The use of Propensity Scores in Observational Studies

Shannon Lane*[1] and Marjorie Rosenberg†[2]

[1]College of St. Benedict
[2]University of Wisconsin-Madison

October 27, 2013

Abstract

Propensity scores have been widely used in efforts to balance the distribution of the covariates in observational studies. This paper uses (**?**) (**?**)the stratification propensity score method and compares it to a regression method. These methods use data from the 2009 Medical Expenditure Panel Study (MEPS) in attempt to balance the covariates and determine differences among the methods.

1 Introduction

To understand the use of propensity scores in observational studies, the 2009 Medical Expenditure Panel Study (MEPS) (Cite). MEPS is a nationwide survey composed of families and individuals, information from their medical providers and employers. The data are on specific health care services used by Americans and how frequently they are used, the cost and the source of payment. Different covariates are used from the data, in order to help identify the question addressed.

Take for example an employer wants to decrease the overall health care expenditures within the employee pool at their company. The employer believes by implementing a weight loss program targeted at overweight employees, the employees would lose weight resulting in the reduction of weight related healthcare costs. The variable Body Mass Index (BMI) is a measurement of obesity. The treatment group in this scenario is individuals with high BMI, and the control group is individuals with healthy BMI. The groups differ in age, sex, education, and other information that could be associated with higher or lower BMI. So how can the employer determine if the program is effective of not?

To determine the success of interventions in health care studies, randomized clinical trials have been the gold standard for years (**?**). To produce reliable results in clinical trials, groups such as a control and treatment are balanced through the randomization of the subjects based on pre-determined covariates. A randomized clinical trial takes a great deal of time, money, and personnel. In hope to reduce these observational studies have become increasingly more researched and understood.

The BMI example is an example of an obser-

*selane@csbsju.edu
†mrosenberg@bus.wisc.edu

vational study. Observational studies are similar to clinical trials but the researcher lacks control over randomizing and balancing the groups (**?**). This leads to potential biases and varying results. In efforts to reduce bias and balance the groups propensity scores are used. Propensity scores are a conditional probability that balances the distribution of the covariates and results in a single scalar (**?**). This allows for comparisons between the control and treated groups to be concluded.

The propensity score is defined by the conditional probability of treatment given baseline covariates

$$e(X) = pr(Z = 1|X), \tag{1}$$

where e(X) is the predicted propensity score. If there are subjects with the same propensity score, then the distribution of the baseline covariates must be the same between the treatment and control subjects (**?**).

Propensity scores have been the leading technique in observational studies to reduce bias and increase precisions (D'Ago). One type of bias discussed in Cook and DeMets' (2008) is selection bias, the subject selects to partake in a study. Propensity scores have the ability to balance the distribution, which reduces selection bias by making groups equal based on baseline covariates.

There are four main propensity score techniques: matching, stratification, inverse weighted treatment probability and covariate regression adjustment. In this paper, the stratification method will be used and analyzed; this method will be further described in the method section.

The purpose of this paper is to be able to compare the stratification propensity score method to a basic regression model using a propensity score as a covariate. In the next section, Method and Data, making the data set, stratification and the regression model will be described and the raw data will be described. The data will be analyzed in the Results section, followed by the Conclusion.

2 Methods

This research was conducted with the use of R, an open-source statistical computing software (cite). The major benefit of open-source is that since it is license free, and the community of users contribute to make it better.

In this paper we specifically compared the natural log of expenditures of obese and non-obese subjects in ten strata, to a regression model of the natural log of expenditures on obese subjects and propensity scores.

Obesity was determined by subjects BMI. In the MEPS data, BMI was calculated based on,

$$BMI = \frac{WeightinPounds}{(HeightinInches)^2} * 703\ (cite). \tag{2}$$

The BMI categories are:

Classification	BMI(kg/m^2)
Underweight	Less Than 18.5
Normal Weight	18.5 – 24.9
Overweight	25.0 – 29.9
Obesity	Greater Than or Equal to 30.0

For the treatment group, subjects are obese, thus are over 30.0 kg/m^2. The control groups is not obese subjects, and are under 29.9 kg/m^2.

For each calendar year, there are two panels in the study. For 2009 there was Panel 13 and

Panel 14. For our research, we only examined Panel 13.

2.1 Making the Data Set

Using the 2009 MEPS data set that contained 36,855 subjects and 1,881 covariates, we decided to reduce the data set and make a new data frame that could aid in the comparison between obese and non-obese subjects.

The new data frame only consisted of individuals in Panel 13 and from the ages 18 to 65. It also contained BMI and Total Expenditure, since BMI is the treatment and Total Expenditure is the expected outcome. Within the new data frame, covariates that were believed to help influence if a subject was obese or not, were also selected. Refer to Appendix A to see all the covariates and their categorization.

All the covariates were recorded into categorical variables and made factors. The natural log of Total Expenditure and Total Family Income were taken. For BMI to be a binary variable, the missing data was removed. This left BMI only to contain obese and non-obese subjects.

For the duration of this paper, MEPS data frame will refer to the data frame created above.

2.2 Stratification Model

The stratification model divides subjects into strata, subgroups. In this research we used ten. The subjects are placed into each strata depending on their estimated propensity score (**?**). Within each stratum, both treated and control subjects are represented (**?**). In terms of our research, obese is the treatment and non-obese is the control. The average outcomes, expenditure, of each groups is then computed.

Using the MEPS data frame, the propensity score was calculated. This was done by running a logistic regression of BMI regressed on the covariates. These propensity scores were sortaed from largest to smallest into ten strata.

Each stratum was then analyzed between the two groups, obese and non-obese. The average difference of expenditures calculated by stratum and then the averages over the strata.

2.3 Regression Model

The OLS regression model is estimated assuming the natural log of expenditures as the dependent variable and using the propensity score as the covariate.

Using a normal regression model, we regressed the natural log of the Total Expenditures on an indivator variable for the treatment group obese, and the propensity scores.

The results from these two models are shown in the Results Section.

2.4 Data

Data is coming!

3 Results

Results will come!

4 Conclusion

Followed by a final conclusion!

5 Acknowledgements

6 Appendix A

Variable Documentation		
Variable	Code Name	Categorization
BMI	BMIM	Unhealthy = BMI Heathy = BMI<29
Geographic	REGION	NORTHEAST = C sachusetts, New Ha York, Pennsylvania mont MIDWEST = India Michigan, Minnes North Dakota, Ohi consin SOUTH = Alabama trict of Columbia, F Louisiana, Marylan olina, Oklahoma, S Texas, Virginia, an WEST = Alaska, orado, Hawaii, Id New Mexico, Orego Wyoming
Metropolitan Statistical Area	MSA	YES NO MISSING
AGE	AGE	NUMERICAL VAL
Sex	SEX	MALE FEMALE
Race	RACE	WHITE BLACK OTHER
Hispanic	HISPAN	YES NO
Family's Total Income	FamilyIncome	NATURAL LOG O
Wears a Seatbelt	SEAT	YES NO MISSING

Family Income as Percent of Poverty Line	POVERTY	LOW MIDDLE HIGH
STUDENT	STUDENT	YES NO MISSING
Employment	EMPLOY	YES NO MISSING
Last Blood Pressure Check	BloodPressure	1-2YRS 3-4YRS >5YRS MISSING
Last Check Up	CHECKUP	1-2YRS 3-4YRS >5YRS MISSING
Last Cholesterol Check	CHOLESTEROL	1-2YRS 3-4YRS >5YRS MISSING
Doctor Advised to eat few high fat or high cholesterol foods	NOFAT	YES NO MISSING
Doctor Advised to Exercise More	EXRCIS	YES NO MISSING
Marital Status	MARRY	YES NOT TOGETHER NEVER MISSING
Spouse	SPOUSE	YES NO MISSING
Years of Education	EDUC	NO = 0YRS HIGH = 1st-12th Grade 1-2YRS COL = 1-2YRS of College 3-4YRS COL = 3-4YRS of College +5YRS COL = Over 5YRS of College MISSING
Total Health Care Expenditure	TotalExp	NATURAL LOG OF NUMERICAL VALUE

Clinical Trial: evaluation of an intervention in l ing statistical methods (?).

Observational Study: a type of research where data el collected from subjects in a treat and compared to a control group

Propensity Score: is a conditional probability that b distribution of covariates and pro gle scalar summary.

Non-randomization: individuals are not randomly a treatment groups.

Covariate: a independent variable in an ob study.

Baseline: the starting value of the individu and later the data will be compa

Bias: influences the data to one side, do for neutral conclusions to be mac

Stratification: subjects in both groups are ranke to their propensity score and sep subgroups based on intervals. T least 5 equal sized subgroups

Clinical Trial: evaluation of an intervention in humans using statistical methods.(1)

Observational Study: a type of research where data elements are collected from subjects in a treatment group and compared to a control group.

Retrospective: individuals or cases whose event has already occurred.

Cross-Sectional: individuals or cases whose event is occurring at a single point in time.

Prospective: individuals or cases whose event is being followed forward in time.

Propensity Score: is a conditional probability that balances the distribution of covariates and produces a single scalar summary.

Confounding: an outside variable that can cause a correlation between the independent and dependent variables.

Non-randomization: individuals are not randomly assigned to treatment groups.

Covariate: a independent variable in an observational study.

Baseline: the starting value of the individual or group and later the data will be compared to.

Bias: influences the data to one side, does not allow for neutral conclusions to be made. (PS Book)

Overt Bias: bias seen in the data at hand.

Hidden Bias: bias not seen in the data at hand but requires information which was neither observed nor recorded.

Selection Bias: bias occurred through being nonrandomized studies.

Counterfactual: an outcome which did not occur.

Average Treatment Effect (ATE): measures the average causal differences in outcomes.

Average effect of Treatment on the Treated (ATT): measures the average causal differences in treated.

Dichotomous: two separate events, mutually exclusive.

Without Replacement: once a subject from the unt matched with a subject fr group, the untreated subjec be matched with any other su treated group.

With Replacement: once a subject from the unt matched with a subject fr group, the untreated subj matched with any other su treated group.

Can use MatchIt "exact" or Matching "Matc

Optimal: matches are made to mini propensity score distance b and untreated subjects.

Can use MatchIt "optimal"

Nearest Neighbor: a treated subject is selected is matched to the untreated s closest propensity score, if t closeness, an untreated subje at random among the group

Can use MatchIt "nearest"

Full Matching: matches one treated subject untreated subject or via vers

Can use MatchIt "full"

Mahalahobis Metric Matching: a treated subject is selected Mahalahobis distance is calc treated subject with the sma selected and removed from d

Can use MatchIt "nearest" distance=malhabobis m.ord

Stratification: subjects in both groups are ing to their propensity score into subgroups based on inte 5equal sized subgroups.

Can use MatchIt "subclass"

Inverse Probability of Treatment Weighting (IPTW): each subject is weighted by i ability of being in a certain g

Can use ipw package

References