Propensity score analysis with the latest SAS/STAT procedures PSMATCH and CAUSALTRT

Yuriy Chechulin, Statistician
Modeling, Advanced Analytics
What is SAS Global Forum

- One of the largest global analytical forums
  - Around 4,000 participants from all over the world
  - Hundreds of simultaneous break-out sessions
    - A real challenge to choose where to go; double and triple bookings for relevant talks
  - Ability to discuss recent methodological developments and ask questions from experts in statistics, machine learning, data mining, etc.

- Selected talks are recorded and available on-line after the Forum
  - [http://video.sas.com/sasgf17/category/videos/user-program](http://video.sas.com/sasgf17/category/videos/user-program)

- Most of the topics presented are available as 5-20 page summary papers on the SAS on-line proceedings website
  - [http://support.sas.com/resources/papers/proceedings17/](http://support.sas.com/resources/papers/proceedings17/)

- While presented topics are implemented in SAS, discussed methods are generalizable to other software implementation if need to be
  - When a given (new) method and its application is well understood, the platform in which it is implemented is usually not so important
  - Great place to learn what’s new and evolving in the field of analytics, statistics, data science, machine learning and data mining
Outline of the Presentation

- Estimating causal effects from observational data using propensity score matching analysis with the latest SAS/STAT procedures PROC PSMATCH and CAUSALTRT
Estimating Causal Effects (1)

- In a randomized study, such as a clinical trials, the subjects are randomly assigned to a treated (exposure) group or a control (non-exposure) group
  - Random assignment ensures that the distribution of the covariates is the same in both groups, and the treatment effect can be estimated by directly comparing the outcomes for the subjects in the two groups

- In contrast, the subjects in an observational study are not randomly assigned to the treated and control groups
  - Confounding can occur if some covariates are related to both the treatment assignment and the outcome
  - Consequently, there can be systematic differences between the treated subjects and the control subjects

- In the presence of confounding, statistical approaches are required that remove the effects of confounding when estimating the effect of treatment
Estimating Causal Effects (2)

- One such approach is regression adjustment, which estimates the treatment effect after adjusting for differences in the baseline covariates
  - However, this approach has practical limitations, as discussed by Austin (2011):
    - model specification, separation of design from analysis, overlap in baseline covariates, rare outcomes covariate requirements

- Propensity score analysis is an alternative approach that circumvents many of these limitations

- The propensity score was defined by Rosenbaum and Rubin (1983) as the probability of assignment to treatment conditional on a set of observed baseline covariates
  - Propensity score analysis minimizes the effects of confounding and provides some of the advantages of a randomized study

- The latest version of SAS/STAT has two new procedures for performing propensity score analysis: PSMATCH and CASUALTRT
A propensity score analysis usually involves a number of steps outlined in the chart (Guo and Fraser 2015).

The outcome variable is intentionally not used in this process, and the variable selection is not related to the observed outcomes (Rubin 2001; Stuart 2010).

Any variables that might have been affected by the treatment should not be included in the process (Rosenbaum and Rubin 1984; Stuart 2010).
Estimating Causal Effects with PROC PSMATCH (1)

- The PSMATCH either computes propensity scores or reads previously-computed propensity scores.
- It provides various methods for using the scores to allow for valid estimation of treatment effect in a subsequent outcome analysis:
  - Inverse probability of treatment weighting
  - Stratification
  - Matching
- For matching, the procedure provides three strategies:
  - Greedy nearest neighbor matching selects the control unit nearest to each treated unit.
  - Optimal matching selects all control units that match each treated unit by minimizing the total absolute difference in propensity score across all matches.
  - Matching with replacement selects the control unit that best matches each treated unit. Each control unit can be matched to more than one treated unit, but it can only be matched to the same treated unit once.
- For all three matching methods, caliper width can be specified.
  - It imposes a restriction on the quality of the matches.
The PSMATCH procedure also provides methods for assessing the balance of baseline covariates and other variables in the treated and control groups after matching, weighting, or stratification

- Differences in the distributions
- Standardized differences
  - For good variable balance, the absolute standardized difference should be less than or equal to 0.25, and the variance ratio should be between 0.5 and 2 (Rubin 2001; Stuart 2010)
  - A smaller threshold of 0.1 for the absolute standardized difference has also been used to indicate meaningful imbalance in the variables (Normand et al. 2001; Mamdani et al. 2005; Austin 2009).
  - Percentage reductions of absolute differences

The output data set created by the PSMATCH procedure serves as input for an appropriate statistical procedure for the outcome analysis

- For example, if matching is used, a simple univariate test or analysis might be sufficient to estimate the treatment effect
Estimating Causal Effects with PROC CAUSALTRT

- The CAUSALTRT procedure provides the following methods to estimate causal effects:
  - Inverse probability weighting methods
  - Regression adjustment
  - Doubly robust methods

- CAUSALTRT can compute standard errors and confidence intervals for the causal effects by both asymptotic and bootstrap methods.

- Modeling the treatment leads to inverse probability weighting methods, and modeling the outcome leads to regression adjustment methods.

- Combined modeling of the treatment and outcome leads to doubly robust methods that can provide unbiased estimates for the treatment effect even if one of the models is misspecified.
PROC PSMATCH example (1)

The following is the example of using propensity score matching to estimate the effect of Mild Traumatic Brain Injury (MTBI) program on health care costs (preliminary first 6 months of the program data)

*Create propensity scores;
proc logistic data=mtbi;
   where MTBI = 1;
   class Age_grp_: Gender Industry_: SIZE_: NOC_: FLANG_: Source_: / param=ref;
   model Period(event='1') = Age_grp_: Gender Industry_: SIZE_: NOC_: FLANG_: Source_:;
   output out= mtbi_cost predicted=prob;
run;
PROC PSMATCH example (2)

*Run propensity score matching (optimal variable ratio method);
ods graphics on;
proc psmatch data=mtbi_cost region=cs;
   class Period Age.grp_: Gender Industry_: SIZE_: NOC_: FLANG_: Source_;;
   psdata treatvar=Period(Treated='1') ps=prob;
   match method=varratio(kmin=1 kmax=3) exact=Gender
                  stat=lps caliper(mult=stddev)=0.7;
   assess lps var=(Age.grp_: Gender Industry_: SIZE_:) / weight=matchwgt varinfo plots=(all);
   output out(obs=match)=mtbi_cost_match lps=_Lps
                  matchid=_MatchID;
run;
PROC PSMATCH example (3)

Output of PSMATCH procedure:

<table>
<thead>
<tr>
<th>Matching Information</th>
</tr>
</thead>
<tbody>
<tr>
<td>Difference Statistic</td>
</tr>
<tr>
<td>Method</td>
</tr>
<tr>
<td>Min Control/Treated Ratio</td>
</tr>
<tr>
<td>Max Control/Treated Ratio</td>
</tr>
<tr>
<td>Caliper (Logit PS)</td>
</tr>
<tr>
<td>Matched Sets</td>
</tr>
<tr>
<td>Matched Obs (Treated)</td>
</tr>
<tr>
<td>Matched Obs (Control)</td>
</tr>
<tr>
<td>Total Absolute Difference</td>
</tr>
</tbody>
</table>
## Output of PSMATCH procedure (continued):

### Standardized Variable Differences (Treated - Control)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Standardized Mean Difference</th>
<th>Mean Difference</th>
<th>Weighted Matched Obs</th>
<th>Standardized Variable Differences (Treated - Control)</th>
<th>Mean Difference</th>
<th>Percent Reduction</th>
<th>Variance Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All Obs</td>
<td>Region Obs</td>
<td>Matched Obs</td>
<td>Weighted Matched Obs</td>
<td>Region Obs</td>
<td>Matched Obs</td>
<td>Weighted Matched Obs</td>
</tr>
<tr>
<td>LPS</td>
<td>0.317</td>
<td>0.305</td>
<td>0.305</td>
<td>0.061</td>
<td>3.9</td>
<td>3.9</td>
<td>80.8</td>
</tr>
<tr>
<td>Age_grp_2</td>
<td>-0.038</td>
<td>-0.037</td>
<td>-0.037</td>
<td>-0.002</td>
<td>3.8</td>
<td>3.8</td>
<td>93.9</td>
</tr>
<tr>
<td>Age_grp_3</td>
<td>0.024</td>
<td>0.024</td>
<td>0.024</td>
<td>0.005</td>
<td>0.7</td>
<td>0.7</td>
<td>78.5</td>
</tr>
<tr>
<td>Age_grp_4</td>
<td>-0.015</td>
<td>-0.014</td>
<td>-0.014</td>
<td>0.015</td>
<td>12.2</td>
<td>12.2</td>
<td>5.0</td>
</tr>
<tr>
<td>Age_grp_5</td>
<td>-0.006</td>
<td>-0.009</td>
<td>-0.009</td>
<td>-0.014</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Age_grp_6</td>
<td>0.044</td>
<td>0.045</td>
<td>0.045</td>
<td>0.012</td>
<td>0.0</td>
<td>0.0</td>
<td>72.7</td>
</tr>
<tr>
<td>GENDER</td>
<td>0.054</td>
<td>0.050</td>
<td>0.050</td>
<td>0.000</td>
<td>6.9</td>
<td>6.9</td>
<td>100.0</td>
</tr>
<tr>
<td>Industry_1</td>
<td>-0.031</td>
<td>-0.029</td>
<td>-0.029</td>
<td>0.000</td>
<td>3.4</td>
<td>3.4</td>
<td>100.0</td>
</tr>
<tr>
<td>Industry_2</td>
<td>0.023</td>
<td>0.024</td>
<td>0.024</td>
<td>0.014</td>
<td>0.0</td>
<td>0.0</td>
<td>36.3</td>
</tr>
<tr>
<td>Industry_3</td>
<td>0.004</td>
<td>0.005</td>
<td>0.005</td>
<td>-0.018</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Industry_4</td>
<td>0.063</td>
<td>0.064</td>
<td>0.064</td>
<td>0.021</td>
<td>0.0</td>
<td>0.0</td>
<td>67.1</td>
</tr>
<tr>
<td>Industry_6</td>
<td>-0.016</td>
<td>-0.014</td>
<td>-0.014</td>
<td>0.004</td>
<td>13.8</td>
<td>13.8</td>
<td>73.7</td>
</tr>
<tr>
<td>Industry_7</td>
<td>0.057</td>
<td>0.058</td>
<td>0.058</td>
<td>0.019</td>
<td>0.0</td>
<td>0.0</td>
<td>66.5</td>
</tr>
<tr>
<td>Industry_8</td>
<td>0.079</td>
<td>0.066</td>
<td>0.066</td>
<td>-0.021</td>
<td>15.6</td>
<td>15.6</td>
<td>73.9</td>
</tr>
<tr>
<td>Industry_9</td>
<td>-0.037</td>
<td>-0.036</td>
<td>-0.036</td>
<td>0.005</td>
<td>2.0</td>
<td>2.0</td>
<td>87.6</td>
</tr>
<tr>
<td>Size_M</td>
<td>0.005</td>
<td>0.001</td>
<td>0.001</td>
<td>0.005</td>
<td>80.4</td>
<td>80.4</td>
<td>0.0</td>
</tr>
<tr>
<td>Size_L</td>
<td>-0.010</td>
<td>-0.010</td>
<td>-0.010</td>
<td>-0.025</td>
<td>1.4</td>
<td>1.4</td>
<td>0.0</td>
</tr>
<tr>
<td>Size_U</td>
<td>-0.025</td>
<td>-0.022</td>
<td>-0.022</td>
<td>-0.002</td>
<td>9.0</td>
<td>9.0</td>
<td>92.2</td>
</tr>
</tbody>
</table>
PROC PSMATCH example (5)

Output of PSMATCH procedure (continued):

![Standardized Variable Differences Graph](image)
Output of PSMATCH procedure (continued):

Distribution of LPS

- Treated (Period = 1)
- Control (Period = 0)

Weighted Matched

Matched

Region

All

Logit of Propensity Score

-2 -1 0 1
Output of PSMATCH procedure (continued):

- Frequency Distribution of Age_grp_2
- Frequency Distribution of GENDER
- Frequency Distribution of Industry_6
- Frequency Distribution of Size_M
Once the data set with the matched observation and corresponding matching weights was created by PSMATCH procedure, usual statistical analysis can be performed on matched data to estimate the causal effects (“apples to apples”)

- Effect on health care costs: t-test

*Costs analysis;

```proc sort data=mtbi_cost_match;
  by descending Period;
run;
ods graphics on;
proc ttest data=mtbi_cost_match test=diff
dist=normal plots=summary order=data;
  class Period;
  weight _matchwgt_
  var HCpmt;
run;```
T-test to estimate the effect on healthcare costs:

The TTEST Procedure
Variable: HCpmt
Weight: _MATCHWGT_ ATT weight in matched observation

<table>
<thead>
<tr>
<th>Period</th>
<th>Method</th>
<th>Mean</th>
<th>95% CL Mean</th>
<th>Std Dev</th>
<th>95% CL Std Dev</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>3782.1</td>
<td>3366.3</td>
<td>4197.8</td>
<td>6428.7</td>
</tr>
<tr>
<td>0</td>
<td></td>
<td>4582.7</td>
<td>4155.2</td>
<td>5010.2</td>
<td>6612.5</td>
</tr>
<tr>
<td>Diff (1-2)</td>
<td>Pooled</td>
<td>-800.6</td>
<td>-1397.9</td>
<td>-203.3</td>
<td>6536.0</td>
</tr>
<tr>
<td>Diff (1-2)</td>
<td>Satterthwaite</td>
<td>-800.6</td>
<td>-1396.6</td>
<td>-204.7</td>
<td>6348.1</td>
</tr>
</tbody>
</table>

Distribution of HCpmt
Summary

- In an observational study, propensity score analysis attempts to replicate the properties of a randomized trial to estimate the treatment effect.
- MTBI program causal effects estimation (preliminary results based on the first 6 months of the program data):
  - Achieved improved outcomes, and the result is statistically significant.
Thank You