

PROPENSITY SCORE MATCHING В PYTHON: ПРАКТИЧЕСКОЕ ПРИМЕНЕНИЕ ДЛЯ БАЛАНСИРОВКИ ГРУПП ПАЦИЕНТОВ В НАБЛЮДАТЕЛЬНЫХ МЕДИЦИНСКИХ ИССЛЕДОВАНИЯХ

М.М. Мелехин¹, М.В. Комарова¹

¹Самарский университет, г. Самара
e-mail: mikhail.melehin@mail.ru

Аннотация

Данная работа посвящена применению метода сопоставления по склонности (PSM) для снижения влияния систематических различий при сравнении групп пациентов с расслоением аорты. Цель – создание сбалансированных групп для корректного сравнения эффектов лечения в наблюдательных исследованиях. В работе представлен практический цикл применения метода Propensity Score Matching в Python для балансировки групп пациентов с расслоением аорты в наблюдательных медицинских исследованиях.

Ключевые слова: Propensity Score Matching, Python, медицинские исследования, балансировка групп, расслоение аорты, биотехнические системы.

Введение

В наблюдательных медицинских исследованиях часто возникает проблема систематических различий между группами пациентов, получающих различные виды лечения или вмешательства. Эти различия, известные как смещение отбора (selection bias), могут существенно искажать результаты анализа и приводить к неверным выводам о причинно-следственных связях. Традиционные методы статистического контроля не всегда способны полностью устранить это смещение, особенно при наличии большого числа ковариат и сложных взаимосвязей между ними [1].

Метод сопоставления по склонности (Propensity Score Matching, PSM) является мощным инструментом для создания сопоставимых групп в нерандомизированных исследованиях. Он позволяет минимизировать систематические различия между группами путем сопоставления наблюдений на основе их вероятности (propensity score) получения лечения, что позволяет более надежно оценивать эффекты вмешательства. Применение PSM особенно актуально в кардиохирургии, где этические и практические ограничения часто не позволяют проводить рандомизированные контролируемые исследования.

Данный доклад посвящен практическому применению PSM в Python для балансировки групп пациентов с расслоением аорты. Целью работы является балансировка ковариат в группах кардиохирургических пациентов, получающих различное лечение.

1. Реализация метода

Propensity Score (PS) рассчитывается с помощью логистической регрессии по формуле (1), где зависимой переменной является принадлежность к группе лечения (например, группа 2), а независимыми переменными – ковариаты, по которым необходимо выровнять группы. Вероятность принадлежности к группе лечения $P(X)$ для каждого пациента i с набором ковариат X_i определяется как [2]:

$$P(X_i) = \frac{1}{1 + e^{-(\beta_0 + \sum_{j=1}^k \beta_j X_{ij})}}, \quad (1)$$

где $P(X_i)$ – вероятность получения лечения для пациента i ;

X_{ij} – значение j -й ковариаты для пациента i ;

β_0 – свободный член (intercept);

β_j – коэффициенты логистической регрессии для j -й ковариаты.

Использовались следующие ковариаты: возраст, распространение расслоения на ветви дуги, мальперфузия, критическое состояние, фракция выброса, размер дуги, пол.

После расчета Propensity Scores выполняется матчинг наблюдений из группы лечения с наблюдениями из контрольной группы на основе ближайшего соседа (nearest neighbor matching). Для обеспечения качества матчинга используется калипер (caliper), который ограничивает максимально допустимую разницу в Propensity Score между сопоставляемыми наблюдениями [3]. В данном исследовании использовался калипер, равный 0,1 от стандартного отклонения Propensity Score, что означает, что для каждого наблюдения из группы лечения ищется ближайшее наблюдение из контрольной группы, разница в propensity score между которыми не превышает 0,1. Математически условие матчинга с калипером можно выразить по формуле (2):

$$|PS_{treatment} - PS_{control}| \leq C, \quad (2)$$

где $C = 0,1 \times SD(PS)$, а $SD(PS)$ – стандартное отклонение Propensity Score.

На рисунке 1 показан алгоритм реализации матчинга, где использовалась библиотека scikit-learn, в частности, класс NearestNeighbors для поиска ближайших соседей на основе рассчитанных Propensity Scores. Дополнительно применялись функции из библиотеки pandas для эффективной работы с данными и numpy для численных операций. Матчинг был выполнен методом ближайшего соседа с соотношением 1:1 без возврата.

```
from sklearn.neighbors import NearestNeighbors
import numpy as np

caliper = 0.1 * ps.std()
nn = NearestNeighbors(n_neighbors=1, metric='euclidean')
nn.fit(control_ps.values.reshape(-1, 1))
distances, indices = nn.kneighbors(treatment_ps.values.reshape(-1, 1))

matches = indices[distances.flatten() <= caliper]
```

Рисунок 1 – Реализация алгоритма матчинга на языке Python

Данный подход позволил сформировать сопоставимые пары пациентов, сохранив при этом приемлемый размер выборки и минимизировав влияние систематических различий между группами.

После выполнения матчинга баланс ковариат между группами оценивается до и после матчинга [4]. Для этого используется стандартизированная разница (Standardized Difference), которая рассчитывается по формуле (3):

$$\text{Standardized, Difference} = \frac{(\text{Mean}_{\text{group2}} - \text{Mean}_{\text{group1}})}{\text{Pooled}_{\text{std}}}, \quad (3)$$

где Mean_G и $\text{Mean}_{\text{group}}$ – средние значения ковариаты в группах 1 и 2 соответственно;

$$\text{Pooled}_{\text{std}} = \sqrt{\frac{(\text{Std}_{\text{group1}}^2 + \text{Std}_{\text{group2}}^2)}{2}} - \text{объединенное стандартное отклонение};$$

Std_G и $\text{Std}_{\text{group}}$ – стандартные отклонения ковариаты в группах 1 и 2 соответственно.

2. Оценка баланса

Критерии оценки баланса:

Стандартизированная разница $< 0,1$ считается показателем хорошего баланса, в то время как $\text{SMD} < 0,25$ является приемлемым балансом.

Численные значения баланса (Standardized Difference) для разных ковариат до и после матчинга показаны в таблице 1:

Таблица 1 – Показатели стандартизированной разницы для набора ковариат

Ковариата	SMD (до матчинга)	SMD (после матчинга)
Возраст	0,31	0,05
Распространение расслоения на ветви дуги	1,07	0,08
Мальперфузия	0,29	0,04
Критическое состояние	0,53	0,08
Фракция выброса	0,50	0,07
Размер дуги	0,61	0,05
Пол	0,20	0,03

Размеры групп и доля исключенных наблюдений:

- размер исходной группы лечения: 150;
- размер исходной группы альтернативного лечения: 125;
- размер сопоставленной группы лечения: 125;
- размер сопоставленной группы альтернативного лечения: 125;
- доля исключенных наблюдений: 25 наблюдений (9,1%).

Заключение

В работе реализован полный цикл метода Propensity Score Matching в Python для балансировки групп пациентов с расслоением аорты. Матчинг проводился методом ближайшего соседа 1:1 с калипером $0,1 \times \text{SD}(\text{PS})$. В результате сопоставления удалось сформировать статистически сбалансированные группы: размер сопоставленных групп составил 125 пациентов в каждой (из исходных 150 и 125). Доля исключённых наблюдений составила 9,1 %. Среднее значение стандартизированной разницы (SMD) по всем ковариатам снизилось с 0,50 (до матчинга) до 0,057 (после матчинга). По всем

исследуемым ковариатам SMD после матчинга не превысил 0,10, что свидетельствует о достижении высокого уровня баланса.

Полученные результаты подтверждают эффективность метода PSM для устранения систематических различий в наблюдательных кардиохирургических данных и позволяют более надёжно оценивать эффекты лечения в нерандомизированных исследованиях.

Список использованных источников

1. Austin, P. C. An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies / P. C. Austin // *Multivariate Behavioral Research*. – 2011. – V. 46, № 3. – P. 399–424.
2. Thoemmes, F. A Systematic Review of Propensity Score Methods in the Social Sciences / F. Thoemmes, E. S. Kim // *Multivariate Behavioral Research*. – 2011. – V. 46, № 1. – P. 90–118.
3. Stuart, E. A. Matching Methods for Causal Inference: A Review and a Look Forward / E. A. Stuart // *Statistical Science*. – 2010. – V. 25, № 1. – P. 1–21.
4. Benedetto, U. Statistical primer: propensity score matching and its alternatives / U. Benedetto // *European Journal of Cardio-Thoracic Surgery*. – 2018. – V. 53, № 6. – P. 1112–1116.
5. Бокерия, Л. А. Расслоение аорты: классификация, диагностика и хирургическое лечение / Л. А. Бокерия, Р. Г. Гудкова // *Анналы хирургии*. – 2018. – № 4. – С. 5–14.
6. King, G. Why propensity scores should not be used for matching / G. King, R. Nielsen // *Political Analysis*. – 2019. – V. 27, № 4. – P. 435–454.
7. Shiba, K. Using propensity scores for causal inference: pitfalls and tips / K. Shiba // *Journal of Clinical Epidemiology*. – 2021. – V. 135. – P. 177–185.
8. Greifer, N. Matching methods for confounder adjustment / N. Greifer // *Statistical Science*. – 2021. – V. 36, № 2. – P. 277–295.

PROPENSITY SCORE MATCHING IN PYTHON: PRACTICAL APPLICATION FOR BALANCING PATIENT GROUPS IN OBSERVATIONAL MEDICAL STUDIES

M.M. Melekhin¹, M.V. Komarova¹

¹*Samara National Research University, Samara
e-mail: mikhail.melehin@mail.ru*

Abstract

This paper applies Propensity Score Matching (PSM) to reduce systematic differences when comparing patient groups with aortic dissection. The goal is to create balanced groups for correct treatment effect comparison in observational studies. The paper presents a practical implementation of the Propensity Score Matching method in Python for balancing groups of patients with aortic dissection in observational medical research.

Keywords: Propensity Score Matching, Python, medical research, group balancing, aortic dissection, biotechnical systems.