SAS Macro to Calculate Standardized Mean Difference

Tong Zhao, Tian Gu, LLX Solutions, LLC

ABSTRACT

The standardized mean difference (SMD) is a measure of effect size between the two groups. It is widely used in the external controlled clinical trials studies, comparing the baseline characteristics of the treated group and the propensity-score (PS) weighted or propensity-score (PS) matched control group. This paper details the methodology for calculating SMD for both continuous and binary variables in such analyses. Additionally, we provide a SAS macro that implements these calculations directly from the underlying formulas, offering researchers a flexible alternative to standard procedures.

INTRODUCTION

The use of external control studies has expanded rapidly in research and regulatory approval. A critical aspect of these studies is assessing the comparability of treated and control groups by evaluating their baseline characteristics. The standardized mean difference (SMD) is widely used to quantify these differences. While some SAS procedures automatically generate SMD, manual calculation is necessary when applying custom weights or alternative samples not supported by the original model.

This paper details the SMD calculation for both continuous and binary variables, focusing on comparisons between treated and control groups after propensity score (PS) weighting or matching. Additionally, we present a macro that computes SMD independently of the PSMATCH procedure in SAS, offering flexibility for customized analyses.

FORMULA

SMD is generally defined as the difference between the treatment groups divided by the standard deviation.

For continuous baseline variables, the formula to derive SMD is

$$d = \frac{(\bar{X}_1 - \bar{X}_2)}{\sqrt{\frac{s_1^2 + s_2^2}{2}}}$$

where $\bar{X}_1$ and $\bar{X}_2$ are the sample means of a baseline variable in each group, and s_1^2 and s_2^2 are the sample variances, respectively (Yang, Dongsheng; Dalton, Jarrod E., 2012).

For binary baseline variables, the formula to derive SMD is

$$d = \frac{(\hat{P}_1 - \hat{P}_2)}{\sqrt{\frac{[\hat{P}_1(1 - \hat{P}_1) + \hat{P}_2(1 - \hat{P}_2)]}{2}}}$$

where $\hat{P}_1$ and $\hat{P}_1$ are the proportion or mean of a binary baseline variable in the two groups, respectively (Yang, Dongsheng; Dalton, Jarrod E., 2012).

In propensity score (PS) weighting analyses, the weights are directly incorporated into the numerator of the standardized mean difference (SMD) calculation.

SAS MACRO

The SAS macro leverages standard SAS procedures (PROC FREQ, PROC MEANS, and DATA STEP) to compute standardized mean differences (SMDs). Designed for flexibility, the macro accommodates both continuous and binary variables via a user-specified switch. It processes one variable per execution and outputs the SMD value in a structured dataset.

The complete SAS macro code is provided in the supplementary materials due to space limitations.

```
/**********************************************************************************/
/* Program    : smd.sas
/* Purpose    : SAS macro to calculate the Standardize Mean Difference
/* Usage      : %smd(inds = data, trt = trt, wt = wt, var = var1, num = 1, binary = 0, outds =
out);
/* Version    : 1.0 03DEC2024
/**********************************************************************************/
/* Parameters :
/*     inds:   input SAS dataset name
/*     trt:    grouping variable of treatment (=1) or control (=0), must be coded as 1/0
/*     wt:     weight variable name
/*     var:    variable name
/*     num:    variable type is num (=1) or not (=0)
/*     binary: variable type is binary (=1) or not (=0)
/*     outds:  output SAS dataset name
/**********************************************************************************/
```

CONCLUSION

The standardized mean difference (SMD) for both continuous and binary variables can be computed using our SAS macro, which implements the standard SMD formula through base SAS procedures.

REFERENCES

Yang, Dongsheng; Dalton, Jarrod E.. (2012). A unified approach to measuring the effect size between two groups using SAS. *SAS Global Forum 2012*. Available at 335-2012: Standardized Difference: An Index to Measure the Effect Size between Two Groups

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Tong Zhao
LLX Solutions, LLC
Tong.zhao@llxsolutions.com

Any brand and product names are trademarks of their respective companies.