

Propensity Score Matching

Propensity Score Matching (PSM) is a statistical technique used in observational studies to reduce selection bias when estimating the effect of a treatment, intervention, or exposure. It's especially helpful when randomized controlled trials (RCTs) aren't feasible.

Instead of randomly assigning people to treatment and control groups, PSM tries to make the groups comparable based on observed characteristics.

□ What is a Propensity Score? The propensity score is the probability that a subject receives the treatment given their observed covariates.

Formally:

$$e(X) = P(\text{Treatment} = 1 \mid \text{Covariates} = X)$$

It's typically estimated using logistic regression, but other machine learning models (e.g., random forests, gradient boosting) can also be used.

⚙ Steps in Propensity Score Matching
Estimate the Propensity Score Use a statistical model to calculate the probability of receiving the treatment for each subject based on their covariates.

Match Subjects Pair each treated subject with one or more non-treated subjects with a similar propensity score. Matching methods include:

Nearest neighbor (most common)

Caliper matching (within a certain range)

Kernel matching

Mahalanobis metric matching

Assess Balance Check if covariates are balanced between treated and control groups after matching (e.g., using standardized mean differences).

Estimate Treatment Effect Compare outcomes between matched treated and control groups to estimate the Average Treatment Effect on the Treated (ATT).

□ When to Use PSM? In observational studies where randomization isn't possible.

When you have a binary treatment and want to control for confounding variables.

⚠ Limitations of PSM Only controls for observed covariates, not unmeasured confounders.

Matching quality depends on the model used to estimate the propensity score.

Can discard a lot of data if there are poor matches (reducing statistical power).

Propensity Score Matched Analysis is a statistical technique used in a **observational study** to reduce

selection bias when comparing treated and untreated groups. Since randomized controlled trials (RCTs) are often infeasible due to ethical or practical constraints, PSM mimics some aspects of randomization by ensuring that the treatment and control groups are balanced in terms of observed confounders.

Key Steps in PSM Analysis Estimate the Propensity Score

The propensity score is the probability of receiving the treatment given a set of observed covariates. It is typically estimated using logistic regression, probit regression, or machine learning methods like random forests.

Match Treated and Control Units

Common matching techniques include: Nearest Neighbor Matching (1:1 or 1:k) Caliper Matching (only matches within a certain range) Stratification (grouping by propensity score quantiles) Inverse Probability of Treatment Weighting (IPTW) Assess Balance Between Groups

Before and after matching, check whether covariates are balanced using standardized mean differences (SMDs). Ideally, after matching, SMDs should be close to 0 (< 0.1 is often considered acceptable). Perform Outcome Analysis

Once matched, compare the outcomes between the treated and control groups using statistical tests such as t-tests, chi-square tests, or regression models. Sensitivity Analysis

Since PSM relies on observed covariates, unmeasured confounding can still bias results. Sensitivity analyses (e.g., Rosenbaum bounds) help assess the impact of potential hidden bias. Advantages of PSM Reduces bias from confounding variables. Mimics randomization when RCTs are not feasible. Can improve causal inference in observational studies. Limitations of PSM Cannot adjust for unmeasured confounders. Requires a sufficiently large dataset for good matching. Some data loss occurs if unmatched units are discarded.

Propensity score-matched **analysis** is a statistical **method** used to reduce the effects of **confounding** variables in a **observational study** or quasi-experimental research. When conducting such studies, researchers may not be able to control for all possible variables that could influence the outcome of interest. As a result, the observed associations between variables may be **biased** or misleading.

The **propensity score** is the probability of an individual being assigned to a particular treatment group, given their observed characteristics. In other words, it estimates the likelihood that an individual received a certain treatment based on their **baseline** characteristics. The goal of propensity score-matched analysis is to create comparable groups of treated and control individuals with similar propensity scores, thus approximating a randomized controlled trial where treatment assignment is random.

Here's a step-by-step explanation of how propensity score-matched analysis is typically performed:

Propensity Score Estimation: The first step is to estimate the propensity scores for each individual in the dataset. This is usually done using logistic regression, where the treatment (or exposure) variable is regressed on a set of observed baseline characteristics or covariates.

Matching: After obtaining the propensity scores, individuals in the treated group are matched with individuals in the control group based on their propensity scores. The matching process aims to create pairs or groups of individuals with similar propensity scores.

One-to-One Matching: In this approach, each treated individual is matched with one control individual with the closest propensity score.

Multiple-to-One Matching: In this approach, multiple control individuals are matched to each treated individual to improve the balance of covariates.

Assessing Balance: After the matching process, it is essential to assess the balance of covariates between the treated and control groups. The goal is to achieve similarity in the distribution of covariates after matching.

Outcome Analysis: Once matched, the outcome of interest (e.g., effectiveness of treatment) can be compared between the treated and control groups. Since the propensity scores have balanced the covariates between the groups, the analysis is less biased and more akin to a randomized controlled trial.

Propensity score-matched analysis is a valuable tool in observational studies because it allows researchers to approximate the benefits of randomization and strengthen the validity of their findings. However, it is essential to acknowledge that propensity score matching is just one of many methods used to address confounding, and researchers should always consider the limitations and assumptions associated with this approach. Sensitivity analyses and other robustness checks are often performed to assess the robustness of the results obtained through propensity score-matched analysis.

Propensity score matching (PSM) is a quasi-experimental method in which the [researcher](#) uses statistical techniques to construct an artificial control group by matching each treated unit with a non-treated unit of similar characteristics. Using these matches, the researcher can estimate the impact of an intervention.

Matching, in general, can be a problematic method because it discards units, can change the target estimand, and is nonsmooth, making inference challenging. Using propensity scores to match adds additional problems.

The [propensity score](#) is a concept used in [statistical](#) and epidemiological research to address confounding bias in observational studies. In observational studies, researchers cannot randomly assign individuals to different groups (treatment and control) as they would in a randomized controlled trial. This can lead to potential bias when trying to assess the causal effect of an intervention or exposure on an outcome.

The propensity score is defined as the conditional probability of an individual receiving a particular treatment or exposure given their observed baseline characteristics or covariates. In other words, it quantifies the likelihood that a participant is assigned to a specific treatment group based on their characteristics. The propensity score is denoted by the symbol " $P(X)$ " or " $e(X)$ ", where X represents

the vector of observed covariates.

The process of estimating the propensity score involves constructing a predictive model, typically using logistic regression, where the treatment or exposure variable is the outcome, and the covariates are the predictors. The logistic regression model yields the propensity scores, which range from 0 to 1.

Once the propensity scores are obtained, researchers can use them to control for confounding in several ways:

Propensity Score Matching: As explained in the previous answer, researchers can match individuals with similar propensity scores between treatment and control groups. This creates comparable groups, effectively reducing the impact of confounding variables.

Propensity Score Weighting: Another approach is to assign weights to each participant based on their propensity scores. Participants with extreme propensity scores are given lower weights, which means their contribution to the analysis is reduced.

Stratification: Researchers can divide the study population into strata based on their propensity scores and then analyze the treatment effect within each stratum. This helps ensure that within each stratum, the confounding factors are balanced.

Propensity Score Adjustment: Propensity scores can be included as a covariate in the outcome analysis to adjust for the effect of the treatment while controlling for confounding.

By using propensity scores, researchers aim to make the treatment and control groups more comparable, making the analysis more robust and reducing the bias caused by confounding variables. However, it is important to note that propensity score methods rely on the assumption of no unmeasured confounding, which cannot be tested directly and should be considered when interpreting the results. Sensitivity analyses are often conducted to assess the impact of unmeasured confounding on the study's conclusions.

In observational studies, generalized propensity score (GPS)-based statistical methods, such as [inverse probability weighting](#) (IPW) and [doubly robust](#) (DR) method, have been proposed to estimate the [average treatment effect](#) (ATE) among multiple treatment groups.

Yan et al. investigated the GPS-based statistical methods to estimate treatment effects from two aspects. The first aspect of the investigation is to obtain an optimal GPS estimation method among four competing GPS estimation methods by using a rank aggregation approach. We further examine whether the optimal GPS-based IPW and DR methods would improve the performance for estimating ATE. It is well known that the DR method is consistent if either the GPS or the outcome models are correctly specified. The second aspect of our investigation is to examine whether the DR method could be improved if we ensemble outcome models. To that end, the bootstrap method and rank aggregation method is used to obtain the ensemble optimal outcome model from several competing outcome models, and the resulting outcome model is incorporated into the DR method, resulting in an ensemble DR (enDR) method. Extensive simulation results indicate that the enDR method provides the best performance in estimating the ATE regardless of the method used for estimating GPS. We illustrate our methods using the MarketScan healthcare insurance claims database to examine the treatment effects among three different bones and substitutes used for spinal fusion surgeries. We draw conclusions based on the estimates from the enDR method coupled with the optimal GPS

estimation method ¹⁾.

¹⁾

Yan X, Abdia Y, Datta S, Kulasekera KB, Ugiliweneza B, Boakye M, Kong M. Estimation of average treatment effects among multiple treatment groups by using an ensemble approach. Stat Med. 2019 Jul 10;38(15):2828-2846. doi: 10.1002/sim.8146. Epub 2019 Apr 2. PubMed PMID: 30941812.

From:

<https://neurosurgerywiki.com/wiki/> - **Neurosurgery Wiki**

Permanent link:

https://neurosurgerywiki.com/wiki/doku.php?id=propensity_score_matching

Last update: **2025/04/03 21:32**

