*SAS code with annotations*

|  |  |  |
| --- | --- | --- |
| Action | SAS Code | Annotations |
| Create PS using logistic regression | proc logistic descending data = ps\_est;  title ‘PS Estimation’;  model hop = var1-var3/lackfit outroc = ps\_r;  output out= ps\_p XBETA=ps\_xb STDXBETA= ps\_sdxb PREDICTED = ps\_pred; run; | Using proc logistic, this code creates a PS (ps\_pred) for group by variables 1-3. In this case the group or intervention is HOP; var1-var3 represent any number of variables that are included as predictors in the PS model. Output is saved in data set ps\_p. |
| PS method 1: Matching | data one;  set ps\_p;  ranvar = ranuni(0);  run;  proc sort data = one;  by hop ranvar;  run;  proc transpose data = one out = data1;  by hop;  var ps\_pred hop ranvar StudentID;  run;  data id\_t (rename=(COL1-COL1,049 = tid1-tid1,049));  set data1;  if hop = 1 and \_NAME\_ = 'StudentID';  data ps\_t (rename=(COL1-COL1,049 = tps1-tps1,049));  set data1;  if hop = 1 and \_NAME\_ = 'ps\_pred';  data id\_c (rename=(COL1-COL19,394 = cid1-cid19,394));  set data1;  if hop = 0 and \_NAME\_ = 'StudentID';  data ps\_c (rename=(COL1-COL19,394 = cps1-cps19,394));  set data1;  if hop = 0 and \_NAME\_ = 'ps\_pred';  data all;  merge id\_t ps\_t id\_c ps\_c;  caliper = .05;  array treat\_id ([Potter](#_ENREF_25)) $ tid1-tid1,049;  array ctl\_id {\*} $ cid1-cid19,394;  array treat\_p {\*} tps1-tps1,049;  array ctl\_p {\*} cps1-cps19,394;  array used\_i {\*} used1 - used19,394;  array matched\_t {\*} $ m\_tid1-m\_tid1,049;  array matched\_c {\*} $ m\_cid1-m\_cid19,394; match\_N = 0;  do i = 1 to 1,049;  min\_diff = 1;  best\_match = 0;  do j = 1 to 19,394;  if used\_i[j] = . then do;  if ABS(treat\_p[i] - ctl\_p[j]) < caliper then do;  if ABS(treat\_p[i] - ctl\_p[j]) < min\_diff then do;  min\_diff = ABS(treat\_p[i] - ctl\_p[j]);  best\_match = j;  end;  end;  end;  end;  if best\_match > 0 then do;  match\_N = match\_N + 1;  used\_i[best\_match] = 1;  matched\_t[match\_N] = treat\_id[i];  matched\_c[match\_N] = ctl\_id[best\_match];  end;  end;  set all;  array matched\_t {\*} $ m\_tid1-m\_tid1,049;  array matched\_c {\*} $ m\_cid1-m\_cid19,394;  do match = 1 to match\_N;  Intervention\_IDN = matched\_t[match];  Control\_IDN = matched\_c[match];  output;  end;  keep match intervention\_idn control\_idn;  run; | This provides an example of 1:1 matching without replacement. Of note, after matching, data are paired, and it is recommended that a paired test (i.e., McNemar’s, paired t-test) be used for data analysis.  Data must be transposed by HOP status to facilitate the matching process.  First, a data set is created that includes the identifier and PS for each subject in the intervention group. Note that number of columns is number of observations in intervention group.  This procedure is repeated for the control group.  Using lengthy code, the data sets are merged, and subjects are matched using a caliper width of 0.05. Note that the caliper for matching (here this is 0.05) is specified in the “caliper” line.  A data step is used to create a data set that includes only matched participants. |
| PS method 2: Stratification | proc rank data = ps\_p out= ps\_strataranks groups=5;  var ps\_pred;  ranks ps\_pred\_rank;  run;  data final.strata;  set ps\_strataranks;  run;  proc sort data=final.strata;  by ps\_pred\_rank;  run; | First, subjects are divided into quintiles based on PS.  The data are then sorted by the strata ranks. |
| PS method 3: Weighting | data ps\_weight;  set ps\_p;  if hop = 1 then ps\_weight = 1/ps\_pred;  else ps\_weight = 1/(1-ps\_pred);  run;  proc means noprint data = ps\_weight;  var ps\_weight;  output out = q mean = mn\_wt;  run;  data ps\_weight2;  if \_n\_ = 1 then set q;  retain mn\_wt;  set ps\_weight;  wt2 = ps\_weight/mn\_wt;  run; | First the weights are created.  Then the weights are normalized. Wt2 is the new normalized weight. |
| Examining common support | proc sort data = ps\_p;  by ps\_pred HOP;  proc boxplot data=ps\_p;  symbol width = 2;  plot ps\_pred\*HOP;  run;  proc univariate data=ps\_p noprint;  class HOP;  var ps\_pred;  histogram ps\_pred;  run; | A boxplot and histogram are then used to examine common support. The common support can be considered the areas where the propensity scores overlap. A limited common support means that PS will have limited ability to improve the analysis. |

\*SAS codes are from SAS Global Forum, Paper 314-2012 ([Lanehart et al., 2012](#_ENREF_19)).

Lanehart, R. E., de Gil, P. R., Kim, E. S., Bellara, A. P., Kromrey, J. D., & Lee, R. S. (2012, April). *Propensity score analysis and assessment of propensity score approaches using SAS® procedures.* Paper presented at the SAS Global Forum: Statistics and Data Analysis, Orlando, FL. Retrieved from http://support.sas.com/resources/papers/proceedings12/314-2012.pdf