучить и научиться применять модели с фиксированными эффектами и модели со случайными эффектами
- изучить и научиться применять модели со смешанными эффектами
- изучить и научиться применять методы анализа паттернов

Если Вы чувствуете, что фиксированный эффект смешался со случайным, то Вам могут помочь следующие люди:

Преподаватели

Дарья Сальникова (e-mail: darysalnikova@yandex.ru;
dsalnikova@hse.ru;

web-page: <https://www.hse.ru/staff/salnikova>)

Алексей Мячин (amyachin@hse.ru;

web-page: <https://www.hse.ru/org/persons/69178868>)

Наш ассистент

Полина Ревина (e-mail: pvrevina@edu.hse.ru; web-page:
<https://www.hse.ru/staff/previna>)

Что Вас ожидает

Формат занятий

Лекции (новый материал + обсуждения) + семинары (решение задач, практика в Python, RStudio)

Типы заданий

- Домашние задания
- Проверочные работы
- Эссе
- Презентация эссе
- Экзамен

Правила прохождения дистанции (1)

Заветная формула

$$\text{Итог} = 0.2 \times \text{ДЗ} + 0.15 \times \text{Сам.} + 0.25 \times \text{ЭССЕ} + 0.1 \times \text{Презентация} + 0.3 \times \text{Экзамен}$$

Домашние задания

Опоздание в пределах одного дня (24 часа) – штраф 1 балл. При опоздании более чем на 2 дня (48 ч. после дедлайна) за данное домашнее задание выставляется оценка 0.

НО для получения содержательных комментариев можно сдать ВСЕГДА (конечно, не позднее окончания курса).

Правила прохождения дистанции (2)

Проверочные работы

В случае пропуска студентом проверочной работы за нее выставляется оценка 0.

Как возместить пропущенные проверочные работы?

Дополнительные практические задания (сдать на проверку) + активность на семинарах

Что сдавать?

ЭССЕ

- Файл с текстом эссе. Лучше всего в .pdf
- Python-notebook / R скрипт (проверьте на реплицируемость без ошибок)
- База данных

Как организовать свой творческий порыв (1)

ЭССЕ (примерно 20000 – 30000 знаков с пробелами)

Как организовать свой творческий порыв (1)

ЭССЕ (примерно 20000 – 30000 знаков с пробелами)

- Abstract (кратко о чем Ваша работа, основные результаты, можно упомянуть о данных и методах)

Как организовать свой творческий порыв (1)

ЭССЕ (примерно 20000 – 30000 знаков с пробелами)

- Abstract (кратко о чем Ваша работа, основные результаты, можно упомянуть о данных и методах)
- Введение
 - ▶ Ключевой вопрос
 - ▶ Кто еще задавался таким вопросом? Кратко – об их находках.
 - ▶ Чем ответы предыдущих исследователей Вам не угодили? (неполные / есть ошибки – неточности)

Как организовать свой творческий порыв (1)

ЭССЕ (примерно 20000 – 30000 знаков с пробелами)

- Abstract (кратко о чем Ваша работа, основные результаты, можно упомянуть о данных и методах)
- Введение
 - ▶ Ключевой вопрос
 - ▶ Кто еще задавался таким вопросом? Кратко – об их находках.
 - ▶ Чем ответы предыдущих исследователей Вам не угодили? (неполные / есть ошибки – неточности)
- Гипотезы (формулируются эмпирически тестируемые предположения, НЕ ПУТАТЬ со статистическими гипотезами, множить гипотезы не стоит – не более 3)

Как организовать свой творческий порыв (2)

ЭССЕ (примерно 20000 – 30000 знаков с пробелами)

Как организовать свой творческий порыв (2)

ЭССЕ (примерно 20000 – 30000 знаков с пробелами)

- Данные и методы
 - ▶ Откуда данные? Опишите источники
 - ▶ Опишите данные: структура данных, размер выборки, описательные статистики, как измерены Ваши переменные и погрузите в контекст исследования: какой концепт они отражают
 - ▶ Обоснование методов, обсуждение альтернативных вариантов
 - ▶ Спецификация моделей, какую задачу решает оценивание той или иной модели

Как организовать свой творческий порыв (3)

ЭССЕ (примерно 20000 – 30000 знаков с пробелами)

Как организовать свой творческий порыв (3)

ЭССЕ (примерно 20000 – 30000 знаков с пробелами)

- Результаты

- ▶ В табличной форме, соответствующей зарубежным публикационным стандартам
- ▶ Описание результатов в «сухой» форме (просто то, что получилось, без критической оценки – оставьте это до раздела Discussion)
- ▶ Robustness checks (проверки результатов на устойчивость)

Как организовать свой творческий порыв (3)

ЭССЕ (примерно 20000 – 30000 знаков с пробелами)

- Результаты

- ▶ В табличной форме, соответствующей зарубежным публикационным стандартам
- ▶ Описание результатов в «сухой» форме (просто то, что получилось, без критической оценки – оставьте это до раздела Discussion)
- ▶ Robustness checks (проверки результатов на устойчивость)

- Обсуждение результатов и заключение

- ▶ Резюме полученных основных результатов и проверок на устойчивость
- ▶ Насколько соответствуют результаты предшествующей литературе, если есть расхождения, то почему
- ▶ Ограничения Вашего исследования

