WeightIt

estimand

the desired estimand. For binary and multi-category treatments, can be "ATE", "ATT", "ATC", and, for some methods, "ATO", "ATM", or "ATOS". The default for both is "ATE". This argument is ignored for continuous treatments. See the individual pages for each method for more information on which estimands are allowed with each method and what literature to read to interpret these estimands.

focal

when multi-category treatments are used and ATT weights are requested, which group to consider the "treated" or focal group. **This group will not be weighted, and the other groups will be weighted to be more like the focal group.** If specified, estimand will automatically be set to "ATT".

Methods:

**cbps: Covariate Balancing Propensity Score Weighting**

In general, this method relies on estimating propensity scores using generalized method of moments and then converting those propensity scores into weights using a formula that depends on the desired estimand

Description

Multinomial Treatments

**For multinomial treatments with three or four categories and when the estimand is the ATE, this method estimates the propensity scores and weights** using one call to CBPS::CBPS(). For multinomial treatments with three or four categories or when the estimand is the ATT, this method estimates the propensity scores and weights using multiple calls to CBPS::CBPS(). The following estimands are allowed: ATE and ATT. The weights are taken from the output of the CBPS fit objects.

----------------------------------------------------------------------

**ps: Propensity Score Weighting Using Generalized Linear Models**

In general, this method relies on estimating propensity scores with a parametric generalized linear model and then converting those propensity scores into weights using a formula that depends on the desired estimand. For binary and multinomial treatments, a binomial or multinomial regression model is used to estimate the propensity scores as the predicted probability of being in each treatment given the covariates. For ordinal treatments, an ordinal regression model is used to estimate generalized propensity scores. For continuous treatments, a generalized linear model is used to estimate generalized propensity scores as the conditional density of treatment given the covariate

Multinomial Treatments

For multinomial treatments, the propensity scores are estimated using multinomial regression from one of a few functions depending on the requested link: for logit ("logit") and probit ("probit") links, mlogit::mlogit() from the mlogit package is used; for the Bayesian probit ("bayes.probit") link, MNP::mnp() from the MNP package is used; and for the biased-reduced multinomial logistic regression ("br.logit"), brglm2::brmultinom() from the brglm2 package is used. **If the treatment variable is an ordered factor, MASS::polr() from the MASS package is used to fit ordinal regression** unless link = "br.logit", in which case brglm2::bracl() from brglm2 is used. Any of the methods allowed in the method argument of polr() can be supplied to link. **The following estimands are allowed: ATE, ATT, ATC**, ATO, and ATM. The weights for each estimand are computed using the standard formulas or those mentioned above. Weights can also be computed using marginal mean weighting through stratification for the ATE, ATT, and ATC. See get\_w\_from\_ps() for details.

**gbm: Propensity Score Weighting Using Generalized Boosted Models**

In general, this method relies on estimating propensity scores using generalized boosted modeling and then converting those propensity scores into weights using a formula that depends on the desired estimand. The algorithm involves using a balance-based or prediction-based criterion to optimize in choosing the value of tuning parameters (the number of trees and possibly others). The method relies on the gbm package.

Multinomial Treatments

For multinomial treatments, this method estimates the propensity scores using gbm::gbm.fit() with distribution = "multinomial" and then selects the optimal tuning parameter values using the method specified in the stop.method argument. **The following estimands are allowed: ATE, ATT, ATC**, ATO, and ATM. The weights are computed from the estimated propensity scores using get\_w\_from\_ps(), which implements the standard formulas. Weights can also be computed using marginal mean weighting through stratification for the ATE, ATT, and ATC. See get\_w\_from\_ps() for details.

**The Multiple Propensity Score as Control for Bias in the Comparison of More Than Two Treatment Arms An introduction From a Case Study in Mental Health Marieke Dingena Spreeuwenberg 2010**

“For nominal treatments, as is often the case in mental health research, Imbens12 suggested the use of multiple PS, defined as the conditional probability of receiving a par ticular level of the treatment given a set of observed pretreatment variables. The multiple PS can be estimated with a multinomial logistic or probit regression. Here, for each subject, the probability of receiving each treatment category given the observed covariates is estimated”

“In the case when treatment categories are defined by an ordinal value, such as treatment dosage, ordinal logistic regression can be used as an alternative estimation method.18'3”

**A Tutorial on Propensity Score Estimation for Multiple Treatments Using Generalized Boosted Models Daniel F. McCaffrey 2013**

“. GBM estimation involves an iterative process with multiple regression trees to capture complex and nonlinear relationships between treatment assignment and the pretreatment covariates without over-fitting the data”

“The choice of estimand depends on the substantive questions a study hopes to address and the population that is the target of the treatment. A study can estimate both ATE and ATT, but one or the other typically is better suited for any particular situation. The ATEs are more likely to be of interest compared with ATTs if every treatment potentially might be offered to every member of the population.”

“Conversely, if the research question focuses on the effectiveness of one treatment program t ″ then the ATT, μt′,t′ − μt′,t″, would be of interest because it measures the relative effectiveness of programs t′ and t″ on the population receiving program t’ ”

“Multinomial logistic regression is a very commonly used approach to modeling the relationship between covariates and outcomes that take on a small number of discrete values, like assignment to one of three treatment conditions, and has been proposed for estimating propensity scores with multiple treatments. It models the probability that an outcome (e.g., treatment assignment, Ti ) equals each of its possible values as a function of a linear combination of the covariates and their products and cross products:”

“The challenge for propensity score estimation is choosing the correct set of interactions and polynomial terms among the covariates to capture any nonlinearities in their relationship to treatment assignment. There is no standard method for model selection in the context of estimating propensity scores for IPTW for multiple treatments. There is not even a standard method for estimating propensity scores for weighting with two treatments”

“GBM can mitigate both these challenges to using multinomial regression. It has automated variable selection [5] and in simulations has proven to provide more stable weights than parametric models”

“4.1.2. Using GBM to Estimate ATEs of Multiple Treatments—Given that GBM has proven to be effective in studies of two treatments, we propose an extension of the method for estimating propensity score weights when there are more than two treatments. Specifically, when interest lies in estimating the ATEs, we propose using GBM in the following fashion to obtain weights:”

**“**. The ATE weights that resulted from fitting the multinomial logistic model were very unstable and yielded very poor balance. Trimming the weights [19] or removing records with outlier weights improved balance but it remained inferior to the balance obtain using GBM. For ATT, the methods were more comparable but GBM continued to provide better balance. We note that 10% of the records in the data had missing values on one or more covariates. GBM automatically adds indicators for missing values and includes them in the model. Multinomial logistic regression requires a manual fix, such imputation, creating missing data indicators, or dropping incomplete records. When applying the parametric approach, we dropped incomplete records for this illustrative analysis.”

**“**In estimating the multiple treatment propensity score weights, a powerful machine learning method, GBM, was used to obtain robust propensity score weights with better balance properties than a simple parametric model (namely the multinomial logistic) did.”

**Utilizing Propensity Score Methods for Ordinal Treatments and Prehospital Trauma Studies**

**Greene, Thomas J.   The University of Texas School of Public Health ProQuest Dissertations Publishing,  2017. 10681743.**

“As a non parametric alternative, one could utilize a Generalized Boosted Model (GBM), an iterative tree-based machine learning algorithm, to estimate the probability of treatment assignment for binary or multiple treatments”

“Although they have not been well-studied and have not to our knowledge been studied for ordinal treatments, GBM and other non parametric propensity modeling approaches provide notable benefits over parametric models”

“Since GBM is tree-based, there is no issue with excluding individuals with missing covariates data since missing values are accommodated by the model. Variable selection as well as the decision to include higher order and interaction terms are automatically incorporated into the propensity model.”

2.4 Simulation Results

“Stratification and GBM weighting result in the lowest standard errors overall, across all settings. ”

“In general, and as expected, as the severity of departure from proportional odds [assumption] increases, mean bias and standard error increase for all methods (though very marginally for stratification), and coverage probabilities tend to decrease within each method. The most drastic decrease in coverage probability with severe departures from proportional odds is seen with weighting, reaching only 79%. Table 2.2 suggests that stratification and GBM weighting are most robust to model violations”

**A machine learning compatible method for ordinal propensity score stratification and matching Thomas J. Greene 2020**

“Although machine learning techniques that estimate propensity scores for observational studies with multivalued treatments have advanced rapidly in recent years, the development of propensity score adjustment techniques has not kept pace. While machine learning propensity models provide numerous benefits, they do not produce a single variable balancing score that can be used for propensity score stratification and matching. This issue motivates the development of a flexible ordinal propensity scoring methodology that does not require parametric assumptions for the propensity model. The proposed method fits a one-parameter power function to the cumulative distribution function (CDF) of the generalized propensity score (GPS) vector resulting from any machine learning propensity model, and is henceforth called the GPS-CDF method. The estimated parameter from the GPS-CDF method, ã, is a scalar balancing score that can be used to group similar subjects in outcome analyses. Specifically, subjects who received different levels of the treatment are stratified or matched based on their ã value to produce unbiased estimates of the average treatment effect (ATE)”

“Although nonrandomized observational studies with ordinal treatments are commonly encountered in public health research including mental health, substance use, and program evaluation,1,2 propensity scoring methods for ordinal treatments remain underdeveloped. Generally, ordinal treatments refer to treatment settings with three or more treatment levels with a defined ordering (eg, treatment dosages, levels of an environmental exposure). Since binary propensity scoring methods are very well established, researchers often disregard the ordered nature of the treatment and simply dichotomize the ordinal treatment or evaluate it as a multinomial treatment”

“The only currently well-established propensity score model for ordinal treatments uses the proportional odds (PO) model to model treatment assignment as a function of treatment-related baseline covariates.1,7 However, the PO assumption is easily violated in real-world data. One notable example is the work of Cavasos-Rehg et al (2014), that sought to analyze the effect of ordinal smoking reduction on mood and anxiety disorders, alcohol use, and drug use.2 The PO propensity model did not satisfy the proportional odds assumption—thus the authors were forced to classify the exposure as multinomial, use a multinomial specific propensity model, and ignore the natural ordering of the exposure variable. The clear limitations that result from propensity model misspecification in the ordinal treatment setting motivates the need for more robust ordinal propensity scoring methods”

“Flexible machine learning models, such as the well-studied generalized boosted model (GBM),25-29 are rapidly replacing traditional logistic regression-based methods as the preferred tool for estimating propensity scores in observational studies since machine learning methods have the ability to handle high-dimensional covariate spaces, automatically select higher order terms and interactions, and down-weight uninformative covariates in the propensity model. Additionally, they do not require a priori assumptions about the true underlying form of the propensity model, and are more robust to propensity model misspecification that logistic regression methods.

Furthermore, in the spirit of the covariate balancing propensity score (CBPS) of Imai and Ratkovic (2014), machine learning methods can automatically produce a GPS vector with optimal balancing score properties, resulting in more precise estimates of treatment effect.14,29 Although machine learning methods have numerous benefits, they do not naturally produce a scalar balancing score. Therefore, their application has been limited within multivalued treatment propensity score stratification and matching, instead focusing primarily on propensity adjustment using IPTW. As the popularity of multivalued propensity score estimation using machine learning grows, stratification and matching methods for ordinal treatments need to be adapted”

“Currently generalized propensity scores estimated from nonparametric models are implemented using IPTW.25,26,29,37,61-63 There are several reasons to continue developing multivalued treatments stratification and matching methods rather than solely relying on IPTW. …. Therefore, in situations likely to produce large highly variable weights, it would be preferable to estimate propensity scores using a nonparametric machine learning algorithm, then adjust for these scores using a stratification or matching procedure”

“This article shows the GPS-CDF method is a flexible, straightforward, and intuitive method of removing covariate imbalance in observational studies with ordinal treatments. The approach does not rely on the proportional odds model; in fact it can be used with any parametric or nonparametric propensity model. The GPS-CDF method provides many opportunities for future research including extensions to continuous and multinomial treatments, applications to public health, genetics, and electronic health records datasets. Hopefully, continued development in the field of multivalued treatments propensity scoring will encourage researchers to utilize these methods in practice.”