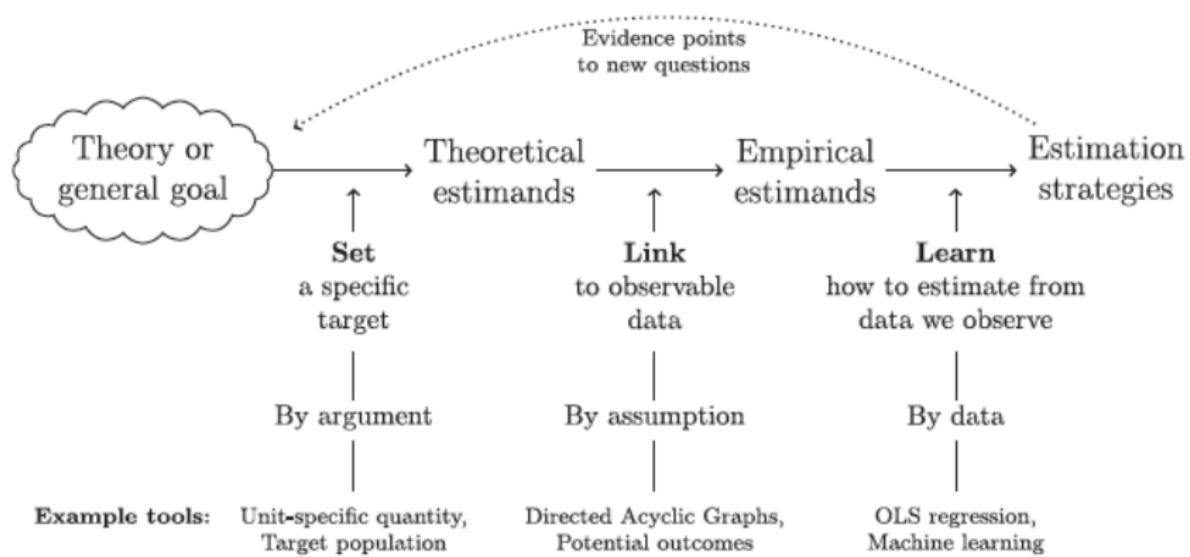


# Регрессионный анализ социально-экономических процессов

Применение регрессионного анализа в исследованиях

19 января 2026

# Источник: Lundberg et al., 2021



# Теоретический параметр

Выдержка из Lundberg et al., 2021:

«We introduce a term for the goal stated outside the model—**the theoretical estimand**— which has two components. The first is **a unit-specific quantity**, which could be a realized outcome (whether person  $i$  is employed), a potential outcome (whether person  $i$  would be employed if they received job training), or a difference in potential outcomes (the effect of job training on the employment of person  $i$ ). The unit-specific quantity clarifies whether the research goal is causal, and if so, what counterfactual intervention is being considered. The second component of the theoretical estimand is **the target population**: over whom or what do we aggregate that unit-specific quantity?»

# Структуры данных

- Пространственные данные:  
в таких данных представлены наблюдения нескольких ( $N$ ) пространственных единиц анализа за один и тот же временной период

# Структуры данных

- Пространственные данные:  
в таких данных представлены наблюдения нескольких ( $N$ ) пространственных единиц анализа за один и тот же временной период
- Временные ряды:  
данные одной пространственной единицы анализа представлены за несколько ( $T$ ) временных периодов

# Структуры данных

- Пространственные данные:  
в таких данных представлены наблюдения нескольких ( $N$ ) пространственных единиц анализа за один и тот же временной период
- Временные ряды:  
данные одной пространственной единицы анализа представлены за несколько ( $T$ ) временных периодов
- Панельные данные:  
данные нескольких ( $N$ ) пространственных единиц представлены за несколько ( $T$ ) временных периодов (есть и пространственная, и времененная перспектива)

# Пространственные данные: пример

id	age	spending_score	city	is_employed
1	37	42	London	True
2	51	78	Paris	False
3	22	15	Moscow	True

# Временные ряды: пример

Quarter	GDP Growth (%)	Inflation (%)	Interest Rate (%)
2023 Q1	2.1	4.2	1.5
2023 Q2	1.8	4.5	1.8
2023 Q3	2.3	4.8	2.0
2023 Q4	2.0	5.2	2.2

# Панельные данные: пример

Company	Year	Revenue (\$K)	Profit Margin (%)	Country
1	2021	987.23	18.34	USA
1	2022	1054.67	12.56	USA
1	2023	912.45	21.78	USA
2	2021	1201.89	8.91	Germany
2	2022	1156.34	14.23	Germany
2	2023	1245.67	9.45	Germany
3	2021	845.12	19.67	Japan
3	2022	901.45	16.89	Japan
3	2023	878.90	22.34	Japan

# TSCS-данные

Кроме того, в англоязычной литературе Вы можете встретить такое понятие, как Time-Series Cross-Sectional Data. Здесь имеет место

- специфика обозначения панельных данных в разных дисциплинах
- длительность временного периода: как правило, TSCS-данными называют данные с относительно небольшим числом пространственных единиц и продолжительным временным рядом. Здесь в большей степени фокусируемся на таких сюжетах, как (не)стационарность данных и моделирование автокорреляции в явном виде

# Примерная структура описания регрессионного анализа и его результатов (1)

- Исследовательский вопрос и гипотезы: в частности, указать теоретический параметр (theoretical estimand) и оцениваемый параметр на доступных данных (empirical estimand)
- Данные и методы
  - ▶ Источники данных
  - ▶ Опишите данные: структура данных, размер выборки, сбалансированность данных, описательные статистики, как измерены Ваши переменные и погрузите в контекст исследования: какой концепт они отражают
  - ▶ Обоснование методов, обсуждение альтернатив
  - ▶ Спецификация моделей (набор переменных, функциональная форма связи), какую задачу решает оценивание той или иной модели

- Диагностики модели (можно в приложении) и используемые корректировки
- Результаты:
  - ▶ в табличной форме
  - ▶ описание результатов в «сухом» формате (сначала просто то, что получилось, без критической оценки)
  - ▶ проверки результатов на устойчивость (к примеру, устойчивость к изменению спецификации модели, изменениям в данных, методах оценивания, бутстррап, плацебо-тесты)
  - ▶ обсуждение результатов и ограничения проведенного анализа