

Машинное обучение в экономике

Эффекты воздействия

Потанин Богдан Станиславович

доцент, старший научный сотрудник, кандидат экономических наук

2025–2026

- Основные понятия:

- Потенциальные исходы.
- Эффект воздействия, средний эффект воздействия, локальный средний эффект воздействия и условный средний эффект воздействия.
- Допущения о независимости, об условной независимости и об экзогенности инструмента.

- Методы оценивания:

- Оценивание средних эффектов воздействия с помощью условных математических ожиданий, взвешивания на обратные условные вероятности и метода с двойной устойчивостью.
- Оценивание локального среднего эффекта воздействия с помощью двойного машинного обучения.
- Оценивание условных средних эффектов воздействия с помощью S-learner, T-learner, X-learner и метода трансформации классов.

Повторение теории вероятностей

Некоторые важные результаты

- Рассмотрим случайные векторы Y , X и Z с конечными математическими ожиданиями, а также функцию $g(\cdot)$, такую, что математическое ожидание $g(X)$ существует.
- При введенных предположках справедливы следующие формулы.

$$E[E(Y|X)] = E[Y] \quad E[g(X)Y|X] = g(X)E[Y|X]$$

$$E[g(X)Y] = E[g(X)E(Y|X)] \quad E[E(Y|X, Z)|X] = E[Y|X]$$

- Если случайные величины X_1, \dots, X_n одинаково распределены и имеют конечное математическое ожидание, то $E(X_i) = E(X_j)$ при любых $i, j \in \{1, \dots, n\}$, а значит:

$$E(\bar{X}_n) = E\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{1}{n} \sum_{i=1}^n E(X_i) = \frac{1}{n} \left(\underbrace{E(X_1) + \dots + E(X_n)}_{n \text{ раз складываем } E(X_i)} \right) = E(X_i)$$

Кроме того, в силу **закона больших чисел** $\bar{X}_n \xrightarrow{P} E(X_i)$.

- Если $g(\cdot)$ – непрерывная функция и X_1, \dots, X_n последовательность случайных векторов такая, что $X_n \xrightarrow{P} x$, где $x \in R^m$, то по **теореме о непрерывном отображении** (СМТ) выполняется $g(X_n) \xrightarrow{P} g(x)$.
- В частности, по **теореме Slutsky** если $X_n \xrightarrow{P} x$ и $Y_n \xrightarrow{P} y$, где $x, y \in R^m$, то:

$$X_n + Y_n \xrightarrow{P} x + y \quad X_n - Y_n \xrightarrow{P} x - y$$

$$X_n \times Y_n \xrightarrow{P} x \times y \quad X_n / Y_n \xrightarrow{P} x / y, \text{ где } y \neq 0$$

Эффект воздействия

Определение

- Предположим, что одновременно существуют два гипотетически **режима** (counterfactual states) целевой переменной, обозначаемых Y_{0i} и Y_{1i} . Но в данных мы наблюдаем только один из них, в зависимости от значения бинарной **переменной воздействия** (treatment) T_i .

$$Y_i = \begin{cases} Y_{1i}, & \text{если } T_i = 1 \\ Y_{0i}, & \text{если } T_i = 0 \end{cases} = T_i Y_{1i} + (1 - T_i) Y_{0i}$$

- Эффект воздействия** (treatment effect) TE_i на Y_i определяется как:

$$TE_i = Y_{1i} - Y_{0i}$$

- Например, Y_i может отражать факт совершения покупки клиентом, а T_i – факт наличия персонального предложения по бесплатной доставке товара на дом.
- Фундаментальная проблема причинно-следственного анализа** (fundamental problem of casual inference) – на практике мы не можем посчитать эффект воздействия TE_i , поскольку не способны одновременно наблюдать оба потенциальных исхода: Y_{0i} и Y_{1i} .
- Например, мы не знаем, как клиент мог бы повести себя и в случае наличия Y_{1i} и в случае отсутствия Y_{i0} предложения о доставке T_i , поскольку он либо получает это предложение (воздействие) $T_i = 1$, либо нет $T_i = 0$.

Средний эффект воздействия

Определение

- На практике часто рассматривается **средний эффект воздействия**:

$$ATE = E(TE_i) = E(Y_{1i} - Y_{0i}) = E(Y_{1i}) - E(Y_{0i})$$

- Представим, что Y_i отражает зарплату индивида, а T_i - факт наличия высшего образования.
- Тогда Y_{1i} отражает зарплату индивида в случае, когда у него есть высшее образование (экзогенно заданное), а Y_{0i} - в случае, когда у него нет высшего образования.
- У каждого индивида одновременно существуют две гипотетические зарплаты Y_{0i} и Y_{1i} , но мы наблюдаем лишь одну из них Y_i , в зависимости от наличия $T_i = 1$ или отсутствия $T_i = 0$ высшего образования, поэтому мы не можем посчитать эффект воздействия $TE_i = Y_{1i} - Y_{0i}$.
- Однако, при определенных предпосылках мы можем оценить средний эффект воздействия АТЕ, который можно интерпретировать как среднюю разницу в заработных платах среди людей в состояниях, когда у них есть высшее образование и в случаях, когда у них нет высшего образования.
- Для простоты дальнейшего изложения будем придерживаться допущения о том, что Y_{ji} , Y_{jt} , T_i и T_t независимы при любых $i \neq t$, где $j \in \{0, 1\}$.

Средний эффект воздействия

Условное распределение потенциальных исходов

- Для того, чтобы корректно оценивать и интерпретировать АТЕ, крайне важно различать безусловные Y_{1i} , Y_{0i} и условные $(Y_{1i}|T_i = 1)$, $(Y_{0i}|T_i = 0)$ распределения потенциальных исходов.
- Рассмотрим очень простой пример, в котором индивиды без высшего образования всегда зарабатывают две тысячи рублей $Y_{0i} = 2$, а с высшим образованием с равной вероятностью зарабатывают одну или три тысячи рублей $P(Y_{1i} = 1) = P(Y_{1i} = 3) = 0.5$.
- В данном случае легко рассчитать эффекты воздействия и средний эффект воздействия:

$$TE_i = Y_{1i} - Y_{0i} = Y_{1i} - 2 = \begin{cases} 3 - 2, & \text{если } Y_{1i} = 3 \\ 1 - 2, & \text{если } Y_{1i} = 1 \end{cases} = \begin{cases} 1, & \text{если } Y_{1i} = 3 \\ -1, & \text{если } Y_{1i} = 1 \end{cases}$$

$$ATE = E(Y_{1i}) - E(Y_{0i}) = (0.5 \times 1 + 0.5 \times 3) - 2 = 2 - 2 = 0$$

- Предположим, что индивид получает высшее образование $T_i = 1$, только если оно позволяет увеличить доходы $Y_{1i} > Y_{0i}$. Тогда разница в условных на уровень образования ожидаемых зарплатах не будет отражать средний эффект воздействия:

$$E(Y_{1i}|T_i = 1) - E(Y_{0i}|T_i = 0) = E(Y_{1i}|Y_{1i} > Y_{0i}) - E(Y_{0i}|Y_{1i} < Y_{0i}) = 3 - 2 = 1 \neq ATE$$

- В данном случае $E(Y_{1i}|T_i = 1)$ и $E(Y_{0i}|T_i = 0)$ отражают ожидаемые зарплаты среди людей фактически получивших и не получивших высшее образование соответственно.

Допущение о независимости

Формулировка

- Из предыдущего примера мы заключили, что в общем случае:

$$ATE = E(Y_{1i}) - E(Y_{0i}) \neq E(Y_{1i}|T_i = 1) - E(Y_{0i}|T_i = 0)$$

- Однако, на практике в некоторых частных случаях мы можем предположить, что ожидаемый исход не зависит от воздействия, то есть $E(Y_{1i}|T_i = 1) = E(Y_{1i})$ и $E(Y_{0i}|T_i = 0) = E(Y_{0i})$.
- Например, для измерения среднего эффекта воздействия вакцины на излечение от болезни, пациентов случайным образом распределяют между группой воздействия, получающей лекарство, и контрольной группой, принимающей плацебо.
- В результате то, излечится ли (гипотетически) пациент при принятии вакцины Y_{1i} и без нее Y_{0i} не зависит от того, получит ли он ее фактически $T_i = 1$, поскольку она выдается случайным образом.
- Введем **допущение о независимости** (mean independence):

$$E(Y_{1i}|T_i = 1) = E(Y_{1i}) \quad E(Y_{0i}|T_i = 0) = E(Y_{0i})$$

- Это допущение обычно соблюдается в рамках контролируемых случайных экспериментов.
- При выполнении допущения о независимости:

$$ATE = E(Y_{1i}|T_i = 1) - E(Y_{0i}|T_i = 0)$$

Допущение о независимости

Оценивание среднего эффекта воздействия

- Предположим, что в данных n_1 наблюдений попали в **группу воздействия** $T_i = 1$, а n_0 наблюдений оказались в **контрольной группе** $T_i = 0$.
- При соблюдении допущения о независимости вследствие закона больших чисел получаем состоятельные оценки:

$$\hat{E}(Y_{1i}) = \frac{1}{n_1} \sum_{i: T_i=1} Y_{1i} \qquad \hat{E}(Y_{0i}) = \frac{1}{n_0} \sum_{i: T_i=0} Y_{0i}$$

- Тогда по теореме Слуцкого состоятельная оценка среднего эффекта воздействия может быть получена как:

$$\widehat{ATE} = \hat{E}(Y_{1i}) - \hat{E}(Y_{0i}) = \frac{1}{n_1} \sum_{i: T_i=1} Y_{1i} - \frac{1}{n_0} \sum_{i: T_i=0} Y_{0i}$$

- Например, если среди $n = 100$ пациентов $n_0 = 60$ получили плацебо и $n_1 = 40$ получили вакцину, причем среди получивших плацебо излечились 15, а среди получивших вакцину 20, то:

$$\widehat{ATE} = (1/40) \times 20 - (1/60) \times 15 = 0.5 - 0.25 = 0.25$$

- **Вывод** – согласно полученной оценке в среднем вакцина повышает вероятность исцеления на 0.25.

Допущение о независимости

Технический комментарий о числе наблюдений в группе воздействия и контрольной группе

- При оценивании $E(Y_{1i})$ и $E(Y_{0i})$ для простоты изложения ранее и далее **предполагается**, что n_1 и n_0 являются константами.
- Однако, за пределами контролируемых экспериментов размеры группы воздействия и контрольной группы, как правило, являются случайными величинами:

$$n_1 = \sum_{i=1}^n T_i \quad n_0 = \sum_{i=1}^n 1 - T_i$$

- Воспользуемся законом больших чисел (в числителе и знаменателе) и теоремой Слуцкого:

$$\begin{aligned} \hat{E}(Y_{1i}) &= \frac{1}{n_1} \sum_{i: T_i=1} Y_{1i} = \frac{\frac{1}{n} \sum_{i=1}^n T_i Y_{1i}}{\frac{1}{n} \sum_{i=1}^n T_i} \xrightarrow{p} \frac{E(T_i Y_{1i})}{E(T_i)} = \\ &= \frac{E(Y_{1i}|T_i=1)P(T_i=1) + 0 \times P(T_i=0)}{P(T_i=1)} = E(Y_{1i}|T_i=1) \end{aligned}$$

- По аналогии нетрудно показать, что $\hat{E}(Y_{0i}) \xrightarrow{p} E(Y_{0i}|T_i=0)$, откуда при допущении о независимости следует состоятельность введенной ранее оценки:

$$\widehat{ATE} = \hat{E}(Y_{1i}) - \hat{E}(Y_{0i}) \xrightarrow{p} E(Y_{1i}|T_i=1) - E(Y_{0i}|T_i=0) = \underbrace{E(Y_{1i}) - E(Y_{0i})}_{\text{допущение о независимости}} = ATE$$

Допущение о независимости

AB-тестирование

- Часто под **AB-тестированием** понимается проверка гипотезы $H_0 : ATE = 0$ при допущении о независимости.
- Например, представим, что клиентская база продавца составляет 1000 человек. Из них он случайным образом отобрал 100 и предоставил им специальное предложение, согласно которому при покупке телефона они получают в подарок наушники.
- Из 100 человек, получивших предложение, покупку совершили 50, а из оставшихся 900 покупку осуществили 360 человек.
- Оценим средний эффект воздействия, то есть насколько в среднем возросла вероятность покупки благодаря предоставлению предложения:

$$\widehat{ATE} = 50/100 - 360/900 = 0.1$$

- Протестируем гипотезу $H_0 : ATE = 0$ против альтернативы $H_1 : ATE > 0$ с помощью теста о разнице долей, тестовая статистика которого, при верной нулевой гипотезе, в асимптотике имеет стандартное нормальное распределение:

$$Z = \frac{0.1}{\sqrt{0.410 \times (1 - 0.410)(1/100 + 1/900)}} \approx 1.93 \quad p\text{-value} = 1 - \Phi(Z) \approx 0.03$$

Допущение об условной независимости

Формулировка

- **Проблема** – если допущение о независимости не соблюдается, то введенная ранее оценка \widehat{ATE} окажется несостоятельной, поскольку:

$$\widehat{ATE} \xrightarrow{P} E(Y_{1i}|T_i = 1) - E(Y_{0i}|T_i = 0) \neq \underbrace{E(Y_{1i}) - E(Y_{0i})}_{ATE}$$

- Обычно оно нарушается в неконтролируемых экспериментах, например, когда имеется самоотбор в число тех, кто решил принять участие в новой программе лояльности магазина или вследствие конкурсного отбора в университеты.
- **Решение** – ввести менее сильное допущение, чем допущение о независимости.
- Введем **допущение об условной независимости** (conditional mean independence), согласно которому:

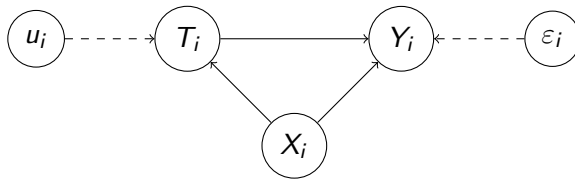
$$E(Y_{1i}|X_i, T_i = 1) = E(Y_{1i}|X_i) \quad E(Y_{0i}|X_i, T_i = 0) = E(Y_{0i}|X_i)$$

- Например, мы можем предположить, что при прочем равном возрасте и интеллекте X_i информация о полученном уровне образования T_i не влияет на наши ожидания по поводу того, сколько бы индивид зарабатывал в случаях, когда у него есть Y_{1i} и когда у него нет Y_{0i} высшего образования.
- В приведенной формулировке допущение об условной независимости также подразумевает **предпосылку об общем носителе**, согласно которой $0 < P(T_i = 1|X_i) < 1$. В противном случае X_i может принимать такие значения, при которых события $T_i = 1$ и $T_i = 0$ невозможны.

Допущение об условной независимости

Интерпретация

- Обычно допущение об условной независимости соблюдается, когда X_i отражает все факторы, которые могут быть статистически связаны и с T_i , и с Y_{ji} .



- Предполагается, что связь между T_i и Y_i обусловлена наблюдаемыми в данных переменными X_i , именуемыми **смешивающими** (confounders).
- Прерывистыми линиями отображены связи T_i и Y_i с агрегированными ненаблюдаемыми переменными u_i и ε_i .

Допущение об условной независимости

Пример

- Представим, что мы оцениваем эффект воздействия высшего образования T_i на зарплату Y_i .
- В состоянии без высшего образования индивид зарабатывает $Y_{0i} = 1$.
- Зарплата при наличии высшего образования описывается как:

$$Y_{1i} = 1 + X_i \implies TE_i = 1 + X_i - 1 = X_i \implies ATE = E(X_i) = 0.5$$

Где X_i отражает любовь индивида к обучению, причем $P(X_i = 0) = P(X_i = 1) = 0.5$.

- Вероятность получения образования зависит **только** от любви к обучению:

$$P(T_i = 1|X_i = 1) = 0.75 \quad P(T_i = 1|X_i = 0) = 0.5$$

- Допущение о независимости не соблюдается, так как $E(Y_{1i})$ и $E(Y_{1i}|T_i = 1)$ не совпадают:

$$E(Y_{1i}) = 1 + E(X_i) = 1 + 0.5 = 1.5$$

$$E(Y_{1i}|T_i = 1) = 1 + 1 \times P(X_i = 1|T_i = 1) + 0 \times P(X_i = 0|T_i = 1) =$$

$$1 + \frac{P(T_i = 1|X_i = 1)P(X_i = 1)}{P(T_i = 1|X_i = 1)P(X_i = 1) + P(T_i = 1|X_i = 0)P(X_i = 0)} = 1 + \frac{0.75 \times 0.5}{0.75 \times 0.5 + 0.5 \times 0.5} = 1.6$$

- Однако, соблюдается допущение об условной независимости, поскольку Y_{0i} константа, а также:

$$E(Y_{1i}|X_i, T_i = 1) = 1 + E(X_i|X_i, T_i = 1) = 1 + X_i = E(Y_{1i}|X_i)$$

Допущение об условной независимости

Оценивание с помощью условных математических ожиданий

- При соблюдении допущения об условной независимости:

$$\begin{aligned} \text{ATE} &= E(Y_{1i}) - E(Y_{0i}) = E(E(Y_{1i}|X_i) - E(Y_{0i}|X_i)) = \\ &= E(E(Y_{1i}|X_i, T_i = 1) - E(Y_{0i}|X_i, T_i = 0)) = \\ &= E(E(T_i Y_{1i} + (1 - T_i) Y_{0i}|X_i, T_i = 1) - E(T_i Y_{1i} + (1 - T_i) Y_{0i}|X_i, T_i = 0)) = \\ &= E(E(Y_i|X_i, T_i = 1) - E(Y_i|X_i, T_i = 0)) \implies \frac{1}{n} \sum_{i=1}^n E(Y_i|X_i, T_i = 1) - E(Y_i|X_i, T_i = 0) \xrightarrow{p} \text{ATE} \end{aligned}$$

- **Проблема** – условные математические ожидания $E(Y_i|X_i, T_i)$ неизвестны.
- **Решение** – поскольку Y_i всегда наблюдается в данных, то можно получить состоятельные оценки $\hat{E}(Y_i|X_i, T_i)$, например, **методами машинного обучения**.
- В итоге средний эффект воздействия оценивается как:

$$\widehat{\text{ATE}} = \frac{1}{n} \sum_{i=1}^n \hat{E}(Y_i|X_i, T_i = 1) - \hat{E}(Y_i|X_i, T_i = 0)$$

Допущение об условной независимости

Оценивание с помощью взвешивания на обратные условные вероятности

- Обратим внимание, что при соблюдении допущения об условной независимости:

$$\begin{aligned} E(T_i Y_i / P(T_i = 1 | X_i) | X_i) &= \underbrace{E(Y_{1i} / P(T_i = 1 | X_i) | X_i, T_i = 1) P(T_i = 1 | X_i)}_{\text{формула полной вероятности, второе слагаемое обнуляется}} = E(Y_{1i} | X_i) \implies \\ \implies E(T_i Y_i / P(T_i = 1 | X_i)) &= E(E(T_i Y_i / P(T_i = 1 | X_i) | X_i)) = E(E(Y_{1i} | X_i)) = E(Y_{1i}) \end{aligned}$$

- По аналогии можно показать, что при допущении об условной независимости:

$$E((1 - T_i) Y_i / (1 - P(T_i = 1 | X_i))) = E(Y_{0i})$$

- В итоге получаем:

$$ATE = E(T_i Y_i / P(T_i = 1 | X_i)) - E((1 - T_i) Y_i / (1 - P(T_i = 1 | X_i)))$$

- Из полученных результатов следует альтернативный способ оценивания АТЕ, именуемый **оценкой с помощью взвешивания на обратные условные вероятности** (inverse probability weighting):

$$\widehat{ATE} = \frac{1}{n} \sum_{i=1}^n \frac{T_i Y_i}{\hat{P}(T_i = 1 | X_i)} - \frac{(1 - T_i) Y_i}{1 - \hat{P}(T_i = 1 | X_i)}$$

- Преимущество** – достаточно с помощью **методов машинного обучения** оценить $\hat{P}(T_i = 1 | X_i)$.

Допущение об условной независимости

Двойная устойчивость

- Мы рассмотрели два способа оценивания АТЕ при допущении об условной независимости, первый из которых опирается на оценки $E(Y_i|X_i, T_i)$, а второй – на оценки $P(T_i = 1|X_i)$.
- Проблема** – точность оценок каждого из этих способов зависит от точности оценок соответствующих условных математических ожиданий или вероятностей. Если они оценены неточно, то и итоговая оценка АТЕ также будет неточной.
- Решение** – обеспечить **двойную устойчивость**, то есть совместить оба способа, чтобы оценка АТЕ оказывалась состоятельной, если по крайней мере один из них дает состоятельную оценку.

$$\begin{aligned}\widehat{ATE} &= \\&= \underbrace{\frac{1}{n} \sum_{i=1}^n \hat{g}_{1i} - \hat{g}_{0i}}_{\text{стремится к АТЕ если } \hat{E}(Y_i|X_i, T_i) \text{ состоятельна}} + \underbrace{\frac{1}{n} \sum_{i=1}^n \frac{T_i(Y_i - \hat{g}_{1i})}{\hat{g}_{Ti}} - \frac{(1 - T_i)(Y_i - \hat{g}_{0i})}{1 - \hat{g}_{Ti}}}_{\text{стремится к 0 если } \hat{E}(Y_i|X_i, T_i) \text{ состоятельна}} = \\&= \underbrace{\frac{1}{n} \sum_{i=1}^n \frac{T_i Y_i}{\hat{g}_{Ti}} - \frac{(1 - T_i) Y_i}{1 - \hat{g}_{Ti}}}_{\text{стремится к АТЕ если } \hat{P}(T_i = 1|X_i) \text{ состоятельна}} + \underbrace{\frac{1}{n} \sum_{i=1}^n \frac{(T_i - \hat{g}_{Ti}) \hat{g}_{0i}}{1 - \hat{g}_{Ti}} - \frac{(T_i - \hat{g}_{Ti}) \hat{g}_{1i}}{\hat{g}_{Ti}}}_{\text{стремится к 0 если } \hat{P}(T_i = 1|X_i) \text{ состоятельна}} \\&\quad \hat{g}_{1i} = \hat{E}(Y_i|X_i, T_i = 1) \quad \hat{g}_{0i} = \hat{E}(Y_i|X_i, T_i = 0) \quad \hat{g}_{Ti} = \hat{P}(T_i = 1|X_i)\end{aligned}$$

Допущение об условной независимости

Двойное машинное обучение

- Средний эффект воздействия можно также оценить с помощью DML метода, рассмотрев уравнение:

$$Y_i = g(T_i, X_i) + \varepsilon_i \quad ATE = E(g(1, X_i) - g(0, X_i))$$

- Допущение об условной независимости можно сформулировать как $E(\varepsilon_i | X_i, T_i) = 0$.

- Рассмотрим вклад, удовлетворяющий условию ортогональности по Нейману:

$$\psi = \frac{T_i(Y_i - g_1(X_i))}{g_T(X_i)} - \frac{(1 - T_i)(Y_i - g_0(X_i))}{1 - g_T(X_i)} + g_1(X_i) - g_0(X_i) - ATE$$

$$g_1(X_i) = E(Y_i | X = x_i, T_i = 1), \quad g_0(X_i) = E(Y_i | X = x_i, T_i = 0), \quad g_T(X_i) = P(T_i = 1 | X_i)$$

- Решая $E(\psi) = 0$ для АТЕ, а также подставляя оценки (полученные с помощью машинного обучения) неизвестных функций и применяя кросс-фиттинг, получаем оценку, обладающую свойством двойной устойчивости:

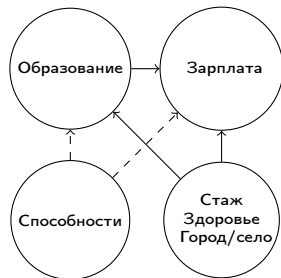
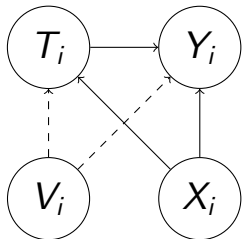
$$\widehat{ATE} = \frac{1}{n} \sum_{i=1}^n \frac{T_i(Y_i - \hat{g}_1^{(q_i)}(X_i))}{\hat{g}_T^{(q_i)}(X_i)} - \frac{(1 - T_i)(Y_i - \hat{g}_0^{(q_i)}(X_i))}{1 - \hat{g}_T^{(q_i)}(X_i)} + \hat{g}_1^{(q_i)}(X_i) - \hat{g}_0^{(q_i)}(X_i)$$

- Функции $\hat{g}_1^{(k)}$, $\hat{g}_0^{(k)}$ и $\hat{g}_T^{(k)}$ оцениваются с помощью машинного обучения на данных, не вошедших в k -ю из K выборок. Также, $q_i = k$, если наблюдение i не вошло в k -ю выборку.

Допущение об экзогенности инструментальной переменной

Эндогенность

- **Проблема** – на практике допущение об условной независимости часто нарушается вследствие **эндогенности**, из-за чего описанные ранее методы обычно дают несостоятельные оценки эффектов воздействия.



- Эндогенность обычно возникает из-за наличия ненаблюдаемых характеристик V_i , одновременно влияющих и на целевую переменную Y_i , и на вероятность воздействия $T_i = 1$.
- **Решение** – адаптировать метод инструментальных переменных.

Допущение об экзогенности инструментальной переменной

Локальный средний эффект воздействия

- Рассмотрим случай, когда эндогенный регрессор T_i и инструментальная переменная Z_i являются бинарными переменными, отражающими, например, факт наличия высшего образования у индивида и его родителей соответственно. Кроме того, временно проигнорируем все остальные признаки X_i .
- По аналогии с Y_{1i} и Y_{0i} введем гипотетические состояния переменной воздействия T_{1i} и T_{0i} .

$$T_i = \begin{cases} T_{1i}, & \text{если } Z_i = 1 \\ T_{0i}, & \text{если } Z_i = 0 \end{cases} = Z_i T_{1i} + (1 - Z_i) T_{0i}$$

- Допустим **экзогенность инструмента** - вектор $(Y_{1i}, Y_{0i}, T_{1i}, T_{0i})$ и инструмент Z_i независимы.
- Выделим четыре **группы** индивидов:

	$T_{1i} = 1$	$T_{1i} = 0$
$T_{0i} = 1$	Всегда согласные (always takers)	Отрицатели (deniers)
$T_{0i} = 0$		Всегда несогласные (never takers)

- Без введения дополнительных строгих допущений, например, о том что эффект воздействия T_i на Y_i является одинаковым для всех индивидов, в общем случае оценить АТЕ не получится, но можно оценить **локальный средний эффект воздействия LATE**, отражающий АТЕ среди соблюдаателей.

$$\text{LATE} = E(Y_{1i} - Y_{0i} | \underbrace{T_{1i} > T_{0i}}_{\text{соблюдаватели}}) = \frac{E(Y_i | Z_i = 1) - E(Y_i | Z_i = 0)}{\underbrace{P(T_i = 1 | Z_i = 1) - P(T_i = 1 | Z_i = 0)}_{\text{если нет отрицателей, есть соблюдаватели и } Z_i \text{ экзогенна}})$$

- Вывод** – при отсутствии отрицателей, наличии соблюдавателей и экзогенном инструменте Z_i для оценивания LATE достаточно оценить $E(Y_i | Z_i)$ и $P(T_i | Z_i)$, для чего необходимо посчитать соответствующие доли.

Допущение об экзогенности инструментальной переменной

Двойное машинное обучение

- **Проблема** – без использования дополнительных наблюдаемых признаков X_i оценка LATE может иметь достаточно большую дисперсию. Кроме того, инструмент Z_i может коррелировать с некоторыми признаками X_i , что часто приводит к нарушению допущения об экзогенности.
- **Решение** – воспользоваться двойным машинным обучением, позволяющим использовать инструментальную переменную Z_i вместе с признаками X_i :

$$\begin{aligned} Y_i &= g_Y(Z_i, X_i) + \varepsilon_i^{(Y)} & T_i &= g_T(Z_i, X_i) + \varepsilon_i^{(T)} & Z_i &= g_Z(X_i) + \varepsilon_i^{(Z)} \\ g_Y(Z_i, X_i) &= E(Y_i | Z_i, X_i) & g_T(Z_i, X_i) &= E(T_i | Z_i, X_i) & g_Z(X_i) &= E(Z_i | X_i) \\ E(\varepsilon_i^{(Y)} | Z_i, X_i) &= 0 & E(\varepsilon_i^{(T)} | Z_i, X_i) &= 0 & E(\varepsilon_i^{(Z)} | X_i) &= 0 \end{aligned}$$

- Функции g_Y , g_T и g_Z оцениваются с помощью машинного обучения. При выполнении некоторых дополнительных технических условий двойное машинное обучение позволяет получить состоятельную и асимптотически нормальную оценку LATE (выражение оценки опущено для краткости).

$$\text{LATE} = \frac{E(g_Y(1, X_i)) - E(g_Y(0, X_i))}{E(g_T(1, X_i)) - E(g_T(0, X_i))}$$

- **Преимущество 1** – благодаря использованию инструментальных переменных решает проблему эндогенности, являющуюся серьезным препятствием к состоятельному оцениванию эффектов воздействия с использованием обсуждавшихся ранее методов.
- **Преимущество 2** – в отличие от классического линейного метода инструментальных переменных не опирается на допущения о линейной связи между переменными.

Условный средний эффект воздействия

Классические подходы к оцениванию

- Средний эффект воздействия не всегда достаточно информативен для принятия конкретных решений. Поэтому в качестве альтернативы часто оценивают **условный средний эффект воздействия**, что в маркетинге именуется **uplift** моделированием.

$$\text{CATE}_i = E(Y_{1i}|X_i) - E(Y_{0i}|X_i) = \underbrace{E(Y_i|X_i, T_i = 1) - E(Y_i|X_i, T_i = 0)}_{\text{допущение об условной независимости}}$$

- При **допущении об условной независимости** можно получить состоятельную оценку условного среднего эффекта воздействия:

$$\widehat{\text{CATE}}_i = \hat{E}(Y_i|X_i, T_i = 1) - \hat{E}(Y_i|X_i, T_i = 0)$$

- Можно либо оценить $\hat{E}(Y_i|X_i, T_i)$ по всей выборке (**Single-learner/S-learner**), либо отдельно $\hat{E}(Y_i|X_i, T_i = 1)$ и $\hat{E}(Y_i|X_i, T_i = 0)$ по группе воздействия и контрольной группе соответственно (**Two-learner/T-learner**).
- Однако существуют и иные, менее очевидные подходы к оцениванию CATE_i .

Условный средний эффект воздействия

Метод трансформации класса

- Скронструируем псевдоисход:

$$Y_i^* = Y_i \left(\frac{T_i}{P(T_i = 1|X_i)} - \frac{1 - T_i}{1 - P(T_i = 1|X_i)} \right)$$

- Обозначим через \hat{Y}_i^* величину Y_i^* , посчитанную с использованием $\hat{P}(T_i = 1|X_i)$ вместо $P(T_i = 1|X_i)$.
- При **соблюдении допущения об условной независимости** можно показать, по аналогии с тем, как это было сделано с оцениванием АТЕ с помощью взвешивания на обратные условные вероятности, что:

$$E(Y_i^*|X_i) = \text{CATE}_i$$

- Следовательно, для оценивания CATE_i можно воспользоваться двухшаговой процедурой, часто именуемой **методом трансформации классов**.
- **Первый шаг** – вычислить $\hat{P}(T_i = 1|X_i)$ и \hat{Y}_i^* .
- **Второй шаг** – оценить $E(\hat{Y}_i^*|X_i)$ и получить $\widehat{\text{CATE}}_i = \hat{E}(\hat{Y}_i^*|X_i)$.

Условный средний эффект воздействия

Интуиция X-learner

- **Проблема** – иногда группа воздействия может включать малое число наблюдений, что осложняет оценивание $\hat{E}(Y_i|X_i, T_i = 1)$ по данным группы воздействия.
- **Решение** – рассмотрим вспомогательную переменную:

$$D_{1i} = Y_{1i} - E(Y_i|X_i, T_i = 0)$$

- Обратим внимание, что при соблюдении **допущения об условной независимости**:

$$E(D_{1i}|X_i) = E(Y_{1i}|X_i) - E(Y_i|X_i, T_i = 0) = E(Y_{1i}|X_i) - E(Y_{0i}|X_i) = \text{CATE}_i$$

- Следовательно, $\hat{E}(D_{1i}|X)$ можно рассматривать как оценку CATE_i , полученную с помощью Y_{1i} и $\hat{E}(Y_i|X_i, T_i = 0)$, то есть без использования неэффективно оцениваемого по малому числу наблюдений группы воздействия $E(Y_i|X_i, T_i = 1)$.
- **Проблема** – в данных отсутствуют наблюдения по D_{1i} .
- **Решение** – рассмотреть $\hat{D}_{1i} = Y_{1i} - \hat{E}(Y_i|X_i, T_i = 0)$ и оценить $\widehat{\text{CATE}}_i = \hat{E}(\hat{D}_{1i}|X_i)$.

Условный средний эффект воздействия

Алгоритм X-learner

- **Первый шаг** – по аналогии с T-learner оцениваются условные математические ожидания $E(Y_i|X_i, T_i = 1)$ и $E(Y_i|X_i, T_i = 0)$.

- **Второй шаг** – рассчитываются вспомогательные переменные:

$$\hat{D}_{1i} = Y_{1i} - \hat{E}(Y_i|X_i, T_i = 0) \quad \hat{D}_{0i} = \hat{E}(Y_i|X_i, T_i = 1) - Y_{0i}$$

- **Третий шаг** – оцениваются $E(\hat{D}_{1i}|X_i)$ и $E(\hat{D}_{0i}|X_i)$.

- **Четвертый шаг** – оцениваются условные вероятности попадания в группу воздействия $P(T_i = 1|X_i)$.

- **Пятый шаг** – с помощью взвешивания оценивается условный эффект воздействия:

$$\widehat{CATE}_i = (1 - \hat{P}(T_i = 1|X_i)) \hat{E}(\hat{D}_{1i}|X_i) + \hat{P}(T_i = 1|X_i) \hat{E}(\hat{D}_{0i}|X_i)$$

- **Интуиция взвешивания** – чем меньше наблюдений в группе воздействия, тем больший вес присваивается $\hat{E}(\hat{D}_{1i}|X_i)$, который не зависит от $\hat{E}(Y_i|X_i, T_i = 1)$.

Качество оценивания условных средних эффектов воздействия

Постановка проблемы

- Для измерения качества прогнозирования условных эффектов воздействия хотелось бы использовать метрику:

$$\text{MSE}_0 = \frac{1}{n} \sum_{i=1}^n \left(\text{CATE}_i - \widehat{\text{CATE}}_i \right)^2$$

- **Проблема** – на практике CATE_i неизвестно.
- **Решение** – в качестве косвенной метрики качества прогнозов условных эффектов воздействия использовать метрику качества точности прогнозов исхода:

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n \left(Y_i - \hat{E}(Y_i | X_i, T_i) \right)^2$$

- **Проблема** – модель, точнее оценивающая Y_i , не обязательно будет точнее оценивать CATE_i . Например, даже если оценки $\hat{E}(Y_i | X_i, T_i = 1)$ и $\hat{E}(Y_i | X_i, T_i = 0)$ обладают очень большим, но примерно одинаковым смещением, то их разница, то есть $\widehat{\text{CATE}}_i$, может оказаться весьма точной оценкой, поскольку смещения сократятся.

Качество оценивания условных средних эффектов воздействия

Использование псевдоисходов

- Рассмотрим псевдоисход Y_i^* , такой, что $E(Y_i^*|X_i) = CATE_i$ и соблюдены некоторые дополнительные, технические условия. Например, можно использовать псевдоисход метода трансформации классов:

$$Y_i^* = Y_i \left(\frac{T_i}{P(T_i = 1|X_i)} - \frac{1 - T_i}{1 - P(T_i = 1|X_i)} \right)$$

- Рассмотрим следующую метрику качества:

$$MSE^* = \frac{1}{n} \sum_{i=1}^n \left(\hat{Y}_i^* - \widehat{CATE}_i \right)^2$$

- Можно показать, что при определенных условиях, в частности, больших объемах выборки, MSE^* является достаточно точной аппроксимацией MSE_0 , в том смысле, что MSE^* и MSE_0 будут схожим образом ранжировать качество моделей.
- Существует множество иных подходов к аппроксимации MSE_0 с помощью метрик, зависящих от различных псевдоисходов.

Средний эффект воздействия на подвергшихся воздействию

Формулировка

- Помимо рассмотренных, существует множество иных характеристик потенциальных исходов, оценивание которых представляет содержательный интерес.
- Например, исследователь может быть заинтересован в оценивании **среднего эффекта воздействия на подвергшихся воздействию** (ATET, ATT, ATTE - average treatment effect on treated):

$$ATET = E(Y_{1i} - Y_{0i} | T_i = 1) = E(Y_{1i} | T_i = 1) - E(Y_{0i} | T_i = 1)$$

- Например, если Y_i и T_i отражают зарплату и факт получения высшего образования соответственно, то ATET отражает средний эффект воздействия высшего образования на зарплату среди тех индивидов, которые фактически получили высшее образование $T_i = 1$.
- Общие принципы оценивания этих характеристик крайне схожи с теми, что мы рассматривали ранее. В частности, ATET можно оценивать с помощью двойного машинного обучения, в том числе с применением инструментальных переменных.