ABSTRACT

Mediation analysis is a statistical technique for investigating the extent to which a mediating variable transmits the relation of an independent variable to a dependent variable. Because it is useful in many fields, there have been rapid developments in statistical mediation methods. The most cutting-edge statistical mediation analysis focuses on the causal interpretation of mediated effect estimates. Cause-and-effect inferences are particularly challenging in mediation analysis because of the difficulty of randomizing subjects to levels of the mediator (MacKinnon, 2008). The focus of this paper is how incorporating longitudinal measures of the mediating and outcome variables aides in the causal interpretation of mediated effects. This paper provides useful SAS® tools for designing adequately powered studies to detect the mediated effect. Three SAS macros were developed using the powerful but easy-to-use REG, CALIS, and SURVEYSELECT procedures to do the following: (1) implement popular statistical models for estimating the mediated effect in the pretest-posttest control group design; (2) conduct a prospective power analysis for determining the required sample size for detecting the mediated effect; and (3) conduct a retrospective power analysis for studies that have already been conducted and a required sample to detect an observed effect is desired. We demonstrate the use of these three macros with an example.

INTRODUCTION

Statistical mediation analysis is a common statistical technique used in marketing research, social sciences, epidemiology, and other related fields because it allows researchers to investigate how and through what mechanism two variables are related (MacKinnon, 2008). For example, mediation analysis was used to determine that self-regulatory depletion leads to increased purchase choice through its effect on product-attitude certainty (Wan, Rucker, Tormala, & Clarkson, 2010). Mediation analysis is also central to the understanding of the effects of interventions to reduce substance use behavior (Larimer et al., 2007; MacKinnon et al., 2001; West & Aiken, 1997; Witkiewitz & Bowen, 2010). Despite the advantage of investigating mechanisms of theoretical relations, it is difficult to make cause-and-effect inferences when using mediation analysis (Bullock, Green, & Ha, 2010; Imai, Keele, Tingley, & Yamamoto, 2011; MacKinnon, 2008; MacKinnon & Pirlott, 2015; Robins & Greenland, 1992) but the use of longitudinal measurements of theoretically relevant variables improves the accuracy of mediation estimates.

This paper focuses on the widely-used pre-test post-test randomized groups research design that allows researchers to use longitudinal data to make cause-and-effect inferences in a mediation analysis.

SINGLE MEDIATOR MODEL

Statistical mediation is a type of analysis that allows researchers to test indirect effects of an independent variable on a dependent variable through the independent variable’s effect on the mediating variable (Lazarsfeld, 1955; MacKinnon, 2008; MacKinnon & Dwyer, 1993; Sobel 1990). A mediating variable is a type of third variable that can explain the relation between two variables. When the third variable is conceptualized in a causal sequence between two other variables, it is called a mediator (James & Brett, 1984; MacKinnon, 2008). Statistical mediation is represented by three linear regression equations. Equation 1 represents the total effect of $X$ on $Y$ ($c$ coefficient), Equation 2 represents the effect of $X$ on $M$ ($a$ coefficient), and Equation 3 represents the effect of $X$ on $Y$ adjusted for $M$ ($c’$ coefficient) and the effect of $M$ on $Y$ adjusted for $X$ ($b$ coefficient). Computing the product of $a$ and $b$ linear regression coefficients from Equation 2 and Equation 3, respectively, represents the mediated effect of $X$ on $Y$ through $M$ ($ab$) (See Figure 1A).
\[ Y = i_1 + cX + e_1 \]  
\[ M = i_2 + aX + e_2 \]  
\[ Y = i_3 + c'X + bM + e_3 \]

The mediated effect is not generally interpretable as a causal effect because even if subjects are randomized to levels of the independent variable, \( X \), subjects are not simultaneously randomized to levels of the mediator, \( M \), therefore the mediator to outcome relation may be biased because of unmeasured confounders (Bullock, Green, & Ha, 2010; Imai, et al., 2011; MacKinnon, 2008; MacKinnon & Pirlott, 2015; Robins & Greenland, 1992). Recent work applying the potential outcomes framework for causal inference (Holland, 1986, 1988, Neyman, 1923; Rubin, 1974) to mediation analysis has clarified the assumptions necessary for causal interpretation of mediated effects (Robins & Greenland, 1992; Imai, Keele, Tingley, 2010; Pearl 2001, 2009, 2014; VanderWeele & Vansteelandt, 2009). One way to strengthen the causal inference about mediating processes is through longitudinal studies of the mediator – outcome relation which allow for adjustment of time-invariant confounders (i.e., confounders that do not change over time) and provide evidence of temporal precedence of the mediator to outcome relation (i.e., does \( M \) precede \( Y \) or does \( Y \) precede \( M \), in time) (Cole & Maxwell, 2003; MacKinnon, 1994, 2008; Maxwell & Cole, 2007; Maxwell, Cole, & Mitchell 2011). The simplest longitudinal experimental design that can adjust for time-invariant confounders of the mediator – outcome relation is the pretest-posttest control group design.

**Figure 1.**

**Left Panel** – Single Mediator Model.

**Right panel** – Mediator model with pretest measures of \( M \) and \( Y \).

**PRETEST-POSTTEST CONTROL GROUP DESIGN**

The pretest-posttest control group design is a very popular experimental design that makes use of longitudinal data (e.g., two-waves of data) and consists of collecting pretest measures prior to randomization, randomly assigning units to either a treatment or a control group, and collecting posttest measures after randomization and delivery of an intervention (Bonate, 2000; Shadish, Cook, and Campbell, 2002). This design allows researchers to take into account nuisance variation in the outcome variable leading to more powerful inferential test statistics, more precise confidence intervals, increased...
internal validity, and adjust for time invariant confounders (Huitema, 2011; Maxwell & Delaney, 2004; Shadish et al., 2002; Valente & MacKinnon, 2017).

In the pretest-posttest control group design with a mediating variable, there are pretest scores for both the mediating variable and the outcome variable and the following adjustment techniques can be used including difference scores, residualized change scores, and Analysis of Covariance (ANCOVA). The variables in this design consist of pretest measures of the mediator and outcome ($M_1$ and $Y_1$, respectively), a randomized experimental manipulation ($X$), and posttest measures of the mediator and the outcome ($M_2$ and $Y_2$, respectively). The mediated effect in this design is a product of $a_{m2x}b_{y2m2}$ as seen in Figure 1B.

The macros presented in this project are focused on estimating mediated effects in the pretest–posttest control group design using five different statistical models (for more detail on these models, see Valente & MacKinnon, 2017). All models are estimated using SAS® PROC REG and DATA steps when necessary. These models are presented below with their associated measure of effect size for the mediated effect (i.e., mediated effect estimate divided by the standard deviation of the outcome variable):

**Cross-Sectional Model:**

\[
M_2 = i_6 + a_{m2x}X + e_6
\]

\[
Y_2 = i_7 + c'_{y2x}X + b_{y2m2}M_2 + e_7
\]

Mediated effect = $a_{m2x} \cdot b_{y2m2}$

Effect Size = \(\frac{(a_{m2x} \cdot b_{y2m2})}{\text{std}Y_2}\)

**Difference Score Model:**

\[
\Delta M = M_2 - M_1
\]

\[
\Delta Y = Y_2 - Y_1
\]

\[
\Delta M = i_8 + a_{\Delta}X + e_8
\]

\[
\Delta Y = i_9 + c'_{\Delta}X + b_{\Delta}\Delta M + e_9
\]

Mediated effect = $a_{\Delta} \cdot b_{\Delta}$

Effect Size = \(\frac{(a_{\Delta} \cdot b_{\Delta})}{\text{std}\Delta Y}\)

**Residualized Change Score Model:**

\[
R_M = \text{Observed } M_2 - \text{Predicted } M_2
\]

\[
R_Y = \text{Observed } Y_2 - \text{Predicted } Y_2
\]

\[
R_M = i_{10} + a_{R}X + e_{10}
\]

\[
R_Y = i_{11} + c'_{R}X + b_{R}R_M + e_{11}
\]

Mediated Effect = $a_{R} \cdot b_{R}$

Effect Size = \(\frac{(a_{R} \cdot b_{R})}{\text{std}R_Y}\)

**Analysis of Covariance (ANCOVA):**

\[
M_2 = i_4 + a_{m2x}X + S_{m2m1} M_1 + b_{m2y1} Y_1 + e_4
\]

\[
Y_2 = i_5 + c'_{y2x}X + S_{y2y1} Y_1 + b_{y2m1} M_1 + b_{y2m2} M_2 + e_5
\]

Mediated effect = $a_{m2x} \cdot b_{y2m2}$

Effect Size = \(\frac{(a_{m2x} \cdot b_{y2m2})}{\text{std}Y_2}\)
In addition to the four models described above, path analysis can estimate the same equations as ANCOVA simultaneously using SAS® PROC CALIS with METHOD=FIML for full-information maximum likelihood. This is the preferred option if there are any missing data (Enders, 2010).

Three macros using SAS® 9.4 compute the mediated effect for the pretest – posttest control group design for all five models mentioned above, perform a retrospective power analysis (i.e., determine power achieved for an observed effect size of the mediated effect), and perform a Monte Carlo based prospective power analysis (i.e., used for sample size planning for a new research study).

THE TWO WAVEMED MACRO

The first macro allows a user to enter a dataset of their choice (data =), specify the variables needed for the analysis (x=, m1=, m2=, y1=, y2=), the sample size of the dataset (nobs =), whether they want single sample significance results or percentile bootstrap significance results (boot =), number of bootstrap samples (nboot =), and alpha level (alpha =).

%TWO WAVEMED(data = , x=, m1=, m2=, y1=, y2=, nobs=, alpha=.05, boot=false, nboot=1000);

Single sample results refer to normal theory significance test and asymmetric confidence intervals (i.e., PRODCLIN; MacKinnon, Fritz, Williams, & Lockwood, 2007). The preferred single sample method is the asymmetric confidence intervals as they provide a more accurate test of the mediated effect because the product of two regression coefficients is not normally distributed for finite sample sizes and a range of effect sizes (MacKinnon, Lockwood, & Williams, 2004). When single sample results are requested by the user, the program makes use of the PRODCLIN program (MacKinnon et al., 2007) (instructions for download are included with the SAS code for these macros).

When researchers use this macro, they will get estimates of the mediated effect for each of the five models, t-values, p-values, asymmetric confidence limits, and a measure of effect size. If a researcher wants to get percentile bootstrap results, they need to specify the boot = options and the number of bootstrap sample in the nboot = option. The default is for boot =false and nboot = 1000 and alpha = .05.

<table>
<thead>
<tr>
<th>Model</th>
<th>Estimate</th>
<th>t_value</th>
<th>p_value</th>
<th>LCL</th>
<th>UCL</th>
<th>EffectSize</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANCOVA</td>
<td>6.703</td>
<td>3.713</td>
<td>.00033</td>
<td>2.77950</td>
<td>10.5048</td>
<td>1.57765</td>
</tr>
<tr>
<td>Path analysis</td>
<td>6.703</td>
<td>3.904</td>
<td>.00019</td>
<td>3.07630</td>
<td>9.3677</td>
<td>1.57765</td>
</tr>
<tr>
<td>Diff.Score</td>
<td>2.747</td>
<td>2.850</td>
<td>.00344</td>
<td>1.01348</td>
<td>4.8267</td>
<td>0.63931</td>
</tr>
<tr>
<td>Cross.Sec.</td>
<td>6.547</td>
<td>3.335</td>
<td>.00092</td>
<td>3.25422</td>
<td>10.3581</td>
<td>1.54093</td>
</tr>
</tbody>
</table>

## Estimate of the Mediated effect with percentile bootstrap confidence limits

<table>
<thead>
<tr>
<th>Model</th>
<th>Estimate</th>
<th>LCL95</th>
<th>UCL95</th>
<th>EffectSize</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANCOVA</td>
<td>6.703</td>
<td>3.67233</td>
<td>10.7657</td>
<td>1.64232</td>
</tr>
<tr>
<td>Path analysis</td>
<td>6.703</td>
<td>3.67233</td>
<td>10.7657</td>
<td>1.64232</td>
</tr>
<tr>
<td>Diff.Score</td>
<td>2.747</td>
<td>0.83583</td>
<td>5.0838</td>
<td>0.64881</td>
</tr>
<tr>
<td>Res. Change</td>
<td>6.117</td>
<td>2.23187</td>
<td>9.2076</td>
<td>1.35392</td>
</tr>
<tr>
<td>Cross.Sec.</td>
<td>6.547</td>
<td>3.92044</td>
<td>10.3749</td>
<td>1.60249</td>
</tr>
</tbody>
</table>

### Output 2. Tests of the mediated effect using percentile bootstrap confidence limits with estimates of effect size.

The estimate of the mediated effect and effect size will be identical across boot options but the confidence intervals will vary, albeit usually not substantially. The asymmetric confidence intervals and bootstrap confidence intervals provide a more powerful test of the mediated effect than relying on the t-test and its associated p-value provided in the table (MacKinnon, et al., 2004).

### THE TWOWAVEMONTECARLO MACRO

The second macro is a Monte Carlo power program that allows users to specify a model for empirical power results (model = ), specify the values of the paths corresponding to the general path model (see Figure 1B) (rmy =, sm =, sy =, a =, b =, c =, m2y1 =, y2m1 =), specify the starting sample size (nobs =), the ending sample size (end =), the increment to increase the starting sample size by (incr =) simulation replications (nsim =), whether or not they want percentile bootstrap empirical power (boot =), the number of bootstrap samples (nboot =) and desired alpha level (alpha =). The default is for nsim = 1000, boot = false (which provides single sample methods empirical power) and alpha = .05. If boot = true then the default number of bootstrap samples is nboot = 1000.

```
%TWOWAVEMONTECARLO(model=, boot=false, nboot=1000, alpha=0.05, nobs=100, end=200, incr=50, nsim=1000, rmy=, sm=, sy=, a=, b=, c=, m2y1=, y2m1=);
```

Both single sample methods are produced on the same plot while the percentile bootstrap empirical power results are produced separately. The parameter values for the Monte Carlo macro were chosen based on parameter values used in a recent simulation study (Valente & MacKinnon, 2017). Users may want to use the single sample methods because it will run faster than the percentile bootstrap procedure and will not require as much free disk space to handle the large datasets generated by the bootstrap procedure.
Output 3. Prospective power analysis using normal theory and distribution of a product methods to test the mediated effect.

Output 4. Prospective power analysis using percentile bootstrap confidence intervals to test the mediated effect.
The empirical power plots will appear smoother and continuously increasing as a function of sample size if either the simulation replications or the bootstrap samples are increased. Some researchers recommend using 10,000 simulation replications for Monte Carlo power estimates (Thoemmes, MacKinnon, & Resier 2010).

**THE TWOWAVEPOSTPOWER MACRO**

The third macro allows users to enter a dataset of their choice, specify the model they want empirical power results for (model =) specify the variables needed for the analysis (x=, m1=, m2=, y1=, y2=), specify the sample size (nobs =), simulation replications (nsim =), whether or not they want percentile bootstrap retrospective empirical power (boot =), the number of bootstrap samples (nboot =) and desired alpha level (alpha =). The default is for nsim = 1000, boot = false (which provides single sample methods empirical power) and alpha = .05. If boot = true then the default number of bootstrap samples is nboot = 1000.

%TWOWAVEPOSTPOWER(data=, model=, x=, m1=, m2=, y1=, y2=, nobs=, nsim=1000, boot=false, nboot=1000, alpha=.05);

This macro calculates retrospective empirical power for the mediated effect for any of the five models described above. This macro will allow users to determine the retrospective power and to determine what power would have been had they had a different sample size assuming the sample effect size is the population effect size.

<table>
<thead>
<tr>
<th>Distribution of a Product and Normal Theory Empirical Power</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample Size = 42 Simulations = 1000</td>
</tr>
<tr>
<td>ANCOVA Model for α = .05</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ANCOVA E.S.</th>
<th>PRODCLIN Power</th>
<th>Normal Theory Power</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.57755</td>
<td>0.954</td>
<td>0.954</td>
</tr>
</tbody>
</table>

**Output 5.** Retrospective empirical power analysis using asymmetric confidence intervals for the ANCOVA model estimate of the mediated effect.

<table>
<thead>
<tr>
<th>Percentile Bootstrap Empirical Power</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample Size = 42 Simulations = 1000 Bootstrap Samples = 1000</td>
</tr>
<tr>
<td>ANCOVA Model for α = .05</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ANCOVA E.S.</th>
<th>Percentile Bootstrap Power</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.57755</td>
<td>0.93</td>
</tr>
</tbody>
</table>

**Output 6.** Retrospective empirical power analysis using percentile bootstrap confidence intervals for the ANCOVA model estimate of the mediated effect.

The output presented corresponds to the retrospective power given the ANCOVA estimate of the mediated effect. Retrospective power is typically used when a significance test fails to reject the null.
hypothesis and a researcher wants to know what the power was given the sample estimated effect size to find the effect. A drawback of this approach is the assumption that the sample estimate of effect size is treated as the actual population value of the effect size which is likely not true and can lead to the “power approach paradox” (Hoenig & Heisey, 2001).

CONCLUSION

Researchers can use the three macros presented to estimate the mediated effect in the pretest-posttest control group design, conduct sample size planning for new research studies involving mediated effects in the pretest-posttest control group design and determine retrospective power for observed mediated effects from previous research studies. These macros provide users with tools for all stages of design and analysis from prospective power analyses (TWOWAVEMONTECARLO), to estimating mediated effects in observed data (TWOWAVEMED), and probing retrospective power if prospective power analyses were not originally conducted (TWOWAVEPOSTPOWER). Overall, these macros will provide users with powerful analytical tools for this popular experimental design while harnessing the reliable and efficient use of SAS® PROCs.

REFERENCES


Lazarsfeld, P. F. (1955). Interpretation of statistical relations as a research operation. In P.F. Lazarsfeld & M. Rosenberg (Eds.), The language of social research: A reader in the methodology of social research (pp. 115 – 125). Glencoe, IL: Free Press


**ACKNOWLEDGMENTS**

This research was supported in part by the National Institute on Drug Abuse, Grant No. R37DA09757.

**RECOMMENDED READING**

- *Base SAS® Procedures Guide*
- *SAS® For Dummies®*

**CONTACT INFORMATION**

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

- **Name:** Matthew J. Valente
- **Organization:** Arizona State University
- **Email:** m.valente@asu.edu

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

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