SAS/STAT 15.2®
User's Guide
The PSMATCH Procedure

SAS® Documentation
November 06, 2020

This document is an individual chapter from *SAS/STAT® 15.2 User's Guide*.

The correct bibliographic citation for this manual is as follows: SAS Institute Inc. 2020. *SAS/STAT® 15.2 User's Guide*. Cary, NC: SAS Institute Inc.

SAS/STAT® 15.2 User's Guide

SAS Institute Inc., SAS Campus Drive, Cary, NC 27513-2414

November 2020

Chapter 100
The PSMATCH Procedure

Contents

Overview: PSMATCH Procedure

In a randomized study, such as a randomized controlled trial, 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, such as a retrospective cohort study or a nonrandomized clinical trial, 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.

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 (2011a). Propensity score analysis is an alternative approach that circumvents many of these limitations.

The propensity score was defined by Rosenbaum and Rubin (1983, p. 47) as the probability of assignment to treatment conditional on a set of observed baseline covariates. Propensity score analysis minimizes the effects of confounding and offers some of the advantages of a randomized study. The basis for propensity score methods is the causal effect model introduced by Rubin (1974).

The PSMATCH procedure provides a variety of tools for propensity score analysis. The procedure either computes propensity scores or reads previously computed propensity scores, and it provides the following methods for using the scores to allow for valid estimation of the treatment effect in a subsequent outcome analysis:

- Inverse probability of treatment weighting and ATT weighting (weighting by odds): The procedure computes weights from the propensity scores. These weights can then be incorporated into a subsequent analysis that estimates the effect of treatment.
- Stratification: The procedure creates strata of observations that have similar propensity scores. In a subsequent analysis, the treatment effect can be estimated within each stratum, and the estimates can be combined across strata.

- Matching: The procedure matches each treated unit with one or more control units that have a similar value of the propensity score. In a subsequent analysis, the treatment effect can be estimated by comparing outcomes between treated and control subjects in the matched sample. If the outcome values for a study are not available prior to matching, only the matched units are needed for follow-up. Thus, the cost of the trial is reduced (Stuart 2010, p. 2).

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. The procedure itself does not carry out the outcome analysis, nor does it make use of the outcome variable.

After adequate variable balance has been achieved (as described in the section "Process of Propensity Score Analysis" on page 8179) and assuming that no other confounding variables are associated with both the treatment assignment and the outcome, the output data set that is created by the PSMATCH procedure serves as input for an appropriate statistical procedure for the outcome analysis.

Process of Propensity Score Analysis

A propensity score analysis usually involves the following steps (Guo and Fraser 2015, p. 131):

1. You specify a set of confounding variables that might be related to both the treatment assignment and the outcome.
2. You use this set of variables to fit a logistic regression model and compute propensity scores. The response is the probability of assignment to the treatment group.
3. If you are using weighting, you compute observation weights for estimating the treatment effect in a weighted outcome analysis.
4. If you are using stratification or matching, you specify the support region for observations. Observations outside this region are not included in the stratification or matching.
5. If you are using stratification, you specify the number of strata and create strata of observations that have similar propensity scores.
6. If you are using matching, you specify the distance metric for similarity of observations and the method for creating matched sets of observations. You can also compute weights for matched observations.
7. You assess the balance of variables by comparing the distributions between the treated and control groups.
8. To improve the balance, you can repeat the process with a different set of variables for the logistic regression model, a different region of support for stratification and matching, a different distance metric, or a different matching method.
9. When you are satisfied with the variable balance, you save the output data set for subsequent outcome analysis.

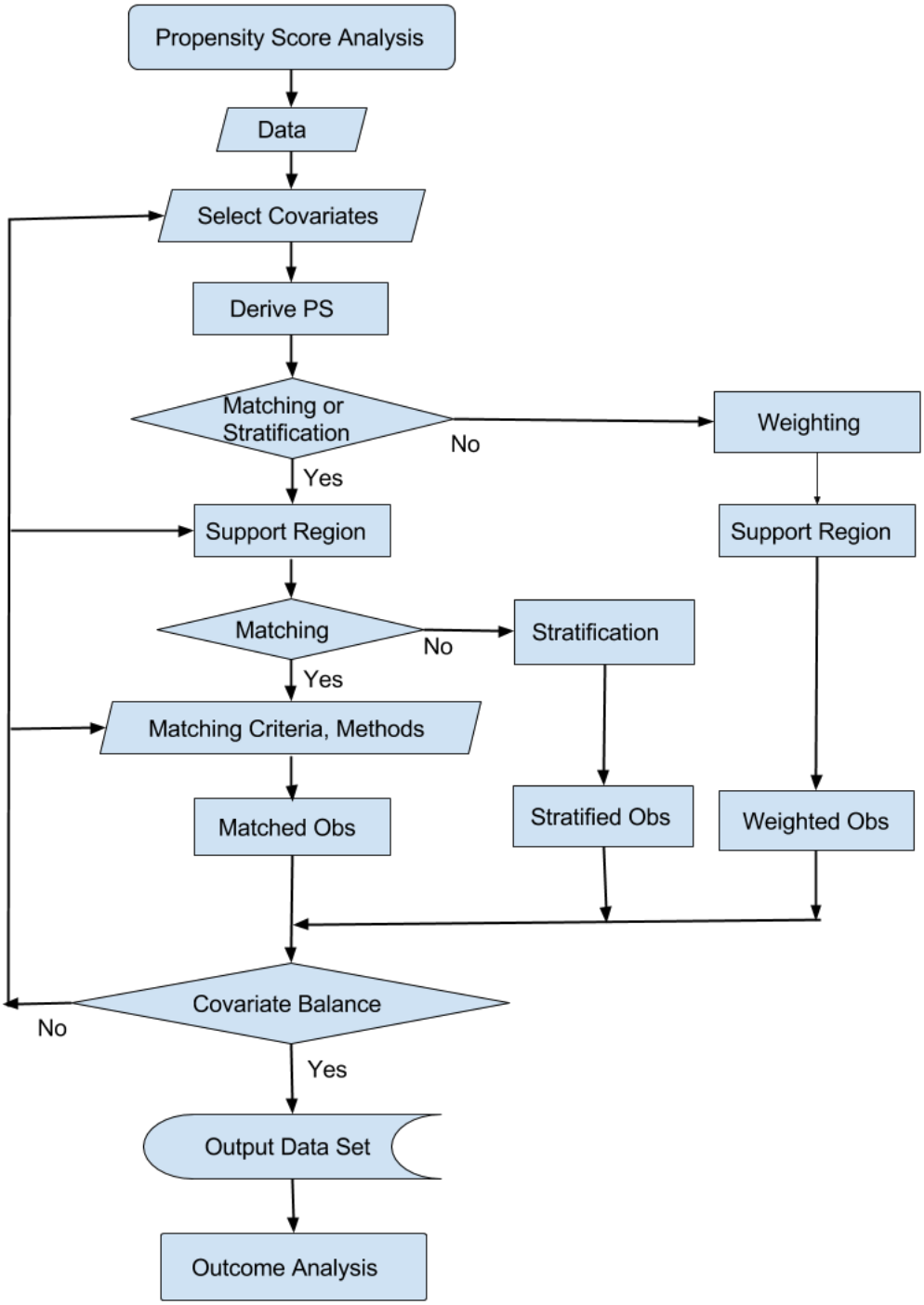
Propensity score analysis assumes that all the variables that affect both the outcome and the treatment assignment have been measured, but this assumption cannot be tested. In some cases, you can perform sensitivity analysis to examine this assumption; see the section "Sensitivity Analysis" on page 8231.

Note that the outcome variable is intentionally not used in these steps, and the selection of variables for the model should be made independently of the observed outcomes (Rubin 2001; Stuart 2010, p. 5). Furthermore, any variables that might have been affected by the treatment should not be included in the process (Rosenbaum and Rubin 1984; Stuart 2010, p. 5).

The flowchart in Figure 100.1 summarizes the steps.

After balance is achieved, you can add the response variable to the output data set that PROC PSMATCH created and perform an outcome analysis that mimics the analysis you would perform with data from a randomized study. For example, if you used matching, a simple univariate test or analysis might be sufficient to estimate the treatment effect.

Figure 100.1 Propensity Score Analysis



Features of the PSMATCH Procedure

You can use the PSMATCH procedure to create propensity scores (PS) for observations from treated and control groups by fitting a binary logistic regression model. Alternatively, you can input propensity scores that have already been created by using a different model or even a different approach such as a tree-based method. For example, you can input propensity scores that have been computed by the LOGISTIC procedure using a binary probit model or by the HPSPLIT procedure using a classification tree.

By default, the PSMATCH procedure uses the propensity scores to computes weights for the observations. Various types of weights are available, depending on whether the outcome analysis will use the weights to estimate the average treatment effect at the population level (ATE) or the average treatment effect for subjects who receive treatment (ATT). For more information about propensity score weighting, see the section "Propensity Score Weighting" on page 8217.

The PSMATCH procedure optionally creates strata of observations that have similar propensity scores. For more information, see the section "Propensity Score Stratification" on page 8220.

The PSMATCH procedure optionally matches observations in the treated and control groups. The procedure provides three strategies for propensity score matching.

- Greedy nearest neighbor matching selects the control unit nearest to each treated unit. Greedy nearest neighbor matching is done sequentially for treated units and without replacement.
- Optimal matching selects all control units that match each treated unit by minimizing the total absolute difference in propensity score across all matches. Optimal matching selects all matches simultaneously and without replacement. Three methods for optimal matching are available: fixed ratio matching, variable ratio matching, and full matching.
- 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, you can specify a caliper width, which imposes a restriction on the quality of the matches. The difference in propensity score between the treated unit and its matching control unit must be less than or equal to the caliper width. For more information about these methods, see the section "Matching Methods" on page 8224.

Matching can be based on the difference in the logit of the propensity score (LPS), as well as the difference in the propensity score (PS). Furthermore, matching can be based on Mahalanobis distance that is computed from a set of continuous covariates (possibly including LPS and LS).

The PSMATCH procedure provides various ways to assess how well the distributions of variables are balanced between the treated and control groups. These variables include the propensity score, the logit of the propensity score, variables used in the logistic regression model, and other variables in the data set. The assessments include the following:

- differences in the distributions of the variables between the treated and control groups after weighting, stratification, and matching
- standardized mean differences in the variables between the treated and control groups after weighting, stratification, and matching

- percentage reductions of absolute differences after weighting, stratification, and matching.

When you use stratification, the differences are also computed within each stratum. For more information about these statistics, see the section "Variable Balance Assessment" on page 8228.

The PSMATCH procedure also provides various plots for assessing balance. These plots include the following:

- bar charts for classification variables
- box plots for continuous variables
- CDF plots for continuous variables
- cloud plots for continuous variables, which are scatter plots in which the points are jittered to prevent overplotting
- cloud plots for inverse probability of treatment weights and ATT weights
- a standardized mean differences plot that summarizes differences between the treated and control groups

When you use stratification, these plots are also produced for each stratum.

The PSMATCH procedures saves propensity scores and weights in an output data set that contains a sample that has been adjusted either by weighting, stratification, or matching. If the sample is stratified, you can save the strata identification in the output data set. If the sample is matched, you can save the matching identification in the output data set.

Provided that the distributions of the variables in the adjusted sample are well balanced between the treated and control groups, the output data set serves as input for a subsequent outcome analysis that incorporates weights or strata or that is based on matched observations. Although the PSMATCH procedure itself does not provide this analysis, many other SAS/STAT procedures can be used for this purpose.

Getting Started: PSMATCH Procedure

This example illustrates the use of the PSMATCH procedure to match observations for individuals in a treatment group with observations for individuals in a control group that have similar propensity scores. The matched observations are saved in an output data set that, with the addition of the outcome variable, can be used to provide an unbiased estimate of the treatment effect.

A pharmaceutical company is conducting a nonrandomized clinical trial to demonstrate the efficacy of a new treatment (Drug_X) by comparing it to an existing treatment (Drug_A). Patients in the trial can choose the treatment that they prefer; otherwise, physicians assign each patient to a treatment. The possibility of treatment selection bias is a concern because it can lead to systematic differences in the distributions of the baseline variables in the two groups, resulting in a biased estimate of treatment effect.

The data set Drugs contains baseline variable measurements for individuals from both treated and control groups. PatientID is the patient identification number, Drug is the treatment group indicator, Gender provides the gender, Age provides the age, and BMI provides the body mass index (a measure of body fat based on height and weight). Typically, more variables are used in a propensity score analysis, but for simplicity only a few variables are used in this example.

Figure 100.2 lists the first 10 observations.

Figure 100.2 Input Drug Data Set

Obs	PatientID	Drug	Gender	Age	BMI
1	284	Drug_X	Male	29	22.02
2	201	Drug_A	Male	45	26.68
3	147	Drug_A	Male	42	21.84
4	307	Drug_X	Male	38	22.71
5	433	Drug_A	Male	31	22.76
6	435	Drug_A	Male	43	26.86
7	159	Drug_A	Female	45	25.47
8	368	Drug_A	Female	49	24.28
9	286	Drug_A	Male	31	23.31
10	163	Drug_X	Female	39	25.34

Note that the Drugs data set does not contain a response variable, because the response variable is not used in a propensity score analysis. Instead, the response variable is added to the output data set that contains the matched observations, and the combined data set is then used for outcome analysis.

The following statements invoke the PSMATCH procedure and request optimal matching to match observations for patients in the treatment group with observations for patients in the control group:

```
ods graphics on;
proc psmatch data=drugs region=cs;
   class Drug Gender;
   psmodel Drug(Treated='Drug_X')= Gender Age BMI;
   match method=optimal(k=1) exact=Gender distance=lps caliper=0.25
         weight=none;
   assess lps allcov / plots=(barchart boxplot);
   output out(obs=match)=Outgs lps=_Lps matchid=_MatchID;
```

```
run;
```

The CLASS statement specifies the classification variables. The PSMODEL statement specifies the logistic regression model that creates the propensity score for each observation, which is the probability that the patient receives Drug_X. The Drug variable is the binary treatment indicator variable and TREATED='Drug_X' identifies Drug_X as the treated group. The Gender, Age, and BMI variables are included in the model because they are believed to be related to the assignment.

The REGION= option specifies which observations are used in stratification and matching. In this example, matching is requested by the MATCH statement, and the REGION=CS option requests that only those observations whose propensity scores (or equivalently, logits of propensity scores) lie in the common support region be used for matching. The common support region is defined as the largest interval that contains propensity scores for subjects in both groups. By default, the region is extended by 0.25 times a pooled estimate of the common standard deviation of the logit of the propensity score. For more information, see the description of the EXTEND= option on page 8193.

The MATCH statement specifies the criteria for matching. The DISTANCE=LPS option (which is the default) requests that the logit of the propensity score be used to compute differences between pairs of observations. The METHOD=OPTIMAL(K=1) option (which is the default) requests optimal matching of one control unit to each unit in the treated group in order to minimize the total within-pair difference, The EXACT=GENDER option forces the treated unit and its matched control unit to have the same value of the Gender variable.

The CALIPER=0.25 option specifies the caliper requirement for matching. This means that for a match to be made, the difference in the logits of the propensity scores for pairs of individuals from the two groups must be less than or equal to 0.25 times the pooled estimate of the common standard deviation of the logits of the propensity scores.

The "Data Information" table in Figure 100.3 displays information about the input and output data sets, the numbers of observations in the treated and control groups, the lower and upper limits for the propensity score support region, and the numbers of observations in the treated and control groups that fall within the support region. Of the 373 observations in the control group, 351 fall within the support region.

Figure 100.3 Data Information

The PSMATCH Procedure

Data Information	
Data Set	WORK.DRUGS
Output Data Set	WORK.OUTGS
Treatment Variable	Drug
Treated Group	Drug_X
All Obs (Treated)	113
All Obs (Control)	373
Support Region	Extended Common Support
Lower PS Support	0.050244
Upper PS Support	0.683999
Support Region Obs (Treated)	113
Support Region Obs (Control)	351

The "Propensity Score Information" table in Figure 100.4 displays summary statistics for propensity scores by treatment group on the basis of all observations, support region observations, and matched observations. When you specify the METHOD=OPTIMAL(K=1) option, all matched observations have the same weight—that is,

each matched unit has a weight of 1. Therefore, all the propensity score summary statistics would remain unchanged if you applied these weights to the matched observations. In the example, the WEIGHT=NONE option suppresses the display of summary statistics for weighted matched observations.

Figure 100.4 Propensity Score Information

Propensity Score Information

	Treated (Drug = Drug_X)					Control (Drug = Drug_A)					Treated - Control
Observations	N	Mean	Standard Deviation	Minimum	Maximum	N	Mean	Standard Deviation	Minimum	Maximum	Mean Difference
All	113	0.3108	0.1325	0.0602	0.6411	373	0.2088	0.1320	0.0202	0.6858	0.1020
Region	113	0.3108	0.1325	0.0602	0.6411	351	0.2176	0.1267	0.0510	0.6824	0.0932
Matched	113	0.3108	0.1325	0.0602	0.6411	113	0.3082	0.1310	0.0619	0.6824	0.0025

The "Matching Information" table in Figure 100.5 displays the matching criteria, the number of matched sets, the numbers of matched observations in the treated and control groups, and the total absolute difference in the logit of the propensity score for all matches.

Figure 100.5 Matching Information

Matching Information

Distance Metric	Logit of Propensity Score
Method	Optimal Fixed Ratio Matching
Control/Treated Ratio	1
Caliper (Logit PS)	0.191862
Matched Sets	113
Matched Obs (Treated)	113
Matched Obs (Control)	113
Total Absolute Difference	2.941871

The ASSESS statement produces a table and plots that summarize differences in specified variables between treated and control groups. As specified by the LPS and ALLCOV options, these variables are the logit of the propensity score (LPS) and all the covariates in the PSMODEL statement: Gender, Age, and BMI. For a binary classification variable (Gender), the difference is in the proportion of the first ordered level (Female).

The "Standardized Mean Differences" table, shown in Figure 100.6, displays standardized mean differences for all observations, observations in the support region, and matched observations. The WEIGHT=NONE option suppresses the display of differences for weighted matched observations. Note that when one control unit is matched to each treated unit, the weights are all 1 for matched treated and control units and the results are identical for weighted matched observations and matched observations.

Figure 100.6 Standardized Mean Differences

The PSMATCH Procedure

Standardized Mean Differences (Treated - Control)						
Variable	Observations	Mean Difference	Standard Deviation	Standardized Difference	Percent Reduction	Variance Ratio
Logit Prop Score	All	0.63997	0.767449	0.83389		0.6517
	Region	0.54546		0.71074	14.77	0.8314
	Matched	0.01056		0.01375	98.35	1.0155
Age	All	-4.09509	6.079104	-0.67363		0.7076
	Region	-3.49368		-0.57470	14.69	0.8000
	Matched	0.16814		0.02766	95.89	1.1262
BMI	All	0.73930	1.923178	0.38441		0.8854
	Region	0.63257		0.32892	14.44	0.9288
	Matched	0.12425		0.06461	83.19	1.1967
Gender	All	-0.02482	0.496925	-0.04994		0.9892
	Region	-0.01651		-0.03323	33.46	0.9922
	Matched	0.00000		0.00000	100.00	1.0000

Standard deviation of All observations used to compute standardized differences

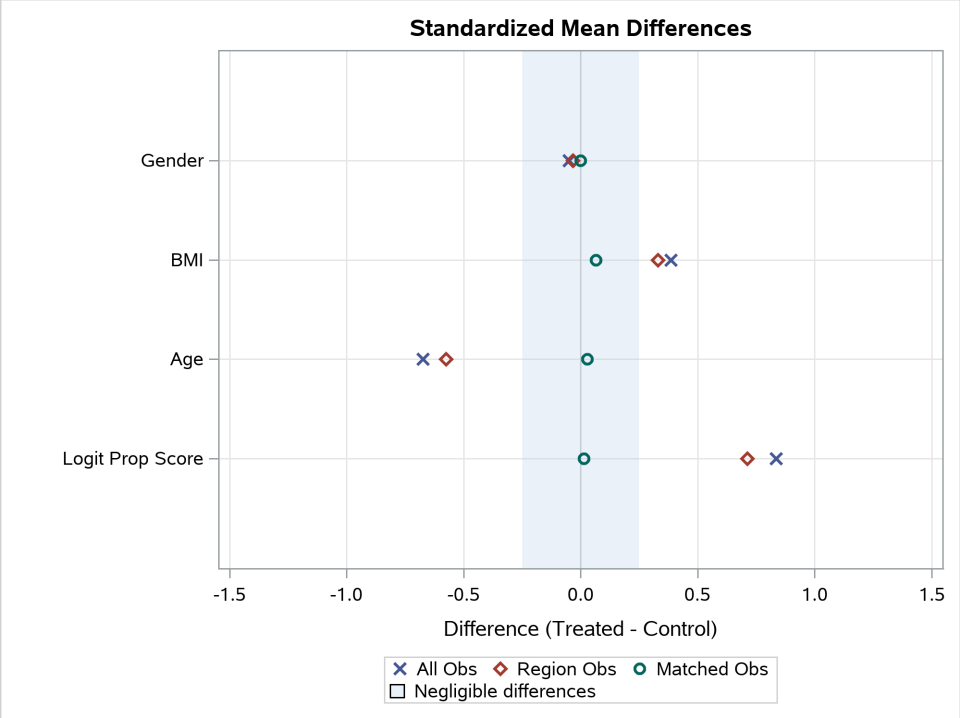
By default, the standard deviations of the variables, pooled across the treated and control groups, are computed based on all observations. The pooled standard deviations are then used to compute standardized mean differences based on all observations, observations in the support region, and matched observations. You can request a different standard deviation with the STDDEV= option. In Figure 100.6 the standardized mean differences are significantly reduced in the matched observations. The largest of these differences in absolute value is 0.0646, which is less than the upper limit of 0.25 recommended by Rubin (2001, p. 174) and Stuart (2010, p. 11). However, many authors use an upper limit of 0.10 (Normand et al. 2001; Mamdani et al. 2005; Austin 2009).

The treated-to-control variance ratios between the two groups are between 1 and 1.1967 for all variables in the matched observations, which is within the recommended range of 0.5 to 2 (Rubin 2001, p. 174).

Note that the standardized mean difference for Gender is 0 in the matched observations because EXACT=GENDER is specified in the MATCH statement.

By default, when ODS Graphics is enabled, the PSMATCH procedure displays a standardized mean differences plot for the variables that are specified in the ASSESS statement, as shown in Figure 100.7.

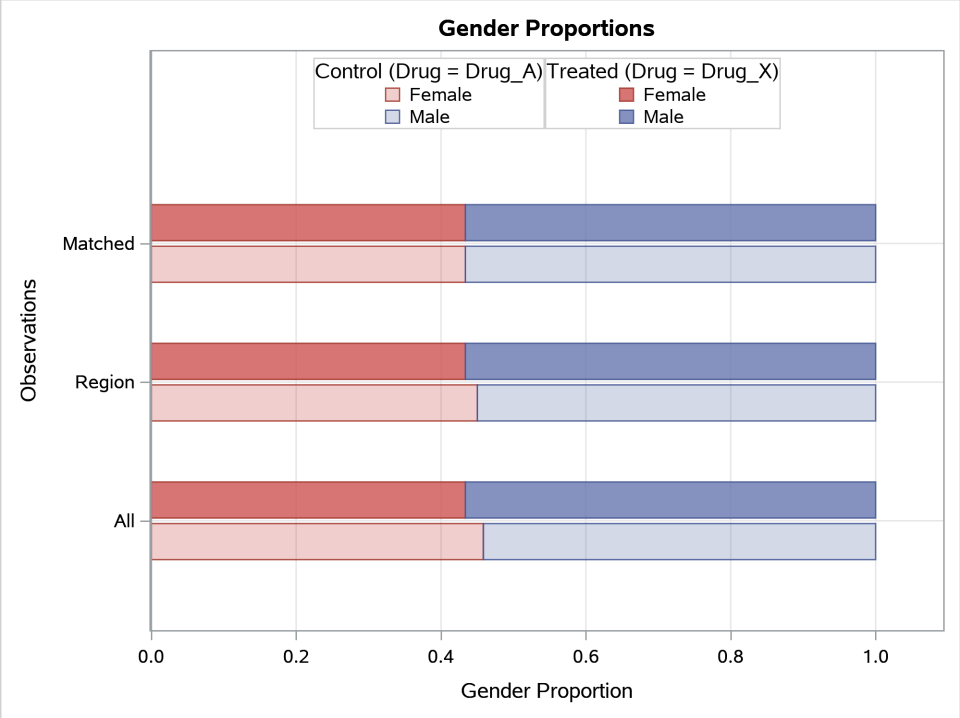
Figure 100.7 Standardized Mean Differences Plot



The "Standardized mean Differences Plot" displays the standardized mean differences in the "Standardized Mean Differences" table in Figure 100.6. All differences for the matched observations are within the recommended limits of –0.25 and 0.25, which are indicated by the shaded area. Again, note that many authors use limits of –0.10 and 0.10. (Normand et al. 2001; Mamdani et al. 2005; Austin 2009). You can use the PLOTS=STDDIFFPLOT(REF=) option to specify the limits for the shaded area.

The PLOTS=BARCHART option requests bar charts that compare the treated and control group distributions of binary classification variables that are specified in the ASSESS statement. The bar chart that is created for Gender is shown in Figure 100.8. The chart displays proportions by default, and it provides comparisons based on all observations, observations in the support region, and matched observations. The distributions of Gender are identical for matched observations because EXACT=GENDER is specified in the MATCH statement.

Figure 100.8 Gender Bar Chart



The PLOTS=BOXPLOT option requests box plots for the logit of the propensity score (LPS) and for the continuous variables that are specified in the ASSESS statement, as shown in Figure 100.9, Figure 100.10, and Figure 100.11. The box plots show good variable balance for the matched observations.

Figure 100.9 LPS Box Plot

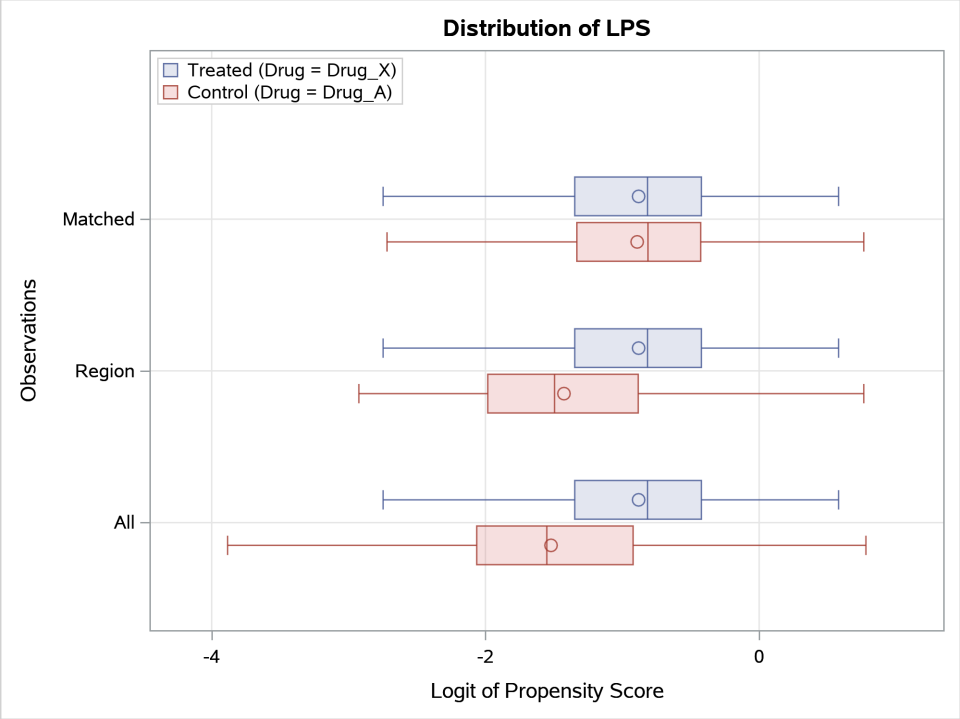


Figure 100.10 Age Box Plot

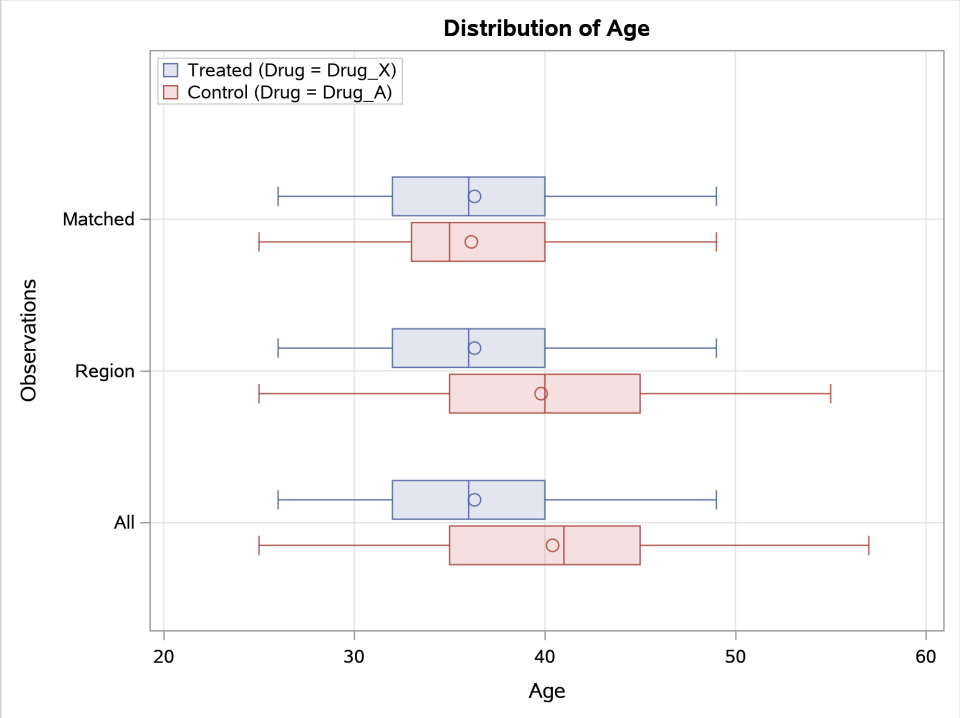
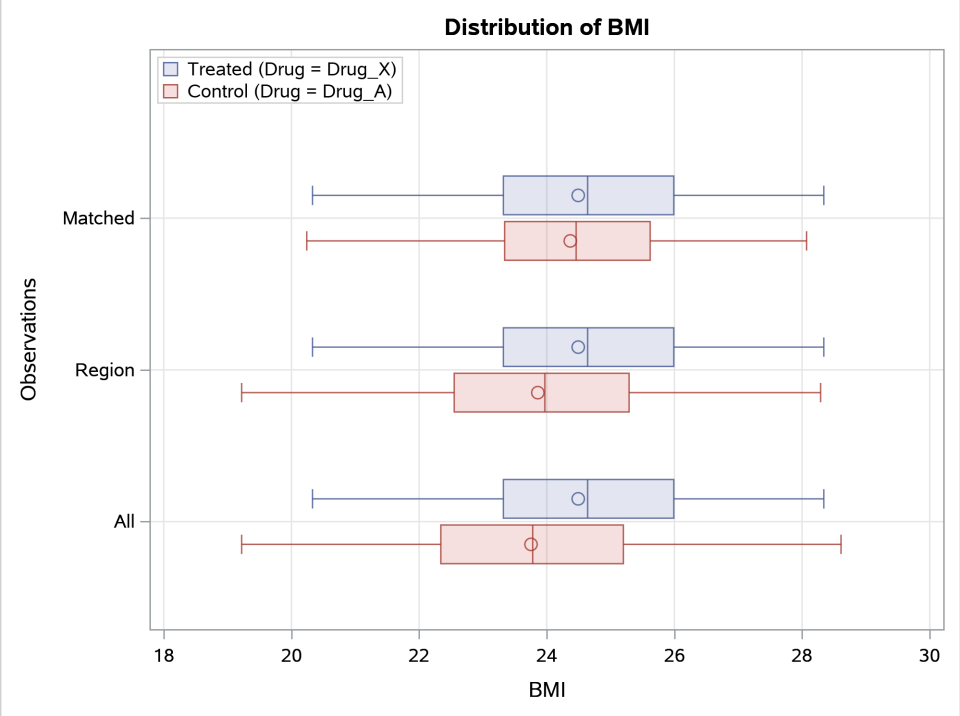


Figure 100.11 BMI Box Plot



Because the matched observations in this example exhibit good balance, you can output them for subsequent

outcome analysis. In situations where you are not satisfied with the balance, you can do one or more of the following to improve the balance: you can select another set of variables for the propensity score model, you can modify the specification of the propensity score model (for example, by introducing nonlinear terms for the continuous variables or by adding interactions), you can modify the matching criteria, or you can choose another matching method.

The OUT(OBS=MATCH)= option in the OUTPUT statement creates an output data set named Outgs that contains the matched observations. By default, this data set includes the variable _PS_ (which provides the propensity score) and the variable _MATCHWGT_ (which provides matched observation weights). The weight for each treated unit is 1. The weight for each matched control unit is also 1 because one control unit is matched to each treated unit. The LPS=_LPS option adds a variable named _LPS that provides the logit of the propensity score, and the MATCHID=_MatchID option adds a variable named _MatchID that identifies the matched sets of observations.

The following statements list the observations in the first five matched sets, as shown in Figure 100.12.

```
proc sort data=outgs out=outgs1;
   by _MatchID;
run;

proc print data=outgs1(obs=10);
   var PatientID Drug Gender Age BMI _PS_ _LPS _MatchWgt_ _MatchID;
run;
```

Figure 100.12 Output Data Set with Matching Numbers

Obs	PatientID	Drug	Gender	Age	BMI	_PS_	_Lps	_MATCHWGT_	_MatchID
1	213	Drug_A	Female	49	23.24	0.06187	-2.71892	1	1
2	89	Drug_X	Female	44	20.75	0.06023	-2.74745	1	1
3	141	Drug_A	Female	43	20.55	0.06401	-2.68256	1	2
4	323	Drug_X	Female	46	22.22	0.06763	-2.62375	1	2
5	420	Drug_A	Male	45	22.08	0.08801	-2.33814	1	3
6	217	Drug_X	Male	49	23.96	0.08772	-2.34185	1	3
7	234	Drug_X	Female	41	21.11	0.08904	-2.32538	1	4
8	290	Drug_A	Female	40	20.57	0.08778	-2.34104	1	4
9	320	Drug_X	Female	46	24.17	0.10323	-2.16184	1	5
10	473	Drug_A	Female	45	23.76	0.10464	-2.14670	1	5

After the responses for the trial are observed and added to the matched data set Outgs, you can estimate the treatment effect by carrying out the same type of outcome analysis on Outgs that you would have used with the original data set Drugs (augmented with responses) as if it were a randomized trial (Ho et al. 2007, p. 223). This assumes that no other confounding variables are associated with both the response variable and the treatment group indicator Drug.

Syntax: PSMATCH Procedure

The following statements are available in the PSMATCH procedure:

```
PROC PSMATCH <options> ;
    ASSESS <ALLCOV> <LPS> <PS> <VAR=(var-list)> </ assess-options> ;
    BY variables ;
    CLASS variables ;
    FREQ variable ;
    ID variable ;
    MATCH <options> ;
    OUTPUT OUT <(OBS=obs-value)>=SAS-data-set <keyword=name <keyword=name . . . >> ;
    PSDATA TREATVAR=treatvar <(trt-option)> ps-option ;
    PSMODEL treatvar <(trt-option)> = <effects> </ WEIGHT=weight> ;
    PSWEIGHT <options> ;
    STRATA <options> ;
```

The PROC PSMATCH statement invokes the PSMATCH procedure. The CLASS statement and either a PSMODEL or PSDATA statement are required. If a PSMODEL statement is specified, the CLASS statement must precede the PSMODEL statement.

The MATCH, PSWEIGHT, and STRATA statements perform matching, weighting, and stratification, respectively. Only one of these statements can be specified.

The following sections describe the PROC PSMATCH statement and then describe the other statements in alphabetical order.

PROC PSMATCH Statement

The PROC PSMATCH statement invokes the PSMATCH procedure. Table 100.1 summarizes the options available in the PROC PSMATCH statement.

Table 100.1 Summary of PROC PSMATCH Options

Option	Description
DATA=	Specifies the input data set
REGION=	Specifies the support region of observations for stratification and matching

DATA=*SAS-data-set*

names the input SAS data set. If the propensity scores are to be derived from this data set, you must also include a PSMODEL statement to specify the binary logistic model. Otherwise, a PSDATA statement is required to identify the variable that contains either the propensity scores or the logits of the propensity scores. If you do not specify this option, the procedure uses the most recently created SAS data set.

REGION=*region* **<(***region-options***)>**
specifies an interval region of propensity scores (or equivalently, logits of propensity scores) that determines which observations are used in stratification and matching. Only those observations whose propensity scores lie in the region are used in stratification and matching. This option also determines which observations are included in the output data set if you specify the OUT(OBS=REGION) option in the OUTPUT statement (even when the STRATA and MATCH statements are not specified). By default, REGION=TREATED if you specify a MATCH statement, and REGION=ALLOBS otherwise.

You can specify the following *regions* along with their *region-options*:

REGION=ALLOBS <(*region-options***)>**
selects all available observations. You can specify the following *region-options* to select observations whose propensity scores lie in a specified range:

PSMIN=*pmin*
specifies the minimum propensity score in the support region, where $pmin \geq 0$. Observations whose propensity scores are less than *pmin* are excluded from the support region. By default, PSMIN=0, so that observations with small propensity scores are not excluded.

PSMAX=*pmax*
specifies the maximum propensity score in the support region, where $pmax \leq 1$. Observations whose propensity scores are greater than *pmax* are excluded from the support region. By default, PSMAX=1, so that observations with large propensity scores are not excluded.

You can also use the PSMIN= and PSMAX= options to exclude observations that have extreme propensity scores from the output data set.

REGION=CS <(*ext-option***)>**
selects observations whose propensity scores (or equivalently, logits of propensity scores) lie in the region of common support for the treated and control groups. This region is the largest interval that contains propensity scores (or logits of propensity scores) for subjects in both groups. The lower endpoint of the region is the larger of the minimum propensity scores (or logits of propensity scores) for the two groups. The upper endpoint is the smaller of the maximum propensity scores (or logits of propensity scores) for the two groups.

You can specify the following *ext-option*:

EXTEND <(*type-options***)> =** p **<(LOWER=**p_l **UPPER=**p_u**)>**
extends the lower and upper ends of the common support region for the support region by p, where $p \geq 0$. By default, EXTEND=0.25.

You can use the following *type-options* to prescribe the extension requirement:

DISTANCE=LPS | PS
specifies the type of the distance that is used to extend the support region.

LPS extends the region by using the logit of the propensity score.

PS extends the region by using the propensity score.

By default, DISTANCE=LPS.

MULT=ONE | STDDEV
specifies the multiplier for the extension *p* to extend the support region.

ONE extends the region by *p*.

STDDEV extends the region by *p* times the pooled estimate of the standard deviation of either LPS (DISTANCE=LPS) or PS (DISTANCE=PS), where this estimate is computed as the square root of the average of the variances in the treated and control groups.

By default, MULT=STDDEV.

The DISTANCE= and MULT= *type-options* prescribe the extension requirement as follows:

- EXTEND(DISTANCE=PS MULT=ONE)=*p* extends the specified support region by *p* in propensity score. That is, if (R_l, R_u) denotes the propensity score interval region that is computed from the specified *region*, then the range of the extended support region is given by $(R_l - p,\ R_u + p)$.
- EXTEND(DISTANCE=PS MULT=STDDEV)=*p* extends the specified support region by $p \times \hat{\sigma}_{ps}$, the square root of the average variance of the propensity score in the treated and control groups. That is, if (R_l, R_u) denotes the propensity score interval region that is computed from the specified *region*, then the range of the extended support region is given by $(R_l - p\,\hat{\sigma}_{ps},\ R_u + p\,\hat{\sigma}_{ps})$.
- EXTEND(DISTANCE=LPS MULT=ONE)=*p* extends the specified support region by *p* in the logit of the propensity score.
- EXTEND(DISTANCE=LPS MULT=STDDEV)=*p* extends the specified support region by $p \times \hat{\sigma}_{lps}$, the square root of the average variance of the logit of the propensity score in the treated and control groups.

You can specify one of the following two options to use an extension other than *p*:

LOWER=p_l extends the lower end of the specified region by p_l, where $p_l \geq 0$.

UPPER=p_u extends the upper end of the specified region by p_u, where $p_u \geq 0$.

REGION=TREATED <(*ext-option*)>
selects observations whose propensity scores lie in the region of propensity scores for observations in the treated group.

You can specify the following *ext-option*:

EXTEND <(*type-options*)>= *p* <(LOWER=p_l UPPER=p_u)>
extends the lower and upper ends of the range of treated observations for the support region by *p*, where $p \geq 0$. By default, EXTEND=0.25.

You can use the *type-options* to prescribe the extension requirement, and these are identical to the *type-options* in the REGION=CS option. You can also specify the LOWER=p_l or UPPER=p_u suboption to use an extension other than *p*.

ASSESS Statement

ASSESS < **ALLCOV** > < **LPS** > < **PS** > < **VAR=***(var-list)* > < / *assess-options* > **;**

The ASSESS statement assesses variable differences between the treated and control groups for all observations and for observations in the support region that is specified in the REGION= option.

It also assesses variable differences for matched observations if you specify a MATCH statement, and it assesses variable differences for observations by stratum if you specify a STRATA statement. In addition, the ASSESS statement assesses variable differences for weighted observations provided that the WEIGHT=NONE suboption is not specified.

You can specify the variables for assessment by using the following options:

ALLCOV
: requests an assessment of differences in the covariates that are specified in the PSMODEL statement. These variables must be binary classification variables or continuous variables in the input data set.

LPS
: requests an assessment of differences in the logit of the propensity score.

PS
: requests an assessment of differences in the propensity score.

VAR=(*var-list***)**
: requests an assessment of differences in the specified list of variables. These variables must be binary classification variables or continuous variables in the input data set. These variables can be the variables not specified in the PSMODEL statement.

If none of these options are specified, an assessment of differences in the propensity score is produced by default.

In addition, you can specify various *assess-options* after a slash (/). Table 100.2 summarizes these options.

Table 100.2 ASSESS Statement Options

Option	Description
PLOTS=	Requests plots for assessment of variable balance
STDBINVAR=	Specifies whether to standardize binary variables in the standardized mean differences table and plot
STDDEV=	Specifies the type of standard deviation to be used in the standardized mean difference computation
VARINFO	Displays variable information for the treated and control groups

PLOTS < (*global-option*) > < = *plot-request* >

PLOTS < (*global-option*) > = (*plot-request* < **...** *plot-request* >)

specifies options that control the plots.

You can specify the following *global-options*:

NODETAILS

displays plots for only two sets of observations: the set of all observations and a second set that depends on specified statements and options. This option does not apply to cloud plots.

If you specify a MATCH statement, the second set consists of matched observations (if neither WEIGHT=MATCHWGT nor WEIGHT=MATCHATEWGT is specified) or weighted matched observations (if WEIGHT=MATCHWGT or WEIGHT=MATCHATEWGT is specified). If you specify a STRATA statement, the second set consists of observations in the support region. If you specify neither a MATCH statement nor a STRATA statement, the second set consists of observations in the support region (if neither WEIGHT=ATEWGT nor WEIGHT=ATTWGT is specified) or the set of weighted observations in the support region (if WEIGHT=ATEWGT or WEIGHT=ATTWGT is specified).

ONLY

suppresses the default plots and displays only plots that are specifically requested.

ORIENT=HORIZONTAL | VERTICAL

controls the orientation of the plots.

HORIZONTAL places the lines and boxes horizontally for variable distribution plots, places the bar lengths horizontally for bar charts, places the variable values horizontally for cloud plots, places the standardized mean differences on the horizontal axis for the standardized mean differences plot, and places the graphs in a single column for CDF plots.

VERTICAL places the lines and boxes vertically for variable distribution plots, places the bar lengths vertically for bar charts, places the variable values vertically for cloud plots, places the standardized mean differences on the vertically axis for the standardized mean differences plot, and places the graphs in a single row (side-by-side) for CDF plots.

By default, ORIENT=HORIZONTAL.

You can specify the following *plot-requests*:

ALL

requests all applicable plots for all variables that are specified in the ASSESS statement. These plots include bar charts for binary classification variables; box plots, CDF plots, and cloud plots for continuous variables; and a combined standardized mean differences plot for all variables. If you specify a STRATA statement, then PROC PSMATCH also produces the plots by stratum.

BAR < (< DISPLAY=ALL | (*bar-list*) > < TYPE=FREQ | PROP >) >

BARCHART < (< DISPLAY=ALL | (*bar-list*) > < TYPE=FREQ | PROP >) >

requests comparative bar charts for binary classification variables. You can use the DISPLAY= option to select variables for which bar charts are to be displayed.

DISPLAY=ALL

requests bar charts for binary classification variables that are specified by the ALLCOV or VAR= option.

DISPLAY=(*bar-list*)

specifies a subset of the binary classification variables for which bar charts are to be displayed.

By default, DISPLAY=ALL.

You can use the TYPE= option to select either the frequencies or the proportions to be displayed in the bar charts.

TYPE=FREQ

displays frequencies of levels for the binary classification variable.

TYPE=PROP

displays proportions of levels for the binary classification variable.

By default, TYPE=PROP.

If you specify a STRATA statement, then the bar charts are also displayed by stratum.

BOX < (DISPLAY=ALL | (*box-list*)) >

BOXPLOT < (DISPLAY=ALL | (*box-list*)) >

requests box plots for continuous variables. You can use the DISPLAY= option to select variables for which box plots are to be displayed.

DISPLAY=ALL

requests box plots for all continuous variables that are specified by the ALLCOV or VAR= option. The option also requests box plots for logits of propensity scores if the LPS option is specified and propensity scores if the PS option is specified.

DISPLAY=(*box-list*)

specifies a subset of the continuous variables for which box plots are to be displayed.

By default, DISPLAY=ALL.

If you specify a STRATA statement, then the box plots are also displayed by stratum.

CDF < (DISPLAY=ALL | (*cdf-list*)) >

CDFPLOT < (DISPLAY=ALL | (*cdf-list*)) >

requests cumulative distribution (CDF) plots for continuous variables. You can use the DISPLAY= option to select variables for which CDF plots are to be displayed.

DISPLAY=ALL
requests CDF plots for all continuous variables that are specified by the ALLCOV or VAR= option. The option also requests CDF plots for logits of propensity scores if the LPS is specified and propensity scores if the PS option is specified.

DISPLAY=(*cdf-list*)
specifies a subset of the continuous variables for which CDF plots are to be displayed.

By default, DISPLAY=ALL.

If you specify a STRATA statement, then the CDF plots by stratum are also displayed.

CLOUD < (DISPLAY=ALL | (*cloud-list*)) >

CLOUDPLOT < (DISPLAY=ALL | (*cloud-list*)) >
requests cloud plots for continuous variables. The term cloud plot is used here to refer to scatter plots in which the points have been jittered by adding random noise to prevent overplotting. Jittering typically occurs when a continuous variable (such as age) is rounded to some convenient unit (such as years). You can use the DISPLAY= option to select variables for which cloud plots are to be displayed.

DISPLAY=ALL
requests cloud plots for all continuous variables that are specified by the ALLCOV or VAR= option. The option also requests cloud plots for logits of propensity scores if the LPS option is specified and propensity scores if the PS option is specified.

DISPLAY=(*cloud-list*)
specifies a subset of the continuous variables for which cloud plots are to be displayed.

By default, DISPLAY=ALL.

If you specify a STRATA statement, then cloud plots are also displayed by stratum.

NONE
suppresses all plots.

STDDIFF < (REF=*r*) >

STDDIFFPLOT < (REF=*r*) >
requests a standardized mean differences plot for all variables that are specified by the ALLCOV or VAR= option. The plot also includes logits of propensity scores if the LPS option is specified and propensity scores if the PS option is specified.

The REF=*r* option displays a shaded band that covers standardized mean differences from $-r$ to r, where $r > 0$. If you specify REF=0, the band is not displayed. By default, REF=0.25 (Rubin 2001, p. 174).

If you specify a STRATA statement, then standardized mean difference plots are also displayed by stratum. However, the shaded band is not displayed in the plot because recommended ranges for stratum-specific standardized mean differences are currently not available in the literature.

WGTCLOUD < (REF=*r*) >

WGTCLOUDPLOT < (REF=*r*) >

requests cloud plots for weights. The option is applicable if you specify the WEIGHT=ATEWGT or WEIGHT=ATTWGT option in the ASSESS statement. The term cloud plot is used here to refer to scatter plots in which the points have been jittered by adding random noise to prevent overplotting.

Observations that have large weights can be highly influential. Well-behaved ATE weights should be less than 10 times the expected weight (Stürmer et al. 2014, p. 578). For more information about expected weights, see the section "Propensity Score Weighting" on page 8217. The REF=*r* option displays a reference line at *r* times the expected weight. By default, REF=10.

Weight cloud plots for ATT weights and IPTW-ATE weights display distinct reference lines for weights for observations in the treated and control groups. For example, see Output 100.1.7. Weight cloud plots for stabilized IPTW-ATE weights display a single reference line at *r* because the expected weight is 1.

By default, PLOTS=STDDIFF.

STDBINVAR=YES | NO

specifies whether to display standardized binary variables in the standardized mean differences table and plot.

YES displays standardized binary variables in the standardized mean differences table and plot.

NO does not display standardized binary variables in the standardized mean differences table and plot, and displays raw binary variables in the standardized mean differences plot.

By default, STDBINVAR=YES.

STDDEV=POOLED < (ALLOBS=YES | NO) >

STDDEV=TREATED < (ALLOBS=YES | NO) >

specifies the standard deviation used in computing standardized mean differences.

POOLED uses the pooled standard deviation, which is computed as the square root of the average of the sample variances for the treated group and the sample variance for the control group.

TREATED uses the standard deviation of the variable values in the treated group only.

By default, STDDEV=POOLED.

The ALLOBS= option specifies the set of observations used to compute the variance.

YES uses the sample variances that are derived from all observations to compute the standardized mean differences for all observations, for observations in the support region, for stratified observations, and for matched observations.

NO uses the variance derived from all observations to compute the standardized mean differences for all observations, uses the variance derived from observations in the support region to compute the standardized mean differences for observations in the support region, uses the variance derived from stratified observations to compute the standardized mean differences for stratified observations, and uses the variance derived from matched observations to compute the standardized mean differences for matched observations.

By default, ALLOBS=YES.

VARINFO
requests a variable information table for the treated and control groups.

Options Available in Previous Releases

A number of options in the ASSESS statement from previous releases have been moved to other statements in the PSMATCH procedure. Starting in SAS/STAT 15.1, these options are obsolete in the ASSESS statement and might not be supported in future releases. Use the options in the new statements, as shown in Table 100.3.

Table 100.3 ASSESS Statement Options Moved to New Statements

Option	Description	New Statement
NLARGESTWGT=	Displays observations that have the largest weights	PSWEIGHT
NMATCHMOST=	Displays observations that have the greatest numbers of matches	MATCH
STRATUMWGT=	Specifies the stratum weights to combine statistics across strata	STRATA
WEIGHT=	Specifies the type of weight for matched observations	MATCH
WEIGHT=	Specifies the type of weight for all observations	PSWEIGHT

BY Statement

BY *variables* **;**

You can specify a BY statement in PROC PSMATCH to obtain separate analyses of observations in groups that are defined by the BY variables. When a BY statement appears, the procedure expects the input data set to be sorted in order of the BY variables. If you specify more than one BY statement, only the last one specified is used.

If your input data set is not sorted in ascending order, use one of the following alternatives:

- Sort the data by using the SORT procedure with a similar BY statement.
- Specify the NOTSORTED or DESCENDING option in the BY statement in the PSMATCH procedure. The NOTSORTED option does not mean that the data are unsorted but rather that the data are arranged in groups (according to values of the BY variables) and that these groups are not necessarily in alphabetical or increasing numeric order.
- Create an index on the BY variables by using the DATASETS procedure (in Base SAS software).

CLASS Statement

CLASS *variables* **;**

The required CLASS statement specifies the following input variables, which are used as classification variables:

- the variable to use as the treatment indicator in the PSDATA and PSMODEL statements
- the classification covariates in the logistic model in the PSMODEL statement
- the classification variables that are specified in the VAR= option in the ASSESS statement

If a PSMODEL statement is specified, the CLASS statement must precede the PSMODEL statement. Classification variables can be either character or numeric.

FREQ Statement

FREQ *variable* **;**

The FREQ statement identifies a *variable* that contains the frequency of occurrence of each observation. PROC PSMATCH treats each observation as if it appears n times, where n is the value of the FREQ variable for the observation. The FREQ statement is not allowed if a MATCH statement is specified.

ID Statement

ID *variable* **;**

The ID statement specifies one or more variables whose values identify the observations that are displayed in tables of extreme weights and most matches that are requested by the NLARGESTWGT= option in the PSWEIGHT statement and the NMATCHMOST= option in the MATCH statement.

MATCH Statement

MATCH < *options* > **;**

The MATCH statement matches observations in the control group to observations in the treatment group. The MATCH statement is not allowed if a FREQ statement is specified. The STRATA and PSWEIGHT statements do not apply if a MATCH statement is specified.

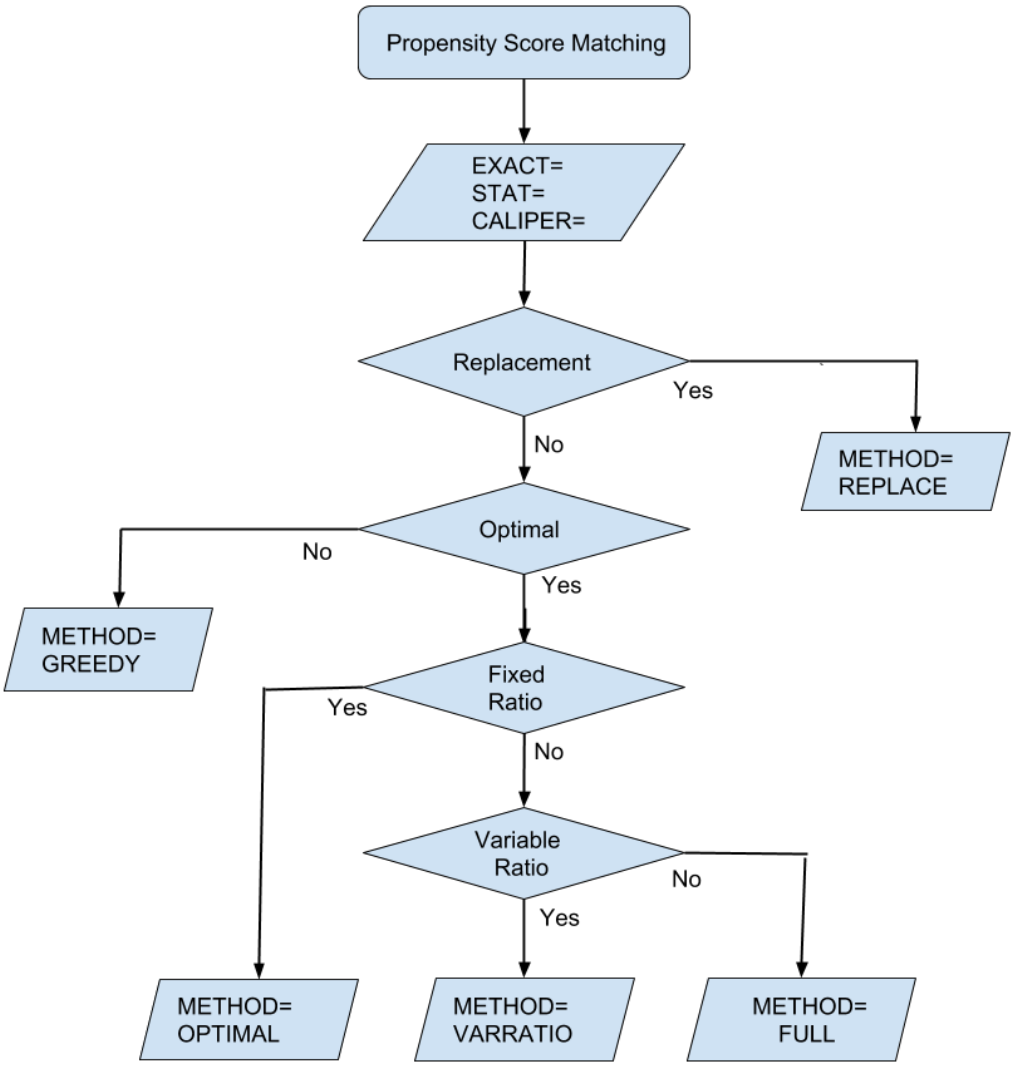
Table 100.4 summarizes the *options* in the MATCH statement.

Table 100.4 MATCH Statement Options

Option	Description
CALIPER=	Specifies the caliper width requirement for matching
DISTANCE=	Specifies the distance for comparing treated units and control units
EXACT=	Requests exact matching for specified classification variables
METHOD=	Specifies the method for matching
NMATCHMOST=	Displays observations that have the greatest numbers of matches
WEIGHT=	Specifies the type of weight for matched observations

The flowchart in Figure 100.13 displays the steps in the propensity score matching process.

Figure 100.13 Propensity Score Matching Options



You can specify the following *options* in the MATCH statement:

CALIPER < (*caliper-options***) > = *r***

specifies the caliper width requirement for matching, where *r* is either missing or greater than 0. The difference in propensity scores (or logits of propensity scores) between the treated unit and its matching control unit must be less than or equal to *r*. If you specify CALIPER=., then the caliper requirement is ignored. By default, CALIPER=0.25 (Rosenbaum and Rubin 1985, p. 37). Austin (2011a) has shown that CALIPER=0.20 is optimal in many settings.

You can use the following two *caliper-options* to prescribe the caliper requirement:

MULT=ONE | STDDEV

specifies the multiplier for the specified caliper width *r*.

ONE uses *r* for the caliper width.

STDDEV uses *r* times the pooled estimate of the standard deviation of the logit of the propensity score (if you specify DISTANCE=LPS) or the propensity score (if you specify DISTANCE=PS), where this estimate is computed as the square root of the average of the variances in the treated and control groups.

By default, MULT=STDDEV.

MAHDISTANCE=LPS | PS

specifies the type of distance to be used in the caliper width computation if you specify the DISTANCE=MAH option in the MATCH statement.

LPS uses the logit of the propensity score.

PS uses the propensity score scale.

By default, MAHDISTANCE=LPS.

If you specify the DISTANCE=LPS or DISTANCE=PS option in the MATCH statement, the specified type of distance is used in the caliper width computation.

DISTANCE=*distance*

specifies the type of distance to be compared when treated units are matched to control units. If you specify the DISTANCE=LPS or DISTANCE=PS option, the specified type of distance is also used in the caliper width computation. By default, DISTANCE=LPS. You can specify the following values for *distance*:

LPS

specifies matching that minimizes the difference between the logits of the propensity scores for the two units.

PS

specifies matching that minimizes the difference between the propensity scores for the two units.

MAH (*var-options* **<** / *mah-options* **>)**
specifies matching that minimizes the Mahalanobis distance between the two units.

You use the following *var-options* to select at least one variable for computing the Mahalanobis distance:

LPS
includes the logit of the propensity score.

PS
includes the propensity score.

VAR=(*var-list***)**
includes variables in the specified *var-list*. These variables must be continuous variables in the input data set.

You can also specify the following *mah-options*:

COV=CONTROL | IDENTITY | POOLED
specifies the type of covariance matrix in the Mahalanobis distance.

CONTROL uses the covariance matrix that is computed from observations in the control group.

IDENTITY uses the identity matrix, and the resulting distance is the Euclidean distance.

POOLED uses the pooled covariance matrix that is computed from observations in the treated group and observations in the control group.

By default, COV=CONTROL.

SQRT=YES | NO
specifies whether to apply the square root transformation to the Mahalanobis distance in the difference computation. This *mah-option* does not affect matching results for greedy nearest neighbor matching or matching with replacement. It affects only results for optimal matching that minimize the total absolute difference.

YES uses the square root of the Mahalanobis distance as the difference between treated and control units.

NO uses the Mahalanobis distance as the difference between treated and control units.

By default, SQRT=YES.

EXACT=*variable* **|** **(***variables***)**
specifies classification *variables* that are to be matched exactly. That is, observations in each matched set must have the same values for these variables. The *variables* must be specified in the CLASS statement.

METHOD=*method* **<(***method-options***)>**
specifies the *method* for the matching. You can specify the following *methods* and *method-options*. By default, METHOD=OPTIMAL.

METHOD=FULL (KMAX=*kmax* **<** *full-options* **>)**
requests optimal full matching. Each treated unit is matched with one or more control units, and each control unit (if matched) is matched with one or more treated units. If the specified total number of control units to be matched is less than the number of available control units, then constrained full matching is performed—that is, not all observations are matched.

You must specify the following suboption:

KMAX=*kmax*
specifies the maximum number of control units to be matched with each treated unit, where $kmax \geq 1$.

You can also specify the following *full-options*:

KMAXTREATED=*kmaxtrt*

KMAXTRT=*kmaxtrt*
specifies the maximum number of treated units for each control, where $kmaxtrt \geq 1$. By default, KMAXTREATED=2.

KMEAN=*kmean*
specifies the average number of control units to be matched with each treated unit. If the resulting number of control units is greater than the number of control units in the support region, the number of control units in the support region is used.

NCONTROL=*m*
specifies the number of control units to be matched. If *m* is greater than the number of control units in the support region, the number of control units in the support region is used.

PCTCONTROL=*p*
specifies the percentage of the total number of control units to be matched. If the resulting number of control units is greater than the total number of control units in the support region, the number of control units in the support region is used.

You can specify only one of the KMEAN=, NCONTROL=, and PCTCONTROL= options for the number of control units in the matched data set. If you do not specify any of the KMEAN=, NCONTROL=, and PCTCONTROL= options, KMEAN= (*kmax* + 1 / *kmaxtrt*) / 2 is used.

METHOD=GREEDY <(K=*k* **ORDER=***order-option***)>**
requests greedy nearest neighbor matching, in which each treated unit is sequentially matched with the *k* nearest control units. Matching depends on the ordering of the treated units, which you can specify in the ORDER= suboption.

You can specify the following suboptions:

K=*k*
specifies the number of matching control units, where $k > 0$, for each treated unit. PROC PSMATCH performs *k* separate loops of matching for treated units. In each loop, the nearest control unit is sequentially matched to each treated unit. By default, K=1 (one control unit for each treated unit).

ORDER=ASCENDING | DESCENDING | RANDOM < (SEED=*number*) >
specifies the ordering of treated units that are used to find the matching control units. You can specify one of the following values:

ASCENDING
orders the treated units in ascending order of the propensity score.

DESCENDING
orders the treated units in descending order of the propensity score.

RANDOM < (SEED=*number*) >
orders the treated units in random order of the propensity score. The SEED= suboption specifies a positive integer to start the pseudorandom number generator. If the SEED= option is not specified, the value is generated from reading the time of day from the computer's clock.

By default, ORDER=DESCENDING.

METHOD=OPTIMAL < (K=*k*) >
requests optimal fixed ratio matching. The K=*k* suboption specifies the number of matching control units, where $k > 0$, for each treated unit. By default, K=1 (one control unit is matched with each treated unit).

METHOD=REPLACE < (K=*k*) >
requests a fixed number *k* of unique matching control units for each treated unit, where the matched control units are selected with replacement. This means that each control unit can be matched to more than one treated unit, but it can only be matched once to the same treated unit. The K=*k* suboption specifies the number of matching control units, where $k > 0$, for each treated unit. By default, K=1 (one control unit is matched with each treated unit).

METHOD=VARRATIO (KMAX=*kmax* < *vr-options* >)
requests optimal variable ratio matching. Each treated unit is matched with one or more control units.

You must specify the following suboption:

KMAX=*kmax*
specifies the maximum number of control units to be matched with each treated unit, where $kmax \geq 1$.

You can also specify the following *vr-options*:

KMEAN=*kmean*
specifies the average number of control units to be matched with each treated unit. If the resulting number of control units is greater than the total number of control units in the support region, the number of control units in the support region is used.

KMIN=*kmin*
specifies the minimum number of control units to be matched with each treated unit. By default, KMIN=1.

NCONTROL=*m*
specifies the total number of control units to be matched. If *m* is greater than the total number of control units in the support region, the number of control units in the support region is used.

PCTCONTROL=*p*
specifies the percentage of total control units to be matched. If the resulting number of control units is greater than the total number of control units in the support region, the number of control units in the support region is used.

You can specify only one of the KMEAN=, NCONTROL=, and PCTCONTROL= options for the number of control units in the matched data set. If you do not specify any of the KMEAN=, NCONTROL=, and PCTCONTROL= options, then KMEAN= (*kmin* + *kmax*) / 2 is used.

NMATCHMOST=*n*
displays a table of the observations that have the greatest numbers of matches, where $n \leq 50$. This option displays observation numbers and numbers of matches for the *n* observations that have the greatest numbers of matches in the treated and control groups. If an ID statement is also specified, the corresponding values of the ID variables are also displayed and serve to identify the observations. The option is not applicable to greedy matching (METHOD=GREEDY) and optimal fixed ratio matching (METHOD=OPTIMAL), where a fixed number of control units are matched to each treated unit. By default, *n*=0 and the table is not displayed.

WEIGHT=ATEWGT | ATTWGT | EQUAL | MATCHATEWGT | MATCHATTWGT | MATCHWGT | NONE
specifies the type of weight for matched observations.

ATEWGT | MATCHATEWGT
weights the treatment group up to the total size of the matched set. That is, in each matched set, the total weight of treated units equals the total number of units in the matched set, and the total weight of control units also equals the total number of units in the matched set. This weighting is available only for an optimal full matching (METHOD=FULL), and it is appropriate for estimating the ATE. For more information about using match weighting to estimate the ATE, see the section "ATE Weighting after Full Matching" on page 8227.

ATTWGT | MATCHATTWGT | MATCHWGT
weights the control group up to the size of the treatment group in the matched set. That is, in each matched set, the total weight of control units equals the number of treated units in the matched set. This weighting is appropriate for estimating the ATT. For more information about using match weighting to estimate the ATT, see the sections "ATT Weighting after Matching without Replacement" on page 8226 and "ATT Weighting after Matching with Replacement" on page 8227.

EQUAL | NONE
uses the same weight for matched observations. That is, each matched unit has a weight of 1, regardless of the treatment group.

By default, WEIGHT=ATTWGT.

OUTPUT Statement

OUTPUT OUT < **(OBS=***obs-value***)** >**=***SAS-data-set* < *keyword***=***name* < *keyword***=***name* . . . > > **;**

The OUTPUT statement specifies the output data set and variables. You must specify the following option:

OUT < (OBS=*obs-value***) >=***SAS-data-set*
names the output data set. The data set also includes the results of matching if you provide the MATCH statement. You can specify one of the following values for *obs-value*:

OBS=ALL requests that the output data set contain all observations.

OBS=REGION requests that the output data set contain only observations in the specified support region.

OBS=MATCH requests that the output data set contain only the matched treated units and control units. This option applies only if you specify the MATCH statement.

By default, OBS=ALL.

Table 100.5 summarizes the *keywords* that are used to create and name the output variables.

Table 100.5 OUTPUT Statement Keywords

Keyword	Description
LPS	Specifies the variable that provides the logit of the propensity score
MATCHID	Specifies the variable that provides identification numbers for the matched units
PS	Specifies the variable that provides the propensity score
STRATA	Specifies the variable that numbers the strata
WEIGHT	Specifies the weight variable

LPS=*name*
creates and names the variable that provides the logit of the propensity score.

MATCHID=*name*
creates and names the variable that provides identification numbers for the matched treated and control units. This suboption applies only if you specify the MATCH statement.

PS=*name*
creates and names the variable that provides the propensity score.

If this option is not specified and the PS= option in the PSDATA statement is also not specified, then the variable that provides the propensity score is automatically created with *name*=_PS_.

STRATA=*name*
creates and names the variable that numbers the strata. The suboption applies only if the STRATA statement is specified.

If this option is not specified but the STRATA statement is specified, then the variable that numbers the strata is automatically created with *name*=_STRATA_.

WEIGHT=*name*
creates and names the weight variable. The output weight variable depends on the specified statements:

- If you specify a MATCH statement, the weight variable contains weights of matched observations as specified by the WEIGHT= option in the MATCH statement.
- If you specify a PSWEIGHT statement, the weight variable contains weights of observations as specified by the WEIGHT= option in the PSWEIGHT statement.

This option does not apply if you specify a STRATA statement.

Keywords Available in Previous Releases

A number of keywords in the OUTPUT statement from previous releases have been replaced by the WEIGHT keyword in the OUTPUT statement with the corresponding new WEIGHT= option in the PSWEIGHT and MATCH statements. These keywords are now obsolete in the OUTPUT statement and might not be supported in future releases.

Table 100.6 summarizes these *keywords* that are used to create and name the output variables.

Table 100.6 OUTPUT Statement Keywords Replaced by WEIGHT Keyword

Keyword	Description	New Keyword
ATEWGT	Specifies the weight variable for ATE weighting	WEIGHT
ATTWGT	Specifies the weight variable for ATT weighting	WEIGHT
MATCHATEWGT	Specifies the weight variable for ATE weighting in matching	WEIGHT
MATCHATTWGT	Specifies the weight variable for ATT weighting in matching	WEIGHT

PSDATA Statement

PSDATA TREATVAR=*treatvar* < **(TREATED**='*level*' | *keyword*) > *ps-option* **;**

You use the PSDATA statement when the DATA= data set contains precomputed propensity scores or logits of propensity scores and you want to base the propensity score analysis on these scores rather than using the PSMODEL statement to specify a logistic regression model for computing the scores. The PSDATA statement specifies the variable in the DATA= data set that is the treatment indicator variable and a variable that contains the propensity scores or logits of the propensity scores. Either the PSMODEL statement or the PSDATA statement is required, and only one can be used.

You must specify the following TREATVAR= option:

TREATVAR=*treatvar* **< (TREATED='***level***' | *keyword*) >**
names the treatment indicator variable, *treatvar*, which must be a binary classification variable that is specified in the CLASS statement.

The TREATED= suboption specifies the treated level for the binary treatment variable. You can specify the value of the treated *level* in quotation marks, or you can specify one of the following *keywords*:

FIRST designates the first ordered level as the treated group.

LAST designates the last ordered level as the treated group.

By default, TREATED=FIRST.

You must also specify one (and only one) of the following *ps-options*:

PS=*name*
names the variable that contains propensity scores, where the variable *name* must be a variable in the DATA= data set.

LPS=*name*
names the variable that contains logits of propensity scores, where the variable *name* must be a variable in the DATA= data set.

PSMODEL Statement

PSMODEL *treatvar* < **(***trt-option***)** > **=** < *effects* > < / **WEIGHT=** *weight* > **;**

The PSMODEL statement specifies the logistic regression model for computing propensity scores. Either the PSMODEL statement or the PSDATA statement is required to provide the propensity scores, and only one can be specified.

The treatment indicator variable *treatvar* must be a binary classification variable that is listed in the CLASS statement, and the *effects* are the explanatory effects, which can include variables, main effects, interactions, and nested effects for the logistic regression model.

You can use the following *trt-option* to specify the treated level for the binary treatment variable:

TREATED='*level***' | *keyword***

models the probability of the specified treated level. You can specify the value of the treated *level* in quotation marks, or you can specify one of the following *keywords*:

FIRST designates the first ordered level as the treated group.

LAST designates the last ordered level as the treated group.

By default, TREATED=FIRST.

You can specify the following option to fit a weighted logistic regression:

WEIGHT=*weight*

specifies a variable that contains the weight of each observation that is used in fitting the logistic regression model to derive the propensity scores. These weights should not be confused with weights that are derived from the propensity scores by the PSMATCH procedure.

PSWEIGHT Statement

PSWEIGHT < *options* > **;**

The PSWEIGHT statement computes weights for the observations on the basis of propensity scores. The PSWEIGHT statement does not apply when you specify either a MATCH statement or a STRATA statement.

Table 100.7 summarizes the *options* in the PSWEIGHT statement.

Table 100.7 PSWEIGHT Statement Options

Option	Description
NLARGESTWGT=	Displays observations that have the largest weights
WEIGHT=	Specifies the type of weight for observations

NLARGESTWGT=*n*

displays a table of the observations that have the most extreme weights, where $n \leq 50$. This option displays observation numbers and weights for the n observations that have the largest weights in the treated and control groups. If you specify an ID statement, the corresponding values of the ID variables are also displayed and serve to identify the observations. By default, n=0 and the table is not displayed.

WEIGHT=ATEWGT | ATTWGT

specifies the type of weight for observations.

ATEWGT < (STABILIZE=YES | NO) >

uses inverse probability of treatment weighting (IPTW). These weights are appropriate for estimation of the ATE. For more information about IPTW-ATE weighting, see the section "Inverse Probability of Treatment Weighting" on page 8217.

YES requests a stabilized inverse probability of treatment weighting.

NO requests an inverse probability of treatment weighting.

By default, STABILIZE=NO.

ATTWGT
uses ATT weighting (also called weighting by odds) to weight the control group up to the treatment group. These weights are appropriate for estimation of the ATT. For more information about ATT weighting, see the section "ATT Weighting" on page 8219.

By default, WEIGHT=ATTWGT.

STRATA Statement

STRATA < *options* > **;**

The STRATA statement divides observations in the support region into strata based on propensity scores, where the support region is specified in the REGION= option in the PROC PSMATCH statement. The STRATA statement does not apply when you specify the MATCH statement.

Table 100.8 summarizes the *options* in the STRATA statement.

Table 100.8 STRATA Statement Options

Option	Description
KEY=	Specifies how to use the observations to construct the strata
NSTRATA=	Specifies the number of strata
STRATUMWGT=	Specifies the type of stratum weights to combine statistics across strata

KEY=TOTAL | TREATED
specifies how to use the observations to construct the strata.

TOTAL requests that each stratum contain approximately the same number of total units, which can be in either the treated group or the control group.

TREATED requests that each stratum contain approximately the same number of units in the treated group.

By default, KEY=TREATED. This option balances the number of treated units across strata, so that a reliable estimate of the treatment effect can be obtained for each stratum. However, a common alternative is to construct strata so that each stratum has the same number of total units. You can request this approach by specifying KEY=TOTAL. In either case, you should examine the number of treated units and the number of control units in each stratum to make sure that a reliable estimate can be obtained for each stratum.

NSTRATA=*n*

specifies the number of strata, where $n \geq 2$. Only observations in the support region are stratified. By default, NSTRATA=5.

STRATUMWGT=TOTAL | TREATED

STRATUMWEIGHT=TOTAL | TREATED

specifies the type of stratum weights that the PSMATCH procedure uses to combine stratum-specific statistics in the assessment of variable balance after stratification. For more information about stratum weights, see the section "Weighting after Stratification" on page 8221.

TOTAL uses the proportions of total units (treated and control units combined) in the strata as the weights. These weights sum up to one and are appropriate for estimating the ATE.

TREATED uses the proportions of treated units in the strata as the weights. These weights sum up to one and are appropriate for estimating the ATT.

By default, STRATUMWGT=TOTAL.

For more information, see the section "Propensity Score Stratification" on page 8220.

Details: PSMATCH Procedure

Output Table Enhancements

This section describes the enhancements in the output tables in this release.

Propensity Score Information Table

An observation is added for weighted observations if you specify a PSWEIGHT statement (as shown in Output 100.1.2), for stratification if you specify a STRATA statement (as shown in Output 100.2.2), and for weighted matched observations if you specify a MATCH statement without the WEIGHT=NONE option (as shown in Output 100.3.2).

In addition, the variable Weight is added for observation weights if you specify a PSWEIGHT statement.

Strata Information Table

The variable StratumWeight is added for stratum weights, as shown in Output 100.2.3.

Observational Studies Contrasted with Randomized Trials

In a randomized study, such as a randomized controlled trial, the subjects are randomly assigned to a treated (exposure) group or a control (nonexposure) group. Random assignment ensures that the distribution of the covariates is the same in both groups, and the treatment effect can be estimated from a direct comparison of 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. The presence of confounding requires statistical approaches that remove the effects of confounding when estimating the effect of treatment.

Observational studies are carried out when it is impractical or unethical to perform a randomized experiment. One example of an observational study is a retrospective cohort study that examines the relationship between a specific disease and a risk factor that occurred in the past; another example is a nonrandomized clinical trial that uses existing data such as control units that are extracted from a registry database.

The approach that the PSMATCH procedure uses and the following terminology are based on the potential outcomes framework for causal inference, which was introduced by Rubin (1974) and Rosenbaum and Rubin (1983). Under this framework, each individual typically has two potential outcomes in an observational study whose goal is to estimate the effect of a treatment:

- $Y(1)$, the outcome that would be observed if the individual receives the treatment.
- $Y(0)$, the outcome that would be observed if the individual does not receive the treatment under identical circumstances to those under which the subject would have received the treatment.

However, only one outcome can be observed.

The treatment effect is defined as $Y(1) - Y(0)$, and the average treatment effect is defined as:

$$\text{ATE} = E\left(Y(1) - Y(0)\right)$$

The average treatment effect for the treated (individuals who actually receive treatment) is defined as:

$$\text{ATT} = E\left(Y(1) - Y(0) \mid T = 1\right)$$

where T denotes the treatment assignment.

In a randomized trial, the potential outcomes $(Y(0), Y(1))$ and the treatment assignment (T) are independent:

$$(Y(0), Y(1)) \perp\!\!\!\perp T$$

Thus, the average treatment effect (ATE) is identical to the average treatment effect for the treated (ATT), which can be expressed as follows and can be estimated from the observed data:

$$E\left(Y(1) \mid T = 1\right) - E(Y(0) \mid T = 0)$$

In an observational study, the potential outcomes $(Y(0), Y(1))$ and the treatment assignment (T) might not be independent. In this case, the ATE and ATT are not the same. Furthermore, outcomes cannot be compared directly to estimate the treatment effect. In particular,

$$\begin{aligned} \text{ATT} &= E\left(Y(1) - Y(0) \mid T = 1\right) \\ &= E\left(Y(1) \mid T = 1\right) - E\left(Y(0) \mid T = 0\right) + E\left(Y(0) \mid T = 0\right) - E\left(Y(0) \mid T = 1\right) \end{aligned}$$

The following term can be estimated from the observed data:

$$E\left(Y(1) \mid T = 1\right) - E\left(Y(0) \mid T = 0\right)$$

However, the selection bias cannot be estimated from the observed data:

$$E\left(Y(0) \mid T = 0\right) - E\left(Y(0) \mid T = 1\right)$$

The selection bias is the average difference in the response that would be observed between individuals in the control group who do not receive treatment and individuals in the treatment group who do not receive treatment. Thus, the usual observed difference between the treated and control groups cannot be used to estimate the treatment effect. For subjects who are not randomly assigned to the treated and control groups, the baseline variables could be related to both the treatment assignment and the outcome, and consequently direct comparison of outcomes could result in biased estimates.

One strategy for correctly estimating the treatment effect is based on the propensity score, which is the conditional probability of the treatment assignment given the observed variables. You use propensity scores to account for confounding by weighting observations, by creating strata of subjects that have similar propensity scores, or by matching control subjects to treated subjects. This is done prior to the outcome analysis and without knowledge of the outcome variable (Rosenbaum and Rubin 1984; Stuart 2010, p. 5). The following section describes the propensity score approach.

Propensity Score Analysis

In a randomized study, the potential outcomes within treatment and control groups are unrelated to treatment assignment because individuals are randomly assigned to the groups. Consequently the treatment assignment given the variables X is strongly ignorable.

Rosenbaum and Rubin (1983) defined treatment assignment to be strongly ignorable when two conditions are met. The first condition (unconfoundedness) states that the potential outcomes $(Y(0), Y(1))$ and the treatment assignment (T) are conditionally independent given the observed baseline variables:

$$(Y(0), Y(1)) \perp\!\!\!\perp T \mid X = x$$

This condition is called the "no unmeasured confounders" assumption because it assumes that all the variables that affect both the outcome and the treatment assignment have been measured. The second condition (probabilistic assignment) states that there is a positive probability that a subject receives each treatment:

$$0 < \Pr(T = 1 \mid X = x) < 1$$

When the treatment assignment in an observational study is assumed to be strongly ignorable, Rosenbaum and Rubin (1983, p. 43) showed that unbiased estimates of average treatment effects can be obtained by conditioning on the propensity score *e(x)*, which is the probability of the treatment assignment conditional on a set of observed variables *X*:

$$e(x) = \Pr(T = 1 \mid X = x)$$

At any value of the propensity score *e(x)*, the difference between the treatment and control means is an unbiased estimate of the average treatment effect at *e(x)*. Consequently, matching on the propensity score and propensity score stratification also produce unbiased estimates of treatment effects (Rosenbaum and Rubin 1983, p. 44).

Furthermore, the propensity score is a balancing score. At each value of the propensity score, the distributions of the variables X are the same in the treated and control groups (Rosenbaum and Rubin 1983, p. 44; Stuart 2010, p. 6). Thus, the treatment assignment *T* and observed variables *X* are conditionally independent given the propensity score Rosenbaum (2010, p. 72):

$$x \perp\!\!\!\perp T \mid e(x)$$

Propensity score analysis attempts to replicate the properties of a randomized trial with respect to the observed variables X. The steps involved in this analysis are described in the section "Process of Propensity Score Analysis" on page 8179.

The following subsections describe the support region and the propensity score methods that are available in the PSMATCH procedure.

Support Region

For stratification and matching, the PSMATCH procedure selects observations whose propensity scores lie in a support region that can be defined in several ways:

- Selecting all available observations. You can request this definition by specifying REGION=ALLOBS in the PROC PSMATCH statement.
- Selecting observations whose propensity scores lie in a specified range. You can request this definition by specifying REGION=ALLOBS and then additionally specifying range options.
- Selecting observations whose propensity scores lie in the region of common support for the propensity scores for observations in the treated and control groups. You can request this definition by specifying REGION=CS. This region can be extended by specifying the EXTEND suboption.
- Selecting observations whose propensity scores lie in the region of propensity scores for observations in the treated group. You can request this definition by specifying REGION=TREATED. This region can be extended by specifying the EXTEND suboption.

In combination with the REGION= option, you can specify the OUT(OBS=REGION) option in the OUTPUT statement to request that only observations in the support region be included in the output data set. You can specify this combination even without the use of stratification or matching. For example, you can use the REGION=ALLOBS(PMSIN=0.1 PSMAX=0.9) option to include only observations whose propensity scores are greater than or equal to 0.1 and less than or equal to 0.9 in the output data set.

Propensity Score Methods

You can use the propensity score methods in the PSMATCH procedure to create an output data set that contains a sample that has been adjusted (either by matching, stratification, or weighting) so that the distributions of the variables are balanced between the treated and control groups. The two groups differ only randomly in their observed or measured variables, as in a randomized study. You can then use the output data set in an outcome analysis to estimate the effect of the treatment.

The following propensity score methods are available in the PSMATCH procedure:

- weighting, which creates weights that are appropriate for estimating the ATE and ATT
- stratification, which creates strata based on propensity scores
- matching, which matches treated units with control units

Note that the outcome variable is not involved in these methods. For more information about these methods, see the sections "Propensity Score Weighting" on page 8217, "Propensity Score Stratification" on page 8220, and "Matching Process" on page 8222.

Propensity Score Weighting

The PSMATCH procedure provides the following methods for weighting observations when matching is not used:

- inverse probability of treatment weighting (IPTW), which is used to estimate the ATE
- stabilized IPTW-ATE weighting, which is used to estimate the ATE
- ATT weighting (also called weighting by odds), which is used to estimate the ATT

If an observation has a propensity score close to 0 or 1, its large IPTW-ATE or ATT weight might incorrectly affect the results in the subsequent weighted outcome analysis. You can use the PSMATCH procedure to examine the observations that have extreme weights.

The PSMATCH procedure also provides methods for weighting matched observations when matching is used (see the section "Weighting after Matching" on page 8226) and for weighting strata when stratification is used (see the section "Weighting after Stratification" on page 8221).

Inverse Probability of Treatment Weighting

Inverse probability of treatment weighting (IPTW) computes the weight for the *j*th observation with propensity score p_j as

$$w_j = \begin{cases} \frac{1}{p_j} & \text{for observations in the treated group} \\ \frac{1}{1-p_j} & \text{for observations in the control group} \end{cases}$$

These weights can be used in an outcome analysis to estimate the average treatment effect,

$$\text{ATE} = E(Y(1) - Y(0))$$

by weighting the two groups up to the full population. For example, for a treated unit with $p_j = 0.25$, the weight is 4, which represents four units in the full population.

Expected IPTW-ATE weights are given by

$$\bar{w} = \begin{cases} \frac{N_t+N_c}{N_t} = \frac{1}{p_t} & \text{for observations in the treated group} \\ \frac{N_t+N_c}{N_c} = \frac{1}{1-p_t} & \text{for observations in the control group} \end{cases}$$

where $p_t = N_t/(N_t + N_c)$ is the proportion of individuals in the treated group.

The PLOTS=WGTCLOUD option in the ASSESS statement requests cloud plots for weights. The plot displays a reference line at r/p_t for observations in the treated group and a reference line at $r/(1-p_t)$ for observations in the control group, where r=10 by default. You can specify a different value for r in the PLOTS=WGTCLOUD(REF=r) option.

You can request inverse probability of treatment weighting by specifying the WEIGHT=ATEWGT option in the PSWEIGHT statement, and then by specifying the WEIGHT= option in the OUTPUT statement to create a variable that contains these weights.

Stabilized IPTW-ATE Weighting

If a treated unit has a propensity score close to 0 or a control unit has a propensity score close to 1, the resulting IPTW-ATE weight can be large. If a few observations have very large weights, the resulting IPTW-ATE estimator has a large variance and is not approximately normally distributed (Robins, Hernan, and Brumback 2000, p. 554).

In order to reduce large variances of this type, Robins, Hernan, and Brumback (2000, p. 554) replace the IPTW-ATE weights with stabilized IPTW-ATE weights:

$$w_j^* = \begin{cases} p_t w_j = \frac{p_t}{p_j} & \text{for observations in the treated group} \\ (1-p_t)w_j = \frac{1-p_t}{1-p_j} & \text{for observations in the control group} \end{cases}$$

where $p_t = N_t/(N_t + N_c)$ is the proportion of individuals in the treated group.

That is, the stabilized IPTW-ATE weights are computed by multiplying the IPTW-ATE weights by the marginal probability of receiving the given treatment. Thus, the expected stabilized IPTW-ATE weight is 1 for observations in the treated group and for observations in the control group.

You can request stabilized inverse probability of treatment weighting by specifying the WEIGHT=ATEWGT(STABILIZE=YES) option in the PSWEIGHT statement, and then by specifying the WEIGHT= option in the OUTPUT statement to create a variable that contains these weights.

Observations that have large weights can be highly influential, and well-behaved stabilized weights should have a mean stabilized weight close to 1 and a maximum stabilized weight less than 10 (Stürmer et al. 2014, p. 578). That is, in each treatment group, ATE weights should have a mean IPTW-ATE weight close to their expected weight and a maximum IPTW-ATE weight less than 10 times their expected weight. For information about these expected weights, see the section "Inverse Probability of Treatment Weighting" on page 8217.

ATT Weighting

ATT weighting (also called weighting by odds) computes the weight for the *j*th observation with propensity score p_j as

$$w_j = \begin{cases} 1 & \text{for observations in the treated group} \\ \frac{p_j}{1-p_j} & \text{for observations in the control group} \end{cases}$$

These weights can be used in an outcome analysis to estimate the following average treatment effect for the treated units (individuals who actually receive treatment) by weighting the control group up to the treated group:

$$\text{ATT} = E(Y(1) - Y(0) \mid T = 1)$$

For example, for a control unit with $p_j = 0.75$, the weight is 3, which represents three units in the treated population.

The expected weight for observations in the control group is given by

$$\bar{w} = \frac{N_t}{N_c} = \frac{p_t}{1 - p_t}$$

where $p_t = N_t/(N_t + N_c)$ is the proportion of individuals in the treated group.

The PLOTS=WGTCLOUD option in the ASSESS statement requests cloud plots for weights. The plot displays a reference line at $r\,p_t/(1 - p_t)$ for observations in the control group, where *r*=10 by default. You can specify a different value for *r* with the PLOTS=WGTCLOUD(REF=*r*) option.

You can request ATT weighting by specifying the WEIGHT=ATTWGT option in the PSWEIGHT statement, and then by specifying the WEIGHT= option in the OUTPUT statement to create a variable that contains these weights.

Large Propensity Score Weights

For IPTW-ATE weighting, if a treated unit has a propensity score close to 0 or a control unit has a propensity score close to 1, the resulting weight can be large. Similarly, for ATT weighting, if a control unit has a propensity score close to 1, the resulting weight can also be large. If a few observations have very large weights, the resulting IPTW-ATE or ATT estimator has a large variance.

You can use the NLARGESTWGT=*n* option to request a table that displays the *n* largest IPTW-ATE or ATT weights in the treated and control groups. You can exclude observations that have extreme weights in the outcome analysis, and the inference is for the resulting subset of observations. You can examine the observations that have extreme weights, find the covariate values that are associated with these extreme weights, and exclude these observations based on covariate values for a more robust interpretation.

You can also specify the PSMIN= and PSMAX= suboptions in the REGION=ALLOBS option in the PROC PSMATCH statement and the OUT(OBS=REGION) option in the OUTPUT statement to exclude observations that have extreme weights from the output data set.

Propensity Score Stratification

Propensity stratification divides the observations into strata that have similar propensity scores, with the objective of balancing the observed variables between treated and control units within each stratum. The treatment effect can then be estimated by combining stratum-specific estimates of treatment effect. Rosenbaum and Rubin (1984, p. 521) show that an adjusted estimate of this type that is based on five strata can remove approximately 90% of the bias in the crude or unadjusted estimate.

The PSMATCH procedure performs stratification when you specify the STRATA statement, which divides the observations contained in the support region into strata (you specify the support region in the REGION= option in the PROC PSMATCH statement).

In general, when observations are stratified, it is common to require the same number of observations in each stratum. However, in the context of propensity score analysis, the number of units in the control group tends to be much larger than the number of units in the treated group. Consequently, this requirement can produce strata for which the number of units in the treated group is insufficient to compute reliable stratum-specific estimates of the treatment effect.

The KEY=TREATED option (which is the default) in the STRATA statement avoids this problem by allocating approximately the same number of treated units to each stratum. Alternatively, you can specify the KEY=TOTAL option to allocate approximately the same number of observations (for either treated or control units) to each stratum. Regardless of the method of allocation, you should ensure that the number of treated units and the number of control units in each stratum are sufficient to estimate the treatment effect.

To assess the variable balance after stratification, you can use the STRATUMWGT= option in the STRATA statement to specify the stratum weights, compute the weighted averages of stratum-specific variable averages in the treated group and in the control group, and then compare the resulting weighted averages between the treated and control groups.

In the outcome analysis, you can use the weighted average of the stratum-specific treatment estimates to estimate the treatment effect. You can estimate the ATT if you weight by the stratum-specific number of treated units, and you can estimate the ATE if you weight by the stratum-specific number of units (treated and control units combined) (Stuart 2010, p. 13; Guo and Fraser 2015, pp. 76–77).

The STRATUMWGT=TOTAL option uses the proportional size of the stratum as the stratum weight. The proportional size is the number of total units (treated and control) in the stratum divided by the total number of units. Stratum weights of this type are appropriate for estimating the ATE. The STRATUMWGT=TREATED option uses the proportional number of treated units as the stratum weight. This number is the number of treated units in the stratum divided by the total number of treated units. Stratum weights of this type are appropriate for estimating the ATT. The following section provides more details about weighting after stratification.

Weighting after Stratification

The STRATA statement creates strata of observations that have similar propensity scores. The NSTRATA= option specifies the number of strata. The KEY=TOTAL option allocates approximately the same number of total units to each stratum, and the KEY=TREATED option allocates the same number of treated units to each stratum.

After stratification, you can use the weighted average of the stratum-specific treatment estimates to estimate the treatment effect in the outcome analysis. The particular weights that you use depend on the estimator (ATE or ATT). Two commonly used stratum weights are weighting by the total units and weighting by the treated units.

The PSMATCH procedure provides the following stratum weights to assess the variable balance after stratification:

- STRATUMWGT=TOTAL (weighting by the number of total units in stratum), which is used to estimate the ATE
- STRATUMWGT=TREATED (weighting by the number of treated units in stratum), which is used to estimate the ATT

Thus, a stratum weight is computed as

$$w_g = \frac{w_{1g}}{\sum_g w_{1g}}$$

where g is the stratum index and

$$w_{1g} = \begin{cases} N_{t(g)} + N_{c(g)} & \text{if STRATUMWGT=TOTAL} \\ N_{t(g)} & \text{if STRATUMWGT=TREATED} \end{cases}$$

where $N_{t(g)}$ is the number of treated units and $N_{c(g)}$ is the number of control units in the gth stratum.

Let $\bar{x}_{\mathrm{t(g)}}$ be the mean of treated units in the gth stratum and $\bar{x}_{\mathrm{c(g)}}$ be the mean of control units in the gth stratum, with corresponding sample variances $V(x_{\mathrm{t(g)}})$ and $V(x_{\mathrm{c(g)}})$. Then the weighted stratum means for the treated and control groups are

$$\bar{x}_{\mathrm{t(S)}} = \sum_g w_g \bar{x}_{\mathrm{t(g)}}$$

$$\bar{x}_{\mathrm{c(S)}} = \sum_g w_g \bar{x}_{\mathrm{c(g)}}$$

The variances of the weighted stratum means $\bar{x}_{\mathrm{t(S)}}$ and $\bar{x}_{\mathrm{c(S)}}$ are then given by

$$V(x_{\mathrm{t(S)}}) = \sum_g w_g^2 V(x_{\mathrm{t(g)}})$$

$$V(x_{\mathrm{c(S)}}) = \sum_g w_g^2 V(x_{\mathrm{c(g)}})$$

These stratum weights are displayed in the "Standardized Mean Differences within Strata" table (see Output 100.2.7). You can use these weights for estimation of the treatment effect in an outcome analysis.

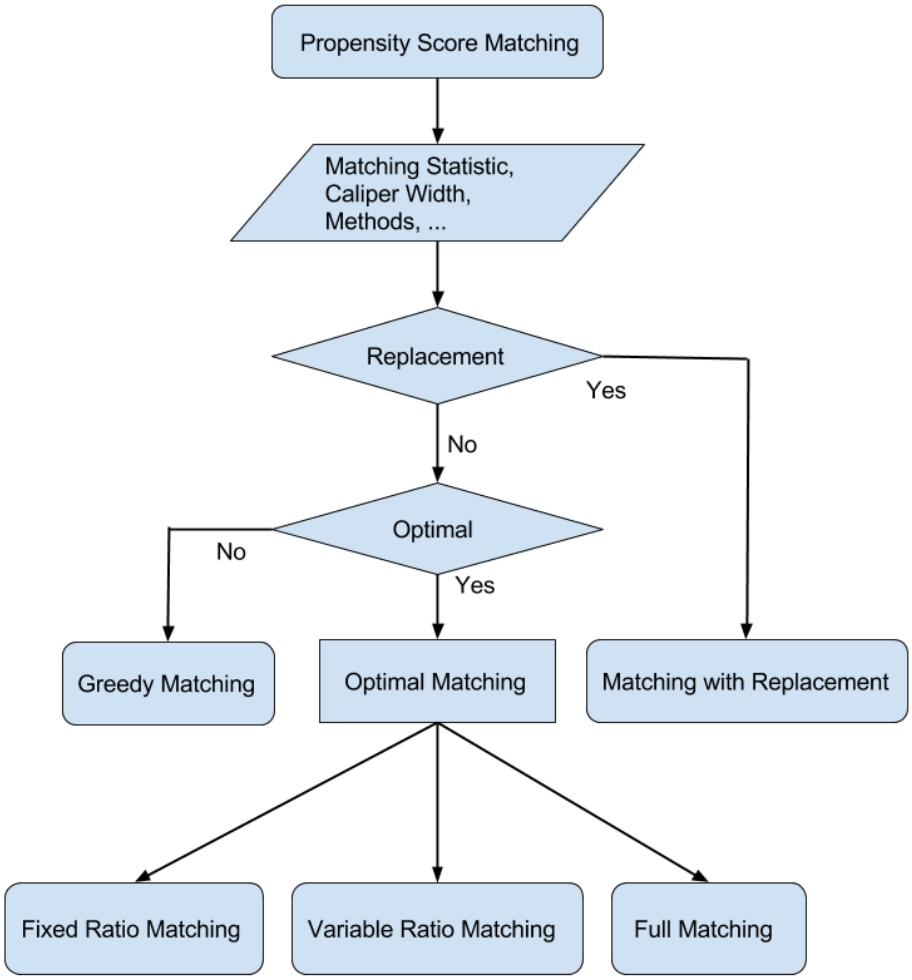
Matching Process

Except for matching with replacement in which multiple control units are matched to each treated unit, propensity score matching creates mutually exclusive sets of observations that have similar propensity scores. Each set has at least one treated unit and at least one control unit. The distribution of observed variables will be similar between treated units and control units in the matched sample.

For propensity score matching, Stuart (2010) reviews matching methods and provides guidance on their use. Austin (2014) provides a detailed comparison of algorithms for matching.

The flowchart in Figure 100.14 summarizes the steps in propensity score matching.

Figure 100.14 Steps in Propensity Score Matching



The PSMATCH procedure provides the following strategies for matching observations in the control group to observations in the treatment group:

- greedy nearest neighbor matching, which sequentially and without replacement selects the control unit

whose propensity score is closest to that of the particular treated unit

- optimal matching, which selects all matches simultaneously and without replacement to minimize the total absolute difference in propensity score across all matches (this approach includes fixed ratio matching, variable ratio matching, and full matching)
- matching with replacement, which selects with replacement the control unit whose propensity score is closest to that of each treated unit

In addition to the propensity score, you can also use the logit of the propensity score and Mahalanobis distance as the matching metric that is used to compare the closeness of two units. For more information, see the section "Matching Methods" on page 8224.

You can use the CALIPER= option in the MATCH statement to request that the difference in the propensity scores for a matched pair be less than or equal to a specified caliper width.

You can request exact matches of the levels of classification variables for treated and control units by specifying the EXACT= option in the MATCH statement.

Matching Metrics

The PSMATCH procedure provides the following metrics or distances for the purpose of matching observations in the treated group with observations in the control group:

- the difference in the logit of the propensity score (DISTANCE=LPS; this is the default)
- the difference in the propensity score (DISTANCE=PS)
- the Mahalanobis distance between sets of continuous variables (DISTANCE=MAH)

You specify the type of distance in the DISTANCE= option in the MATCH statement. Let p_{ti} and p_{cj} be the propensity scores of the *i*th treated unit and the *j*th control unit, respectively. When you specify DISTANCE=PS, matching is based on the absolute difference of propensity scores:

$$\left| p_{ti} - p_{cj} \right|$$

When you specify DISTANCE=LPS, matching is based on the absolute difference of logits of propensity scores:

$$\left| \text{logit}(p_{ti}) - \text{logit}(p_{cj}) \right|$$

When you specify DISTANCE=MAH, two different versions of the Mahalanobis distance (d) can be computed, as specified in the SQRT= suboption of the DISTANCE=MAH option,

$$d = \begin{cases} \sqrt{(\mathbf{X}_{ti} - \mathbf{X}_{cj})' \, \mathbf{V}^{-1} \, (\mathbf{X}_{ti} - \mathbf{X}_{cj})} & \text{if SQRT=YES (this is the default)} \\ (\mathbf{X}_{ti} - \mathbf{X}_{cj})' \, \mathbf{V}^{-1} \, (\mathbf{X}_{ti} - \mathbf{X}_{cj}) & \text{if SQRT=NO} \end{cases}$$

where d is the Mahalanobis distance; $\mathbf{X}$ is the set of variables that include the logit of the propensity score if LPS is specified, the propensity score if PS is specified, and all continuous variables that are specified in the VAR= options; $\mathbf{X}_{ti}$ contains variable values of the ith treated unit; $\mathbf{X}_{cj}$ contains variable values of the jth control unit; and $\mathbf{V}$ is the covariance matrix of $\mathbf{X}$.

The SQRT= option does not affect the results for greedy nearest neighbor matching and matching with replacement; it affects only the results for optimal matching.

Three different covariance matrices can be used to compute the Mahalanobis distance (as specified in the COV= suboption of the DISTANCE=MAH option):

- the covariance matrix that is based on observations in the control group (COV=CONTROL; this is the default)
- the pooled covariance matrix that is based on observations in the treated and control groups (COV=POOLED)
- the identity matrix (COV=IDENTITY), which yields the Euclidean distance.

You can include the propensity score and the logit of the propensity among the variables that are used to compute the Mahalanobis distance. For example, when you specify DISTANCE=MAH(PS VAR=(X1 X2 X3) / COV=POOLED), the PSMATCH procedure computes the Mahalanobis distance between observations in the treated and control groups by using the propensity score and variables X1, X2, and X3. The covariance matrix is the pooled covariance matrix of the treated and control groups.

Matching Methods

When you specify the MATCH statement, the PSMATCH procedure matches observations in the control group to observations in the treatment group by using one of the methods that are described in the following subsections. You can request the method in the METHOD= option.

Greedy Nearest Neighbor Matching

Greedy nearest neighbor matching, requested by the METHOD=GREEDY option, selects the control unit whose propensity score best matches the propensity score of each treated unit. Greedy nearest neighbor matching is done sequentially and without replacement.

The following criteria are available for greedy nearest neighbor matching:

- the number of control units to be matched to each treated unit (you can specify this number in the K= suboption)
- the order of propensity scores of treated units, which can be ascending, descending, or random (you can specify the order in the ORDER= suboption)

Replacement Matching

Replacement matching, requested by the METHOD=REPLACE option, selects with replacement the control unit whose propensity score is closest to the propensity score for each treated unit. You can specify the number of control units to be matched to each treated unit in the K= suboption.

Optimal Matching

Optimal matching selects all matches simultaneously and without replacement to minimize the total absolute difference in propensity score across all matches. You can request the following optimal matching methods:

- fixed ratio matching, requested by the METHOD=OPTIMAL option, which matches a fixed number of control units to each treated unit
- variable ratio matching, requested by the METHOD=VARRATIO option, which matches one or more control units to each treated unit
- full matching, requested by the METHOD=FULL option, which matches each treated unit to one or more control units or matches each control unit to one or more treated units. By additionally specifying the KMEAN=, NCONTROL=, or PCTCONTROL= suboptions, you can request constrained full matching in which the number of matched control units is less than the total number of available controls.

As alternatives to matching on the propensity score, you can match on the logit of the propensity score or use the Mahalanobis distance to match on a set of variables (possibly including the PS or the LPS). All three of these methods minimize the total absolute difference across all matches in the matching metric, which is the total difference in the logit of the propensity score by default.

Table 100.9 lists the suboptions available for optimal matching. The symbol X indicates that the option applies for the specified method.

Table 100.9 Applicable Options for Optimal Matching

METHOD=	K=	KMIN=	KMAX=	KMAXTRT=	KMEAN= NCONTROL= PCTCONTROL=
OPTIMAL	X				
VARRATIO		X	X		X
FULL			X	X	X

- K= specifies the number of control units to be matched to each treated unit.
- KMIN= specifies the minimum number of control units to be matched to each treated unit.
- KMAX= specifies the maximum number of control units to be matched to each treated unit.
- KMAXTRT= specifies the maximum number of treated units to be matched to each matched control unit.

- KMEAN= specifies the average number of control units to be matched to each treated unit.
- NCONTROL= specifies the total number of control units to be matched.
- PCTCONTROL= specifies the percentage of control units to be matched.

You can specify only one of the KMEAN=, NCONTROL=, and PCTCONTROL= options.

Weighting after Matching

If the matched observations show good variable balance after matching, you can perform an outcome analysis to estimate the treatment effect by comparing outcomes between treated and control subjects in the matched sample. Except for the case of one-to-one matching without replacement, the matched observation weights should be used in the balance assessment and in the outcome analysis.

The PSMATCH procedure provides the following methods for weighting after matching:

- ATT weighting, which is used to estimate the ATT
- ATE weighting after full matching, which is used to estimate the ATE

ATT Weighting after Matching without Replacement

ATT weights for use after matching without replacement are computed as

$$w_{gj} = \begin{cases} 1 & \text{for treated units in the } g\text{th matched set} \\ \frac{N_{gt}}{N_{gc}} & \text{for control units in the } g\text{th matched set} \end{cases}$$

where N_{gt} is the number of treated units and N_{gc} is the number of control units in the gth matched set.

That is, in each matched set, each treated unit has a weight of 1 and each control unit has a weight that equals the number of treated units in the matched set divided by the number of control units in the matched set. Thus, with a one-to-one greedy or optimal matching, the weight is 1 for both the treated and control units. Under a different matching algorithm, if the gth matched set contains N_{gt}=1 treated unit and N_{gc}=3 control units, then the weight for each treated unit is 1 and the weight for each control unit is 1/3.

The total weight for the controls is equal to the total number of treated units in each matched group, and the total weight for the matched controls is equal to the total number of matched treated units.

You can request ATT weighting after matching by specifying the MATCHATTWGT option in the MATCH statement, and then by specifying the WEIGHT= option in the OUTPUT statement to create a variable that contains these weights.

ATT Weighting after Matching with Replacement

The PSMATCH procedure creates mutually exclusive sets of units after matching with replacement. In the matched set, each treated unit is connected to all control units either directly or indirectly. For example, assume that the treated group contains units T_1, T_2, ..., and the treated group contains units C_1, C_2, ... If T_1 is matched to C_1 and C_2 and T_2 is matched to C_2 and C_3, then T_1 is connected directly to C_1 and C_2 and is connected indirectly to C_3, Similarly, T_2 is connected directly to C_2 and C_3, and indirectly to C_1.

In each matched set, each treated unit has a weight of 1 and each control unit has a weight that is computed from contributions of its matched treated units. That is, if a treated unit has two matched control units, then each control unit has a weight of 1/2 from this treated unit.

For example, assume that T_1 is matched to C_1 and C_2 and T_2 is matched to C_2 and C_3, and the five units do not have other matches. Then C_1 has a weight of 1/2 (from T_1), C_2 has a weight of 1 (1/2 from T_1 and 1/2 from T_2), and C_3 has a weight of 1/2 (from T_2).

The total weight for the controls is equal to the total number of treated units in each matched group, and the total weight for the matched controls is equal to the total number of matched treated units.

You can request ATT weighting after matching by specifying the MATCHATTWGT option in the MATCH statement, and then by specifying the WEIGHT= option in the OUTPUT statement to create a variable that contains these weights.

ATE Weighting after Full Matching

When optimal full matching is done, ATE weights for use after matching are computed as

$$w_{gj} = \begin{cases} \frac{N_g}{N_{gt}} & \text{for treated units in the } g\text{th matched set} \\ \frac{N_g}{N_{gc}} & \text{for control units in the } g\text{th matched set} \end{cases}$$

where N_{gt} is the number of treated units, N_{gc} is the number of control units, and $N_g = N_{gt} + N_{gc}$ is the total number of units in the gth matched set.

That is, in each matched set, each treated unit has a weight that equals the total number of treated and control units divided by the number of treated units, and each control unit has a weight that equals the total number of treated and control units divided by the number of control units in the matched set. Thus, if a matched set contains N_{gt}=1 treated unit and N_{gc}=3 control units, then the treated unit has a weight of 4 and each control unit has a weight of 4/3; if a matched set contains N_{gt}=2 treated units and N_{gc}=1 control unit, then each treated unit has a weight of 3/2 and the control unit has a weight of 3.

The total weight for the treated units and the total weight for the control units are each equal to the combined number of treated and control units in each matched group. Thus, the total weight for matched treated units and the total weight for matched control units are each equal to the total number of matched units (treated and control units combined).

ATE weighting is available only for full matching (METHOD=FULL) and is appropriate only for unrestricted full matching (that is, when all available control units are matched).

You can request ATT weighting after matching by specifying the MATCHATEWGT option in the MATCH statement, and then by specifying the WEIGHT= option in the OUTPUT statement to create a variable that contains these weights.

Variable Balance Assessment

Propensity score analysis assumes that the true propensity scores are known. When the propensity scores are estimated—as is usually the case in practice—you need to assess how well the distributions of the propensity scores (or their logits) and the adjusted variables are balanced between the treatment group and the control group. PROC PSMATCH provides numerical and graphical tools that you can use to assess the balance of the propensity scores (or their logits) and the adjusted variables that are either continuous or binary.

The ASSESS statement in the PSMATCH procedure provides a variety of statistical measures and graphical displays for comparing these distributions. You can make these assessments for all the observations in the data set, the observations in the support region, or the matched observations (if you specify a MATCH statement).

Two statistical measures for balance assessment are the standardized mean difference between the treatment and control groups and the treated-to-control variance ratio. For good variable balance, the absolute standardized mean difference should be less than or equal to 0.25, and the variance ratio should be between 0.5 and 2 (Rubin 2001, p. 174; Stuart 2010, p. 11). Some authors have applied a smaller threshold of 0.1 to the absolute standardized mean difference (Normand et al. 2001; Mamdani et al. 2005; Austin 2009).

The standardized mean difference is computed by dividing the difference in the means of the variable in the two groups by an estimate of the standard deviation. Two estimates of the standard deviation are available in the PSMATCH procedure:

- the square root of the average of the variances in the treatment and control groups (Rosenbaum and Rubin 1985, p. 37),
- the standard deviation of observations in the treatment group only (Stuart 2010, p. 11)

For binary classification variables, the mean is taken to be the proportion p of units having the first classification level, and the variance is computed as $p(1-p)$ (Austin, Grootendorst, and Anderson 2007, p. 737).

If you specify a STRATA statement, then stratum-specific standardized mean differences are computed for observations in the support region.

The PSMATCH procedure displays the standardized mean differences in plots. You can also request box plots and cloud plots for continuous variables, and bar charts for binary classification variables. These plots are also produced for each stratum if you specify a STRATA statement.

The next three subsections describe how standardized mean differences and treated-to-control variance ratios are computed for all observations, observations in the support region, and matched observations.

Standardized Mean Differences for All Observations

For all observations in the data set, let $\bar{x}_{\mathrm{t(A)}}$ be the mean of the observations in the treatment group and let $\bar{x}_{\mathrm{c(A)}}$ be the mean of the observations in the control group, with corresponding sample variances $V(x_{\mathrm{t(A)}})$ and $V(x_{\mathrm{c(A)}})$. Then the standardized mean difference is

$$d_{\mathrm{(A)}} = \frac{\bar{x}_{\mathrm{t(A)}} - \bar{x}_{\mathrm{c(A)}}}{s_{\mathrm{(A)}}}$$

where the standard deviation is given by

$$s_{(\mathrm{A})} = \begin{cases} \sqrt{\frac{V(x_{\mathrm{t}(\mathrm{A})})+V(x_{\mathrm{c}(\mathrm{A})})}{2}} & \text{if STDDEV=POOLED} \\ \sqrt{V(x_{\mathrm{t}(\mathrm{A})})} & \text{if STDDEV=TREATED} \end{cases}$$

The treated-to-control variance ratio is

$$\frac{V(x_{\mathrm{t}(\mathrm{A})})}{V(x_{\mathrm{c}(\mathrm{A})})}$$

Standardized Mean Differences for Observations in the Support Region

For observations in the support region, let $\bar{x}_{\mathrm{t}(\mathrm{R})}$ be the mean of observations in the treatment group and $\bar{x}_{\mathrm{c}(\mathrm{R})}$ be the mean of observations in the control group, with corresponding sample variances $V(x_{\mathrm{t}(\mathrm{R})})$ and $V(x_{\mathrm{c}(\mathrm{R})})$. Then the standardized mean difference is

$$d_{(\mathrm{R})} = \begin{cases} \frac{\bar{x}_{\mathrm{t}(\mathrm{R})}-\bar{x}_{\mathrm{c}(\mathrm{R})}}{s_{(\mathrm{A})}} & \text{if STDDEV=POOLED(ALLOBS=YES)} \\ & \text{or STDDEV=TREATED(ALLOBS=YES)} \\ \frac{\bar{x}_{\mathrm{t}(\mathrm{R})}-\bar{x}_{\mathrm{c}(\mathrm{R})}}{s_{(\mathrm{R})}} & \text{if STDDEV=POOLED(ALLOBS=NO)} \\ & \text{or STDDEV=TREATED(ALLOBS=NO)} \end{cases}$$

where the standard deviation $s_{(\mathrm{R})}$ is given by

$$s_{(\mathrm{R})} = \begin{cases} \sqrt{\frac{V(x_{\mathrm{t}(\mathrm{R})})+V(x_{\mathrm{c}(\mathrm{R})})}{2}} & \text{if STDDEV=POOLED} \\ \sqrt{V(x_{\mathrm{t}(\mathrm{R})})} & \text{if STDDEV=TREATED} \end{cases}$$

That is, with ALLOBS=YES, the standard deviation that is derived from all observations in the data set is used to compute the standardized mean difference. With ALLOBS=NO, the standard deviation that is derived from observations in the support region is used to compute the standardized mean difference.

The treated-to-control variance ratio is

$$\frac{V(x_{\mathrm{t}(\mathrm{R})})}{V(x_{\mathrm{c}(\mathrm{R})})}$$

The percentage reduction in the standardized mean difference is computed as

$$100 \times \frac{\max(|d_{(\mathrm{A})}| - |d_{(\mathrm{R})}|, 0)}{|d_{(\mathrm{A})}|}$$

Pooled Standardized Mean Differences across the Strata

Let $\bar{x}_{\mathrm{t(S)}}$ be the weighted stratum mean of treated observations, and let $\bar{x}_{\mathrm{c(S)}}$ be the weighted stratum mean of control observations, with corresponding variances $V(x_{\mathrm{t(S)}})$ and $V(x_{\mathrm{c(S)}})$. For information about these statistics, see the section "Weighting after Stratification" on page 8221.

The standardized mean difference is

$$d_{\mathrm{(S)}} = \begin{cases} \frac{\bar{x}_{\mathrm{t(S)}} - \bar{x}_{\mathrm{c(S)}}}{s_{\mathrm{(A)}}} & \text{if STDDEV=POOLED(ALLOBS=YES)} \\ & \text{or STDDEV=TREATED(ALLOBS=YES)} \\ \frac{\bar{x}_{\mathrm{t(S)}} - \bar{x}_{\mathrm{c(S)}}}{s_{\mathrm{(S)}}} & \text{if STDDEV=POOLED(ALLOBS=NO)} \\ & \text{or STDDEV=TREATED(ALLOBS=NO)} \end{cases}$$

where the standard deviation $s_{\mathrm{(S)}}$ is given by

$$s_{\mathrm{(S)}} = \begin{cases} \sqrt{\frac{V(x_{\mathrm{t(S)}}) + V(x_{\mathrm{c(S)}})}{2}} & \text{if STDDEV=POOLED} \\ \sqrt{V(x_{\mathrm{t(S)}})} & \text{if STDDEV=TREATED} \end{cases}$$

The treated-to-control variance ratio is

$$\frac{V(x_{\mathrm{t(S)}})}{V(x_{\mathrm{c(S)}})}$$

The percentage reduction for the standardized mean difference is computed as

$$100 \times \frac{\max(|d_{\mathrm{(A)}}| - |d_{\mathrm{(S)}}|, 0)}{|d_{\mathrm{(A)}}|}$$

Standardized Mean Differences for Matched Observations

Let $\bar{x}_{\mathrm{t(M)}}$ be the mean of matched observations in the treatment group, and let $\bar{x}_{\mathrm{c(M)}}$ be the mean of matched observations in the control group, with corresponding sample variances $V(x_{\mathrm{t(M)}})$ and $V(x_{\mathrm{c(M)}})$. Then the standardized mean difference is

$$d_{\mathrm{(M)}} = \begin{cases} \frac{\bar{x}_{\mathrm{t(M)}} - \bar{x}_{\mathrm{c(M)}}}{s_{\mathrm{(A)}}} & \text{if STDDEV=POOLED(ALLOBS=YES)} \\ & \text{or STDDEV=TREATED(ALLOBS=YES)} \\ \frac{\bar{x}_{\mathrm{t(M)}} - \bar{x}_{\mathrm{c(M)}}}{s_{\mathrm{(M)}}} & \text{if STDDEV=POOLED(ALLOBS=NO)} \\ & \text{or STDDEV=TREATED(ALLOBS=NO)} \end{cases}$$

where the standard deviation $s_{\mathrm{(M)}}$ is given by

$$s_{\mathrm{(M)}} = \begin{cases} \sqrt{\frac{V(x_{\mathrm{t(M)}}) + V(x_{\mathrm{c(M)}})}{2}} & \text{if STDDEV=POOLED} \\ \sqrt{V(x_{\mathrm{t(M)}})} & \text{if STDDEV=TREATED} \end{cases}$$

The treated-to-control variance ratio is

$$\frac{V(x_{\mathrm{t(M)}})}{V(x_{\mathrm{c(M)}})}$$

The percentage reduction for the standardized mean difference is computed as

$$100 \times \frac{\max(|d_{(\mathrm{A})}| - |d_{(\mathrm{M})}|, 0)}{|d_{(\mathrm{A})}|}$$

Sensitivity Analysis

Propensity score analysis assumes that all the confounders (variables that affect both the outcome and the treatment assignment) have been measured. If some confounders are unobserved, individuals that have the same observed covariates might not have the same probability of being assigned to the treated group. The assumption of no unmeasured confounders cannot be verified, so you should analyze the sensitivity of inferences to departures from this assumption. Sensitivity analysis considers how strong the unobserved covariates would have to be in order to negate the conclusion of the study (assuming that the initial analysis found a significant effect of the treatment).

Liu, Kuramoto, and Stuart (2013) describe seven commonly used techniques for sensitivity analysis. Based on the study objectives, these methods are classified into two groups. One group finds the tipping point that negates the statistical significance of the outcome-treatment association (Liu, Kuramoto, and Stuart 2013). The other group (not discussed here) derives the point estimate of the true outcome-treatment association with a 95% confidence interval (Liu, Kuramoto, and Stuart 2013).

Rosenbaum (2010, p. 77) provides a sensitivity analysis based on the odds ratio,

$$\frac{\pi_k/(1-\pi_k)}{\pi_l/(1-\pi_l)}$$

where π_k and π_l are the probabilities that the kth and lth individuals are assigned to the treated group, given that they have the same observed covariates, $x_k = x_l$.

For all kth and lth individuals with $x_k = x_l$, assume that the odds ratio is bounded by

$$\frac{1}{\Gamma} \leq \frac{\pi_k/(1-\pi_k)}{\pi_l/(1-\pi_l)} \leq \Gamma$$

where $\Gamma \geq 1$.

The parameter Γ measures the degree of hidden bias from unobserved confounders. For example, with $\Gamma = 2$,

$$\frac{\pi_k/(1-\pi_k)}{\pi_l/(1-\pi_l)} = 2$$

which indicates that even though they have the same values of the observed covariates, the kth individual is twice as likely as the lth individual to be in the treated group because of hidden bias.

Propensity score analysis assumes that if the kth and lth individuals have the same observed covariates, then $\pi_k = \pi_l$ and $\Gamma = 1$. When an outcome analysis leads to a significant result, Rosenbaum's sensitivity

analysis finds a tipping point, $\Gamma = \gamma$, that negates the conclusion of the study. A large value of γ is evidence that only a large departure from random treatment assignment can negate the conclusion of the study. If $\Gamma = \gamma$ is plausible, the study conclusion is not robust to hidden bias from an unobserved confounder.

For the case of one-to-one matched observations, Rosenbaum (2010, pp. 78–84) provides a sensitivity analysis that is based on paired observations. The following subsection describes this approach.

Sensitivity Analysis on Matched Observations

In a set of one-to-one matched observations, if individuals k and l are in the same matched set, then the probability that individual k is in the treated group and individual l is in the control group is

$$\frac{\pi_k}{\pi_k + \pi_l}$$

and the following equation can be used for sensitivity analysis:

$$\frac{1}{1+\Gamma} \leq \frac{\pi_k}{\pi_k + \pi_l} \leq \frac{\Gamma}{1+\Gamma}$$

If $\Gamma = 1$, then $\pi_k = \pi_l$.

For example, let y_{jt} and y_{jc} be the responses for the treated and control units in the jth matched set. The response is the improvement after treatment, and a positive value indicates a beneficial effect. Let

$$d_j = y_{jt} - y_{jc}$$

be the difference in responses between the treated and control units.

Suppose that a signed rank test is used in the outcome analysis. The signed rank statistic is

$$S = \sum_{j:d_j>0} d_j^+$$

where d_j^+ is the rank of $|d_j|$.

Assume that Γ=1. Then under the hypothesis of no treatment effect, S has mean

$$\mu_0 = \frac{n_t(n_t+1)}{4}$$

where n_t is the number of matched sets and the variance V (assuming that all d_j is distinct) is

$$V_0 = \frac{n_t(n_t+1)(2n_t+1)}{24}$$

When $n_t > 20$, the significance of

$$\frac{S-\mu_0}{\sqrt{V_0}}$$

can be computed from the Student's t distribution with $n_t - 1$ degrees of freedom.

For $\Gamma = \gamma$, S has mean

$$\mu_1 = \frac{\gamma}{1+\gamma}\frac{n_t(n_t+1)}{2} = \frac{2\gamma}{1+\gamma}\mu_0$$

and variance

$$V_1 = \frac{\gamma}{(1+\gamma)^2}\frac{n_t(n_t+1)(2n_t+1)}{6} = \frac{4\gamma}{(1+\gamma)^2}V_0$$

If a signed rank test shows a significantly better benefit in the treated group with $\Gamma = 1$, the sensitivity analysis searches for a tipping point that negates the study conclusion. A study conclusion is robust to hidden bias from the unobserved confounder if an extreme value of Γ is needed to alter the study conclusion.

Example 100.9 illustrates a sensitivity analysis on a set of one-to-one matched observations.

Table Output

By default, the PSMATCH procedure displays the "Data Information" and "Propensity Score Information" tables. If you specify a MATCH statement, the procedure also displays the "Matching Information" table. If you specify a STRATA statement, the procedure also displays the "Strata Information" table.

If you specify the NLARGESTWGT= option in the PSWEIGHT statement, the "Observations with Largest Weights" table is displayed. Similarly, if you specify the NMATCHMOST= option in the MATCH statement, the "Observations with Most Matches" table is also displayed.

If you specify the ASSESS statement, the "Standardized Mean Differences" table is displayed. In addition, if you specify a STRATA statement, the "Standardized Mean Differences within Strata" table is also displayed.

If you specify the VARINFO option in the ASSESS statement, the "Variable Information" table is displayed. In addition, if you specify a STRATA statement, the "Strata Variable Information" table is also displayed.

Data Information

The "Data Information" table displays the names of the input and output data sets, the numbers of observations in the treated group and the control group, and the numbers of observations in the support region that are in the treated group and the control group. The minimum and maximum propensity scores for observations in the support region are also displayed.

Matching Information

The "Matching Information" table displays the matching metric, the matching method, and the caliper width. The table also displays the number of matched sets of observations, the numbers of matched observations in the treated and control groups, and the total absolute difference across all matches.

Observations with Largest Weights

The "Observations with Largest Weights" table displays observations that have the largest weights in the treated and control groups. The table is produced only if WEIGHT= ATTWGT or WEIGHT=ATEWGT is specified in the PSWEIGHT statement. The table displays the observation numbers and their weights. If you also specify an ID statement, the table displays values of the ID variables.

Observations with Most Matches

The "Observations with Most Matches" table displays observations that have the greatest numbers of matches in the treated and control groups. The table is produced only if a MATCH statement is specified and METHOD=GREEDY and METHOD=OPTIMAL are not specified. The table displays the observation numbers and their numbers of matches. If you also specify an ID statement, the table displays the values of the ID variables.

Propensity Score Information

The "Propensity Score Information" table displays descriptive statistics (the number of observations, mean, standard deviation, minimum, and maximum) for the propensity scores of observations in the treated group and the control group. These statistics are computed using all observations, observations in the support region, and matched observations (if you specify a MATCH statement).

Standardized Mean Differences

The "Standardized Mean Differences" table displays statistics that summarize the differences in the variables and the logit propensity score (LPS) between the treated and control groups. These statistics are computed using all observations, observations in the support region, and matched observations (if you specify a MATCH statement).

The statistics include the following:

- the mean difference between observations in the treated and control groups
- the standard deviation that is used to compute the standardized mean difference
- the standardized mean difference, which is the mean difference divided by the standard deviation
- the percentage reduction of the standardized mean difference for observations in the support region, compared with the standardized mean difference of all observations (this statistic is also computed for matched observations if you specify a MATCH statement)
- the treated-to-control ratio of variances between observations in the treated and control groups

For more information about these statistics, see the sections "Standardized Mean Differences for All Observations" on page 8228, "Standardized Mean Differences for Observations in the Support Region" on page 8229, "Pooled Standardized Mean Differences across the Strata" on page 8230, and "Standardized Mean Differences for Matched Observations" on page 8230.

Strata Information

The "Strata Information" table displays descriptive statistics that include the propensity score range, the numbers of observations in the treated group and the control group, and the total number of observations in each stratum.

Strata Standardized Mean Differences

The "Standardized Mean Differences within Strata" table displays the variable difference statistics between the treated and control groups in each stratum.

For each variable, the statistics include the following:

- the mean difference between observations in the treated and control groups
- the standardized mean difference, which is the mean difference divided by the standard deviation that is displayed in the "Standardized Mean Differences" table
- the percentage reduction of the absolute standardized mean difference for observations in the stratum, compared with the absolute standardized mean difference for all observations
- the treated-to-control ratio of variances between observations in the treated and control groups in each stratum

Strata Variable Information

The "Strata Variable Information" table displays descriptive statistics that include the number of observations, variable mean, and standard deviation of the observations in each of the treatment and control groups within each stratum. For continuous variables, the statistics also include the minimum and maximum.

Variable Information

For variables that are specified in the ASSESS statement, the "Variable Information" table displays descriptive statistics that are computed using all observations and observations in each of the treatment and control groups in the support region.

These statistics include the sample size, mean, and standard deviation. For continuous variables, the statistics also include the minimum and maximum. If you specify a MATCH statement, the table also displays descriptive statistics for the matched observations in the treatment and control groups.

ODS Table Names

PROC PSMATCH assigns a name to each table it creates. You must use these names to refer to tables when you use the Output Delivery System (ODS). These names are listed in Table 100.10. For more information about ODS, see Chapter 22, "Using the Output Delivery System."

Table 100.10 ODS Tables Produced by PROC PSMATCH

ODS Table Name	Description	Statements	Option
DataInfo	Data information		
LargestWgtObs	Observations with the largest weights	PSWEIGHT	NLARGESTWGT=
MatchInfo	Matching information	MATCH	
MatchMostObs	Observations with the most matches	MATCH	NMATCHMOST=
PSInfo	Propensity score information		

Table 100.10 *continued*

ODS Table Name	Description	Statements	Option
StdDiff	Standardized mean differences between the treated group and the control group	ASSESS	
StrataInfo	Strata information	STRATA	
StrataStdDiff	Standardized mean differences within strata	ASSESS, STRATA	
StrataVarInfo	Strata variable information	ASSESS, STRATA	VARINFO
VarInfo	Variable information	ASSESS	VARINFO

Graphics Output

This section describes the use of ODS for creating graphics with the PSMATCH procedure. To request these graphs, ODS Graphics must be enabled and you must specify the ASSESS option. In addition, except for the standardized mean differences plot (which is the default), you must use the PLOTS= option in the ASSESS statement to specify the plots. For more information about ODS Graphics, see Chapter 23, "Statistical Graphics Using ODS."

Bar Chart

The PLOTS=BARCHART option displays bar charts for binary classification variables in the treated and control groups for all observations and for observations in the support region. If you specify the MATCH statement, bar charts are also created for matched observations.

Box Plot

The PLOTS=BOXPLOT option displays box plots for continuous variables in the treated and control groups for all observations and for observations in the support region. If you specify the MATCH statement, box plots are also created for matched observations.

CDF Plot

The PLOTS=CDFPLOT option displays cumulative distribution function (CDF) plots for continuous variables in the treated and control groups for all observations and for observations in the support region. If you specify the MATCH statement, CDF plots are also created for matched observations.

Cloud Plot

The PLOTS=CLOUDPLOT option displays cloud plots for continuous variables in the treated and control groups for all observations and for observations in the support region. If you specify the MATCH statement, cloud plots are also created for matched observations. Here the term cloud plot refers to a scatter plot in which the points are jittered by adding random noise to data in the plot in order to prevent overplotting. For example, with a continuous variable and the default ORIENT=HORIZONTAL option, the variable values are

displayed horizontally and the treated and control groups are displayed vertically. The exact variable values are displayed along the horizontal axis, but the points are jittered in the vertical direction.

Standardized Mean Differences Plot

The PLOTS=STDDIFFPLOT option displays a plot of the standardized mean differences for continuous and binary classification variables for all observations and for observations in the support region. If you specify the MATCH statement, plots are also created for matched observations.

Strata Bar Chart

If you specify a STRATA statement, the PLOTS=BARCHART option displays bar charts for binary classification variables in the treated and control groups for the observations in each stratum.

Strata Box Plot

If you specify a STRATA statement, the PLOTS=BOXPLOT option displays box plots for continuous variables in the treated and control groups for the observations in each stratum.

Strata CDF Plot

If you specify a STRATA statement, the PLOTS=CDFPLOT option displays cumulative distribution function (CDF) plots for continuous variables in the treated and control groups for the observations in each stratum.

Strata Cloud Plot

If you specify a STRATA statement, the PLOTS=CLOUDPLOT option displays cloud plots for continuous variables in the treated and control groups for the observations in each stratum. The cloud plot is a scatter plot that is jittered by adding random noise in order to prevent overplotting.

Strata Standardized Mean Differences Plot

If you specify a STRATA statement, the PLOTS=STDDIFFPLOT option displays standardized mean differences plots for continuous and binary classification variables for the observations in each stratum.

Weight Cloud Plot

The PLOTS=WGTCLOUDPLOT option displays cloud plots for weights in the treated and control groups for all observations and for observations in the support region. The cloud plot is also called the jittered scatter plot: it adds random noise to data in the plot in order to prevent overplotting. The option is applicable if you specify a PSWEIGHT statement.

ODS Graphics

Statistical procedures use ODS Graphics to create graphs as part of their output. ODS Graphics is described in detail in Chapter 23, "Statistical Graphics Using ODS."

Before you create graphs, ODS Graphics must be enabled (for example, by specifying the ODS GRAPHICS ON statement). For more information about enabling and disabling ODS Graphics, see the section "Enabling and Disabling ODS Graphics" on page 651 in Chapter 23, "Statistical Graphics Using ODS."

The overall appearance of graphs is controlled by ODS styles. Styles and other aspects of using ODS Graphics are discussed in the section "A Primer on ODS Statistical Graphics" on page 650 in Chapter 23, "Statistical Graphics Using ODS."

PROC PSMATCH assigns a name to each graph it creates. You can use these names to refer to the graphs when you use ODS. To request the graph, ODS Graphics must be enabled and you must specify the ASSESS option. In addition, except for the standardized mean differences plot (which is the default), you must use the PLOTS= option in the ASSESS statement to specify the plots, as indicated in Table 100.11.

Table 100.11 Graphs Produced by PROC PSMATCH

ODS Graph Name	Plot Description	Statements	PLOTS=
BarChart	Binary variable bar chart	ASSESS	BARCHART
BoxPlot	Continuous variable box plot	ASSESS	BOXPLOT
CDFPlot	Continuous variable CDF plot	ASSESS	CDFPLOT
CloudPlot	Continuous variable cloud plot	ASSESS	CLOUDPLOT
StdDiffPlot	Standardized mean differences plot	ASSESS	STDDIFFPLOT
StrataBarChart	Strata binary variable bar chart	ASSESS, STRATA	BARCHART
StrataBoxPlot	Strata continuous variable box plot	ASSESS, STRATA	BOXPLOT
StrataCDFPlot	Strata continuous variable CDF plot	ASSESS, STRATA	CDFPLOT
StrataCloudPlot	Strata continuous variable cloud plot	ASSESS, STRATA	CLOUDPLOT
StrataStdDiffPlot	Standardized mean differences plot within strata	ASSESS, STRATA	STDDIFFPLOT
WgtCloudPlot	Weight cloud plot	ASSESS, PSWEIGHT	WGTCLOUDPLOT

Examples: PSMATCH Procedure

The examples in this section illustrate some of the methods of propensity analysis that are available in the PSMATCH procedure:

- Propensity score weighting is illustrated in Example 100.1.
- Stratification is illustrated in Example 100.2.
- Optimal variable ratio matching is illustrated in Example 100.3.
- Optimal one-to-one matching is illustrated in Example 100.7, Example 100.8, and Example 100.9. Example 100.7 uses Mahalanobis distance as the matching metric, and Example 100.8 and Example 100.9 use the logit of the propensity score as the matching metric.
- Greedy nearest neighbor matching is illustrated in Example 100.4 and Example 100.5.
- Matching with replacement is illustrated in Example 100.6.

With the exception of Example 100.8, all the examples use propensity scores that are obtained from a binary logistic regression model that is fitted by using the PSMATCH procedure. Example 100.8 illustrates the use of precomputed propensity scores.

The PSMATCH procedure provides a variety of statistical and graphical methods that you can use to assess covariate balance. Because this assessment is an essential aspect of propensity score analysis, the examples emphasize the use of the ASSESS statement.

Although the PSMATCH procedure does not provide outcome analysis, Example 100.5 illustrates an outcome analysis that is carried out after a propensity score analysis. Likewise, Example 100.9 illustrates a sensitivity analysis that accompanies an outcome analysis.

With the exception of Example 100.4, Example 100.5, and Example 100.6, all the examples illustrate situations in which the outcome data are not available at the time that the propensity score analysis is done. In such situations, you might not need to retain the covariate data for all individuals in the study in order to carry out the outcome analysis. For example, if you use the matching method for propensity score analysis, only the matched units are needed for follow-up. Retaining only the matched units reduces the cost of the study (Stuart 2010, p. 2).

Example 100.4, Example 100.5, and Example 100.6 illustrate situations in which the outcome data happen to be available at the time of the propensity score analysis. In such situations, the outcome data should not be used in the analysis (Stuart 2010, p. 2).

For simplicity, the examples in this section involve only a few variables. In practice, propensity score analysis often involves many more variables.

Example 100.1: Propensity Score Weighting

This example illustrates how you can create observation weights that are appropriate for estimating the average treatment effect (ATE) in a subsequent outcome analysis (the outcome analysis itself is not shown here).

The data for this example are observations on patients in a nonrandomized clinical trial. The trial and the Drugs data set that contains the patient information are described in the section "Getting Started: PSMATCH Procedure" on page 8184.

The following statements specify a logistic regression model for obtaining propensity scores, compute observation weights from the propensity scores, request statistics and plots for balance assessment, and save the weights in an output data set:

```
ods graphics on;
proc psmatch data=drugs region=allobs;
   class Drug Gender;
   psmodel Drug(Treated='Drug_X')= Gender Age BMI;
   psweight weight=atewgt nlargestwgt=6;
   assess lps var=(Gender Age BMI)
          / varinfo plots=(barchart boxplot(display=(lps BMI)) wgtcloud);
   id BMI;
   output out(obs=all)=OutEx1 weight=_ATEWgt_;
run;
```

The PSMODEL statement specifies the logistic regression model that creates the propensity score for each observation, which is the probability that the patient receives Drug_X. The CLASS statement specifies the classification variables in the model. The Drug variable is the binary treatment indicator variable, and TREATED='Drug_X' identifies Drug_X as the treated group. The Gender, Age, and BMI variables are included in the model because they are believed to be related to the assignment. The REGION=ALLOBS option specifies that the support region contains all observations. Weights are computed for all observations, regardless of the REGION= option.

The "Data Information" table in Output 100.1.1 displays the numbers of observations in the treated and control groups, the lower and upper limits of the propensity scores for observations in the support region, and the numbers of observations in the treated and control groups that fall within the support region. Because REGION=ALLOBS is specified, the lower and upper limits for of the propensity scores for observations in the support region are the minimum and maximum of the propensity scores for all observations. Consequently, all 373 observations in the control group fall within the support region, and all 133 observations in the treated group fall within the support region.

Output 100.1.1 Data Information

The PSMATCH Procedure

Data Information	
Data Set	WORK.DRUGS
Output Data Set	WORK.OUTEX1
Treatment Variable	Drug
Treated Group	Drug_X
All Obs (Treated)	113
All Obs (Control)	373
Support Region	All Obs
Lower PS Support	0.020157
Upper PS Support	0.685757
Support Region Obs (Treated)	113
Support Region Obs (Control)	373

The "Propensity Score Information" table in Output 100.1.2 displays summary statistics by treatment group for all observations (labeled "All"), for observations in the support region (labeled "Region"), and for weighted observations in the support region (labeled "Weighted"). Because the support region consists of all observations, the first two rows in the table are identical. The WEIGHT=ATEWGT option in the PSWEIGHT statement displays summary statistics for ATE weighted observations.

Output 100.1.2 Propensity Score Information

Propensity Score Information

	Treated (Drug = Drug_X)						Control (Drug = Drug_A)					
Observations	**N**	**Weight**	**Mean**	**Standard Deviation**	**Minimum**	**Maximum**	**N**	**Weight**	**Mean**	**Standard Deviation**	**Minimum**	**Maximum**
All	113		0.3108	0.1325	0.0602	0.6411	373		0.2088	0.1320	0.0202	0.6858
Region	113		0.3108	0.1325	0.0602	0.6411	373		0.2088	0.1320	0.0202	0.6858
Weighted	113	460.45	0.2454	0.1268	0.0602	0.6411	373	489.59	0.2381	0.1496	0.0202	0.6858

Propensity Score Information	
Observations	**Treated - Control Mean Difference**
All	0.1020
Region	0.1020
Weighted	0.0073

The ASSESS statement produces tables and plots, shown in Output 100.1.3 through Output 100.1.5 and in Output 100.1.7 through Output 100.1.10, that summarize differences in the distributions of specified variables between treated and control groups. As requested by the LPS and VAR= options, these variables are the logit of the propensity score and the data variables Gender, Age, and BMI. Differences are summarized for all observations and for observations in the support region. Again, these two sets of differences are identical because REGION=ALLOBS is specified. The WEIGHT=ATEWGT option in the PSWEIGHT statement requests that differences in the variables also be summarized for the weighted observations. By comparing the differences for weighted observations to the differences for observations in the support region, you can

assess how well weighting improves the balance for each variable.

The VARINFO option requests the "Variable Information" table, shown in Output 100.1.3, which displays variable summary statistics and differences between the treated and control groups for all observations (labeled "All"), for observations in the support region (labeled "Region"), and for weighted observations (labeled "Weighted"). For the binary classification variable (Gender), the difference is in the proportion of the first ordered level (Female).

Output 100.1.3 Variable Information

The PSMATCH Procedure

Variable Information

		Treated (Drug = Drug_X)					
Variable	Observations	N	Weight	Mean	Standard Deviation	Minimum	Maximum
Logit Prop Score	All	113		-0.88062	0.681761	-2.74745	0.58035
	Region	113		-0.88062	0.681761	-2.74745	0.58035
	Weighted	113	460.45	-1.25406	0.741386	-2.74745	0.58035
Age	All	113		36.30973	5.534114	26.00000	49.00000
	Region	113		36.30973	5.534114	26.00000	49.00000
	Weighted	113	460.45	38.59813	5.773228	26.00000	49.00000
BMI	All	113		24.49257	1.863797	20.33000	28.34000
	Region	113		24.49257	1.863797	20.33000	28.34000
	Weighted	113	460.45	24.03522	1.896607	20.33000	28.34000
Gender	All	113		0.43363	0.495575		
	Region	113		0.43363	0.495575		
	Weighted	113	460.45	0.47335	0.499289		

Variable Information

		Control (Drug = Drug_A)						Treated - Control
Variable	Observations	N	Weight	Mean	Standard Deviation	Minimum	Maximum	Mean Difference
Logit Prop Score	All	373		-1.52059	0.844486	-3.88386	0.78036	0.63997
	Region	373		-1.52059	0.844486	-3.88386	0.78036	0.63997
	Weighted	373	489.59	-1.35103	0.894234	-3.88386	0.78036	0.09698
Age	All	373		40.40483	6.579103	25.00000	57.00000	-4.09509
	Region	373		40.40483	6.579103	25.00000	57.00000	-4.09509
	Weighted	373	489.59	39.32670	6.771606	25.00000	57.00000	-0.72857
BMI	All	373		23.75327	1.980778	19.22000	28.61000	0.73930
	Region	373		23.75327	1.980778	19.22000	28.61000	0.73930
	Weighted	373	489.59	23.95492	2.004019	19.22000	28.61000	0.08030
Gender	All	373		0.45845	0.498270			-0.02482
	Region	373		0.45845	0.498270			-0.02482
	Weighted	373	489.59	0.45479	0.497952			0.01856

The statistics in Output 100.1.3 are identical for all observations and for observations in the support region because REGION=ALLOBS is specified.

As indicated in the column labeled Weight, the total weight of the treated units is 460.45 and the total weight of the control units is 489.59, which are close to 486, the total number of units. The weights are ATE weights

because WEIGHT=ATEWGT is specified in the PSWEIGHT statement. For information about ATE weights, see the section "Inverse Probability of Treatment Weighting" on page 8217.

Note that in comparison to the unweighted means, the weighted means for the control units are closer in absolute value to the corresponding weighted means for the treated units.

The "Standardized Mean Differences" table, shown in Output 100.1.4, displays standardized mean differences in the variables between the treated and control groups, based on all observations, on observations in the support region, and on weighted observations.

Output 100.1.4 Standardized Mean Differences

Standardized Mean Differences (Treated - Control)

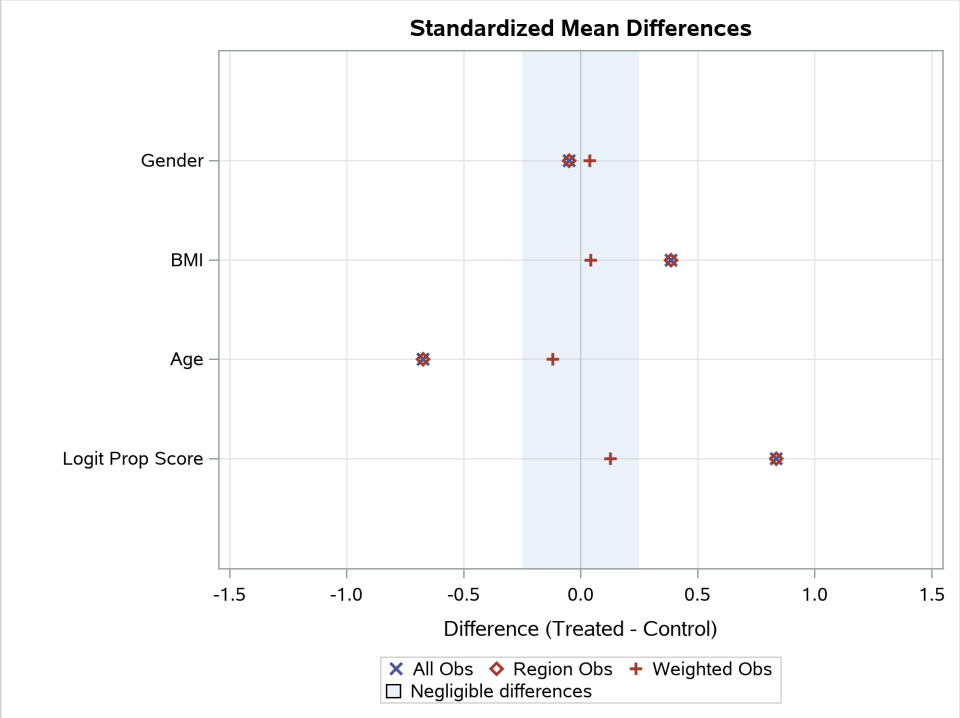
Variable	Observations	Mean Difference	Standard Deviation	Standardized Difference	Percent Reduction	Variance Ratio
Logit Prop Score	**All**	0.63997	0.767449	0.83389		0.6517
	Region	0.63997		0.83389	0.00	0.6517
	Weighted	0.09698		0.12636	84.85	0.6874
Age	**All**	-4.09509	6.079104	-0.67363		0.7076
	Region	-4.09509		-0.67363	0.00	0.7076
	Weighted	-0.72857		-0.11985	82.21	0.7269
BMI	**All**	0.73930	1.923178	0.38441		0.8854
	Region	0.73930		0.38441	0.00	0.8854
	Weighted	0.08030		0.04175	89.14	0.8957
Gender	**All**	-0.02482	0.496925	-0.04994		0.9892
	Region	-0.02482		-0.04994	0.00	0.9892
	Weighted	0.01856		0.03735	25.21	1.0054

Standard deviation of All observations used to compute standardized differences

The standardized mean differences based on weighted observations are significantly reduced; the largest of these differences is 0.12637 in absolute value, which is less than the upper limit of 0.25 that is recommended by Rubin (2001, p. 174) and Stuart (2010, p. 11). The treated-to-control variance ratios between the two groups are within the recommended range of 0.5 to 2. The percentage of reduction in variable mean difference is 0 for observations in the support region because REGION=ALLOBS is specified.

The PSMATCH procedure displays a standardized mean differences plot, shown in Output 100.1.5, for the variables that are specified in the ASSESS statement.

Output 100.1.5 Standardized Mean Differences Plot



The "Standardized Mean Differences Plot" displays the differences that are shown in the "Standardized Mean Differences" table in Output 100.1.4. All differences for the weighted observations are within the recommended limits of –0.25 and 0.25, which are indicated by the shaded area.

The NLARGESTWGT=6 option displays the "Observations with Largest Weights" table, shown in Output 100.1.6, which lists the observations that have the six largest weights in the treated and control groups.

Output 100.1.6 Observations with Largest Weights

Observations with Largest IPTW-ATE Weights

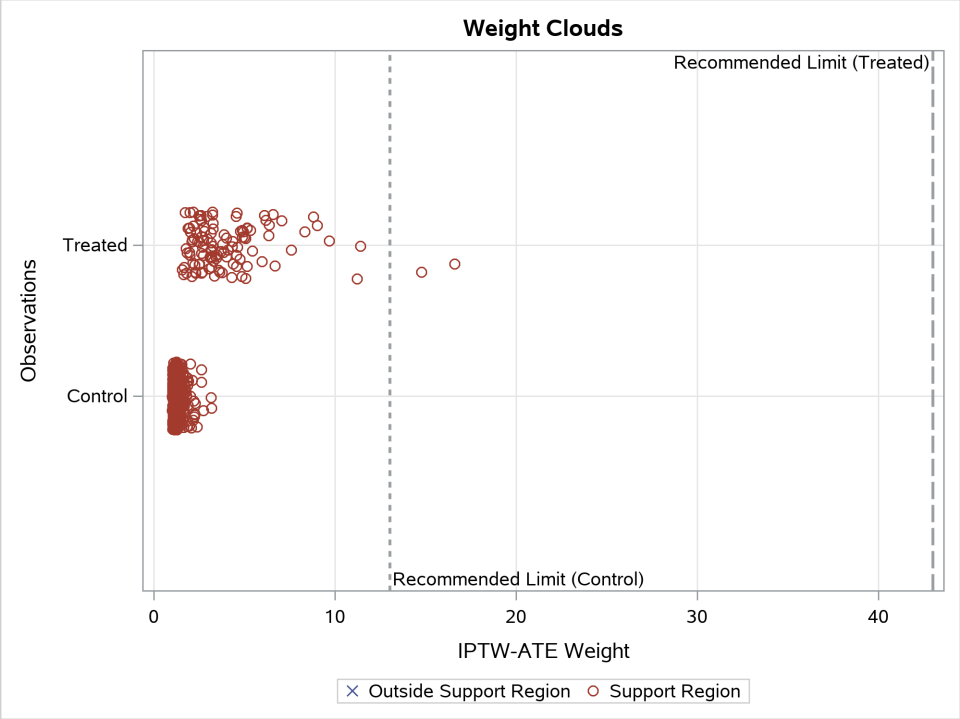
Treated (Drug = Drug_X) Expected Weight = 4.3009				Control (Drug = Drug_A) Expected Weight = 1.3029			
Observation	BMI	Weight	Scaled Weight	Observation	BMI	Weight	Scaled Weight
202	20.75	16.60	3.86	**317**	28.61	3.18	2.44
479	22.22	14.79	3.44	**134**	28.07	3.15	2.42
250	23.96	11.40	2.65	**437**	25.76	2.74	2.10
227	21.11	11.23	2.61	**417**	26.81	2.62	2.01
274	24.17	9.69	2.25	**446**	27.75	2.62	2.01
174	23.56	9.02	2.10	**81**	27.20	2.40	1.84

In the table, the scaled weights (which are the weights divided by their expected weights) are also displayed for ease of comparison. For more information about the expected weights in the treated and control group, see the section "Propensity Score Weighting" on page 8217.

The PLOTS=WGTCLOUD option displays a cloud plot for the stabilized weights, which is shown in Output 100.1.7. This plot is called a cloud plot because the points are jittered in the vertical direction in order

to avoid overplotting.

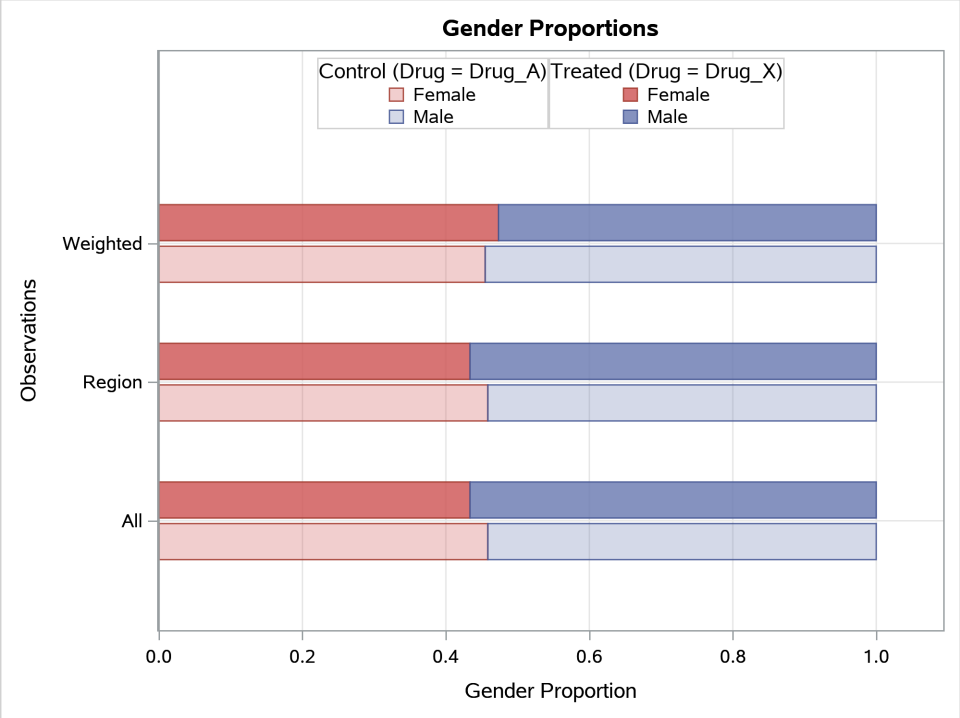
Output 100.1.7 Weight Cloud Plot



By default, the plot displays reference lines that represent 10 times the expected ATE weights in the treated and control groups. For information about these average weights, see the section "Inverse Probability of Treatment Weighting" on page 8217.

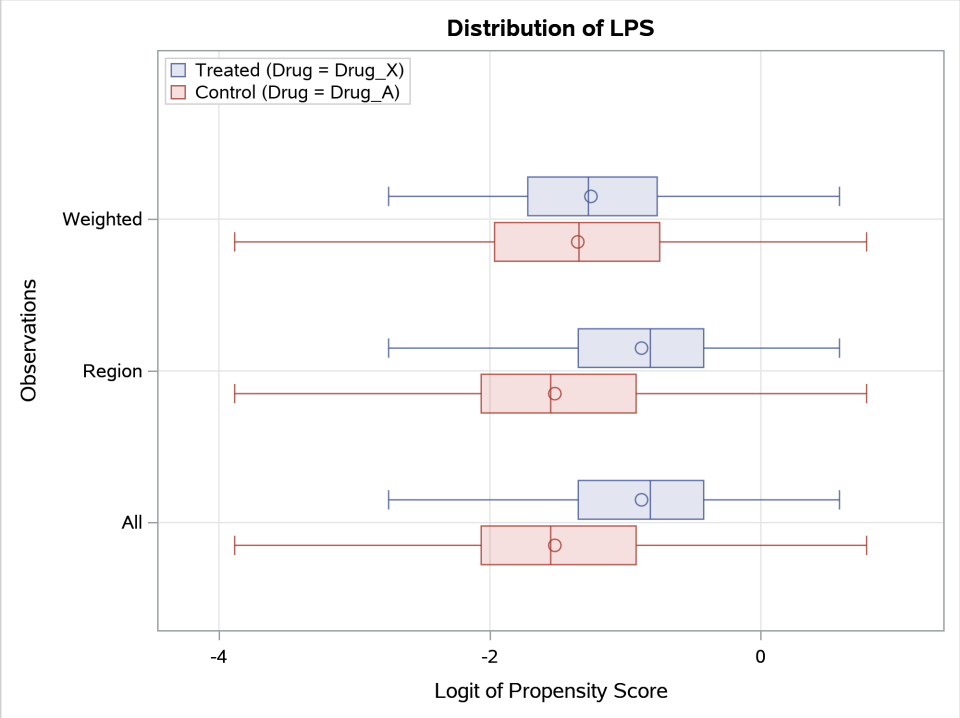
The PLOTS=BARCHART option displays a bar chart for each classification variable that is specified in the ASSESS statement. As shown in Output 100.1.8, the bar chart shows the distributions of Gender based on all observations, on observations in the support region, and on weighted observations. By default, the bar chart displays the proportions of levels of Gender. Weighting the observations makes a slight improvement in the balance between males and females.

Output 100.1.8 Gender Bar Chart

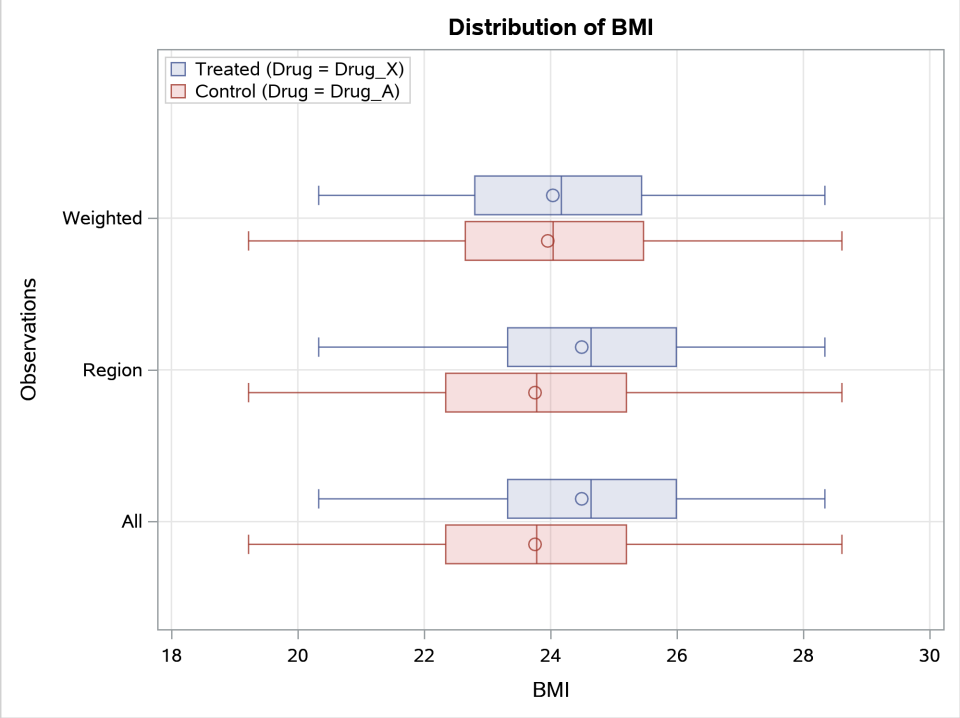


The PLOTS=BOXPLOT(DISPLAY=(LPS BMI)) option displays box plots for LPS and BMI, as shown in Output 100.1.9 and Output 100.1.10, respectively. These plots compare the distributions of the variables for the treated and control groups. Weighting the observations makes a good improvement in the balance between males and females.

Output 100.1.9 LPS Box Plot



Output 100.1.10 BMI Box Plot



Because there is good balance in the weighted distributions of the variables Gender, Age, and BMI, the

observations and their weights can be saved in an output data set for use in a subsequent outcome analysis.

In situations where you are not satisfied with the variable balance, you can do one or more of the following to improve the balance: you can select another set of variables to fit the propensity score model, you can modify the specification of the propensity score model by using nonlinear terms for the continuous variables or by adding interactions (Rosenbaum and Rubin 1984), or you can choose another propensity score method (such as matching).

The OUT(OBS=ALL)=OutEx1 option in the OUTPUT statement creates an output data set, OutEx1, that contains all available observations. The following statements list the first 10 observations in OutEx1, as shown in Output 100.1.11.

```
proc print data=OutEx1(obs=10);
   var PatientID Drug Gender Age BMI _ps_ _AteWgt_;
run;
```

Output 100.1.11 Output Data Set with ATE Weights

Obs	PatientID	Drug	Gender	Age	BMI	_PS_	_ATEWgt_
1	284	Drug_X	Male	29	22.02	0.36444	2.74397
2	201	Drug_A	Male	45	26.68	0.22296	1.28694
3	147	Drug_A	Male	42	21.84	0.11323	1.12768
4	307	Drug_X	Male	38	22.71	0.19733	5.06767
5	433	Drug_A	Male	31	22.76	0.35311	1.54586
6	435	Drug_A	Male	43	26.86	0.27263	1.37482
7	159	Drug_A	Female	45	25.47	0.14911	1.17523
8	368	Drug_A	Female	49	24.28	0.07780	1.08437
9	286	Drug_A	Male	31	23.31	0.38341	1.62182
10	163	Drug_X	Female	39	25.34	0.24995	4.00073

By default, the output data set includes the variable _PS_, which provides the propensity score. The weight for each treated unit is computed as $1 / p$ and the weight for each control unit is computed as $1 / (1 - p)$, where p is the propensity score.

After the responses for the trial are observed, they can be added to the data set OutEx1 as the starting point for an outcome analysis. Assuming that no other confounding variables are associated with both the response variable and the treatment group indicator Drug, you can estimate the ATE by performing a weighted version of the outcome analysis that you would have used to estimate the treatment effect if the original data set had resulted from a randomized trial.

Example 100.2: Propensity Score Stratification

This example illustrates how you can stratify observations based on their propensity scores, so that the stratified observations can be used to estimate the treatment effect in a subsequent outcome analysis (the outcome analysis is not shown here).

The data for this example are observations on patients in a nonrandomized clinical trial. The trial and the Drugs data set that contains the patient information are described in the section "Getting Started: PSMATCH Procedure" on page 8184.

The following statements create five strata that are based on propensity scores:

```
ods graphics on;
proc psmatch data=drugs region=allobs;
   class Drug Gender;
   psmodel Drug(Treated='Drug_X')= Gender Age BMI;
   strata nstrata=5 key=treated stratumwgt=total;
   assess ps var=(Gender BMI)
          / varinfo plots=(barchart cdfplot);
   output out(obs=all)=OutEx2;
run;
```

The PSMODEL statement specifies the logistic regression model that creates the propensity score for each observation, which is the probability that the patient receives Drug_X. The CLASS statement specifies the classification variables in the model. The Drug variable is the binary treatment indicator variable, and TREATED='Drug_X' identifies Drug_X as the treated group. The Gender, Age, and BMI variables are included in the model because they are believed to be related to the assignment.

The PSMATCH procedure stratifies the observations whose propensity scores lie in the support region that is specified in the REGION= option. The REGION=ALLOBS option requests that all observations be stratified.

The STRATA statement creates strata of observations based on their propensity scores. The NSTRATA=5 option (which is the default) stratifies the observations into five strata and the KEY=TREATED option (which is the default) requests that each stratum contain approximately the same number of treated observations.

The "Data Information" table in Output 100.2.1 displays the numbers of observations in the treated and control groups, the lower and upper limits of the propensity scores for observations in the support region, and the numbers of observations in the treated and control groups that fall within the support region. Because REGION=ALLOBS is specified, the lower and upper limits of the propensity scores for observations in the support region are simply the minimum and maximum of the propensity scores for all observations. Likewise, all 373 observations in the control group fall within the support region.

Output 100.2.1 Data Information

The PSMATCH Procedure

Data Information	
Data Set	WORK.DRUGS
Output Data Set	WORK.OUTEX2
Treatment Variable	Drug
Treated Group	Drug_X
All Obs (Treated)	113
All Obs (Control)	373
Support Region	All Obs
Lower PS Support	0.020157
Upper PS Support	0.685757
Support Region Obs (Treated)	113
Support Region Obs (Control)	373
Number of Strata	5

The "Propensity Score Information" table in Output 100.2.2 displays summary statistics for the treated and control groups. Statistics are computed for all observations, for observations in the support region, and for strata. The first two rows of statistics are identical because REGION=ALLOBS is specified.

Output 100.2.2 Propensity Score Information

Propensity Score Information

	Treated (Drug = Drug_X)					Control (Drug = Drug_A)					Treated - Control
Observations	**N**	**Mean**	**Standard Deviation**	**Minimum**	**Maximum**	**N**	**Mean**	**Standard Deviation**	**Minimum**	**Maximum**	**Mean Difference**
All	113	0.3108	0.1325	0.0602	0.6411	373	0.2088	0.1320	0.0202	0.6858	0.1020
Region	113	0.3108	0.1325	0.0602	0.6411	373	0.2088	0.1320	0.0202	0.6858	0.1020
Strata	5	0.2426	0.0212	0.1404	0.5121	5	0.2318	0.0227	0.1159	0.5312	0.0107

Note that for strata, the displayed statistics refer to the strata rather than to the observational units. These statistics are the number of strata, the weighted mean of the stratum means, the standard deviation of this weighted mean, and the minimum and maximum of these stratum means. The weights for computing the weighted mean are the stratum weights (as shown in the last column of Output 100.2.3), which are proportions of total units (treated and control units combined) in strata by the default STRATUMWGT=TOTAL option.

When you specify a STRATA statement, the "Strata Information" table, which is shown in Output 100.2.3, displays the following information for each stratum: the minimum and maximum propensity scores, the number of observations in the treatment group, the number of observations in the control group, the total number of observations, and the stratum weight.

Output 100.2.3 Strata Information

Strata Information

Stratum Index	Propensity Score Range		Frequencies Treated	Control	Total	Stratum Weight
1	0.0202	0.1944	22	209	231	0.475
2	0.1967	0.2613	23	59	82	0.169
3	0.2619	0.3223	23	38	61	0.126
4	0.3259	0.4342	23	41	64	0.132
5	0.4379	0.6858	22	26	48	0.099

The table shows that each stratum contains approximately the same number of observations for treated units, as requested by the KEY=TREATED option. In addition, there are enough control units in each stratum to ensure a reliable estimate of the treatment effect for this stratum, even though the propensity score distributions in the treated and control groups are different.

The ASSESS statement produces tables and plots, shown in Output 100.2.4 through Output 100.2.14, that summarize differences in the distributions of specified variables between treated and control groups. As requested by the PS and VAR= options, these variables are the propensity score and the data variables Gender and BMI. By default, differences are summarized for all observations and for observations in the support region. Again, these two sets of differences are identical because REGION=ALLOBS is specified. When you specify a STRATA statement, differences after stratification are also displayed.

The VARINFO option requests the "Variable Information" table, shown in Output 100.2.4, which displays variable summary statistics and mean differences between the treated and control groups for all observations (labeled "All"), for observations in the support region (labeled "Region"), and for the stratified observations (labeled "Strata"). The first two sets of statistics and mean differences are identical because REGION=ALLOBS is specified. For the binary classification variable (Gender), the difference is in the proportion of the first ordered level (Female).

Output 100.2.4 Variable Information

The PSMATCH Procedure

Variable Information

		Treated (Drug = Drug_X)					Control (Drug = Drug_A)				
Variable	Observations	N	Mean	Standard Deviation	Minimum	Maximum	N	Mean	Standard Deviation	Minimum	Maximum
Prop Score	All	113	0.31077	0.132467	0.06023	0.64115	373	0.20880	0.131969	0.02016	0.68576
	Region	113	0.31077	0.132467	0.06023	0.64115	373	0.20880	0.131969	0.02016	0.68576
	Strata	5	0.24256	0.021157	0.14040	0.51209	5	0.23182	0.022710	0.11589	0.53118
BMI	All	113	24.49257	1.863797	20.33000	28.34000	373	23.75327	1.980778	19.22000	28.61000
	Region	113	24.49257	1.863797	20.33000	28.34000	373	23.75327	1.980778	19.22000	28.61000
	Strata	5	24.08101	0.954962	23.50091	25.69500	5	23.89531	1.019772	23.17541	25.73077
Gender	All	113	0.43363	0.495575			373	0.45845	0.498270		
	Region	113	0.43363	0.495575			373	0.45845	0.498270		
	Strata	5	0.44921	0.270442			5	0.45232	0.270450		

Variable Information

Variable	Observations	Treated - Control Mean Difference
Prop Score	All	0.10197
	Region	0.10197
	Strata	0.01074
BMI	All	0.73930
	Region	0.73930
	Strata	0.18570
Gender	All	-0.02482
	Region	-0.02482
	Strata	-0.00311

For each variable, the row labeled "Strata" displays the number of strata, the weighted mean of the stratum means, and the standard deviation of this weighted mean, where the weights are computed as proportions of total units (treated and control units combined) in strata by the default STRATUMWGT=TOTAL option. The row also displays the minimum and maximum of the variable averages within the strata.

For each variable, the "Standardized Mean Differences" table in Output 100.2.5 displays the standardized mean differences between the treated and control groups for all observations, for observations in the support region, and for the stratified observations. The sections "Weighting after Stratification" on page 8221 and "Pooled Standardized Mean Differences across the Strata" on page 8230 explain how the statistics are computed for the stratified observations.

Output 100.2.5 Standardized Mean Differences

Standardized Mean Differences (Treated - Control)						
Variable	Observations	Mean Difference	Standard Deviation	Standardized Difference	Percent Reduction	Variance Ratio
Prop Score	All	0.10197	0.132218	0.77124		1.0076
	Region	0.10197		0.77124	0.00	1.0076
	Strata	0.01074		0.08121	89.47	0.8678
BMI	All	0.73930	1.923178	0.38441		0.8854
	Region	0.73930		0.38441	0.00	0.8854
	Strata	0.18570		0.09656	74.88	0.8769
Gender	All	-0.02482	0.496925	-0.04994		0.9892
	Region	-0.02482		-0.04994	0.00	0.9892
	Strata	-0.00311		-0.00627	87.45	0.9999
Standard deviation of All observations used to compute standardized differences						

When you specify a STRATA statement, the ASSESS statement also produces stratum-specific versions of tables and plots that summarize differences in the distributions of the specified variables between treated and control groups.

In addition to the "Variable Information" table shown in Output 100.2.4, the VARINFO option in the ASSESS statement produces the "Strata Variable Information" table, shown in Output 100.2.6, which displays variable summary statistics and mean differences between the treated and control groups for the observations in each stratum.

Output 100.2.6 Strata Variable Information

The PSMATCH Procedure

Strata Variable Information												
		Treated (Drug = Drug_X)					Control (Drug = Drug_A)					Treated - Control
Variable	Stratum Index	N	Mean	Standard Deviation	Minimum	Maximum	N	Mean	Standard Deviation	Minimum	Maximum	Mean Difference
Prop Score	1	22	0.14040	0.041360	0.06023	0.19436	209	0.11589	0.043859	0.02016	0.19413	0.02451
	2	23	0.22199	0.019418	0.19674	0.25936	59	0.22821	0.018395	0.19734	0.26130	-0.00622
	3	23	0.29986	0.018811	0.26350	0.32230	38	0.29457	0.017541	0.26186	0.32156	0.00529
	4	23	0.38087	0.026077	0.32668	0.43418	41	0.37055	0.030646	0.32594	0.43421	0.01032
	5	22	0.51209	0.058200	0.43793	0.64115	26	0.53118	0.071893	0.44120	0.68576	-0.01910
BMI	1	22	23.50091	1.751203	20.33000	26.11000	209	23.17541	1.917237	19.24000	27.85000	0.32550
	2	23	23.65304	1.794401	20.43000	26.66000	59	23.87847	1.951062	19.22000	27.68000	-0.22543
	3	23	24.70783	1.764444	20.85000	27.56000	38	24.10816	1.698325	20.24000	27.60000	0.59967
	4	23	24.91522	1.950177	20.98000	28.34000	41	24.93585	1.484916	22.37000	28.29000	-0.02064
	5	22	25.69500	1.130338	23.32000	28.06000	26	25.73077	1.337953	23.41000	28.61000	-0.03577
Gender	1	22	0.45455	0.497930			209	0.50718	0.499948			-0.05263
	2	23	0.56522	0.495728			59	0.32203	0.467256			0.24318
	3	23	0.39130	0.488042			38	0.44737	0.497222			-0.05606
	4	23	0.43478	0.495728			41	0.39024	0.487805			0.04454
	5	22	0.31818	0.465770			26	0.50000	0.500000			-0.18182

The "Standardized Mean Differences within Strata" table in Output 100.2.7 is a stratum-specific version of the

"Standardized Mean Differences" table in Output 100.2.5; it displays the variable mean differences, standardized mean differences, percentage reductions, ratios of variances for the observations, and stratum weights in each stratum. In Output 100.2.7, the standardized mean difference is the variable mean difference divided by the standard deviation shown in the "Standardized Mean Differences" table; the percentage reduction compares the standardized mean difference with the standardized mean difference of all observations.

The stratum weight is the number of treated units in each stratum divided by the combined number of treated units, as specified by the STRATUMWGT=TREATED option.

Output 100.2.7 Standardized Mean Differences within Strata

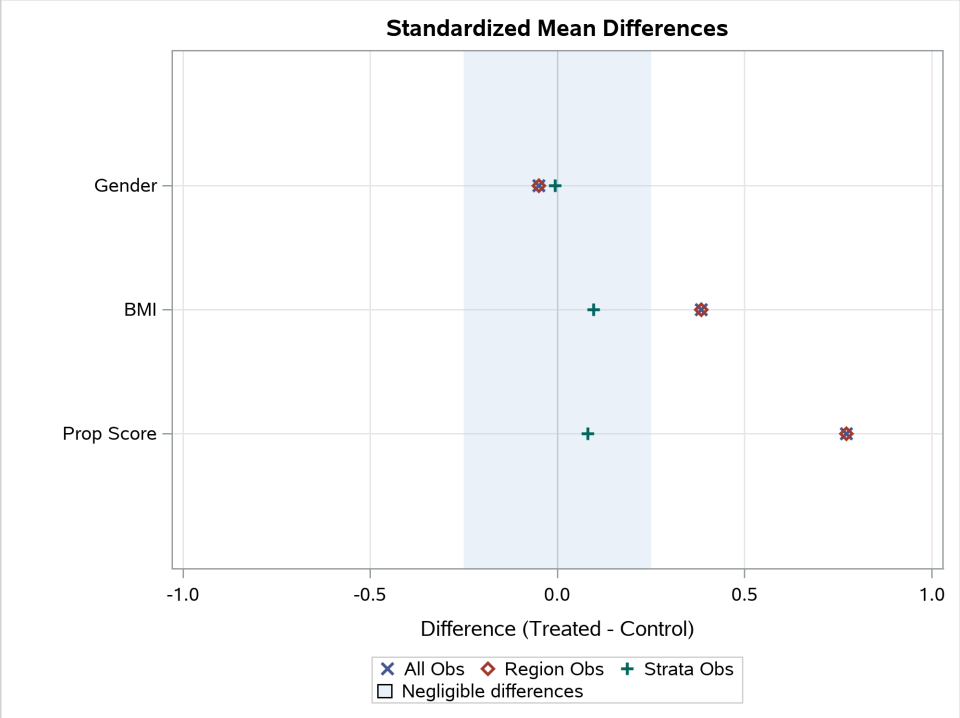
Standardized Mean Differences (Treated - Control) within Strata

Variable	Stratum Index	Mean Difference	Standardized Difference	Percent Reduction	Variance Ratio	Stratum Weight
Prop Score	**1**	0.02451	0.18537	75.96	0.8893	0.475
	2	-0.00622	-0.04703	93.90	1.1143	0.169
	3	0.00529	0.04003	94.81	1.1500	0.126
	4	0.01032	0.07803	89.88	0.7241	0.132
	5	-0.01910	-0.14443	81.27	0.6553	0.099
BMI	**1**	0.32550	0.16925	55.97	0.8343	0.475
	2	-0.22543	-0.11722	69.51	0.8459	0.169
	3	0.59967	0.31181	18.89	1.0794	0.126
	4	-0.02064	-0.01073	97.21	1.7248	0.132
	5	-0.03577	-0.01860	95.16	0.7137	0.099
Gender	**1**	-0.05263	-0.10591	0.00	0.9919	0.475
	2	0.24318	0.48938	0.00	1.1256	0.169
	3	-0.05606	-0.11282	0.00	0.9634	0.126
	4	0.04454	0.08963	0.00	1.0328	0.132
	5	-0.18182	-0.36589	0.00	0.8678	0.099

Note that a zero percentage reduction is displayed for Gender in each stratum because its standardized mean difference in the stratum (in absolute value) is larger than the standardized mean difference of all observations (0.04994 in absolute value).

Output 100.2.8 displays a standardized mean differences plot for the variables that are specified in the ASSESS statement.

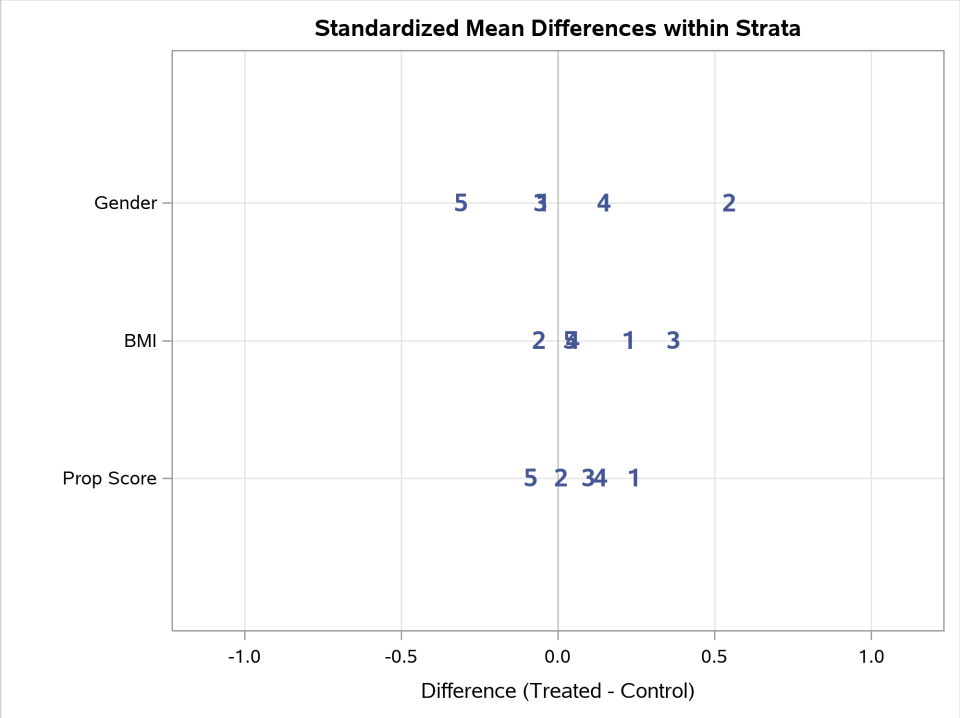
Output 100.2.8 Standardized Mean Differences Plot



In addition to differences based on all observations and on observations in the support region (which are identical), this plot displays differences based on combining estimates across strata, which are much smaller. For more information about these differences, see the sections "Weighting after Stratification" on page 8221 and "Pooled Standardized Mean Differences across the Strata" on page 8230.

Output 100.2.9 displays a plot of the standardized mean differences for each of the five strata.

Output 100.2.9 Standardized Mean Differences within Strata Plot

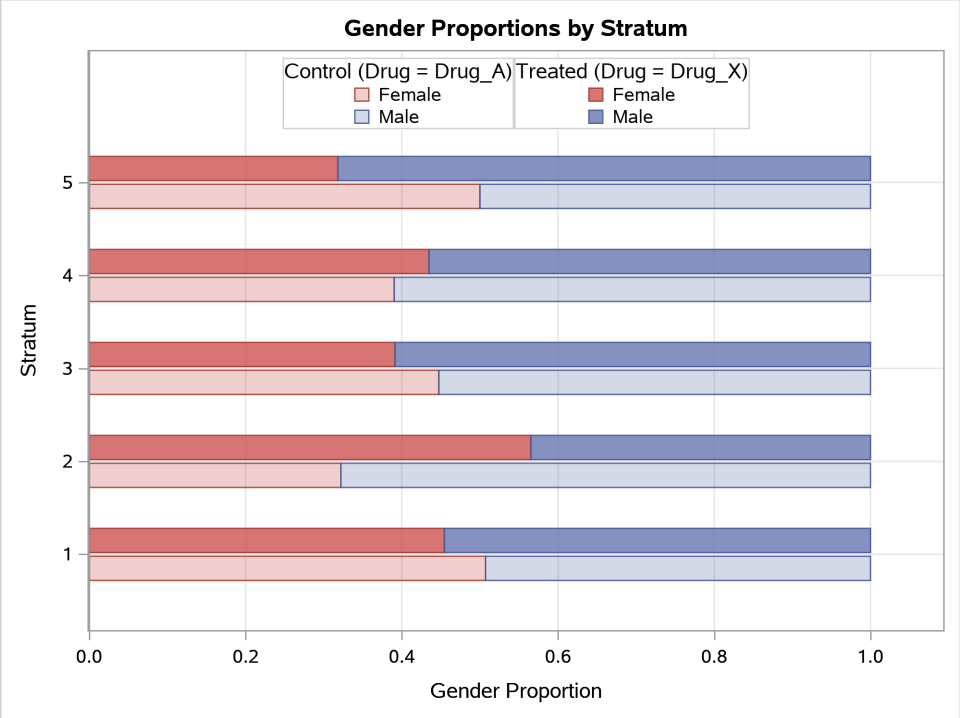


Note that recommended ranges for stratum-specific standardized mean differences are currently not available in the literature.

The "Standardized Mean Differences within Strata" plot corresponds to the "Standardized Mean Differences within Strata" table in Output 100.2.9. The plot reveals larger differences in Stratum 2 and Stratum 5 for Gender.

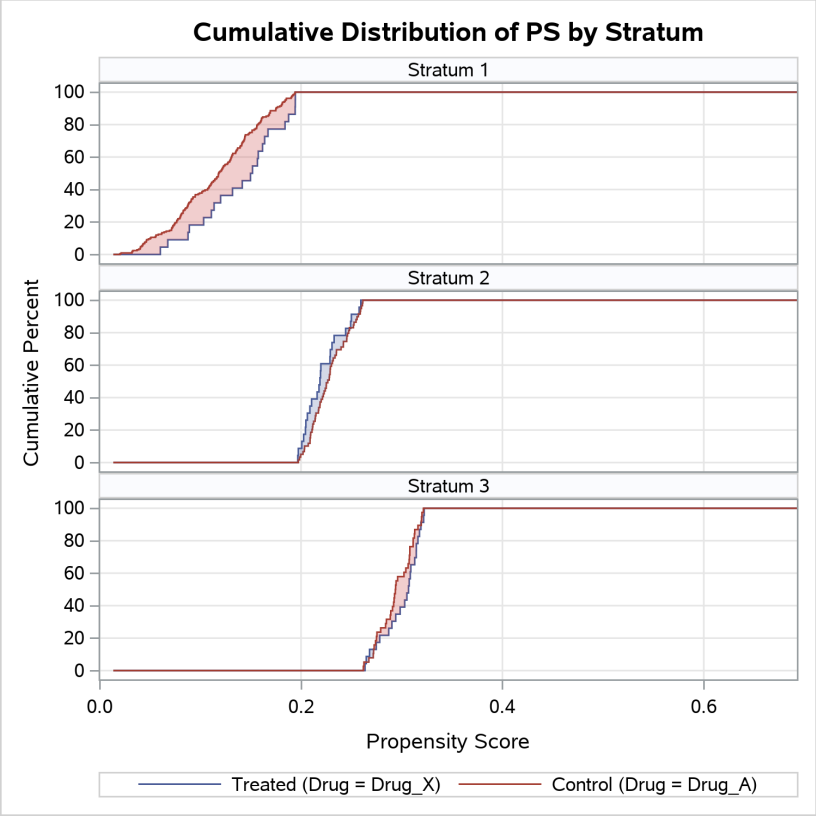
The PLOTS=BARCHART option displays stratum-specific bar charts for the distributions of classification variables in the treated and control groups, as shown in Output 100.2.10 for Gender. Here the largest differences in the distributions occur in Stratum 2 and Stratum 5.

Output 100.2.10 Gender Strata Bar Chart

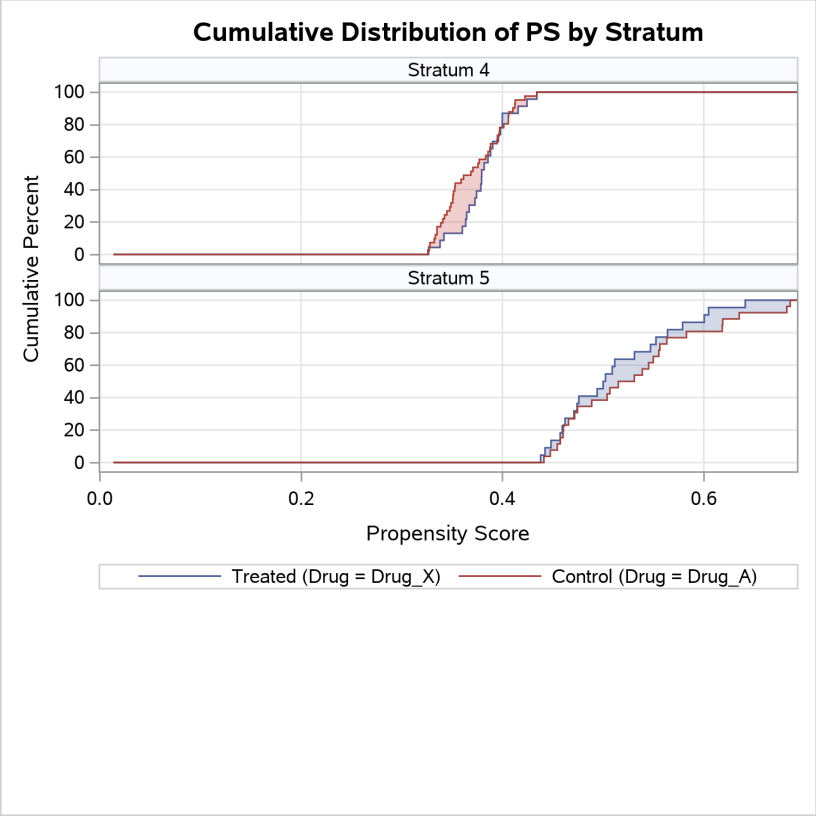


The PLOTS=CDFPLOT option displays stratum-specific CDF plots for the continuous variables in the treated and control groups, as shown in Output 100.2.11 and Output 100.2.12 for PS and in Output 100.2.13 and Output 100.2.14 for BMI.

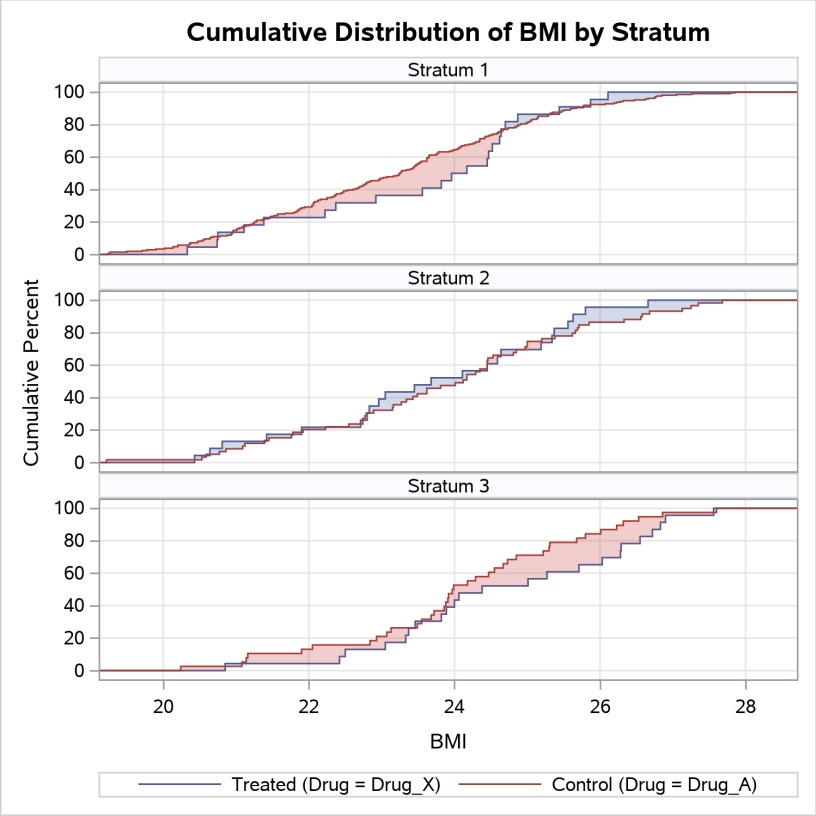
Output 100.2.11 PS Strata CDF Plot



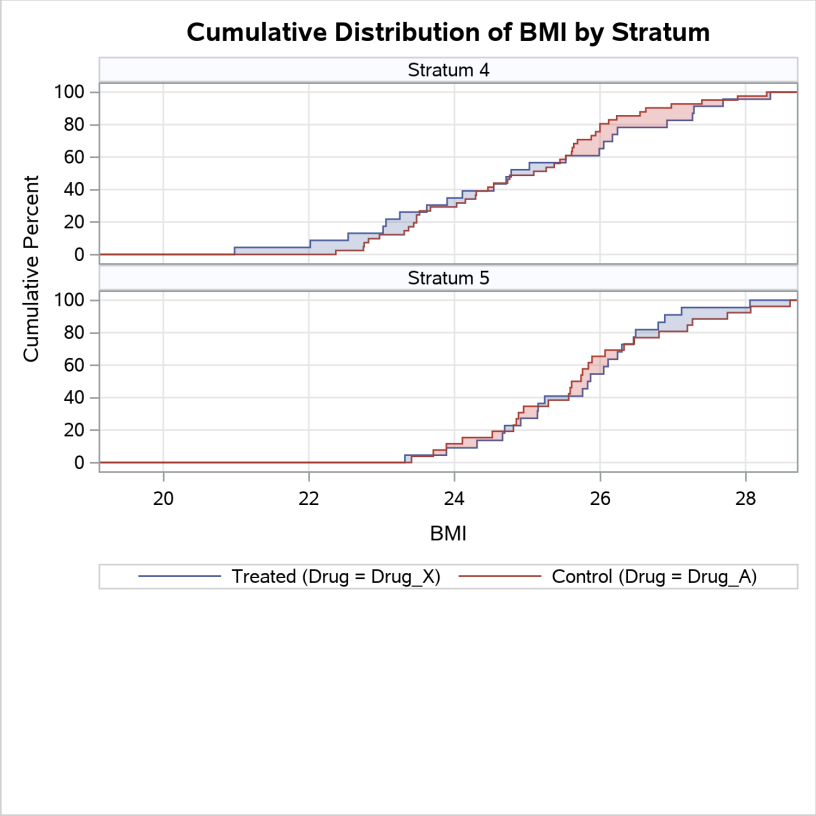
Output 100.2.12 PS Strata CDF Plot



Output 100.2.13 BMI Strata CDF Plot



Output 100.2.14 BMI Strata CDF Plot



The plots show the differences in the distributions in strata. Here, the largest differences in the distributions of propensity score occur in Stratum 1 (lower values in the control group) and in Stratum 5 (higher values in the treated group)

Because stratification results in good balance for the variables in this example, as shown in Output 100.2.5 and Output 100.2.8, the stratified observations can be saved in an output data set for use in a subsequent outcome analysis.

In situations where you are not satisfied with the variable balance, you can do one or more of the following to improve the balance: you can select another set of variables to fit the propensity score model, you can modify the specification of the propensity score model (for instance, by using nonlinear terms for the continuous variables or by adding interactions), you can increase the number of strata, or you can choose another propensity score method (such as matching).

The OUT(OBS=ALL)=OutEx2 option in the OUTPUT statement creates an output data set named OutEx2 that contains all observations. The following statements list the first 10 observations in OutEx2, which are shown in Output 100.2.15:

```
proc print data=OutEx2(obs=10);
   var PatientID Drug Gender Age BMI _ps_ _Strata_;
run;
```

Output 100.2.15 Output Data Set with Strata

Obs	PatientID	Drug	Gender	Age	BMI	_PS_	_STRATA_
1	284	Drug_X	Male	29	22.02	0.36444	4
2	201	Drug_A	Male	45	26.68	0.22296	2
3	147	Drug_A	Male	42	21.84	0.11323	1
4	307	Drug_X	Male	38	22.71	0.19733	2
5	433	Drug_A	Male	31	22.76	0.35311	4
6	435	Drug_A	Male	43	26.86	0.27263	3
7	159	Drug_A	Female	45	25.47	0.14911	1
8	368	Drug_A	Female	49	24.28	0.07780	1
9	286	Drug_A	Male	31	23.31	0.38341	4
10	163	Drug_X	Female	39	25.34	0.24995	2

By default, the output data set includes the variable _PS_, which provides the propensity score, and the variable _STRATA_, which identifies the stratum.

After the responses for the trial are observed, they can be added to the data set OutEx2 as the starting point for an outcome analysis. Assuming that no other confounding variables are associated with both the response variable and the treatment group indicator Drug, you can estimate the treatment effect within each stratum and combine these estimates across strata to estimate the overall treatment effect (Stuart 2010, pp. 13–14). Note that the same stratum weights, as specified in the STRATUMWGT= option in the assessment, should be used in the outcome analysis.

Example 100.3: Optimal Variable Ratio Matching

This example illustrates how you can perform optimal variable ratio matching of observations in a control group with observations in a treatment group, so that the matched observations can be used to estimate the treatment effect in a subsequent outcome analysis. The outcome analysis itself is not shown here.

The data for this example are observations on patients in a nonrandomized clinical trial. The trial and the Drugs data set that contains the patient information are described in the section "Getting Started: PSMATCH Procedure" on page 8184.

The following statements request optimal variable ratio matching to match each observation for patients in the treatment group with a variable number of observations for patients in the control group:

```
ods graphics on;
proc psmatch data=drugs region=treated(extend(distance=ps mult=one)=0.025);
   class Drug Gender;
   psmodel Drug(Treated='Drug_X')= Gender Age BMI;
   match distance=ps method=varratio(kmin=1 kmax=4) exact=(Gender) caliper=.
         nmatchmost=5;
   assess ps var=(Gender Age BMI)
          / stddev=pooled(allobs=no) plots(orient=vertical nodetails);
   id BMI;
   output out(obs=match)=OutEx3 matchid=_MatchID;
run;
```

The PSMODEL statement specifies the logistic regression model that creates the propensity score for each observation, which is the probability that the patient receives Drug_X. The CLASS statement specifies

the classification variables in the model. The Drug variable is the binary treatment indicator variable, and TREATED='Drug_X' identifies Drug_X as the treated group. The Gender, Age, and BMI variables are included in the model because they are believed to be related to the assignment.

The PSMATCH procedure matches only those observations whose propensity scores lie in the support region that you specify with the REGION= option. Here the option REGION=TREATED requests that only those observations whose propensity scores lie in the region defined by the treated observations be used in matching. The suboption EXTEND(DISTANCE=PS MULT=ONE)=0.025 requests that this region be extended by 0.025 in propensity score.

The MATCH statement specifies the criteria for matching. The DISTANCE=PS option requests that the propensity score be used in computing differences between pairs of observations. The METHOD=VARRATIO(KMIN=1 KMAX=4) option requests optimal variable ratio matching of one to four control units to each unit in the treated group in order to minimize the total absolute difference in propensity scores across all matches.

The default average number of control units that are matched to each treated unit is computed as the mean of the KMIN= and KMAX= values, so an average of two control units are matched to each treated unit. The EXACT=GENDER option requests that the treated unit and its matched control unit have the same value of Gender. The CALIPER=. option ignores the caliper requirement for matching.

The "Data Information" table in Output 100.3.1 displays the numbers of observations in the treated and control groups, the lower and upper limits of the propensity scores for observations in the support region, and the numbers of observations in the treated and control groups that fall within the support region. Of the 373 observations in the control group, 366 fall within the support region. By definition, all 113 of the observations in the treated group fall within the support region.

Output 100.3.1 Data Information

The PSMATCH Procedure

Data Information	
Data Set	WORK.DRUGS
Output Data Set	WORK.OUTEX3
Treatment Variable	Drug
Treated Group	Drug_X
All Obs (Treated)	113
All Obs (Control)	373
Support Region	Extended Treated Group
Lower PS Support	0.035231
Upper PS Support	0.666148
Support Region Obs (Treated)	113
Support Region Obs (Control)	366

The "Propensity Score Information" table in Output 100.3.2 displays summary statistics by treatment group for all observations, for observations in the support region, and for matched observations. The three sets of summary statistics for the treated group are identical because REGION=TREATED is specified. Because the default WEIGHT=ATTWGT is used in the MATCH statement, the table also displays summary statistics of weighted matched observations by treatment group.

Output 100.3.2 Propensity Score Information

Propensity Score Information												
	Treated (Drug = Drug_X)						Control (Drug = Drug_A)					
Observations	N	Weight	Mean	Standard Deviation	Minimum	Maximum	N	Weight	Mean	Standard Deviation	Minimum	Maximum
All	113		0.3108	0.1325	0.0602	0.6411	373		0.2088	0.1320	0.0202	0.6858
Region	113		0.3108	0.1325	0.0602	0.6411	366		0.2087	0.1267	0.0371	0.6351
Matched	113		0.3108	0.1325	0.0602	0.6411	283		0.2450	0.1214	0.0510	0.6351
Weighted Matched	113	113.00	0.3108	0.1325	0.0602	0.6411	283	113.00	0.3006	0.1323	0.0510	0.6351

Propensity Score Information	
Observations	Treated - Control Mean Difference
All	0.1020
Region	0.1021
Matched	0.0658
Weighted Matched	0.0102

By default (or if you specify WEIGHT=ATTWGT in the MATCH statement), each treated unit receives a weight of 1 and each control unit receives a weight that is computed as the number of treated units divided by the number of control units in the matched set. For example, if three control units are matched to a treated unit in a matched set, then each control unit receives a weight of 1/3. Thus, the total weights of the weighted matched observations in the two treatment groups are the same. These weights are used to compute the standardized mean differences. For more information about these weights, see the section "Weighting after Matching" on page 8226.

The "Matching Information" table in Output 100.3.3 displays the matching criteria, the number of matched sets, the numbers of matched observations in the treated and control groups, and the total absolute difference in the propensity scores for all matches. Note that with an average of two and a half control units to each treated unit, 283 control units are matched.

Output 100.3.3 Matching Information

Matching Information	
Distance Metric	Propensity Score
Method	Optimal Variable Ratio Matching
Min Control/Treated Ratio	1
Max Control/Treated Ratio	4
Matched Sets	113
Matched Obs (Treated)	113
Matched Obs (Control)	283
Total Absolute Difference	3.644643

The ASSESS statement produces a table and a plot that summarize differences in the distributions of specified variables between treated and control groups for all observations, for observations in the support region, and for matched observations.

The "Standardized Mean Differences" table, shown in Output 100.3.4, displays standardized mean differences

in the variables between the treated and control groups for all observations, for observations in the support region, and for matched observations. As requested by the PS and VAR= options, the variables that are listed in the table are the propensity score and the variables Gender, Age, and BMI. For the binary classification variable (Gender), the difference is in the proportion of the first ordered level (Female).

Output 100.3.4 Standardized Mean Differences

The PSMATCH Procedure

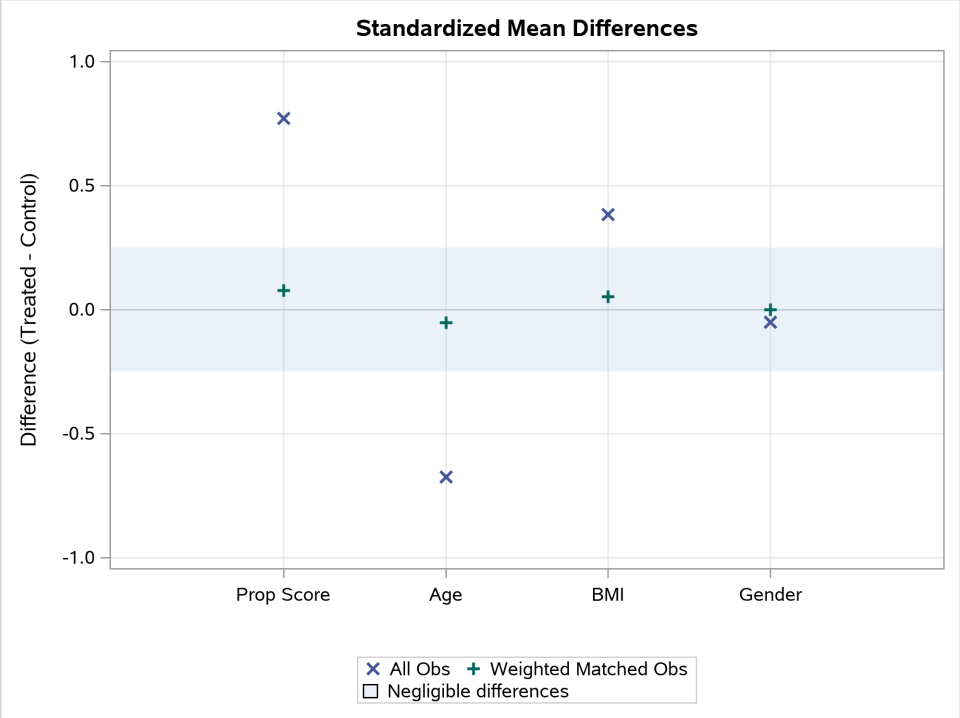
Standardized Mean Differences (Treated - Control)						
Variable	Observations	Mean Difference	Standard Deviation	Standardized Difference	Percent Reduction	Variance Ratio
Prop Score	**All**	0.10197	0.132218	0.77124		1.0076
	Region	0.10210	0.129634	0.78757	0.00	1.0924
	Matched	0.06579	0.127033	0.51787	32.85	1.1915
	Weighted Matched	0.01019	0.132397	0.07696	90.02	1.0021
Age	**All**	-4.09509	6.079104	-0.67363		0.7076
	Region	-3.98262	6.002096	-0.66354	1.50	0.7393
	Matched	-2.19910	5.693934	-0.38622	42.67	0.8951
	Weighted Matched	-0.29720	5.612599	-0.05295	92.14	0.9460
BMI	**All**	0.73930	1.923178	0.38441		0.8854
	Region	0.72871	1.902421	0.38305	0.36	0.9227
	Matched	0.41500	1.879869	0.22076	42.57	0.9665
	Weighted Matched	0.09697	1.822647	0.05320	86.16	1.0957
Gender	**All**	-0.02482	0.496925	-0.04994		0.9892
	Region	-0.02266	0.496832	-0.04560	8.69	0.9899
	Matched	-0.02574	0.496963	-0.05179	0.00	0.9889
	Weighted Matched	0.00000	0.495575	0.00000	100.00	1.0000

Note that a zero percentage reduction is displayed for Gender in the matched observation because its standardized mean difference (0.05179 in absolute value) is larger than the standardized mean difference of all observations (0.04994 in absolute value).

The standardized mean differences are significantly reduced in the matched observations, the standardized mean differences are less than the recommended upper limit of 0.25, and the variance ratios between the two groups are within the recommended range of 0.5 to 2.

The "Standardized Mean Differences" plot is shown in Output 100.3.5.

Output 100.3.5 Standardized Mean Differences Plot



The "Standardized Mean Differences" plot displays the standardized mean differences that are listed in the "Standardized Mean Differences" table in Output 100.3.4. When you specify the ORIENT=VERTICAL option, the standardized mean differences are plotted on the vertical axis. All differences for the matched observations are within the recommended limits of –0.25 and 0.25, which are indicated by the shaded area.

The NMATCHMOST=5 option requests the "Observations with Most Matches" table, which is shown in Output 100.3.6, and displays the five observations that have the most matches in the treated and control groups.

Output 100.3.6 Observations with the Most Matches

Observations with Most Matches

Treated (Drug = Drug_X)

Observation	BMI	Matched Control
202	20.75	4
479	22.22	4
250	23.96	4
227	21.11	4
274	24.17	4

Because matching results in good balance for the variables in this example, the matched observations can be saved in an output data set for use in a subsequent outcome analysis.

In situations where you are not satisfied with the variable balance, you can do one or more of the following to improve the balance: you can select another set of variables to fit the propensity score model, you can modify

the matching criteria, or you can choose another matching method.

The OUT(OBS=MATCH)=OutEx3 option in the OUTPUT statement creates an output data set, OutEx3, that contains the matched observations. The following statements list the observations in the first two matched sets, as shown in Output 100.3.7:

```
proc sort data=OutEx3 out=OutEx3a;
   by _MatchID;
run;

proc print data=OutEx3a(obs=10);
   var PatientID Drug Gender Age BMI _PS_ _MATCHWGT_ _MatchID;
run;
```

Output 100.3.7 Output Data Set with Optimal Variable Ratio Matches

Obs	PatientID	Drug	Gender	Age	BMI	_PS_	_MATCHWGT_	_MatchID
1	141	Drug_A	Female	43	20.55	0.064010	0.25	1
2	213	Drug_A	Female	49	23.24	0.061866	0.25	1
3	89	Drug_X	Female	44	20.75	0.060231	1.00	1
4	311	Drug_A	Female	49	22.80	0.056086	0.25	1
5	104	Drug_A	Female	46	20.95	0.050951	0.25	1
6	137	Drug_A	Female	45	22.04	0.072149	0.25	2
7	158	Drug_A	Female	48	23.64	0.075028	0.25	2
8	245	Drug_A	Female	52	25.32	0.071559	0.25	2
9	40	Drug_A	Female	42	20.65	0.072654	0.25	2
10	323	Drug_X	Female	46	22.22	0.067625	1.00	2

By default, the output data set includes the variable _PS_ (which provides the propensity score) and the variable _MATCHWGT_ (which provides matched observation weights). The weight for each treated unit is 1. Because METHOD=VARRATIO(KMIN=1 KMAX=4) is specified in the MATCH statement, one, two, three, or four control units are matched to each treated unit; so the weight for each matched control unit is 1, 1/2, 1/3, or 1/4. The MATCHID=_MatchID option creates a variable named _MatchID that identifies the matched sets of observations.

After the responses for the trial are observed, they can be added to the data set OutEx3 as the starting point for an outcome analysis. Assuming that no other confounding variables are associated with both the response variable and the treatment group indicator Drug, you can estimate the treatment effect from the matched observations by performing a weighted version of the outcome analysis that you would have used to estimate the treatment effect if the original data set had resulted from a randomized trial.

Example 100.4: Greedy Nearest Neighbor Matching

This example illustrates how you can perform greedy matching of observations in a control group with observations in a treatment group, so that the matched observations can be used to estimate the treatment effect in a subsequent outcome analysis. An outcome analysis is not shown here but is discussed in Example 100.5.

At the completion of a school year, a school administrator asks whether taking a music class caused an improvement in the grade point averages (GPAs) of students. The reasoning behind this question is that learning to read and perform music might improve general reading ability, concentration, and memory.

The data set School contains information about students that is available at the end of the school year. StudentID is the student identification number, Music indicates whether the student took a music class, Gender provides the gender of the student, and Absence is the percentage of absences. Output 100.4.1 lists the first 10 observations.

Output 100.4.1 Input School Data Set

Obs	StudentID	Music	Gender	Absence
1	18	No	Female	3.71
2	61	No	Male	2.08
3	95	No	Female	2.54
4	41	No	Male	3.01
5	19	Yes	Female	0.08
6	51	No	Female	1.20
7	110	No	Male	2.21
8	87	No	Female	2.30
9	103	No	Female	3.08
10	175	No	Female	1.12

In this example, the outcome variable GPA for the students was available at the end of the year. However, following recommended practice (Stuart 2010, p. 2), the values of GPA are not used in the propensity score analysis that is described in this example. Instead, the variable GPA is reserved for the outcome analysis, which is carried out on the output data set that is created by the PSMATCH procedure after it has been augmented with the values of GPA. See Example 100.5 for an illustration of an outcome analysis.

The following statements request greedy nearest neighbor matching to sequentially match each observation for students in the treatment group (those who took music) with one observation for students in the control group (those who did not take music):

```
ods graphics on;
proc psmatch data=School region=treated;
   class Music Gender;
   psmodel Music(Treated='Yes')= Gender Absence;
   match distance=lps method=greedy(k=1) exact=Gender caliper=0.5
         weight=none;
   assess lps var=(Gender Absence)
          / stddev=pooled(allobs=no) stdbinvar=no plots(nodetails)=all;
   output out(obs=match)=OutEx4 matchid=_MatchID;
run;
```

The PSMODEL statement specifies the logistic regression model that creates the propensity score for each student, which is the probability that the student enrolled in the music class. The Music variable is the binary treatment indicator variable and TREATED='Yes' identifies Yes as the treated group. The Gender and Absence variables are included in the model because they are believed to be related to enrolling in the music class. The CLASS statement specifies the classification variables.

The PSMATCH procedure matches only those observations whose propensity scores lie in the support region that you specify in the REGION= option. Here the REGION=TREATED option requests that only those observations whose propensity scores lie in the region defined by the treated observations be used in matching. By default, the region is extended by 0.25 times the pooled estimate of the common standard deviation of the logits of the propensity scores.

The MATCH statement requests matching and specifies the criteria for matching. The DISTANCE=LPS option (which is the default) requests that the logit of the propensity score be used in computing differences between pairs of observations. The METHOD=GREEDY(K=1) option requests greedy nearest neighbor matching in which one control unit is matched with each unit in the treated group; this produces the smallest within-pair difference among all available pairs with this treated unit. The EXACT=GENDER option requests that the treated unit and its matched control unit have the same value of the Gender variable. The CALIPER=0.5 option specifies the caliper requirement for matching. Units are matched only if the difference in the logits of the propensity scores for pairs of units from the two groups is less than or equal to 0.5 times the pooled estimate of the standard deviation.

The "Data Information" table, shown in Output 100.4.2, displays the numbers of observations in the treated and control groups, the lower and upper limits of the propensity scores for observations in the support region, and the numbers of observations in the treated and control groups that fall within the support region. Of the 140 observations in the control group, 132 fall within the support region.

Output 100.4.2 Data Information

The PSMATCH Procedure

Data Information	
Data Set	WORK.SCHOOL
Output Data Set	WORK.OUTEX4
Treatment Variable	Music
Treated Group	Yes
All Obs (Treated)	60
All Obs (Control)	140
Support Region	Extended Treated Group
Lower PS Support	0.079911
Upper PS Support	0.530945
Support Region Obs (Treated)	60
Support Region Obs (Control)	132

The "Propensity Score Information" table, shown in Output 100.4.3, displays summary statistics for the treatment and control groups. These statistics are computed for all observations, for observations in the support region, and for matched observations. The three sets of statistics are identical for the treated group because REGION=TREATED is specified.

Output 100.4.3 Propensity Score Information

Propensity Score Information											
	Treated (Music = Yes)					Control (Music = No)					Treated - Control
Observations	N	Mean	Standard Deviation	Minimum	Maximum	N	Mean	Standard Deviation	Minimum	Maximum	Mean Difference
All	60	0.3472	0.0963	0.0928	0.4902	140	0.2798	0.1251	0.0263	0.4887	0.0675
Region	60	0.3472	0.0963	0.0928	0.4902	132	0.2940	0.1141	0.0832	0.4887	0.0533
Matched	60	0.3472	0.0963	0.0928	0.4902	60	0.3402	0.0986	0.0928	0.4887	0.0070

The "Matching Information" table in Output 100.4.4 displays the matching criteria, the number of matched sets, the numbers of matched observations in the treated and control groups, and the total absolute difference in the logits of the propensity scores for all matches.

Output 100.4.4 Matching Information

Matching Information	
Distance Metric	Logit of Propensity Score
Method	Greedy Matching
Control/Treated Ratio	1
Order	Descending
Caliper (Logit PS)	0.326669
Matched Sets	60
Matched Obs (Treated)	60
Matched Obs (Control)	60
Total Absolute Difference	2.946274

The ASSESS statement produces tables and plots that summarize differences in the distributions of specified variables between treated and control groups for all observations, for observations in the support region, and for matched observations. You can use these results to assess how well matching achieves a balance in the distributions of these variables. As requested by the LPS and VAR= options, the variables are the logit of the propensity score and the covariates Gender and Absence. The WEIGHT=NONE option suppresses the display of differences for weighted matched observations. When PROC PSMATCH matches one control unit to each treated unit, it assigns a weight of 1 for all matched treated and control units, so the results are identical for weighted matched observations and matched observations.

The "Standardized Mean Differences" table, shown in Output 100.4.5, displays standardized mean differences in the variables between the treated and control groups, which are computed for all observations, for observations in the support region, and for matched observations. For the binary classification variable (Gender), the computed difference is in the proportion of the first ordered level (Female).

Output 100.4.5 Standardized Mean Differences

The PSMATCH Procedure

Standardized Mean Differences (Treated - Control)						
Variable	Observations	Mean Difference	Standard Deviation	Standardized Difference	Percent Reduction	Variance Ratio
Logit Prop Score	All	0.40769	0.653338	0.62402		0.3810
	Region	0.28487	0.556440	0.51195	17.96	0.6138
	Matched	0.03305	0.489082	0.06758	89.17	0.9698
Absence	All	-0.69807	1.136767	-0.61409		0.3550
	Region	-0.48623	0.973273	-0.49958	18.65	0.5561
	Matched	-0.05833	0.836446	-0.06974	88.64	0.9374
Gender	All	-0.04524				
	Region	-0.03485			22.97	
	Matched	0.00000			100.00	

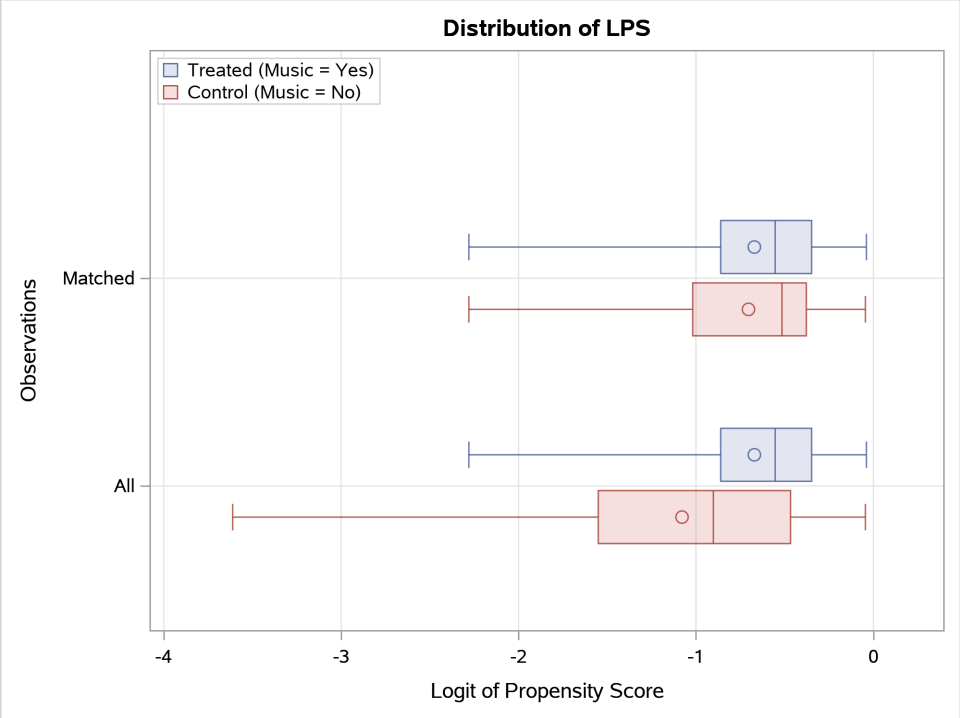
The standardized mean differences are significantly reduced in the matched observations, and the largest of these differences is 0.07015 in absolute value, which is less than the recommended upper limit of 0.25 (Rubin 2001, p. 174; Stuart 2010, p. 11). The treated-to-control variance ratios are 0.9380 and 0.9701 in the matched observations, which are within the recommended range of 0.5 to 2. Because EXACT=GENDER is specified in the MATCH statement, the standardized mean difference for Gender is 0 in the matched observations.

When you specify PLOTS=ALL, the PSMATCH procedure creates all applicable plots. Output 100.4.6 displays a plot of the standardized mean differences in Gender, Absence, and the logit of the propensity score for all observations and matched observations. Because the NODETAILS option is specified, the comparison of observations in the support region is not displayed except for the cloud plots.

Output 100.4.6 Standardized Mean Differences Plot

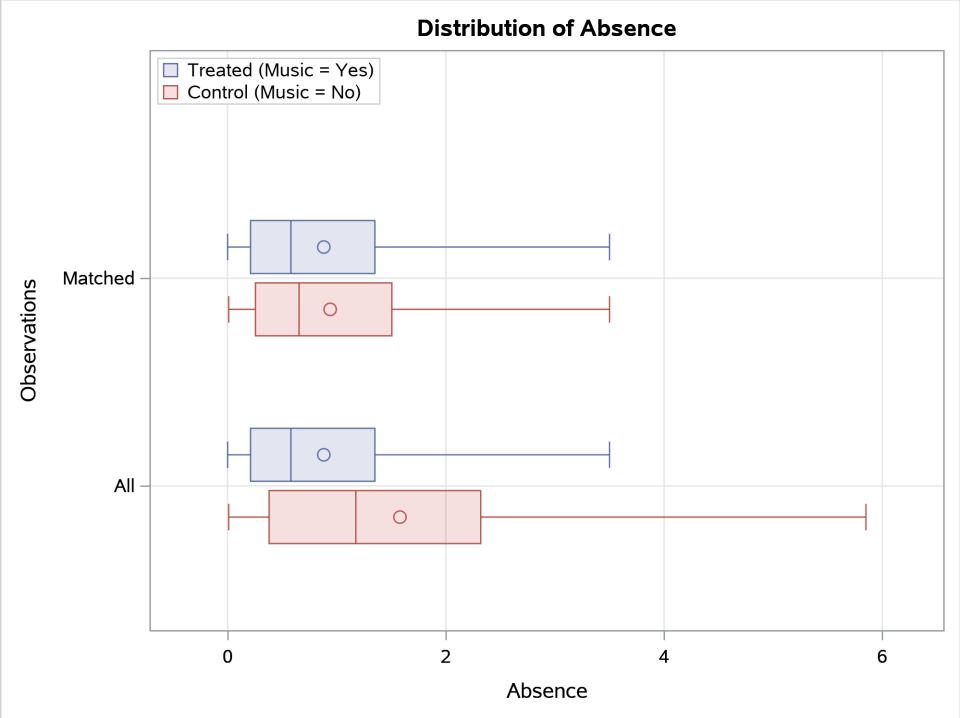
Output 100.4.7 displays box plots that compare the distributions of the logit propensity score for units in the treated and control groups, based on all observations, on observations in the support region, and on matched observations. Note that the two distributions are well balanced for matched observations.

Output 100.4.7 LPS Box Plot



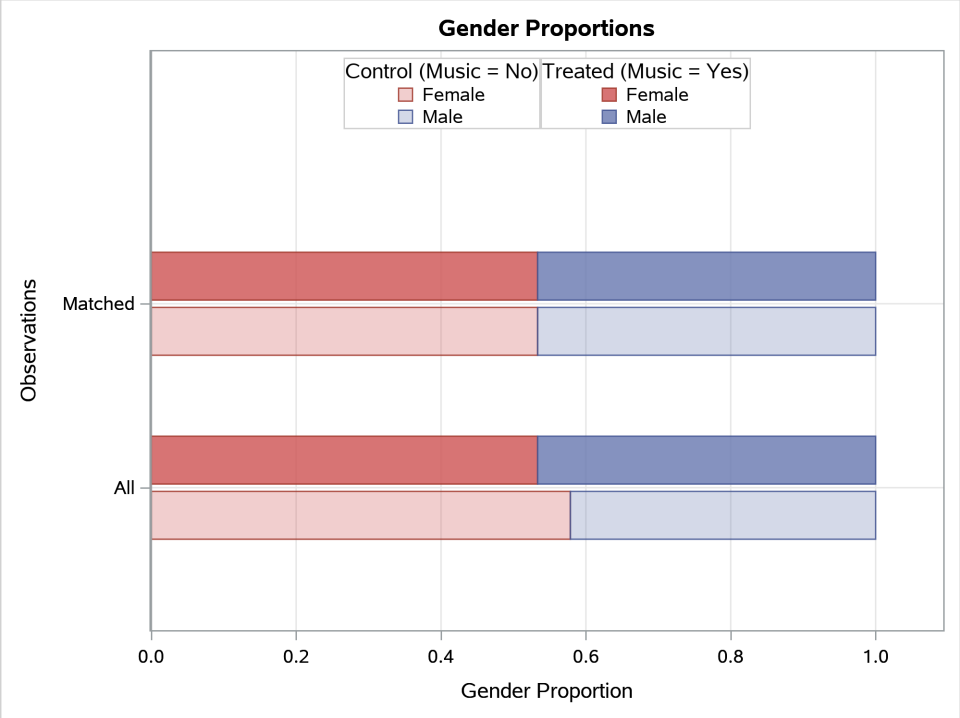
Output 100.4.8 displays box plots that compare the distributions of Absence for units in the treated and control groups, based on all observations and on matched observations. Again, note that the two distributions are well balanced for matched observations.

Output 100.4.8 Absence Box Plot



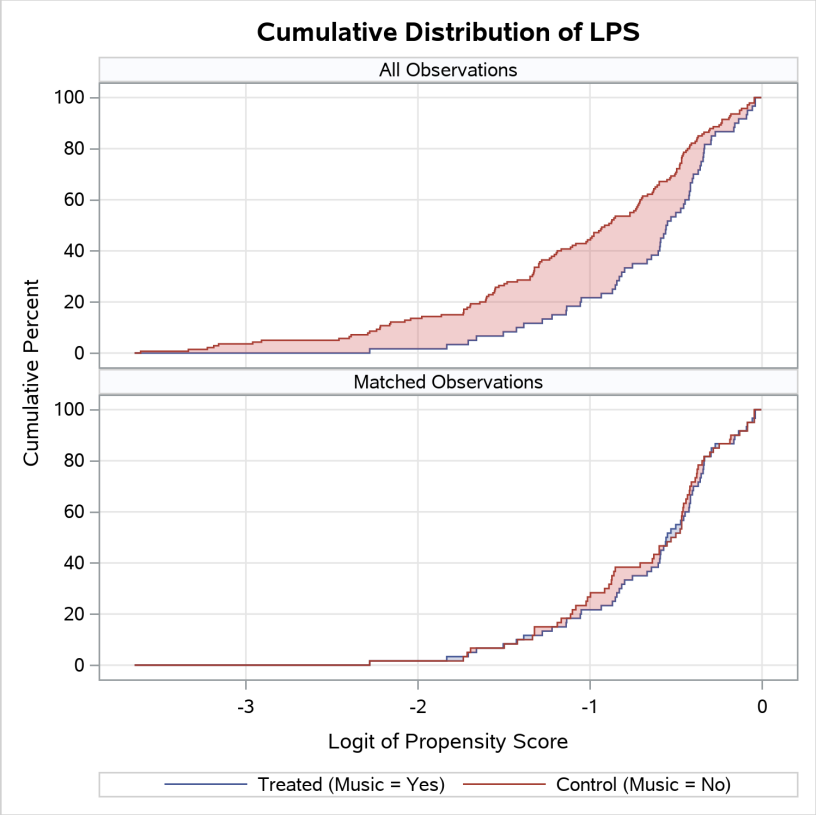
Output 100.4.9 displays bar charts that compare the distributions of Gender for units in the treated and control groups, based on all observations and on matched observations. Again, note that the two distributions are well balanced for matched observations.

Output 100.4.9 Gender Bar Chart



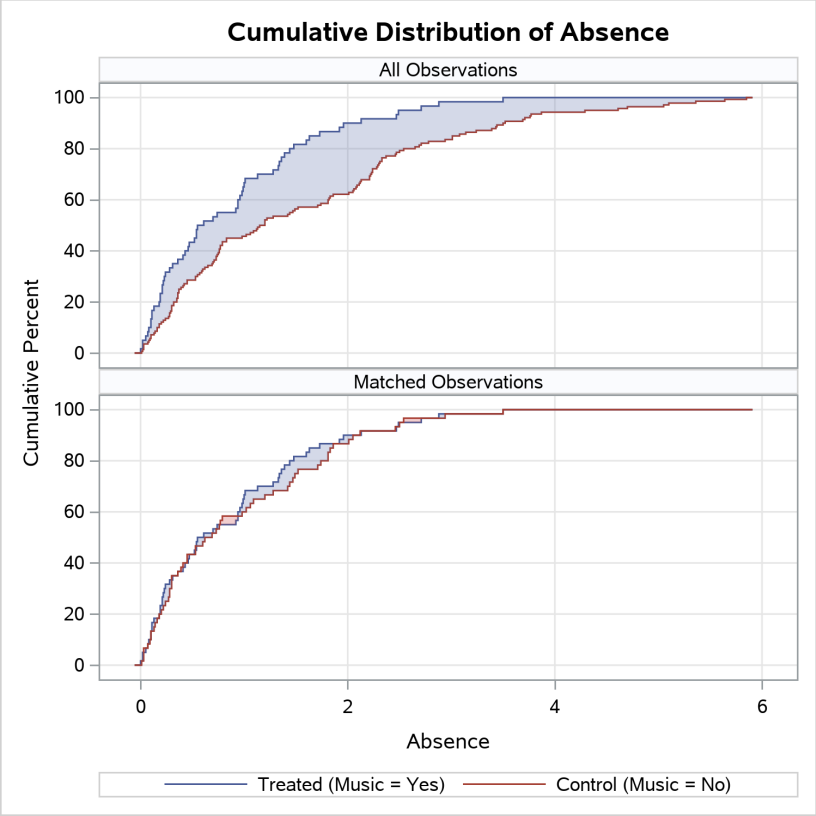
Output 100.4.10 displays a cumulative distribution function (CDF) plot that compares the CDFs of the logit of the propensity score (LPS) for observations in the treated and control groups, based on all observations and on matched observations.

Output 100.4.10 CDF Plot for Logit Propensity Score



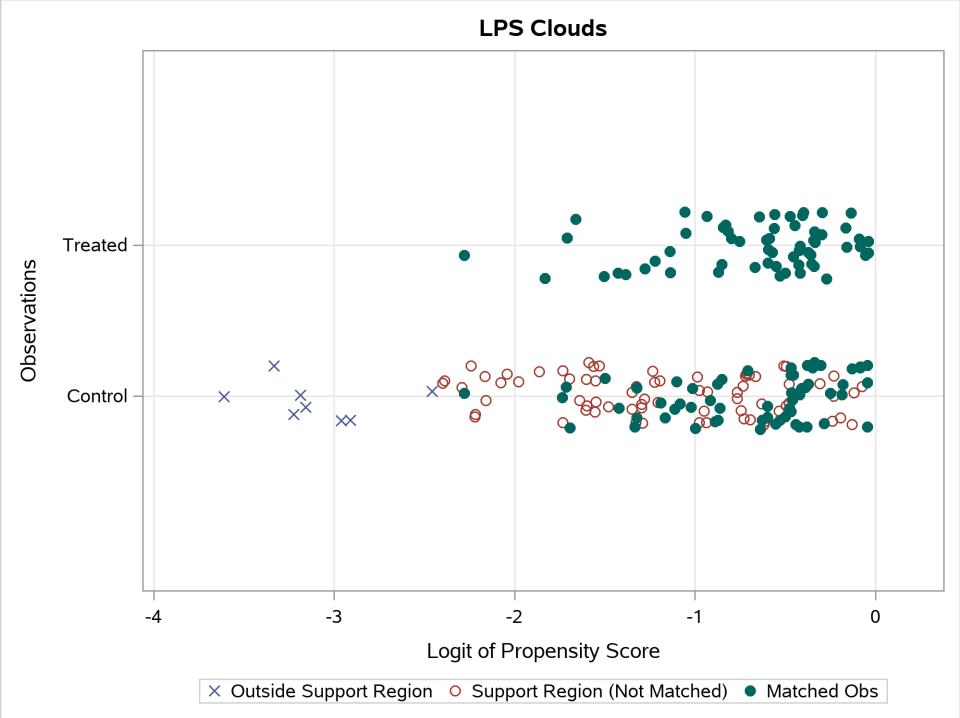
Output 100.4.11 displays a plot that compares the CDFs of Absence for observations in the treated and control groups.

Output 100.4.11 CDF Plot for Absence



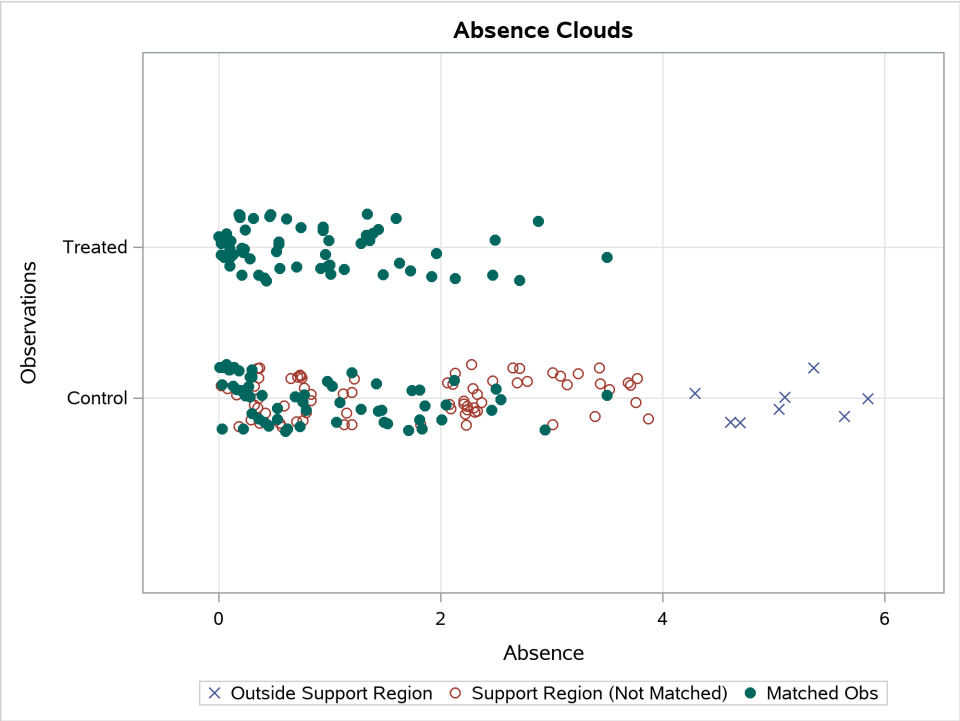
Output 100.4.12 displays a cloud plot that compares the values of the logit of the propensity score (LPS) for observations in the treated and control groups, based on all observations and on matched observations. The points are jittered in the vertical direction to avoid overplotting.

Output 100.4.12 LPS Cloud Plot



Output 100.4.13 displays a cloud plot that compares the values of Absence for observations in the treated and control groups.

Output 100.4.13 Absence Cloud Plot



Note that the NODETAILS option does not apply to the cloud plots.

Because matching results in good balance for the variables in this example, the matched observations can be saved in an output data set for use in a subsequent outcome analysis.

In situations where you are not satisfied with the variable balance, you can do one or more of the following to improve the balance: you can select another set of variables to fit the propensity score model, you can modify the matching criteria, or you can choose another matching method.

The OUT(OBS=MATCH)=OutEx4 option in the OUTPUT statement creates an output data set, OutEx4, that contains the matched observations. The following statements list the observations in the first five matched sets, as shown in Output 100.4.14:

```
proc sort data=OutEx4 out=OutEx4a;
   by _MatchID;
run;

proc print data=OutEx4a(obs=10);
   var StudentID Music Gender Absence _PS_ _MATCHWGT_ _MatchID;
run;
```

Output 100.4.14 Output Data Set with Matching Numbers

Obs	StudentID	Music	Gender	Absence	_PS_	_MATCHWGT_	_MatchID
1	156	Yes	Male	0.02	0.49015	1	1
2	173	No	Male	0.03	0.48874	1	1
3	142	Yes	Male	0.02	0.49015	1	2
4	105	No	Male	0.03	0.48874	1	2
5	64	No	Male	0.03	0.48874	1	3
6	98	Yes	Male	0.05	0.48590	1	3
7	89	No	Male	0.10	0.47883	1	4
8	182	Yes	Male	0.10	0.47883	1	4
9	30	No	Male	0.10	0.47883	1	5
10	115	Yes	Male	0.11	0.47742	1	5

By default, the output data set includes the variable _PS_ (which provides the propensity score) and the variable _MATCHWGT_ (which provides matched observation weights). The MATCHID=_MatchID option creates a variable named _MatchID that identifies the matched sets of observations.

If you assume that no other confounding variables are associated with both the GPA and the music class indicator Music, you can add the GPAs for the students to the data set OutEx4 and perform an outcome analysis on that data set to estimate the effect of music class.

Example 100.5: Outcome Analysis after Matching

This example illustrates how you can carry out an outcome analysis of observations that have been matched as the result of a propensity score analysis.

The data set School, described in Example 100.4, contains information about students that is available at the end of the school year.

The data set Grades contains information about the student grades. StudentID is the student identification number, and GPA is the GPA for the student.

The following statements combine the two data sets and list the 10 observations in the combined SchoolGrades data set, as shown in Output 100.4.14:

```
proc sort data=School out=School1;
   by StudentID;
run;

proc sort data=Grades out=Grades1;
   by StudentID;
run;

data SchoolGrades;
   merge School1 Grades1;
   by StudentID;
run;

proc print data=SchoolGrades(obs=10);
   var StudentID Music Gender Absence GPA;
run;
```

Output 100.5.1 SchoolGrades Data Set

Obs	StudentID	Music	Gender	Absence	GPA
1	1	Yes	Male	1.39	3.99
2	2	No	Female	0.71	3.94
3	3	No	Male	4.29	3.32
4	4	Yes	Female	2.49	3.78
5	5	No	Female	0.02	4.10
6	6	No	Female	0.32	4.12
7	7	No	Female	0.20	4.28
8	8	Yes	Female	0.21	4.40
9	9	No	Female	0.53	3.96
10	10	No	Male	2.78	3.14

For comparison with the outcome analyses that are performed on matched observations later in this example, the following steps perform a *t* test for the effect of music class on GPA using the original (unmatched) data:

```
proc ttest data=SchoolGrades;
   class Music;
   var GPA;
run;
```

The table in Output 100.5.2 shows that the effect of music class is significantly different from 0.

Output 100.5.2 *t* Test for Difference

The TTEST Procedure

Variable: GPA

Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	198	-3.43	0.0007
Satterthwaite	Unequal	148.02	-3.85	0.0002

Although the *t* test shows a significant effect, this effect might be related to the student's gender or absence record. The following regression analysis controls for these effects:

```
proc glm data=SchoolGrades;
   class music(ref='No') gender;
   model GPA= music gender absence / solution;
run;
```

The parameter estimates table in Output 100.5.3 shows that the effect of music class has a *p*-value of 0.1089, which is larger than the *p*-values in Output 100.5.2.

Output 100.5.3 Music Class Effect Estimate

The GLM Procedure

Dependent Variable: GPA

Parameter		Estimate		Standard Error	t Value	Pr > |t|
Intercept		3.903163559	B	0.03913163	99.74	<.0001
Music	**Yes**	0.066558065	B	0.04132470	1.61	0.1089
Music	**No**	0.000000000	B	.	.	.
Gender	**Female**	0.059468770	B	0.03699850	1.61	0.1096
Gender	**Male**	0.000000000	B	.	.	.
Absence		-0.152888978		0.01483823	-10.30	<.0001

However, the regression adjustment requires a sufficient overlap between the covariate distributions for students who took music and students who did not take music. You can ensure a sufficient covariate overlap by performing a propensity score analysis that uses greedy nearest neighbor matching.

The following statements request this analysis:

```
proc psmatch data=School region=treated;
   class Music Gender;
   psmodel Music(Treated='Yes')= Gender Absence;
   match distance=lps method=greedy(k=1) exact=Gender caliper=0.5
         weight=none;
   output out(obs=match)=OutEx4 matchid=_MatchID;
run;
```

These statements are identical to the PROC PSMATCH statements in Example 100.4, except that the ASSESS statement is not used here.

The OUT(OBS=MATCH)=OutEx4 option creates an output data set, OutEx4, that contains the matched observations. Output 100.5.4 displays the observations in the first four matched sets, as shown in Output 100.5.4:

Output 100.5.4 Output Data Set with Matching Numbers

Obs	StudentID	Music	Gender	Absence	_PS_	_MATCHWGT_	_MatchID
1	156	Yes	Male	0.02	0.49015	1	1
2	173	No	Male	0.03	0.48874	1	1
3	142	Yes	Male	0.02	0.49015	1	2
4	105	No	Male	0.03	0.48874	1	2
5	64	No	Male	0.03	0.48874	1	3
6	98	Yes	Male	0.05	0.48590	1	3
7	89	No	Male	0.10	0.47883	1	4
8	182	Yes	Male	0.10	0.47883	1	4

By default, the output data set includes the variable _PS_ (which provides the propensity score) and the variable _MATCHWGT_ (which provides matched observation weights). The MATCHID=_MatchID option creates a variable named _MatchID that identifies the matched sets of observations. Because the data set OutEx4 contains only the matched observations for a one-to-one matching, all the matched observation weights are 1 and can therefore be omitted from the outcome analysis. For more information about the matched observation weights, see the section "Weighting after Matching" on page 8226.

If you assume that no other confounding variables are associated with both GPA and Music, you can add the GPAs for the students to the data set OutEx4 and perform an outcome analysis of GPA on that data set to estimate the effect of the music class. The following statements combine the two data sets:

```
proc sort data=OutEx4 out=OutEx4b;
   by StudentID;
run;

data OutEx4Grades;
   merge OutEx4b Grades1;
   by StudentID;
run;
```

The following statements use a *t* test to estimate the effect of music class from the matched observations:

```
proc ttest data=OutEx4Grades;
   class Music;
   var GPA;
run;
```

The *t* test in Output 100.5.5 has a a large *p*-value of 0.4974, which shows that the effect of the music class is not significant.

Output 100.5.5 *t* Test for Difference

The TTEST Procedure

Variable: GPA

Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	118	-0.68	0.4974
Satterthwaite	Unequal	117.43	-0.68	0.4974

The following regression analysis of the matched observations controls for the effects of gender and absence:

```
proc glm data=OutEx4Grades;
   class music(ref='No') gender;
   model GPA= music gender absence / solution;
run;
```

The "Parameter Estimates" table in Output 100.5.6 shows that the effect of music class has a large *p*-value of 0.5647.

Output 100.5.6 Music Class Effect Estimate

The GLM Procedure

Dependent Variable: GPA

Parameter		Estimate		Standard Error	t Value	Pr > |t|
Intercept		3.919862471	B	0.04782202	81.97	<.0001
Music	**Yes**	0.026056453	B	0.04511875	0.58	0.5647
Music	**No**	0.000000000	B	.	.	.
Gender	**Female**	0.052862308	B	0.04521048	1.17	0.2447
Gender	**Male**	0.000000000	B	.	.	.
Absence		-0.121889379		0.02719279	-4.48	<.0001

In summary, the outcome analyses that are based on matched observations reach a different conclusion than the outcome analyses that are based on the original data.

Example 100.6: Matching with Replacement

This example illustrates how you can perform matching with replacement of observations in a control group with observations in a treatment group, so that the matched observations can be used to estimate the treatment effect in a subsequent outcome analysis.

The data for this example are observations on students in a school. The School data set, which contains information about the students, is described in Example 100.4. The following statements request matching with replacement to match observations for students in the treatment group with observations for students in the control group:

```
ods graphics on;
proc psmatch data=School region=allobs(psmin=0.05);
   class Music Gender;
   psmodel Music(Treated='Yes')= Gender Absence;
   match method=replace(k=1) distance=ps exact=Gender caliper=.
         nmatchmost=6;
   assess ps var=(Gender Absence);
   output out(obs=match)=outex6 matchid=_MatchID;
run;
```

The PSMODEL statement specifies the logistic regression model that creates the propensity score for each student, which is the probability that the student enrolled in a music class. The CLASS statement specifies the classification variables. The Music variable is the binary treatment indicator variable, and TREATED='Yes' identifies Yes as the treated group. The Gender and Absence variables are included in the model because they are believed to be related to enrolling in the music class.

The PSMATCH procedure matches only those observations whose propensity scores lie in the support region that you specify with the REGION= option. Here the REGION=ALLOBS(PSMIN=0.05) option requests that all available observations whose propensity scores are greater than or equal to 0.05 be used for matching.

The MATCH statement requests matching and specifies the criteria for matching. The DISTANCE=PS option requests that the propensity score be used to compute differences between pairs of observations. The METHOD=REPLACE(K=1) option requests matching with replacement in which each treated unit is matched to the closest control unit.

The EXACT=GENDER option requests that the treated unit and its matched control unit have the same value of the Gender variable. The CALIPER=. option ignores the caliper requirement for matching.

The "Data Information" table in Output 100.6.1 displays the numbers of observations in the treated and control groups, the lower and upper limits of the propensity scores for observations in the support region, and the numbers of observations in the treated and control groups that fall within the support region. Of the 140 observations in the control group, 134 fall within the support region.

Output 100.6.1 Data Information

The PSMATCH Procedure

Data Information	
Data Set	WORK.SCHOOL
Output Data Set	WORK.OUTEX6
Treatment Variable	Music
Treated Group	Yes
All Obs (Treated)	60
All Obs (Control)	140
Support Region	PS Bounded Obs
Lower PS Support	0.05
Upper PS Support	0.490152
Support Region Obs (Treated)	60
Support Region Obs (Control)	134

The "Propensity Score Information" table in Output 100.6.2 displays summary statistics by treatment group for all observations, for observations in the support region, and for matched observations. Because the default

WEIGHT=ATTWGT is specified in the PSWEIGHT statement, the table also displays summary statistics for weighted matched observations.

Output 100.6.2 Propensity Score Information

Propensity Score Information												
	Treated (Music = Yes)						Control (Music = No)					
Observations	N	Weight	Mean	Standard Deviation	Minimum	Maximum	N	Weight	Mean	Standard Deviation	Minimum	Maximum
All	60		0.3472	0.0963	0.0928	0.4902	140		0.2798	0.1251	0.0263	0.4887
Region	60		0.3472	0.0963	0.0928	0.4902	134		0.2906	0.1166	0.0517	0.4887
Matched	60		0.3472	0.0963	0.0928	0.4902	41		0.3350	0.1036	0.0928	0.4887
Weighted Matched	60	60.00	0.3472	0.0963	0.0928	0.4902	41	60.00	0.3458	0.0959	0.0928	0.4887

Propensity Score Information	
Observations	Treated - Control Mean Difference
All	0.0675
Region	0.0567
Matched	0.0122
Weighted Matched	0.0014

Note that the number of matched control units (41) is less than the number of matched treated units (60). When matching is done with replacement, a control unit can be matched with more than one treated unit.

The "Matching Information" table in Output 100.6.3 displays the matching criteria, the number of matched sets, the numbers of matched observations in the treated and control groups, and the total absolute difference in the propensity scores for all matches. In this example, 41 control units are matched to 60 treated units.

Output 100.6.3 Matching Information

Matching Information	
Distance Metric	Propensity Score
Method	Replacement Matching
Control/Treated Ratio	1
Matched Sets	41
Matched Obs (Treated)	60
Matched Obs (Control)	41
Total Absolute Difference	0.263828

The ASSESS statement produces tables and plots that summarize differences in the distributions of the specified variables between treated and control groups for all observations, for observations in the support region, and for matched observations. You can use these results to assess how well the matching achieves a balance in the distributions of these variables. As requested by the PS and VAR= options, the variables are the propensity score and the covariates Gender and Absence.

The "Standardized Mean Differences" table displays standardized mean differences in the variables between the treated and control groups, as shown in Output 100.6.4.

Output 100.6.4 Standardized Mean Differences

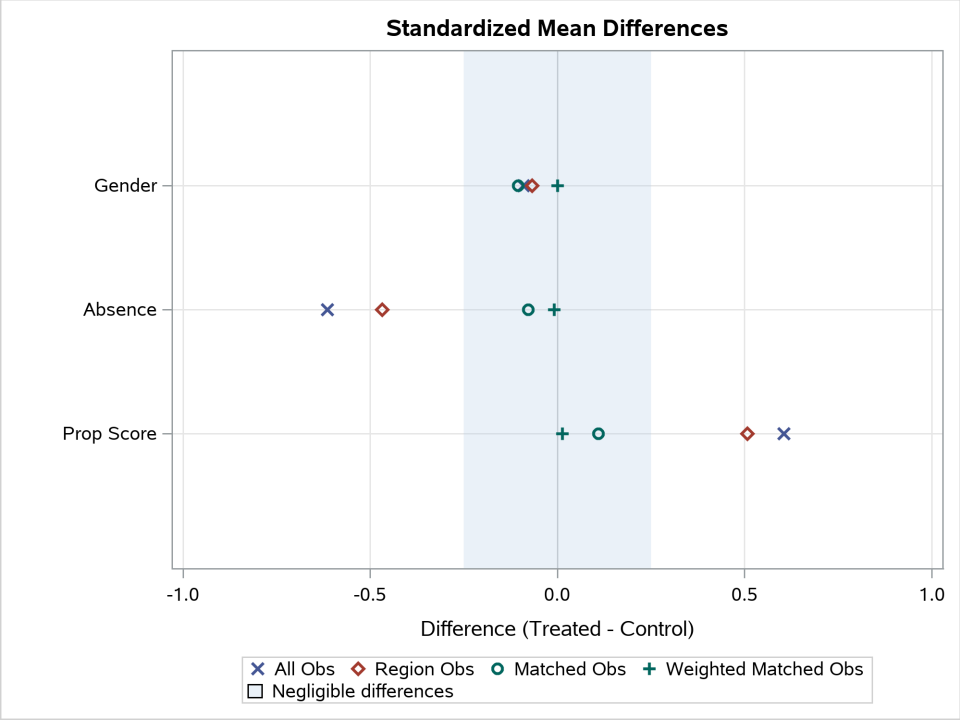
The PSMATCH Procedure

Standardized Mean Differences (Treated - Control)						
Variable	Observations	Mean Difference	Standard Deviation	Standardized Difference	Percent Reduction	Variance Ratio
Prop Score	All	0.06749	0.111638	0.60451		0.5930
	Region	0.05667		0.50759	16.03	0.6824
	Matched	0.01220		0.10926	81.93	0.8652
	Weighted Matched	0.00141		0.01261	97.91	1.0083
Absence	All	-0.69807	1.136767	-0.61409		0.3550
	Region	-0.53223		-0.46820	23.76	0.5052
	Matched	-0.08899		-0.07828	87.25	0.8202
	Weighted Matched	-0.01033		-0.00909	98.52	1.0049
Gender	All	-0.04524	0.496344	-0.09114		1.0208
	Region	-0.03383		-0.06816	25.22	1.0138
	Matched	-0.05203		-0.10483	0.00	1.0254
	Weighted Matched	0.00000		0.00000	100.00	1.0000
Standard deviation of All observations used to compute standardized differences						

Note that a zero percentage reduction is displayed for Gender in the matched observation because its standardized mean difference (0.10483, in absolute value) is larger than the standardized mean difference of all observations (0.09114 in absolute value).

The PSMATCH procedure displays a standardized differences plot, shown in Output 100.6.5, for the variables that are specified in the ASSESS statement.

Output 100.6.5 Standardized Mean Differences Plot



All differences for the matched observations are within the recommended limits of –0.25 and 0.25, which are indicated by the shaded area.

The NMATCHMOST=6 option requests a table of the six observations in the control group that have the most matches, which is shown in Output 100.6.6. The table does not display the observations that have the most matches in the treated group because each treated unit is matched only once (K=1).

Output 100.6.6 Observations with the Most Matches

Observations with Most Matches	
Control (Music = No)	
Observation	Matched Treated
124	4
99	3
123	3
113	2
140	2
101	2

Because matching results in good balance for the variables in this example, the matched observations can be saved in an output data set for use in a subsequent outcome analysis.

In situations where you are not satisfied with the variable balance, you can do one or more of the following to improve the balance: you can select another set of variables to fit the propensity score model, you can modify the matching criteria, or you can choose another matching method.

The OUT(OBS=MATCH)=OutEx6 option in the OUTPUT statement creates an output data set, OutEx6, that contains the matched observations. The following statements list the observations in the last four matched sets, which are shown in Output 100.6.7.

```
proc sort data=OutEx6 out=OutEx6a;
   by descending _MatchID;
run;

proc print data=OutEx6a(obs=10);
   var StudentID Music Gender Absence _PS_ _MATCHWGT_ _MatchID;
run;
```

Output 100.6.7 Output Data Set of Matched Observations with Replacement

Obs	StudentID	Music	Gender	Absence	_PS_	_MATCHWGT_	_MatchID
1	156	Yes	Male	0.02	0.49015	1	41
2	142	Yes	Male	0.02	0.49015	1	41
3	173	No	Male	0.03	0.48874	2	41
4	64	No	Male	0.03	0.48874	1	40
5	98	Yes	Male	0.05	0.48590	1	40
6	89	No	Male	0.10	0.47883	2	39
7	115	Yes	Male	0.11	0.47742	1	39
8	182	Yes	Male	0.10	0.47883	1	39
9	130	No	Male	0.18	0.46753	1	38
10	104	Yes	Male	0.19	0.46612	1	38

By default, the output data set includes the variable _PS_ (which provides the propensity score) and the variable _MATCHWGT_ (which provides matched observation weights). The MATCHID=_MatchID option creates a variable named _MatchID that identifies the matched sets of observations.

If you assume that no other confounding variables are associated with both the GPA and the music class indicator Music, you can add the GPAs for the students to the data set OutEx6 and perform an outcome analysis of GPA on this data set to estimate the music class effect.

Example 100.7: Mahalanobis Distance Matching

This example illustrates how you can perform Mahalanobis distance matching of observations in a control group with observations in a treatment group, so that the matched observations can be used to estimate the treatment effect in a subsequent outcome analysis. The outcome analysis itself is not shown here.

The data for this example are observations on patients in a nonrandomized clinical trial. The trial and the Drugs data set that contains the patient information are described in the section "Getting Started: PSMATCH Procedure" on page 8184.

The following statements request optimal matching based on Mahalanobis distances between patients in the treatment group and patients in the control group:

```
ods graphics on;
proc psmatch data=drugs region=cs;
   class Drug Gender;
   psmodel Drug(Treated='Drug_X')= Gender Age BMI;
   match method=optimal(k=1) exact=Gender
         distance=mah(lps var=(Age BMI)) caliper=.
         weight=none;
   assess lps var=(Gender Age BMI);
   output out(obs=match)=OutEx7 matchid=_MatchID;
run;
```

The PSMODEL statement specifies the logistic regression model that creates the propensity score for each observation, which is the probability that the patient receives Drug_X. The CLASS statement specifies the classification variables in the model. The Drug variable is the binary treatment indicator variable, and TREATED='Drug_X' identifies Drug_X as the treated group. The Gender, Age, and BMI variables are included in the model because they are believed to be related to the assignment.

The PSMATCH procedure matches only those observations whose propensity scores lie in the support region that you specify with the REGION= option. Here the REGION=CS option requests that only those observations whose propensity scores (or equivalently, logits of propensity scores) lie in the common support region be used for matching. The common support region is the largest interval that contains propensity scores (or equivalently, logits of propensity scores) for both treated and control observations. By default, the region is extended by 0.25 times the pooled estimate of the common standard deviation of the logits of the propensity scores.

The MATCH statement specifies the criteria for matching. The DISTANCE=MAH(LPS VAR=(AGE BMI)) option requests that Mahalanobis distance be used in computing differences between pairs of observations, and that this distance be derived from the logit of the propensity score and the Age and BMI variables. The METHOD=OPTIMAL(K=1) option (which is the default) requests optimal matching of one control unit to each unit in the treated group in order to minimize the total within-pair difference. The EXACT=GENDER

option requests that the treated unit and its matched control unit have the same value of Gender. The CALIPER=. option ignores the caliper requirement for matching.

The "Data Information" table in Output 100.7.1 displays the numbers of observations in the treated and control groups, the lower and upper limits of propensity scores for observations in the support region, and the numbers of observations in the treated and control groups that fall within the support region. Of the 373 observations in the control group, 351 fall within the support region.

Output 100.7.1 Data Information

The PSMATCH Procedure

Data Information	
Data Set	WORK.DRUGS
Output Data Set	WORK.OUTEX7
Treatment Variable	Drug
Treated Group	Drug_X
All Obs (Treated)	113
All Obs (Control)	373
Support Region	Extended Common Support
Lower PS Support	0.050244
Upper PS Support	0.683999
Support Region Obs (Treated)	113
Support Region Obs (Control)	351

The "Propensity Score Information" table in Output 100.7.2 displays summary statistics by treatment group for all observations, for observations in the support region, and for matched observations.

Output 100.7.2 Propensity Score Information

Propensity Score Information

	Treated (Drug = Drug_X)					Control (Drug = Drug_A)					Treated - Control
Observations	**N**	**Mean**	**Standard Deviation**	**Minimum**	**Maximum**	**N**	**Mean**	**Standard Deviation**	**Minimum**	**Maximum**	**Mean Difference**
All	113	0.3108	0.1325	0.0602	0.6411	373	0.2088	0.1320	0.0202	0.6858	0.1020
Region	113	0.3108	0.1325	0.0602	0.6411	351	0.2176	0.1267	0.0510	0.6824	0.0932
Matched	113	0.3108	0.1325	0.0602	0.6411	113	0.3053	0.1337	0.0640	0.6824	0.0055

The "Matching Information" table in Output 100.7.3 displays the matching criteria, the number of matched sets, the numbers of matched observations in the treated and control groups, and the total Mahalanobis difference for all matches.

Output 100.7.3 Matching Information

Matching Information	
Distance Metric	Mahalanobis Distance
Mahalanobis Covariance	Control Group
Method	Optimal Fixed Ratio Matching
Control/Treated Ratio	1
Matched Sets	113
Matched Obs (Treated)	113
Matched Obs (Control)	113
Total Absolute Difference	20.05243

The ASSESS statement produces tables and plots that summarize differences in the distributions of the specified variables between treated and control groups for all observations, for observations in the support region, and for matched observations. As requested by the LPS and VAR= options, the variables that are listed in the table are the logit of the propensity score and the variables Gender, Age, and BMI. The WEIGHT=NONE option suppresses the display of differences for the weighted matched observations. Note that when one control unit is matched with each treated unit, the weights are all 1 for matched treated and control units, and the results are identical for the weighted matched observations and the matched observations.

The "Standardized Mean Differences" table, shown in Output 100.7.4, displays standardized mean differences in the variables between the treated and control groups for all observations, for observations in the support region, and for matched observations. For a binary classification variable (Gender), the difference is in the proportion of the first ordered level (Female).

Output 100.7.4 Standardized Mean Differences

The PSMATCH Procedure

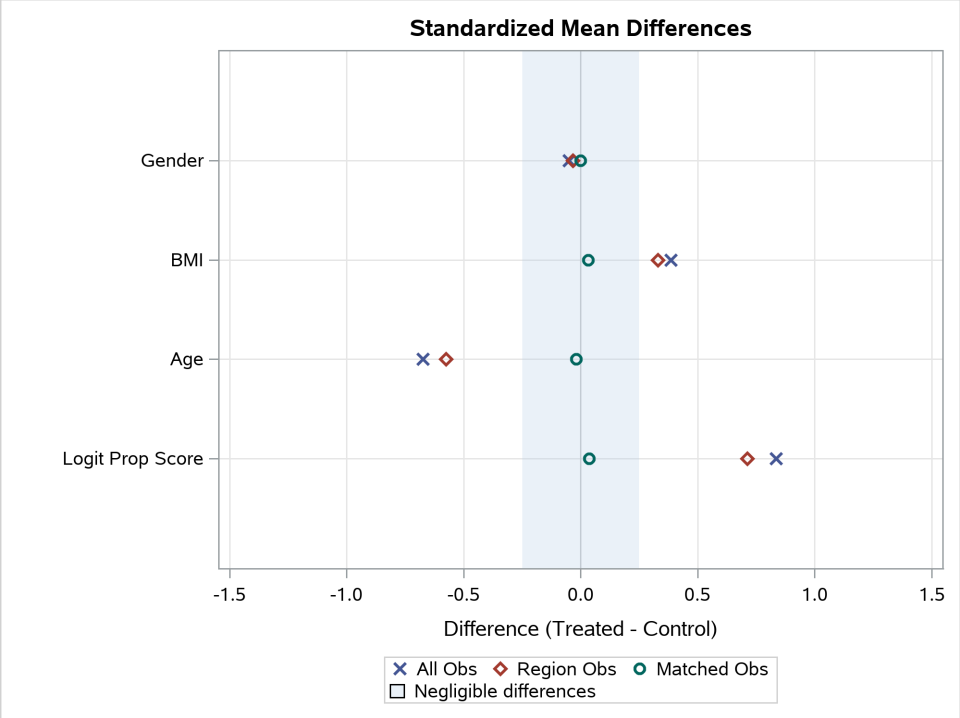
Standardized Mean Differences (Treated - Control)						
Variable	**Observations**	**Mean Difference**	**Standard Deviation**	**Standardized Difference**	**Percent Reduction**	**Variance Ratio**
Logit Prop Score	**All**	0.63997	0.767449	0.83389		0.6517
	Region	0.54546		0.71074	14.77	0.8314
	Matched	0.02757		0.03592	95.69	0.9813
Age	**All**	-4.09509	6.079104	-0.67363		0.7076
	Region	-3.49368		-0.57470	14.69	0.8000
	Matched	-0.11504		-0.01892	97.19	1.0007
BMI	**All**	0.73930	1.923178	0.38441		0.8854
	Region	0.63257		0.32892	14.44	0.9288
	Matched	0.06186		0.03216	91.63	1.0176
Gender	**All**	-0.02482	0.496925	-0.04994		0.9892
	Region	-0.01651		-0.03323	33.46	0.9922
	Matched	0.00000		0.00000	100.00	1.0000
Standard deviation of All observations used to compute standardized differences						

The standardized mean differences are significantly reduced in the matched observations. The largest of these differences is 0.03592 in absolute value, which is less than the recommended upper limit of 0.25. The treated-to-control variance ratios between the two groups are also in the recommended range of 0.5 to 2.

Because EXACT=GENDER is specified in the MATCH statement, the standardized mean difference for Gender is 0 in the matched observations.

The PSMATCH procedure displays a standardized mean differences plot, shown in Output 100.7.5, for the variables that are specified in the ASSESS statement.

Output 100.7.5 Standardized Mean Differences Plot



The plot displays the standardized mean differences that are listed in the "Standardized Mean Differences" table in Output 100.7.4. All differences for the matched observations are within the recommended limits of –0.25 and 0.25, which are indicated by the shaded area.

Because matching results in good balance for the variables in this example, the matched observations can be saved in an output data set for use in a subsequent outcome analysis.

In situations where you are not satisfied with the variable balance, you can do one or more of the following to improve the balance: you can select another set of variables to fit the propensity score model, you can modify the matching criteria, or you can choose another matching method.

The OUT(OBS=MATCH)=OutEx7 option in the OUTPUT statement creates an output data set, OutEx7, that contains the matched observations. The following statements list the observations in the first four matched sets, which are shown in Output 100.7.6:

```
proc sort data=OutEx7 out=OutEx7a;
   by _MatchID;
run;

proc print data=OutEx7a(obs=8);
    var PatientID Drug Gender Age BMI _ps_ _MatchWgt_ _MatchID;
run;
```

Output 100.7.6 Output Data Set with Mahalanobis Distance Matches

Obs	PatientID	Drug	Gender	Age	BMI	_PS_	_MATCHWGT_	_MatchID
1	141	Drug_A	Female	43	20.55	0.064010	1	1
2	89	Drug_X	Female	44	20.75	0.060231	1	1
3	137	Drug_A	Female	45	22.04	0.072149	1	2
4	323	Drug_X	Female	46	22.22	0.067625	1	2
5	429	Drug_A	Male	49	24.00	0.088477	1	3
6	217	Drug_X	Male	49	23.96	0.087716	1	3
7	111	Drug_A	Female	41	21.01	0.087140	1	4
8	234	Drug_X	Female	41	21.11	0.089042	1	4

By default, the output data set includes the variable _PS_ (which provides the propensity score) and the variable _MATCHWGT_ (which provides matched observation weights). The weight for each treated unit is 1. Because K=1 is specified in the METHOD= option in the MATCH statement, one control unit is matched to each treated unit, and so the weight for each matched control unit is also 1. The MATCHID=_MatchID option creates a variable named _MatchID that identifies the matched sets of observations.

After the responses for the trial are observed, they can be added to the data set OutEx7 as the starting point for an outcome analysis. Assuming that no other confounding variables are associated with both the response variable and the treatment group indicator Drug, you can estimate the treatment effect from the matched observations by performing an outcome analysis that you would have used to estimate the treatment effect if the original data set had resulted from a randomized trial.

Example 100.8: Matching with Precomputed Propensity Scores

The PSMATCH procedure provides the capability for fitting a binary logistic regression model that is used to compute propensity scores for matching. However, there might be situations in which you have already computed the propensity scores—for example, by using other procedures in SAS/STAT software that perform logistic regression. This example illustrates optimal matching with precomputed propensity scores that are provided in the input data set for PROC PSMATCH.

The data for this example are observations on patients in a nonrandomized clinical trial. The trial and the Drugs data set that contains the patient information are described in the section "Getting Started: PSMATCH Procedure" on page 8184.

The following statements use the LOGISTIC procedure to derive propensity scores:

```
ods select none;
proc logistic data=drugs;
   class Drug Gender;
   model Drug(Event='Drug_X')= Gender Age BMI / link=cloglog;
   output out=drug1 p=pscore;
run;
ods select all;
```

The LINK=CLOGLOG option fits the complementary log-log model and derives propensity scores that are used in the PSMATCH procedure. The option is used just to demonstrate that, other than the logit link that is provided in the PSMATCH procedure, you can use a different model to derive propensity scores and then input these propensity scores in the PSMATCH procedure.

The output data set Drug1 is constructed from the data set Drugs and contains the PScore variable for propensity scores.

Output 100.8.1 lists the first 10 observations.

Output 100.8.1 Data Set with Propensity Scores

Obs	PatientID	Drug	Gender	Age	BMI	pscore
1	284	Drug_X	Male	29	22.02	0.35498
2	201	Drug_A	Male	45	26.68	0.21794
3	147	Drug_A	Male	42	21.84	0.12261
4	307	Drug_X	Male	38	22.71	0.19821
5	433	Drug_A	Male	31	22.76	0.34298
6	435	Drug_A	Male	43	26.86	0.26261
7	159	Drug_A	Female	45	25.47	0.15077
8	368	Drug_A	Female	49	24.28	0.08713
9	286	Drug_A	Male	31	23.31	0.37211
10	163	Drug_X	Female	39	25.34	0.24005

The following statements request optimal matching to match patients in the treatment group to patients in the control group:

```
ods graphics on;
proc psmatch data=Drug1 region=cs;
   class Drug Gender;
   psdata treatvar=Drug(Treated='Drug_X') ps=pscore;
   match method=optimal(k=1) exact=Gender distance=lps caliper=0.5
         weight=none;
   assess lps var=(Gender Age BMI);
   output out(obs=match)=OutEx8 lps=_Lps matchid=_MatchID;
run;
```

The PSMODEL statement is not used in this example because the propensity scores are provided in Drug1. Instead, the PSDATA statement is used to identify the binary treatment variable and the propensity score variable in Drug1. The CLASS statement specifies the classification variables. The PS= option specifies pscore as the propensity score variable. The TREATVAR=DRUG option specifies Drug as the binary treatment indicator variable, and TREATED='Drug_X' identifies Drug_X as the treated group.

The PSMATCH procedure matches only those observations whose propensity scores lie in the support region that you specify with the REGION= option. Here the REGION=CS option requests that only those observations whose propensity scores (or equivalently, logits of propensity scores) lie in the common support region be used for matching. The common support region is the largest interval that contains propensity scores (or equivalently, logits of propensity scores) for both treated and control observations. By default, the region is extended by 0.25 times the pooled estimate of the common standard deviation of the logits of the propensity scores.

The MATCH statement specifies the criteria for matching. The DISTANCE=LPS option (which is the default) requests that the logit of the propensity score be used in computing differences between pairs of observations. The METHOD=OPTIMAL(K=1) option (which is the default) requests optimal matching of one control unit to each unit in the treated group in order to minimize the total within-pair difference. The EXACT=GENDER option requests that the treated unit and its matched control unit have the same value of the Gender variable. The CALIPER=0.5 option requests that a match be made only if the difference in the logits of the propensity

scores for pairs of individuals is less than or equal to 0.5 times the pooled estimate of the common standard deviation of the logits of the propensity scores.

The "Data Information" table in Output 100.8.2 displays the numbers of observations in the treated and control groups, the lower and upper limits for the propensity scores of observations in the support region, and the numbers of observations in the treated and control groups that fall within the support region. Of the 373 observations in the control group, 352 fall within the support region.

Output 100.8.2 Data Information

The PSMATCH Procedure

Data Information	
Data Set	WORK.DRUG1
Output Data Set	WORK.OUTEX8
Treatment Variable	Drug
Treated Group	Drug_X
All Obs (Treated)	113
All Obs (Control)	373
Support Region	Extended Common Support
Lower PS Support	0.060563
Upper PS Support	0.698199
Support Region Obs (Treated)	113
Support Region Obs (Control)	352

The "Propensity Score Information" table in Output 100.8.3 displays summary statistics by treatment group for all observations, for observations in the support region, and for matched observations.

Output 100.8.3 Propensity Score Information

Propensity Score Information

	Treated (Drug = Drug_X)					Control (Drug = Drug_A)					Treated - Control
Observations	**N**	**Mean**	**Standard Deviation**	**Minimum**	**Maximum**	**N**	**Mean**	**Standard Deviation**	**Minimum**	**Maximum**	**Mean Difference**
All	113	0.3040	0.1287	0.0715	0.6594	373	0.2089	0.1255	0.0295	0.7135	0.0952
Region	113	0.3040	0.1287	0.0715	0.6594	352	0.2146	0.1177	0.0606	0.6519	0.0894
Matched	113	0.3040	0.1287	0.0715	0.6594	113	0.2984	0.1215	0.0723	0.6519	0.0056

The "Matching Information" table in Output 100.8.4 displays the matching criteria, the number of matched sets, the numbers of matched observations in the treated and control groups, and the total absolute difference in the logits of the propensity scores for all matches.

Output 100.8.4 Matching Information

Matching Information	
Distance Metric	Logit of Propensity Score
Method	Optimal Fixed Ratio Matching
Control/Treated Ratio	1
Caliper (Logit PS)	0.356051
Matched Sets	113
Matched Obs (Treated)	113
Matched Obs (Control)	113
Total Absolute Difference	3.616259

The ASSESS statement produces tables and plots that summarize differences in the distributions of the specified variables between treated and control groups for all observations, for observations in the support region, and for matched observations. As requested by the LPS and VAR= options, the variables that are listed in the table are the logit of the propensity score and the variables Gender, Age, and BMI. The WEIGHT=NONE option suppresses the display of differences for the weighted matched observations. When one control unit is matched to each treated unit, the weights are all 1 for matched treated and control units, so the results for weighted matched observations and matched observations are identical.

The "Standardized Mean Differences" table displays standardized mean differences in the variables between the treated and control groups. For a binary classification variable (Gender), the difference is in the proportion of the first ordered level (Female).

Output 100.8.5 Standardized Mean Differences

The PSMATCH Procedure

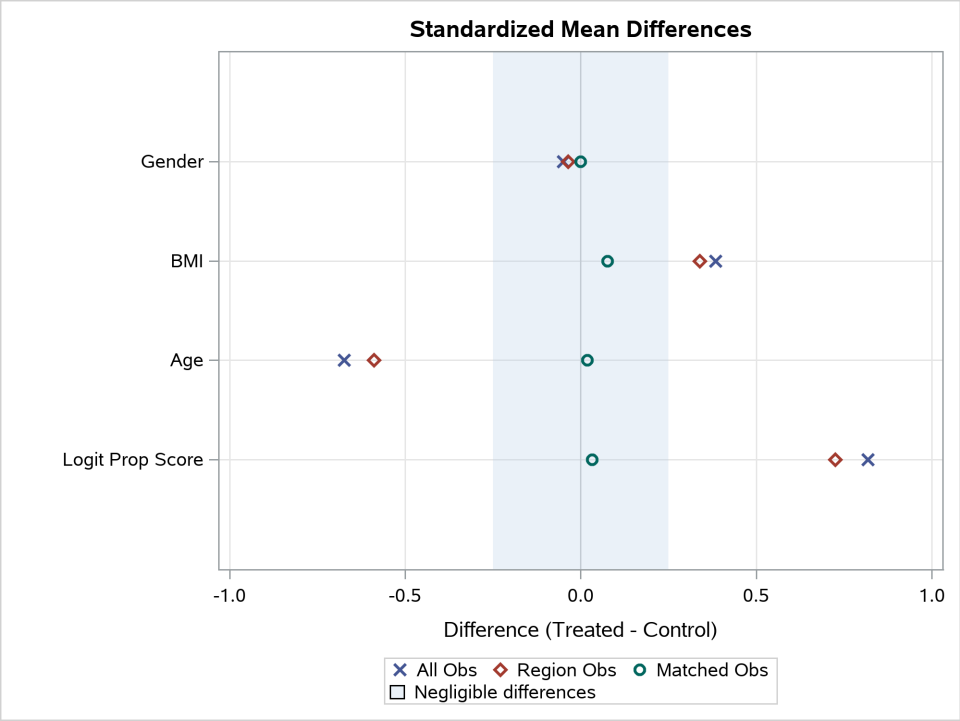
Standardized Mean Differences (Treated - Control)						
Variable	**Observations**	**Mean Difference**	**Standard Deviation**	**Standardized Difference**	**Percent Reduction**	**Variance Ratio**
Logit Prop Score	**All**	0.58239	0.712102	0.81785		0.7177
	Region	0.51613		0.72480	11.38	0.9052
	Matched	0.02268		0.03184	96.11	1.0929
Age	**All**	-4.09509	6.079104	-0.67363		0.7076
	Region	-3.58515		-0.58975	12.45	0.7928
	Matched	0.11504		0.01892	97.19	1.0143
BMI	**All**	0.73930	1.923178	0.38441		0.8854
	Region	0.65089		0.33845	11.96	0.9394
	Matched	0.14619		0.07602	80.23	1.3509
Gender	**All**	-0.02482	0.496925	-0.04994		0.9892
	Region	-0.01808		-0.03638	27.16	0.9916
	Matched	0.00000		0.00000	100.00	1.0000
Standard deviation of All observations used to compute standardized differences						

The standardized mean differences are significantly reduced in the matched observations, and the largest of these differences is 0.076 in absolute value, which is less than the recommended upper limit of 0.25. The treated-to-control variance ratios between the two groups are between 1 and 1.3509 for all variables in the matched observations, which is within the recommended range of 0.5 to 2. Because both EXACT=GENDER and METHOD=OPTIMAL are specified in the MATCH statement, the standardized mean difference for

Gender is 0 in the matched observations.

The PSMATCH procedure displays a standardized mean differences plot, as shown in Output 100.8.6, for the variables that are specified in the ASSESS statement.

Output 100.8.6 Standardized Mean Differences Plot



The "Standardized Mean Differences Plot" displays the standardized mean differences that are listed in the "Standardized Mean Differences" table in Output 100.8.5. All differences for the matched observations are within the recommended limits of –0.25 and 0.25, which are indicated by the shaded area.

Because matching results in good balance for the variables in this example, the matched observations can be saved in an output data set for use in a subsequent outcome analysis.

In situations where you are not satisfied with the variable balance, you can do one or more of the following to improve the balance: you can select another set of variables to fit the propensity score model, you can modify the matching criteria, or you can choose another matching method.

The OUT(OBS=MATCH)=OutEx8 option in the OUTPUT statement creates an output data set, OutEx8, that contains the matched observations. The following statements list the observations in the first five matched sets, as shown in Output 100.8.7:

```
proc sort data=OutEx8 out=OutEx8a;
   by _MatchID;
run;

proc print data=OutEx8a(obs=10);
   var PatientID Drug Gender Age BMI pscore _LPS _MatchWgt_ _MatchID;
run;
```

Output 100.8.7 Output Data Set With Optimal Matches

Obs	PatientID	Drug	Gender	Age	BMI	pscore	_Lps	_MATCHWGT_	_MatchID
1	213	Drug_A	Female	49	23.24	0.07234	-2.55123	1	1
2	89	Drug_X	Female	44	20.75	0.07152	-2.56356	1	1
3	245	Drug_A	Female	52	25.32	0.08090	-2.43015	1	2
4	323	Drug_X	Female	46	22.22	0.07822	-2.46677	1	2
5	429	Drug_A	Male	49	24.00	0.09865	-2.21228	1	3
6	217	Drug_X	Male	49	23.96	0.09796	-2.22013	1	3
7	234	Drug_X	Female	41	21.11	0.09887	-2.20987	1	4
8	66	Drug_A	Female	48	24.53	0.09927	-2.20531	1	4
9	183	Drug_A	Female	45	23.62	0.10931	-2.09786	1	5
10	320	Drug_X	Female	46	24.17	0.11056	-2.08507	1	5

By default, the output data set includes the variable _PS_ (which provides the propensity score) and the variable _MATCHWGT_ (which provides matched observation weights). The weight for each treated unit is 1. Because K=1 is specified in the METHOD=OPTIMAL option in the MATCH statement, one control unit is matched to each treated unit, so the weight for each matched control unit is also 1. The LPS=_LPS option creates a variable named _LPS (which provides the logit of the propensity score) and the MATCHID=_MatchID option creates a variable named _MatchID (which identifies the matched sets of observations).

After the responses for the trial are observed, they can be added to the data set OutEx8 as the starting point for an outcome analysis. Assuming that no other confounding variables are associated with both the response variable and the treatment group indicator Drug, you can estimate the treatment effect from the matched observations by performing an outcome analysis that you would have used to estimate the treatment effect if the original data set had resulted from a randomized trial.

Example 100.9: Sensitivity Analysis after One-to-One Matching

This example illustrates how you can analyze sensitivity to the assumption of no unobserved confounders after performing one-to-one matching with the PSMATCH procedure. For a detailed description of this analysis, see the section "Sensitivity Analysis" on page 8231.

A pharmaceutical company conducts a nonrandomized clinical trial to demonstrate the efficacy of a new treatment (Drug_X) to decrease the low-density lipoprotein (LDL) by comparing it to an existing treatment (Drug_A). The data set Drugs, which is described in "Getting Started: PSMATCH Procedure" on page 8184, contains baseline variable measurements for individuals from the treated and control groups.

Output 100.9.1 lists the first eight observations.

Output 100.9.1 Input Drugs Data Set

Obs	PatientID	Drug	Gender	Age	BMI
1	1	Drug_X	Male	29	22.02
2	2	Drug_A	Male	45	26.68
3	3	Drug_A	Male	42	21.84
4	4	Drug_X	Male	38	22.71
5	5	Drug_A	Male	31	22.76
6	6	Drug_A	Male	43	26.86
7	7	Drug_A	Female	45	25.47
8	8	Drug_A	Female	49	24.28

The possibility of treatment selection bias is a concern in the analysis of the results. Patients in the trial can choose the treatment that they prefer; otherwise, physicians assign each patient to a treatment. This could lead to systematic differences in the distributions of the baseline variables in the two groups, resulting in a biased estimate of the treatment effect. Propensity score analysis that is based on matching offers an alternative that addresses this problem by balancing the distributions of the variables.

The following statements request optimal matching of observations for patients in the treatment group with observations for patients in the control group:

```
proc psmatch data=drugs region=cs;
   class Drug Gender;
   psmodel Drug(Treated='Drug_X')= Gender Age BMI;
   match method=optimal(k=1) exact=Gender distance=lps caliper=0.25
         weight=none;
   output out(obs=match)=Outgs lps=_Lps matchid=_MatchID;
run;
```

The statements are identical to those in "Getting Started: PSMATCH Procedure" on page 8184, except that the ASSESS statement is not used here. The MATCH statement requests optimal matching of one control unit to each unit in the treated group in order to minimize the total within-pair difference.

The OUT(OBS=MATCH)=Outgs option in the OUTPUT statement creates an output data set, Outgs, that contains the matched observations.

After the trial, the data set Cholesterol contains the LDL information for the matched observations. PatientID is the patient identification number, and the response variable LDL is the decrease in LDL, measured in milligrams per deciliter of blood (mg/dl).

The following statements combine the two data sets and list the eight observations in the combined Cholesterol data set, which are shown in Output 100.9.2:

```
proc sort data=Outgs out=Outgs1;
   by PatientID;
run;

proc sort data=Cholesterol out=Cholesterol1;
   by PatientID;
run;

data OutEx9a;
   merge Outgs1 Cholesterol1;
```

```
   by PatientID;
run;

proc print data=OutEx9a(obs=8);
   var PatientID Drug Gender Age BMI LDL _MatchID;
run;
```

Output 100.9.2 Output Data Set with LDL Decreases

Obs	PatientID	Drug	Gender	Age	BMI	LDL	_MatchID
1	1	Drug_X	Male	29	22.02	6.54	74
2	3	Drug_A	Male	42	21.84	-5.66	7
3	4	Drug_X	Male	38	22.71	5.52	24
4	5	Drug_A	Male	31	22.76	7.26	76
5	9	Drug_A	Male	31	23.31	2.64	82
6	10	Drug_X	Female	39	25.34	4.77	43
7	13	Drug_X	Female	32	24.78	4.25	84
8	18	Drug_X	Male	34	26.30	0.68	99

The following statements compute the differences in LDL between the treated and control units in each matched set:

```
proc sort data=OutEx9a out=OutEx9b;
   by _MatchID Drug;
run;

proc transpose data=OutEx9b out=OutEx9c;
   by _MatchID;
   var LDL;
run;

data OutEx9c;
   set OutEx9c;
   Diff= Col2 - Col1;
   drop Col1 COl2;
run;
```

Output 100.9.3 lists the differences in LDL decrease in the first four matched sets.

Output 100.9.3 LDL Differences in Matched sets

Obs	_MatchID	_NAME_	Diff
1	1	LDL	3.25
2	2	LDL	2.44
3	3	LDL	6.34
4	4	LDL	-1.51

The following statements perform a signed rank test, and the results are shown in Output 100.9.4.

```
ods select TestsForLocation;
proc univariate data=OutEx9c;
   var Diff;
```

```
   ods output TestsForLocation=LocTest;
run;
```

Output 100.9.4 Tests for Location

The UNIVARIATE Procedure
Variable: Diff

Tests for Location: Mu0=0				
Test		**Statistic**	**p Value**	
Student's t	**t**	2.663999	**Pr > |t|**	0.0089
Sign	**M**	9.5	**Pr >= |M|**	0.0900
Signed Rank	**S**	859.5	**Pr >= |S|**	0.0131

The "Tests for Location" table shows that there is a significant decrease in LDL at the 0.025 level for patients in the treated group.

Propensity score analysis assumes that all confounders (variables that affect both the outcome and the treatment assignment) have been measured. However, this assumption cannot be verified. When there are unobserved covariates, individuals that have the same observed covariates might not have the same probability of being assigned to the treated group. If you assume that all confounders have been measured, you should examine the sensitivity of inferences to departures from the assumption.

Based on the approach described in the section "Sensitivity Analysis on Matched Observations" on page 8232, the signed rank statistic is

$$S = \sum_{j:d_j>0} d_j^+$$

Note that this statistic is not centered, unlike the signed rank statistic that is computed by PROC UNIVARIATE and is shown in Output 100.9.4:

$$\sum_{j:d_j>0} d_j^+ - \frac{n_t(n_t+1)}{4}$$

The following statements compute the signed rank statistic:

```
data SgnRank;
   set LocTest;
   nPairs=113;
   if (Test='Signed Rank');
   SgnRank= Stat + nPairs*(nPairs+1)/4;
   keep nPairs SgnRank;
run;
```

Output 100.9.5 displays the signed rank statistic.

Output 100.9.5 Signed Rank Statistic

Obs	nPairs	SgnRank
1	113	4080

Using this statistic, the following statements compute and display p-values for signed rank tests that correspond to Γ values that range from 1 to 1.5.

```
data Test1;
   set SgnRank;
   mean0     = nPairs*(nPairs+1)/2;
   variance0 = mean0*(2*nPairs+1)/3;

   do Gamma=1 to 1.5 by 0.05;
      mean     = Gamma/(1+Gamma) * mean0;
      variance = Gamma/(1+Gamma)**2 * variance0;
      tTest    = (SgnRank - mean) / sqrt(variance);
      pValue   = 1 - probt(tTest, nPairs-1);
      output;
   end;
run;

proc print data=Test1;
run;
```

Output 100.9.6 p-Values for Γ Values from 1 to 1.5

Obs	nPairs	SgnRank	mean0	variance0	Gamma	mean	variance	tTest	pValue
1	113	4080	6441	487369	1.00	3220.50	121842.25	2.46233	0.00766
2	113	4080	6441	487369	1.05	3299.05	121769.77	2.23797	0.01360
3	113	4080	6441	487369	1.10	3373.86	121565.96	2.02529	0.02261
4	113	4080	6441	487369	1.15	3445.19	121249.18	1.82309	0.03548
5	113	4080	6441	487369	1.20	3513.27	120835.29	1.63034	0.05292
6	113	4080	6441	487369	1.25	3578.33	120338.02	1.44615	0.07546
7	113	4080	6441	487369	1.30	3640.57	119769.32	1.26976	0.10340
8	113	4080	6441	487369	1.35	3700.15	119139.55	1.10049	0.13674
9	113	4080	6441	487369	1.40	3757.25	118457.74	0.93774	0.17520
10	113	4080	6441	487369	1.45	3812.02	117731.79	0.78101	0.21822
11	113	4080	6441	487369	1.50	3864.60	116968.56	0.62981	0.26505

Output 100.9.6 shows that at the tipping point Γ=1.15, the p-value is 0.0355, which is larger than the Type I error level of 0.025. Thus the study conclusion is reversed if for two individuals k and l in the same matched set, the probability that individual k is in the treated group and l is in the control group is

$$\frac{\pi_k}{\pi_k + \pi_l} = \frac{\Gamma}{1+\Gamma} = 0.535$$

If Γ=1.15 represents only a small departure from random treatment assignment (Γ=1), the study conclusion is not robust to hidden bias from an unobserved confounder.

References

Austin, P. C. (2007). "The Performance of Different Propensity Score Methods for Estimating Marginal Odds Ratios." *Statistics in Medicine* 26:3078–3094.

Austin, P. C. (2009). "Balance Diagnostics for Comparing the Distribution of Baseline Covariates between Treatment Groups in Propensity-Score Matched Samples." *Statistics in Medicine* 28:3083–3107.

Austin, P. C. (2011a). "An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies." *Multivariate Behavioral Research* 46:399–424.

Austin, P. C. (2011b). "Optimal Caliper Widths for Propensity-Score Matching When Estimating Differences in Means and Differences in Proportions in Observational Studies." *Pharmaceutical Statistics* 10:150–161.

Austin, P. C. (2014). "A Comparison of 12 Algorithms for Matching on the Propensity Score." *Statistics in Medicine* 33:1057–1069.

Austin, P. C., Grootendorst, P., and Anderson, G. M. (2007). "A Comparison of the Ability of Different Propensity Score Models to Balance Measures Variables between Treated and Untreated Subjects: A Monte Carlo Study." *Statistics in Medicine* 26:734–753.

Austin, P. C., and Stuart, E. A. (2015a). "Moving towards Best Practice When Using Inverse Probability of Treatment Weighting (IPTW) Using the Propensity Score to Estimate Causal Treatment Effects in Observational Studies." *Statistics in Medicine* 34:3661–3679.

Austin, P. C., and Stuart, E. A. (2015b). "The Performance of Inverse Probability of Treatment Weighting and Full Matching on the Propensity Score in the Presence of Model Misspecification When Estimating the Effect of Treatment on Survival Outcomes." *Statistical Methods in Medical Research* 26:1654–1670. http://smm.sagepub.com/content/early/2015/04/30/0962280215584401.full.pdf+html.

Cole, S. R., and Hernán, M. A. (2008). "Constructing Inverse Probability Weights for Marginal Structural Models." *American Journal of Epidemiology* 168:656–664.

Crump, R. K., Hotz, V. J., Imbens, G. W., and Mitnik, O. A. (2009). "Dealing with Limited Overlap in Estimation of Average Treatment Effects." *Biometrika* 96:187–199.

Faries, D. E., Leon, A. C., Haro, J. M., and Obenchain, R. L., eds. (2010). *Analysis of Observational Health Care Data Using SAS*. Cary, NC: SAS Institute Inc.

Guo, S., and Fraser, M. W. (2015). *Propensity Score Analysis: Statistical Methods and Applications*. 2nd ed. Thousand Oaks, CA: Sage Publications.

Hansen, B. B. (2004). "Full Matching in an Observational Study of Coaching for the SAT." *Journal of the American Statistical Association* 99:609–618.

Hill, J., and Reiter, J. P. (2006). "Interval Estimation for Treatment Effects Using Propensity Score Matching." *Statistics in Medicine* 25:2230–2256.

Ho, D., Imai, K., King, G., and Stuart, E. A. (2007). "Matching as Nonparametric Preprocessing for Reducing Model Dependence in Parametric Causal Inference." *Political Analysis* 15:199–236.

Imbens, G. W., and Rubin, D. B. (2015). *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction*. New York: Cambridge University Press.

Liu, W., Kuramoto, S. J., and Stuart, E. A. (2013). "An Introduction to Sensitivity Analysis for Unobserved Confounding in Non-experimental Prevention Research." *Prevention Science* 14:570–580.

Lunceford, J. K., and Davidian, M. (2004). "Stratification and Weighting via the Propensity Score in Estimation of Causal Treatment Effects: A Comparative Study." *Statistics in Medicine* 23:2937–2960.

Mamdani, M., Sykora, K., Li, P., Normand, S. L., Streiner, D. L., Austin, P. C., Rochon, P. A., and Anderson, G. M. (2005). "Reader's Guide to Critical Appraisal of Cohort Studies: 2. Assessing Potential for Confounding." *BMJ* 330:960–962.

Normand, S.-L. T., Landrum, M. B., Guadagnoli, E., Ayanian, J. Z., Ryan, T. J., Cleary, P. D., and McNeil, B. J. (2001). "Validating Recommendations for Coronary Angiography Following Acute Myocardial Infarction in the Elderly: A Matched Analysis Using Propensity Scores." *Journal of Clinical Epidemiology* 54:387–398.

Pan, W., and Bai, H., eds. (2015). *Propensity Score Analysis: Fundamentals and Developments*. New York: Guilford Press.

Robins, J. M., Hernan, M. A., and Brumback, B. (2000). "Marginal Structural Models and Causal Inference in Epidemiology." *Epidemiology* 11:550–560.

Rosenbaum, P. R. (2010). *Design of Observational Studies*. New York: Springer-Verlag.

Rosenbaum, P. R., and Rubin, D. B. (1983). "The Central Role of the Propensity Score in Observational Studies for Causal Effects." *Biometrika* 70:41–55.

Rosenbaum, P. R., and Rubin, D. B. (1984). "Reducing Bias in Observational Studies Using Subclassification on the Propensity Score." *Journal of the American Statistical Association* 79:516–524.

Rosenbaum, P. R., and Rubin, D. B. (1985). "Constructing a Control Group Using Multivariate Matched Sampling Methods That Incorporate the Propensity Score." *American Statistician* 39:33–38.

Rubin, D. B. (1974). "Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies." *Journal of Educational Psychology* 66:688–701.

Rubin, D. B. (1980a). "Bias Reduction Using Mahalanobis-Metric Matching." *Biometrics* 36:293–298.

Rubin, D. B. (1980b). "Comment on D. Basu, 'Randomization Analysis of Experimental Data: The Fisher Randomization Test'." *Journal of the American Statistical Association* 75:591–593.

Rubin, D. B. (1990). "Comment: Neyman (1923) and Causal Inference in Experiments and Observational Studies." *Statistical Science* 5:472–480.

Rubin, D. B. (2001). "Using Propensity Scores to Help Design Observational Studies: Application to the Tobacco Litigation." *Health Services and Outcomes Research Methodology* 2:169–188.

Rubin, D. B. (2005). "Causal Inference Using Potential Outcomes: Design, Modeling, Decisions." *Journal of the American Statistical Association* 100:322–331.

Stuart, E. A. (2007). "Estimating Causal Effects Using School-Level Data Sets." *Educational Researcher* 36:187–198.

Stuart, E. A. (2010). "Matching Methods for Causal Inference: A Review and a Look Forward." *Statistical Science* 25:1–21.

Stuart, E. A., and Ialongo, N. S. (2010). "Matching Methods for Selection of Subjects for Follow-Up." *Multivariate Behavioral Research* 45:746–765.

Stuart, E. A., Marcus, S. M., Horvitz-Lennon, M. V., Gibbons, R. D., and Normand, S.-L. T. (2009). "Using Non-experimental Data to Estimate Treatment Effects." *Psychiatric Annals* 39:719–728.

Stürmer, T., Wyss, R., Glynn, R. J., and Brookhart, M. A. (2014). "Propensity Scores for Confounder Adjustment When Assessing the Effects of Medical Interventions Using Nonexperimental Study Designs." *Journal of Internal Medicine* 275:570–580.

Yue, L. Q., Lu, N., and Xu, Y. (2014). "Designing Premarket Observational Comparative Studies Using Existing Data as Controls: Challenges and Opportunities." *Journal of Biopharmaceutical Statistics* 24:994–1010.

Subject Index

Syntax Index