

# Регрессионный анализ: панельные данные и каузальность

Введение: план действий и правила игры

7 сентября 2021

# К чему стремимся: в каком направлении мы движемся

- выйти в два измерения: как пространственное, так и временное (привет панельным данным!)
- выйти хорошо вооруженными (подходящими методами для работы с панельными данными)
- практиковать разборчивость (понимать, когда и за какой инструмент – инструмент в широком смысле – хвататься)
- переосмыслить свои работы (в рамках курса Вас ждет творческая работа – эссе) в терминах каузального вывода

# В связи с этим нужно быть готовым к:

- моделям с самыми разными эффектами (фиксированными, случайными, смешанными)
- спагетти (первый график-спагетти уже сегодня!)
- подаче под «соусом» каузального вывода
- не статистикой единой: обсуждению дизайну исследования

# Наш план действий

- ❶ Фейерверк эффектов для анализа панельных данных:
  - фиксированные эффекты
  - случайные эффекты
  - смешанные эффекты
- ❷ Источники эндогенности
- ❸ Рисуем идеальную картинку для каузального вывода: экспериментальный дизайн, считаем «голыми руками»
- ❹ Каузальный вывод: ключевые понятия, логика, ограничения. Сопровождение: картинки DAG
- ❺ Подключаем тяжелую артиллерию: инструменты **IV**
- ❻ We **DiD** it: Посчитаем разности, а потом разность разностей
- ❼ Заключение: Презентуем эссе и готовимся к экзамену

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

Тот же самый человек, что ведет занятия

Дарья Сальникова

(e-mail: [dsalnikova@hse.ru](mailto:dsalnikova@hse.ru);

[darysalnikova@yandex.ru](mailto:darysalnikova@yandex.ru);

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

Наш ассистент

Вероника Янчук (e-mail: [vayanchuk@edu.hse.ru](mailto:vayanchuk@edu.hse.ru))

# Что Вас ожидает

## Формат занятий

Семинары (лекции + дискуссии) и практические занятия (решение задач и отработка в RStudio)

## Типы заданий

- Домашние задания (×3)
- Проверочные работы (×3)
- Контрольная работа (×1)
- Активность – работа на практических занятиях
- Эссе (×1)
- Экзамен (×1)

# Правила прохождения дистанции (1)

## Заветная формула

$$\text{Итог} = 0.15 \times \text{ДЗ} + 0.15 \times \text{Проверочные} + 0.15 \times \text{КР} + 0.1 \times \text{Активность} + 0.15 \times \text{ЭССЕ} + 0.3 \times \text{Экзамен}$$

## 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 декабря 2021).

# Выполнение домашних заданий

Объем НЕ имеет значения и НЕ влияет на оценку,  
НО влияет ПОЛНОТА и КОРРЕКТНОСТЬ ответов



# Выполнение домашних заданий

Объем НЕ имеет значения и НЕ влияет на оценку,  
НО влияет ПОЛНОТА и КОРРЕКТНОСТЬ ответов

- Результаты (в табличной форме)
- Описание результатов текстом и ответы на поставленные вопросы
- Напишите небольшое заключение по результатам анализа: основные выводы, ограничения
- Код в R должен быть реплицируемым

# Правила прохождения дистанции (2)

## Проверочные работы

Closed-book. В случае пропуска по уважительной подтвержденной причине вес проверочной переносится на экзамен, в противном случае при пропуске выставляется оценка 0.

## Как частично возместить пропущенные проверочные работы?

Бонусные задания

## Контрольная работа

Closed-book. В случае пропуска по уважительной подтвержденной причине вес переносится на экзамен.

# Правила прохождения дистанции (3)

## Экзамен

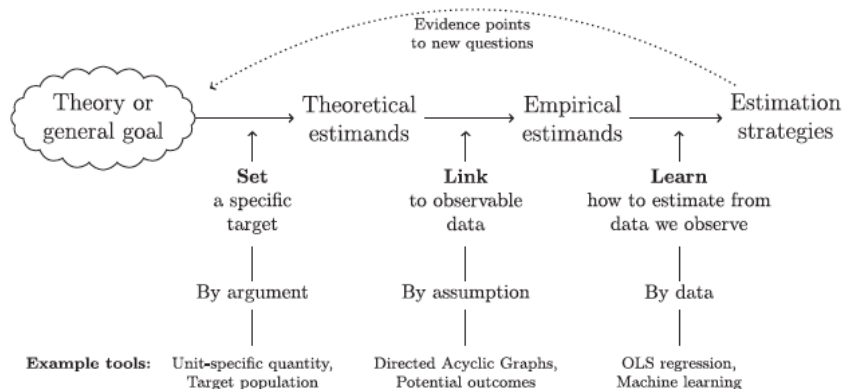
Экзамен проходит в формате closed-book. В экзаменационной работе могут встретиться задания повышенной сложности или задания, только частично разобранные на семинарах, верное решение которых предполагается для возможности получения оценок 9 и 10. Процент таких заданий не превышает 20%.

# Эссе: когда сдавать и что сдавать?

Deadline: 23.59 5 декабря 2021

- Файл с текстом эссе (примерно 15000 – 20000 знаков с пробелами). Лучше всего в .pdf
- R-код (проверьте на реплицируемость без ошибок)
- Массив данных

# Схема из Lundberg et al., 2021



# 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. Определить theoretical estimand
  - 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. Определить theoretical estimand
  - 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, выделите empirical estimand
  - ▶ 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

