

SAS® for Bayesian Mediation Analysis

Miočević Milica, Arizona State University; David P. MacKinnon, Arizona State University

ABSTRACT

Recent statistical mediation analysis research focuses on using Bayesian methods for several reasons including the ability to incorporate prior information about model parameters, ease of constructing intervals for simple and complex models, and exact inference for small samples without restrictive assumptions of conventional mediation analysis (Yuan & MacKinnon, 2009). The focus of this paper is how the Bayesian framework offers a solution to mediation analysis with small samples; namely, using diffuse prior distributions when there is no prior literature on the phenomenon being studied, and incorporating prior information into the analysis when there is existing knowledge about the expected magnitude of mediation effects. The SAS MCMC procedure allows researchers to use one of two simple and effective methods to incorporate their prior knowledge into the statistical analysis and obtain the posterior probabilities for quantities of interest, such as the mediated effect.

This paper presents four examples of using SAS PROC MCMC to analyze a single mediator model using: (1) diffuse prior information for each regression coefficient in the model, (2) informative prior distributions for each regression coefficient, (3) diffuse prior distribution for the covariance matrix of variables in the model, and (4) informative prior distribution for the covariance matrix.

INTRODUCTION

Statistical mediation analysis is common in business, social sciences, epidemiology, and other related fields, because it helps explain how and why two variables are related (see MacKinnon, 2008). A mediating variable (M) is a variable that transmits or carries the effect of an independent variable (X) to a dependent variable (Y). Mediation analysis can be used to investigate how product presentation affects liking of the product, which then affects purchase of the product. Mediation analysis can also be used to identify the mechanism by which a health intervention changes norms that then change health behavior. There are many motivations for mediation analysis including the investigation of the process through which X affects Y, explanation of when and why the relation of X and Y occurs in different settings and populations, identification of key ingredients in interventions, and to interpret why interventions achieve beneficial effects and why they do not (Judd & Kenny, 1981; MacKinnon, 1994).

Research on mediation analysis methods is an active area of research. This paper focuses on four methods to apply Bayesian statistics to the single mediator model using PROC MCMC.

SINGLE MEDIATOR MODEL

The goal of many research projects is to identify and describe a relation between two variables, X and Y. Sometimes a third variable can improve the understanding of the relation between two variables. When a third variable is intermediate between X and Y in a causal chain, it is called a mediator (James & Brett, 1984; MacKinnon, 2008). More specifically, mediators are operationally defined as variables that transmit the influence that one variable (X) exerts on another (Y). The simplest mediation model is the single mediator model and consists of three variables: the independent variable (X) related to the mediator (M), which is related to the dependent variable (Y) (MacKinnon, 2008). The single mediator model is portrayed in Figure 1 and can be described using the following three regression equations:

$$Y = i_1 + cX + e_1 \quad (1)$$

$$M = i_2 + aX + e_2 \quad (2)$$

$$Y = i_3 + bM + c'X + e_3 \quad (3)$$

where c represents the total effect of X on Y, c' represents the effect of X on Y adjusted for the effect of the mediator M, b measures the relation between the mediator M and the dependent variable Y adjusted for the independent variable X, and a measures the relation between X and M. The intercepts are i_1 , i_2 , and i_3 , and it is assumed that the three error terms, e_1 , e_2 , and e_3 follow a normal distribution with a mean of zero and variance σ_1^2 , σ_2^2 and σ_3^2 , respectively.

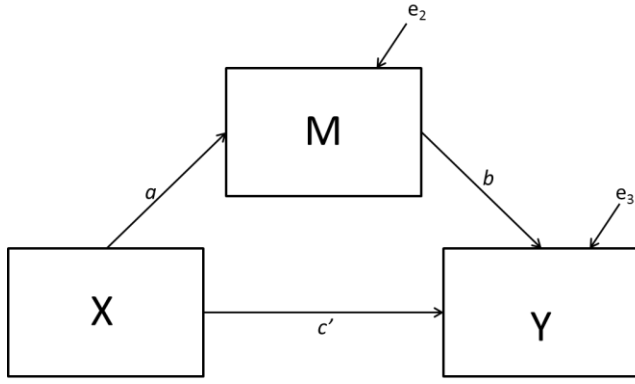


Figure 1. Single Mediator Model

In the single mediator model with continuous M and Y, the mediated effect can be computed in two ways:

1. By multiplying coefficients a and b from Equations 2 and 3, respectively
2. By subtracting coefficient c' (Equation 3) from the coefficient c (Equation 1)

The first method for computing the mediated effect is more general, and the remainder of this paper will use the product of the coefficients method for calculating the mediated effect.

BAYESIAN STATISTICS

The primary distinction between Bayesian and frequentist philosophies lies in their respective application of the probability concept. In the frequentist framework, parameters are treated as fixed, and only the data are assigned a distribution. In Bayesian methods, the data and parameters are assigned distributions. Thus, in the Bayesian framework, the prior information, the data, and the final estimate are all in distribution form. Bayes Theorem is expressed with the following formula:

$$p(\theta | data) = \frac{p(\theta, data)}{p(data)} = \frac{p(\theta)p(data | \theta)}{p(data)} \quad (4)$$

Where $p(\theta | data)$ represents the posterior distribution, $p(\theta)$ is the prior distribution placed on all unknown parameters in the model, $p(data | \theta)$ is the conditional distribution of the data given the parameter, and $p(data)$ is a constant with respect to the parameter of interest, and can thus be omitted in order to produce a simpler way to compute a quantity proportional to the posterior distribution:

$$p(\theta | data) \propto p(\theta)p(data | \theta). \quad (5)$$

Just like frequentist methods, Bayesian methods offer both point and interval estimates of parameters. One form of interval estimates in the Bayesian framework is called credibility intervals. As opposed to frequentist confidence intervals for which one can say that in the long run $(1-\alpha)\%$ confidence intervals (with α being the desired Type I error rate specified by the researcher) will contain the true value of the parameter being estimated, the Bayesian credibility intervals can be interpreted in terms of probability. In other words, as long as the assumptions of the model hold, a Bayesian 95% credibility interval means that there is 95% probability that the parameter value is included in the interval, as opposed to “95% confidence”.

BAYESIAN MEDIATION

We describe two ways to perform a Bayesian mediation analysis and obtain credible intervals (Yuan & MacKinnon, 2009; Enders, Fairchild, & MacKinnon, 2012). In this project, they will be referred to as the coefficients and covariance Bayesian methods. The coefficients Bayesian method is based on the Yuan and MacKinnon (2009) approach, and its implementation starts with the assignment of prior distributions to the parameters in the model, which are i_2 , i_3 , a , b , c' , σ_M^2 and σ_Y^2 (intercepts in the equations 2 and 3 above, the regression coefficients used to compute the mediated effect, the direct path between the independent and dependent variables accounting for the mediated effect, and the error variances of the mediator and dependent variable, respectively). Yuan and MacKinnon (2009) recommend normal distributions as priors for the intercepts and regression coefficients and an inverse gamma distribution (IG) with parameters m (for shape) and n (for scale) as a prior distribution for the variance parameters. Assigning small values to m and n gives the prior distribution a large variance, thus reducing its effect on the shape of the posterior distribution.

If conjugate prior distributions are used, it is sometimes possible to solve for the parameters of the posterior distribution analytically (Greenland, 2006). When it is not possible to obtain a solution analytically, one can use Markov Chain Monte Carlo, a computer-intensive method for estimating probability distributions, in order to empirically approximate the posterior distributions for the parameters of interest (Metropolis, 1987; Gill, 2008). The 95% credibility intervals are then obtained by taking the 0.025 and 0.975 quantiles from the posterior distribution of the parameter of interest.

The covariance Bayesian approach to estimating the mediated effect uses the covariance matrix of the variables X, M, and Y (Enders, Fairchild & MacKinnon, 2012). A prior distribution for the covariance matrix is modeled as an inverse Wishart distribution (a multivariate generalization of the chi-squared distribution) with hyperparameters df_p (the degrees of freedom) and Λ_p (the sum of squares and cross products matrix) that define the center and spread of the distribution (respectively). The prior distribution for the covariance matrix is then written as:

$$p(\Sigma) \sim W^{-1}(df_p, \Lambda_p^{-1}) \quad (6)$$

If a covariance matrix from a prior study is available, the following conversion can be applied to it in order to obtain a prior distribution for the model:

$$\Lambda_p = (N_p - 1)\Sigma_p \quad (7)$$

In the above formula, Σ_p is the covariance matrix and N_p is the sample size from the previous study. The df_p value determines the influence of the prior distribution on the posterior distribution; the larger this value, the more weight is given to the prior distribution. Assigning a value of 20, for example, to df_p would be the same as saying that the prior distribution contributes 20 data points to the analysis (Enders, Fairchild, & MacKinnon, 2012). If X, M, and Y are normally distributed, the posterior distribution of a covariance matrix is also an inverse Wishart denoted as:

$$p(\Sigma | Y) \sim W^{-1}(df, \Lambda^{-1}) \quad (8)$$

where the degrees of freedom are a sum of the degrees of freedom from the prior distribution and the data, and the sum of squares and cross product matrix is the sum of the inverse lambda matrices of the prior distribution and the data. The regression coefficients a and b can be obtained from the covariance matrix where t subscript indicates the covariance matrix element (such as s_{XM} indicating the covariance between X and M):

$$a = \frac{s_{XM}}{s_X^2} \quad (9)$$

$$b = \frac{(s_{MY}s_X^2 - s_{XM}s_{XY})}{(s_X^2s_M^2 - s_{XM}^2)} \quad (10)$$

and their product is computed at each draw from the posterior distribution, thus creating a posterior distribution for the mediated effect from which the credible interval is estimated.

BAYESIAN MEDIATION USING PROC MCMC

Both Bayesian mediation methods described above can be implemented using the MCMC procedure in SAS, and both methods can be applied with diffuse and informative prior distributions for the parameters. The same data set will be used to illustrate all four methods. The only difference between the diffuse and informative prior distributions conditions is that for the diffuse prior distribution conditions, the available prior information will be ignored and the analysis will be done as if no prior information was available. For more examples of using SAS PROC MCMC, see Chen (2009) and Chen (2011).

A simulated data set from MacKinnon (2008) chapter 3 has been modified and is used to illustrate how PROC MCMC can be used for Bayesian mediation. The independent variable, X, is room temperature. The mediator, M, is reported thirst, and the dependent variable Y is water consumption. The observed data set contains 30 participants and the data set used to obtain prior information contains 20 participants.

PRIOR DISTRIBUTIONS FOR REGRESSION COEFFICIENTS

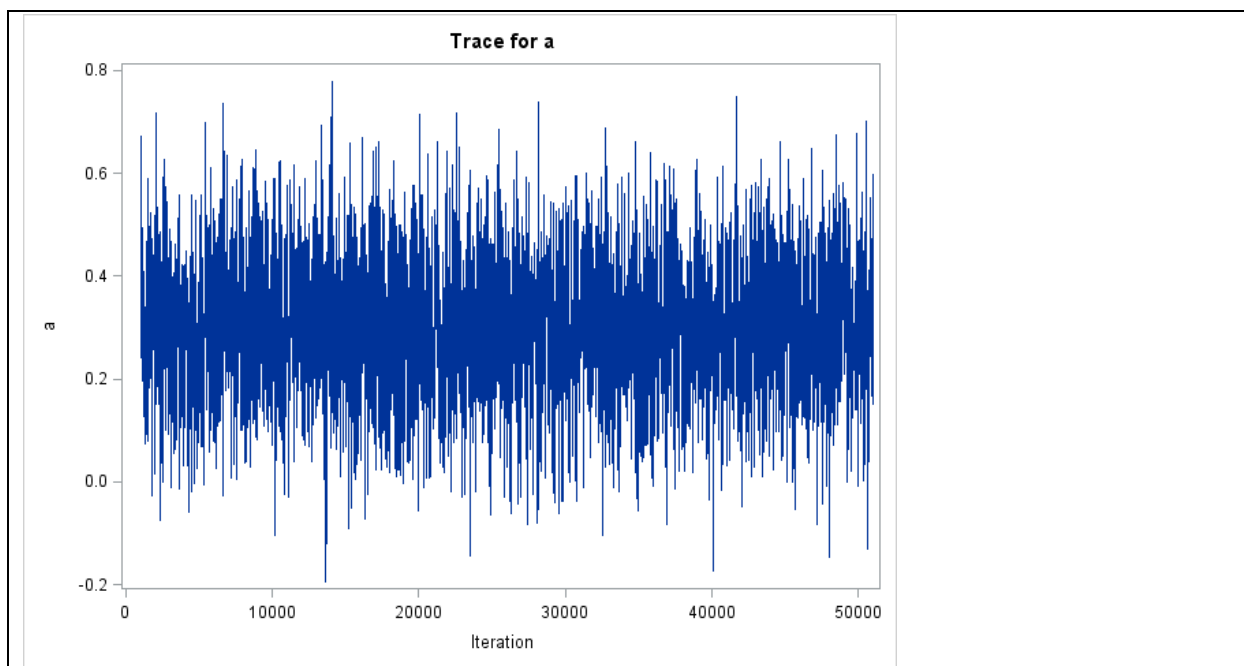
Diffuse prior distributions for regression coefficients

When there is no prior information about the phenomenon being studied, researchers assign diffuse prior distributions to parameters in the model. Regression intercepts and coefficients are usually assigned normal distributions. In order to reflect a lack of knowledge about the parameter, the prior distribution is given a large variance. Precision is the inverse of variance; thus, in order to reflect a lack of knowledge about a parameter, one can also assign to it a prior distribution with low precision. PROC MCMC allows for the spread of the normal prior distributions to be specified in terms of either variance or precision; in the code below, the precision parameterization is used for the prior distributions of the regression coefficients and intercepts. Variances of the mediator and dependent variable can be specified using gamma or inverse-gamma distributions; here, the gamma parametrization with `iscale` is used because of the simplicity of formulating a prior distribution with a mean of 1.

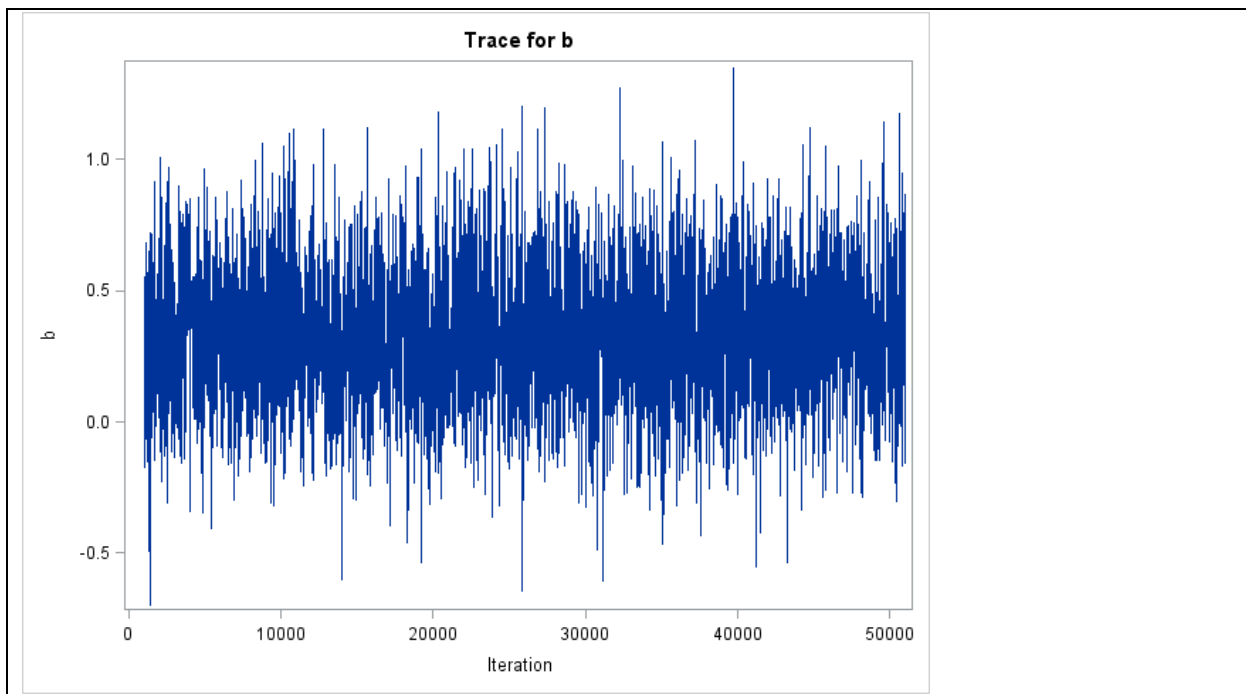
Thus, a Bayesian mediation model with a diffuse prior distribution can be estimated using the following code:

```
proc mcmc data=a outpost=out nmc=50000 thin=5 seed=2 mchistory=detailed
plots(unpack)=trace;
parms a 0 b 0 cpr 0 int2 0 int3 0;
parms sigmaem 1 sigmaey 1;
prior a b cpr int2 int3~ normal (mean=0, prec=1e-6);
prior sigmaem sigmaey ~ gamma(shape=0.01, iscale=0.01);
mum=int2+a*X;
muy=int3+b*M+cpr*X;
model M~ n(mum, prec=sigmaem);
model Y~ n(muy, prec=sigmaey);
run;
```

The first line of code specifies `NMC=50000` draws from the posterior distribution, and `THIN=5` specifies that every fifth draw should be kept. The two `PARMS` statements assign starting values to all parameters in the model, while the two `PRIOR` statements assign prior distributions to the parameters. Subsequent statements specify the model predicting the means of variables `M` and `Y` that are then used in the model statements. The second line of code specifies that only the trace plots are of interest for diagnosing convergence; however, the user can also request autocorrelation function plots, as well as kernel density plots of the posterior distribution for a given parameter. Trace plots represent sample draws versus iteration number, and are used to assess whether the Markov chain has converged (reached its stationary distribution). The trace plots below indicate good mixing for parameters `a` and `b`, and chains that mix well tend to converge sooner. Given the mixing observed in the trace plots, one can proceed to draw inferences from the posterior distributions of the parameters `a` and `b`. For more on convergence diagnostics, see Gelman, Carlin, Stern, and Rubin (2013) and Kruschke (2010).



Output 1. Diagnostic plot for parameter `a`



Output 2. Diagnostic plot for parameter b

The results of this analysis will yield summaries of posterior distributions for parameters a , b , cpr (c' in the previous sections), $i2$, $i3$, σ_{maem} , σ_{maey} .

Posterior Summaries						
Parameter	N	Mean	Standard Deviation	Percentiles		
				25%	50%	75%
a	10000	0.3045	0.1266	0.2202	0.3054	0.3881
b	10000	0.3344	0.2448	0.1710	0.3338	0.4973
cpr	10000	0.3044	0.1674	0.1967	0.3047	0.4146
int2	10000	-18.4542	8.8695	-24.2478	-18.5557	-12.5823
int3	10000	-19.2764	11.4522	-26.9016	-19.2176	-11.9639
sigmaem	10000	1.4574	0.3853	1.1824	1.4276	1.6957
sigmaey	10000	0.9551	0.2601	0.7687	0.9330	1.1157

Posterior Intervals					
Parameter	Alpha	Equal-Tail Interval		HPD Interval	
a	0.050	0.0476	0.5514	0.0566	0.5600
b	0.050	-0.1456	0.8122	-0.1586	0.7934
cpr	0.050	-0.0282	0.6280	-0.0340	0.6208
int2	0.050	-35.7658	-0.4642	-37.1545	-1.9319
int3	0.050	-41.3729	3.5397	-41.1682	3.6676
sigmaem	0.050	0.7991	2.2946	0.7338	2.2062
sigmaey	0.050	0.5185	1.5337	0.4840	1.4795

Output 3. Summaries of posterior distributions for the parameters in the model

However, the computation of the mediated effect will require additional code that computes the mediated effect for each iteration and orders them by magnitude across iterations in order to obtain the desired percentiles:

```
data ab; set out;
ab=a*b;
run;
proc sort data=ab out=percentiles;
by ab;
run;
proc univariate data=percentiles ;
var ab;
output out=p pctlpre=P_ pctlpts=2.5 50 97.5;
run;
proc print data=p;
run;
```

The code above gives the user the 95% central credibility interval for the mediated effect and the median value of the mediated effect.

P_2_5	P_50	P_97_5
-0.041839	0.089018	0.30858

Table 1. Lower 95% central credibility limit, median value of the mediated effect , and upper 95% credibility limit for the regression coefficients method with diffuse prior distributions

Given that zero is between the two credibility limits, one can conclude that the mediated effect is not statistically significant.

By changing the upper and lower percentiles, users can obtain central credibility intervals with a confidence coefficient of 99%, 90%, etc. For example, the code

```
output out=p pctlpre=P_ pctlpts=5 50 95;
run;
```

outputs the limits for the 90% central credibility interval, and the median value of the mediated effect between the two limits.

Informative prior distributions for regression coefficients

In this analysis, prior information is available from a previous study. The values obtained from running a frequentist regression analysis in a previous study are summarized in Table 1:

Regression coefficient	Value (standard error)
<i>a</i>	0.35135 (0.30464)
<i>b</i>	0.50068 (0.18577)
<i>cpr</i> (<i>c'</i>)	-0.09483 (0.24882)
<i>int2</i> (<i>i₂</i>)	-21.38514 (21.44807)
<i>int3</i> (<i>i₃</i>)	8.59907 (17.36533)

Table 2. Results from a PROC REG analysis from a previous study

The code is now changed from the previous analysis in order to incorporate the available information about the parameters in the mediation model:

```
proc mcmc data=sample outpost=outinfo nmc=50000 thin=5 seed=2 mchistory=detailed
plots(unpack)=trace;
parms a 0 b 0 cpr 0 int2 0 int3 0;
parms sigmaem 1 sigmaey 1;
prior a~ normal (mean=0.35135, sd=0.30464);
prior b~ normal (mean=0.50068, sd=0.18577 );
prior cpr~ normal (mean=-0.09483, sd=0.24882);
prior int2~ normal (mean=-21.38514, sd=21.44807);
prior int3~ normal (mean=8.59907, sd=17.36533);
prior sigmaem ~ gamma(shape=0.01, iscale=0.01);
prior sigmaey ~ gamma(shape=0.01, iscale=0.01);
mum=int2+a*X;
muy=int3+b*M+cpr*X;
model M~ n(mum, prec=sigmaem);
model Y~ n(muy, prec=sigmaey);
run;
```

The code above is still insufficient to obtain the credibility interval and median value of the mediated effect. The SORT and UNIVARIATE procedures can be used to obtain these values, as described above. The results of the Bayesian mediation analysis with informative priors are presented in Table 3:

P_2_5	P_50	P_97_5
0.0340095698	0.1414053641	0.3051573554

Table 3. Lower 95% central credibility limit, median value of the mediated effect, and upper 95% credibility limit for the regression coefficients method with informative prior distributions

Given that zero is not in the credibility interval, one can conclude that the mediated effect was statistically significant.

PRIOR DISTRIBUTIONS FOR THE COVARIANCE MATRIX

Diffuse prior distribution for the covariance matrix

Another way to do a Bayesian mediation analysis using PROC MCMC is to assign a prior distribution to the covariance matrix of variables X, M, and Y. The vector of means of these three variables is assigned a multivariate normal distribution, while the covariance matrix is assigned an inverse Wishart distribution with the number of observations as the degrees of freedom, and the covariation matrix as the second parameter.

The code below is structured so that the expectation for each covariance between variables is zero, and the variance of each variable is 1:

```
proc mcmc data=sample outpost=output thin=1 seed=2 nmc=50000 mchistory=detailed
plots(unpack)=trace;
array data[3] X M Y;
array mu[3];
array Sigma[3,3];
array mu0[3] (0 0 0);
array Sigma0[3,3] (1000 0 0 0 1000 0 0 0 1000);
array SDIFFUSE[3,3] (2 0 0 0 2 0 0 0 2);
parms mu Sigma;
```

```
prior mu ~ mvn(mu0, Sigma0);
prior Sigma ~ iwish(3, SDIFFUSE);
model data ~ mvn(mu, Sigma);
run;
```

In this case, PROC MCMC gives summary statistics for each element of the mu and Sigma matrices.

Posterior Summaries						
Parameter	N	Mean	Standard Deviation	Percentiles		
				25%	50%	75%
mu1	50000	70.0297	0.2407	69.8709	70.0299	70.1878
mu2	50000	2.8661	0.1735	2.7524	2.8665	2.9799
mu3	50000	2.9982	0.2195	2.8533	2.9994	3.1433
Sigma1	50000	1.7462	0.4856	1.4045	1.6668	1.9965
Sigma2	50000	0.5031	0.2648	0.3247	0.4720	0.6449
Sigma3	50000	0.6779	0.3354	0.4500	0.6351	0.8569
Sigma4	50000	0.5031	0.2648	0.3247	0.4720	0.6449
Sigma5	50000	0.9081	0.2566	0.7289	0.8652	1.0383
Sigma6	50000	0.4282	0.2387	0.2691	0.4004	0.5558
Sigma7	50000	0.6779	0.3354	0.4500	0.6351	0.8569
Sigma8	50000	0.4282	0.2387	0.2691	0.4004	0.5558
Sigma9	50000	1.4268	0.3982	1.1455	1.3614	1.6317

Posterior Intervals					
Parameter	Alpha	Equal-Tail Interval		HPD Interval	
mu1	0.050	69.5549	70.5053	69.5546	70.5049
mu2	0.050	2.5223	3.2081	2.5147	3.1987
mu3	0.050	2.5649	3.4292	2.5647	3.4286
Sigma1	0.050	1.0372	2.9046	0.9148	2.6888
Sigma2	0.050	0.0738	1.1115	0.0369	1.0450
Sigma3	0.050	0.1413	1.4574	0.0778	1.3571

Posterior Intervals					
Parameter	Alpha	Equal-Tail Interval		HPD Interval	
Sigma4	0.050	0.0738	1.1115	0.0369	1.0450
Sigma5	0.050	0.5398	1.5293	0.4893	1.4200
Sigma6	0.050	0.0373	0.9800	-0.00354	0.9182
Sigma7	0.050	0.1413	1.4574	0.0778	1.3571
Sigma8	0.050	0.0373	0.9800	-0.00354	0.9182
Sigma9	0.050	0.8492	2.3874	0.7676	2.2114

Output 4. Summaries of posterior distributions for the parameters in the model

The trace plots, omitted from the output, indicate good mixing for all parameters in the model. The posterior means of the variables X, M, and Y are not of interest in the computation of the mediated effect. Instead, the mediated effect is computed by using the elements from the covariance matrix of X, M, and Y.

In order to obtain the mediated effect, the elements of the posterior distribution of Sigma must be multiplied for each iteration, and then sorted in order to obtain the posterior distribution of *ab* from which the median and central credibility interval will be computed:

```
data abefmD; set output;
abefmD=(Sigma4/Sigma1)*(Sigma1*Sigma6-Sigma4*Sigma3)/(Sigma1*Sigma5-Sigma4**2);
run;
proc sort data=abefmD out=percentilesefmD;
by abefmD;
run;
proc univariate data=percentilesefmD ;
var abefmD;
output out=pefmD pctlpre=P_ pctlpts=2.5 50 97.5;
run;
```

The format of the final table containing the credibility interval and the median value of the posterior distribution of the mediated effect looks identical to the format of the output in the method where regression coefficients were assigned prior distributions. The discrepancy in the numerical results is due to the fact that slightly different information is encoded in the prior distributions for regression coefficients and the covariance matrix. This occurs because of two reasons: (1) the variance and covariances for the variable X are not specified in the method involving prior distributions for regression coefficients, and are specified in the method involving prior distributions for the means and covariance matrix, and (2) the prior information for the overlapping parameters in the two methods has not been equated in this example.

P_2_5	P_50	P_97_5
-0.035865046	0.0769028919	0.2705150395

Table 4. Lower 95% central credibility limit, median value of the mediated effect, and upper 95% credibility limit for the covariance matrix method with diffuse prior distributions

Given that zero is in the credibility interval, one can conclude that the mediated effect was not statistically significant. As with the previous method, it is possible to obtain 99%, 90% and other types of central credibility intervals for the mediated effect by simply changing the values of the lower and upper percentiles SAS is computing.

Informative prior distribution for the covariance matrix

This approach can be used when prior information is available. From the previous study, it is known that the means of X, M, and Y are 70.40, 3.35, and 3.60 (respectively). It is also known that the covariance matrix between the variables X, M, and Y was obtained from a sample of 20, meaning that the degrees of freedom parameter for the

inverse Wishart prior distribution for the covariance matrix of X, M, and Y will equal 20. The observed covariances between X, M, and Y in the previous study were the following:

	X	M	Y
X	0.778947368	0.273684211	0.063157895
M	0.273684211	1.397368421	0.673684211
Y	0.063157895	0.673684211	1.094736842

Table 5. Covariance matrix of X, M, and Y from a previous study

In order to use the prior information about the covariance matrix, it needs to be converted into the sums of squares and cross-products matrix by multiplying each element by 19, the degrees of freedom. The sums of squares and cross-products matrix is then used as the S matrix in the code below.

When prior information about the means and covariance values for the variables X, M, and Y is available, the code changes to reflect what is known about parameters *mu* and *Sigma*:

```
proc mcmc data=sample outpost=output thin=1 seed=2 nmc=50000 mchistory=detailed
plots(unpack)=trace;
array data[3] X M Y;
array mu[3];
array Sigma[3,3];
array mu0[3] (70.40 3.35 3.60);
array Sigma0[3,3] (1000 0 0 0 1000 0 0 0 1000);
array S [3,3]
(14.799999999 5.200000009 1.200000005
5.200000009 26.5512.80000001
1.200000005 12.80000001 20.8);
parms mu Sigma;
prior mu ~ mvn(mu0, Sigma0);
prior Sigma ~ iwish(20 , S);
model data ~ mvn(mu, Sigma);
run;
```

Similar to the prior analysis, in order to obtain the posterior distribution of the mediated effect, one has to multiply the correct elements of the Sigma matrix for each iteration. Based on this posterior distribution of the mediated effect *ab*, one can obtain the median mediated effect and the credibility interval.

P_2_5	P_50	P_97_5
0.0169123879	0.1239834521	0.2946883078

Table 6. Lower 95% central credibility limit, median value of the mediated effect, and upper 95% credibility limit for the covariance matrix method with informative prior distributions

Given that zero is not in the credibility interval, one can conclude that the mediated effect was statistically significant.

CONCLUSION

Depending on the researcher's preference and the available prior information, there are two intuitive ways to run a Bayesian mediation analysis in PROC MCMC, as outlined by this paper. One method uses prior information on parameters of the mediation relations and the other method uses prior information on the covariance matrix of X, M, and Y. Furthermore, PROC MCMC offers an easy way to vary the amount of prior information in the analysis by changing the specifications in the prior statements. Another advantage of PROC MCMC is that the approach could be extended to more complicated models such as models with moderators and covariates. Relatedly, this PROC MCMC approach could be extended for estimation of other mediation quantities including effect size measures and causal mediated effects (Daniels, Roy, Kim, Hogan, & Perri, 2012). PROC MCMC is an intuitive tool for studying and applying Bayesian mediation models.

REFERENCES

- Chen, F. (2009). Bayesian modeling using the MCMC procedure. In *Proceedings of the SAS Global Forum 2008 Conference, Cary NC: SAS Institute Inc.*
- Chen, F. (2011). The RANDOM statement and more: moving on with PROC MCMC. In *Proceedings of the SAS Global Forum 2011 Conference, Cary NC: SAS Institute Inc.*
- Daniels, M.J., Roy, J. A., Kim, C., Hogan, J. Wl, & Perri, M.G. (2012). Bayesian inference for the causal effect of mediation. *Biometrics*, 68, 1028-1036.
- Enders, C. K., Fairchild, A. J., & MacKinnon, D. P. (2013). A Bayesian Approach for Estimating Mediation Effects With Missing Data. *Multivariate Behavioral Research*, 48(3), 340-369.
- Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2013). *Bayesian data analysis*. CRC press.
- Greenland, S. (2006). Bayesian perspectives for epidemiological research: I. Foundations and basic methods. *International Journal of Epidemiology*, 35(3), 765-775.
- James, L. R., & Brett, J. M. (1984). Mediators, moderators, and tests for mediation. *Journal of Applied Psychology*, 69(2), 307-321.
- Judd, C. M., & Kenny, D. A. (1981). Process analysis: Estimating mediation in treatment evaluations. *Evaluation Review*, 5(5), 602-619.
- Kruschke, J. (2010). *Doing Bayesian data analysis: a tutorial introduction with R*. Academic Press.
- MacKinnon DP. Analysis of mediating variables in prevention and intervention research. In: Cazares A, Beatty LA, editors. *Scientific methods in prevention research*. U.S. Government Printing Office; Washington, DC: 1994. pp. 127–153. NIDA Research Monograph 139. DHHS Publication No. 94-3631.
- MacKinnon, D.P. (2008). *Introduction to statistical mediation analysis*. New York, NY: Erlbaum.
- Metropolis, N. (1987). The beginning of the Monte Carlo method. *Los Alamos Sci.* 15 Special Issue 125–130.
- Yuan, Y., & MacKinnon, D. P. (2009). Bayesian mediation analysis. *Psychological Methods*, 14(4), 301-322.

ACKNOWLEDGMENTS

This research was supported in part by the National Institute on Drug Abuse, Grant No. R01DA009757.

RECOMMENDED READING

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

CONTACT INFORMATION

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

Name: Milica Miočević
Organization: Arizona State University
Address: NA
City, State ZIP: NA
Work Phone: NA
Fax: NA
Email: mmiocevi@asu.edu
Web: <http://ripl.faculty.asu.edu/about/current/sascommunity.org>: (Provide your home page or presenter's page)

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.