

Домашнее задание №1

Практическая эконометрика

11 ноября 2022 г.

1 Общие требования

Задание выполняется 2 студентами (можно по одному, но идеально 2), сдаётся на `on.esop` в виде нескольких файлов: оформленный связный текст с ответами на вопросы, таблицами и графиками (если они необходимы по заданию) в формате `pdf` и файл с кодом в `R`. Они должны называться фамилией (фамилиями) студента(ов), например `ivanov_kraskov.pdf` и `ivanov_kraskov.R`. Если вы делаете задание в паре, то достаточно, если на онэкон прикрепит файл кто-то один.

Код в `R` должен работать. Команды `install.packages()` пишите, пожалуйста, в качестве комментариев после знака `#`. В коде в первой строчке в качестве комментария напишите, пожалуйста, авторов.

Оформление отчёта должно быть стандартным для письменных работ (вспомните правила оформления курсовой или проекта по ЭКМ-2). Таблицы не надо вставлять скринами.

Обдумывая решения, можно и нужно задавать вопросы, а вот писать код и ответы нужно самостоятельно, соблюдая академическую этику (не списывать). Одинаковые работы (списанные части текста или кода) от разных коллективов авторов обнуляются.

За задание вы можете получить 20 основных баллов из общей семестровой суммы 100 баллов, а также 6 бонусных баллов сверх суммы БРС. Бонусные задания обозначены символом ★.

Срок сдачи задания - суббота 26 ноября 23:59.

2 Задание

В рамках этой домашней работы мы попробуем написать несколько полезных технических штук в R и повторить часть результатов статьи Banerjee, Abhijit, Esther Duflo, and Richard Hornbeck. 2014. «Bundling Health Insurance and Microfinance in India: There Cannot Be Adverse Selection If There Is No Demand.» *American Economic Review*, 104 (5): 291-97. DOI: 10.1257/aer.104.5.291. Целиком читать статью НЕ ОБЯЗАТЕЛЬНО, но желательно. Она небольшая. А еще она интересна тем, что является одной из работ, которая была отмечена Нобелевским комитетом в 2019 году (Дюфло и Банерджи стали победители Нобелевской премии по экономике в 2019 г.).

Текст статьи дополнительно размещен в соответствующем разделе на `on.esop`. Данные, тест статьи и описание переменных – по ссылке.

В домашней работе вам встретятся задания двух типов: письменные вопросы по содержанию и требующие написать код.

2.1 Общее (4 + 2★ баллов)

1. Текст (1 балл) На какой вопрос отвечают исследователи? Опишите коротко: 1-2 предложения.
2. Текст★ (2 балла) Какой вклад статья вносит в своей области? С какими другими исследовательскими вопросами связано это исследование? (1 балл за описание по пунктам). Знаете ли вы, почему ещё этот вопрос может быть интересен (со ссылками на литературу)? (1 балл в случае интересной и релевантной ссылки на литературу НЕ из обзора данной статьи)
3. Текст (1 балл) Для чего авторы приводят таблицу 1 и какой вывод из неё делают?
4. Код и текст (2 балла) Повторите Таблицу 1 из статьи. Используйте файл с данными `clean_baseline_PUBLISH.dta`. Код должен выдавать таблицу. Вставьте готовую таблицу в текст. Она должна содержательно быть такой же, но выглядеть не обязательно точно так же, как в статье. Достаточно столбцов 2 и 3.

2.2 Рандомизация (5 + 4★ баллов)

Представьте, что вы – помощник исследователей Абхиджита Банерджи, Эстер Дюфло и Ричарда Хорнбека и делаете рандомизацию для эксперимента с нуля:

1. Текст★ (1 балл) Почему авторы проводят рандомизацию на уровне деревень, а не отдельных домохозяйств?
2. Код и текст (2 балла) Разминочная рандомизация без пре-стратификации. Возьмите данные из файла `clean_baseline_PUBLISH.dta`. (`clean_treatment_PUBLISH.dta` тоже дополнительно можно, если он вам поможет, но можно и без него). Сделайте свою рандомизацию любым из способов, которые мы с вами разбирали, просто разделив наблюдения на 2 группы (третмент- и контрольную) так, чтобы в третмент-группу отправлялись столько же деревень, сколько и в статье, а остальные - в контрольную группу. Проверьте качество Вашей рандомизации аналогично с Таблицей 1. Результат опишите в тексте. Вставьте таблицу в текст.
3. Код и текст (2 балла) Сделайте свою рандомизацию без пре-стратификации, устойчивую к перестановке строк (например, на основе шифровки ID деревни), разделив наблюдения на 2 группы (третмент- и контрольную) так, чтобы в третмент-группу отправлялись столько же деревень, сколько и в статье, а остальные - в контрольную группу. Проверьте качество Вашей рандомизации тем же способом, что и в предыдущем пункте. Результат опишите в тексте.
4. Текст (1 балл) Для чего авторам нужна престратификация? Почему это лучше, чем сквозная рандомизация по деревням?
5. Код и текст★ (3 балла) Рандомизация почти-как-в-статье (с престратификацией). Вы делаете рандомизацию для эксперимента с нуля. Сделайте свою рандомизацию с престратификацией по стратам, которые есть в данных. Проверьте качество Вашей рандомизации тем же способом, что и в предыдущем пункте. Результат опишите в тексте.

2.3 Оценка эффекта (11 баллов)

1. Текст (1 балл) В таблице 2 авторы оценивают эффект на пролонгацию кредита. Чем отличаются оценки в столбцах 1 и 2? А в столбцах 3 и 4?
2. Код (4 балла) Воспроизведите результат из таблицы 2 (на часть баллов из этого пункта – хотя бы регрессию 2). Вам могут помочь данные из файлов: `clean_loans_annual_PUBLISH.dta`, `clean_treatment_PUBLISH.dta`, `analysis_sample_PUBLISH.dta`, `clean_endline_PUBLISH.dta`. Нужно будет «сшить» данные из нескольких файлов. Код должен выводить таблицу. Готовую таблицу нужно вставить в текст. По оформлению она может быть не точно такой же, как в в статье.
3. Текст (1 балл) Оценки чего авторы приводят в таблице 3? Какую(-ие) спецификации они для этого используют?

4. Код (4 балла) Воспроизведите таблицу 3. Вам могут помочь данные из файлов: `clean_baseline_PUBLISH.dta`, `clean_loans_annual_PUBLISH.dta`, `clean_endline_PUBLISH.dta`. Нужно будет «сшить» данные из нескольких файлов. Код должен выводить таблицу. Готовую таблицу нужно вставить в текст. По оформлению она может быть не точно такой же, как в статье.
5. Текст (1 балл) Какие основные выводы делают авторы?

P.S. В пунктах, где мы пишем, что вам нужно будет «сшить» данные из нескольких файлов, вам может пригодиться ознакомиться с семейством `join` функций. Это можно сделать по одной из ссылок ниже или на любом другом сайте, который вам нравится:

- Applied Economics with R. Merging the datasets (Hans H. Sievertsen)
- Library of Statistical Techniques. Combining Datasets (Nick Huntington-Klein)
- Introduction to R. Combining datasets (Hans H. Sievertsen)
- An Introduction to Human Resource Analytics Using R. Joining (Merging) Data (David E. Caughlin)