



www.quant.ku.edu

KUANT Guides

Guide No.
KUANT 012.1

Monte Carlo Simulation in Mplus:

Lee. J. (2008)

Preliminaries

Monte Carlo simulation studies are often used to investigate the performance of statistical estimators under various conditions. They can also be used to decide on the sample size needed for a study and to determine power (Muthén & Muthén, 2002).

Mplus provides extensive Monte Carlo simulation facilities for both data generation and analysis. The below example is for a simple SEM model in which two latent variables are measured by each three indicator variables. Given the sample size and the population parameter values, this example estimates the power to reject a zero loading and the Type I error to reject a zero correlation between the two latent variables.

Syntax

TITLE:

```
Monte Carlo Example
```

MONTECARLO:

```
NAMES = i1-i6;  
NOBSERVATIONS = 500;  
NREPS = 2000;  
SEED = 12345;
```

MODEL POPULATION:

```
PA BY i1*.6 i2*.6 i3*.0;  
NA BY i4*.6 i5*.6 i6*.0;  
PA@1 NA@1;  
PA WITH NA*.0;  
i1*.64 i2*.64 i3*.64;  
i4*.64 i5*.64 i6*.64;
```

MODEL:

```
PA BY i1*.6 i2*.6 i3*.0;  
NA BY i4*.6 i5*.6 i6*.0;  
PA@1 NA@1;  
PA WITH NA*.0;  
i1*.64 i2*.64 i3*.64;  
i4*.64 i5*.64 i6*.64;
```

OUTPUT:

```
TECH9;
```

Commands

1. The **MONTECARLO** Command

This command describes the details of a Monte Carlo simulation study.

The **NOBSERVATIONS** option specifies the sample size to be used for data generation and for analysis.

The **NREPS** option specifies the number of replications, that is, the number of samples to draw from a specified population.

The **SEED** option specifies the seed to be used for the random draws.

2. The **MODEL POPULATION** command

This command provides the population parameter values to be used in data generation. Each parameter must be specified by the @ symbol (for fixed parameters) or the * symbol (for free parameters), followed by its population value.

3. The **MODEL** command

This command describes the analysis model. It also provides the population parameter values for computing coverage and starting values in the estimation of the model.

Output

		ESTIMATES			S. E. M. S. E.		95%	% Sig
	Population	Average	Std. Dev.	Average		Cover	Coeff	
PA	BY							
I1		0.600	0.6009	0.0565	0.0584	0.0032	0.955	1.000
I2		0.600	0.5996	0.0597	0.0583	0.0036	0.941	1.000
I3		0.600	0.5976	0.0589	0.0582	0.0035	0.951	1.000
NA	BY							
I4		0.600	0.5979	0.0602	0.0582	0.0036	0.942	1.000
I5		0.600	0.5996	0.0588	0.0583	0.0035	0.952	1.000
I6		0.600	0.6000	0.0607	0.0584	0.0037	0.945	1.000
PA	WITH							
NA		0.000	-0.0004	0.0722	0.0711	0.0052	0.943	0.057

Interpretation

The **Population** column gives the population parameter values given in the MODEL command.

The **Average** column gives the average of the parameter estimates across the replications.

The **Std. Dev.** column gives the standard deviation of the parameter estimates across the replications. When the number of replications is large, this is considered to be the population standard error.

The **S. E. Average** column gives the average of the estimated standard errors across the replications.

The **M. S. E.** column gives the mean square error for each parameter.

The **95% Cover** column gives the proportion of replications for which the 95% confidence interval contains the population parameter value.

The **% Sig Coeff** column gives the proportion of replications for which the null hypothesis that a parameter is equal to zero is rejected at the .05 level. For parameters with population values different from zero, this value is an estimate of power with respect to a single parameter. For parameters with population values equal to zero, this value is an estimate of Type I error.

In this example, the power to reject a (false) zero loading is greater than .80 (i.e., $1 - \beta = 1.00$) but the Type I error to reject a (true) zero correlation is not less than .05 (i.e., $\alpha = 0.057$).