# Using R as a Wrapper in Simulation Studies

Tim F Liao
University of Illinois

## Outline

- Introduction
- Procedure for External Calls
- R Code
- An Example
- Conclusion

## Introduction

- Monte Carlo simulation is a useful method for assessing statistical robustness
- Various specialty statistical software packages are popular among data analysts
- These programs are often limited in their capacities to perform simulation
- R as a general-purpose program that is approaching the status of lingua franca, has much to offer
- R can be used as a wrapper to call external programs to carry out simulation

## Procedure

- A: Initial call to run an external program
  - Run a LCA model to simulate data
  - Estimate a model of simulated data
- B: Collect necessary output
  - Check if output read is indeed output wanted
  - Collect output in a single data matrix
- C: Monte Carlo simulation; repeat A & B a large number of times
- Conduct post-simulation analysis of the output
  - Draw violin plots of parameter estimates
  - Any other analyses

# R Code

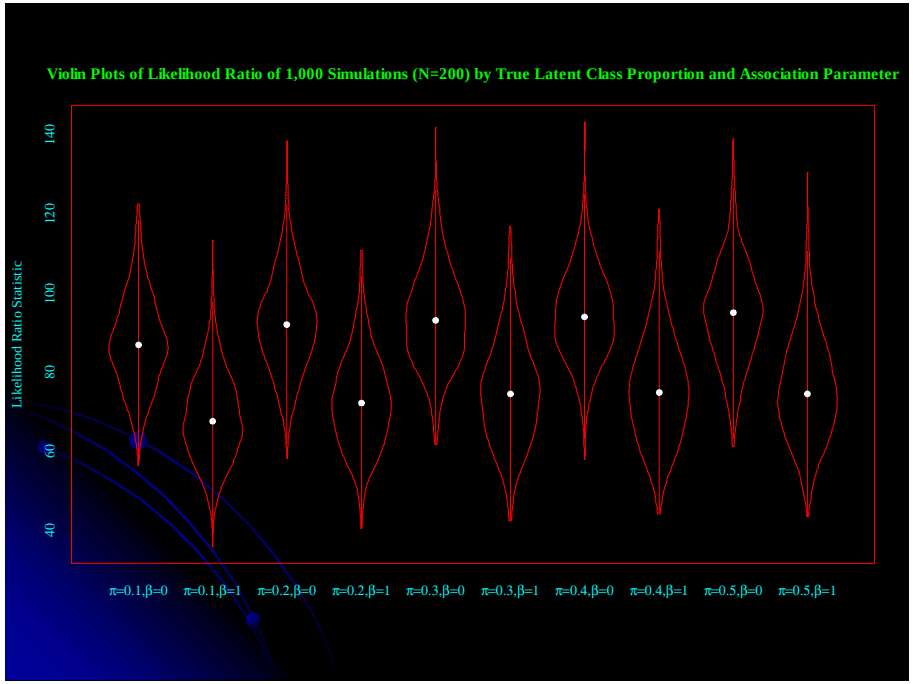
```
#test run of 1000 simulations
T <- 1000
C <- 30
zM<-matrix(rep(0,C*T),nrow=T)
for(i in (1:T)) {
  #the zM simulations
  system("c:/temp/simlem/lem c:/temp/simlem/SimM1a.inp
   c:/temp/simlem/SimM.out")
  system("c:/temp/simlem/lem c:/temp/simlem/Sim5E1.inp
   c:/temp/simlem/Sim5E.out")
  a1<-read.table("c:/temp/simlem/simla.out")[4,]
  … …
  b4<-read.table("c:/temp/simlem/simla.out")[17,]
  b5<-read.table("c:/temp/simlem/simla.out")[21,]
  a <- a1+a2+a3+a4+a5
  b <- b1+b2+b3+b4+b5
  if (a<b) {
    zM[i,1]<-read.table("c:/temp/simlem/simla.out")[2,]
  }
  else {
    zM[i,1]<-read.table("c:/temp/simlem/simla.out")[3,]
  }
 … …
}
```

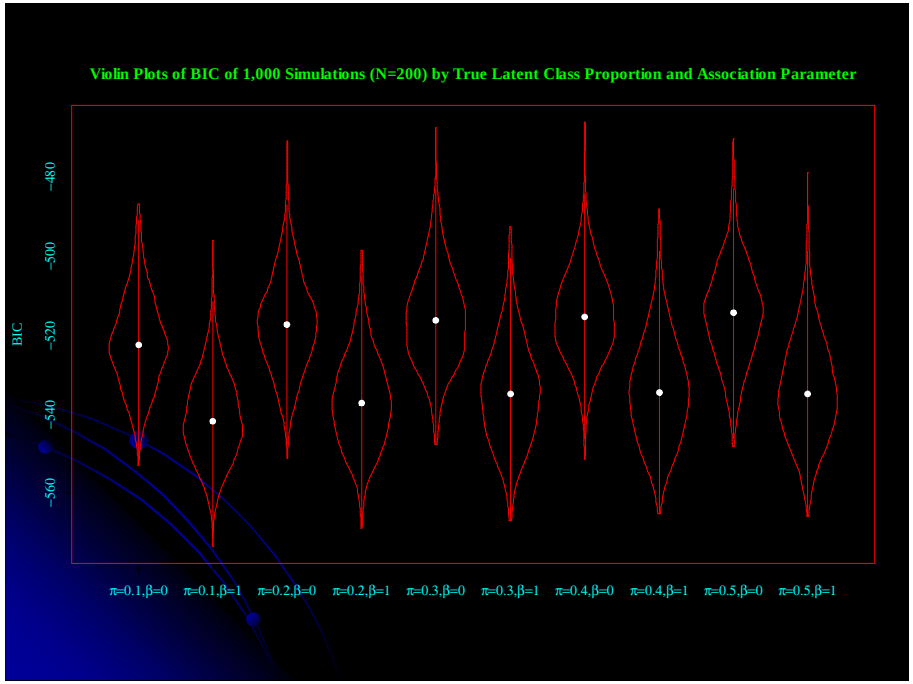
# Example: Latent Class Analysis

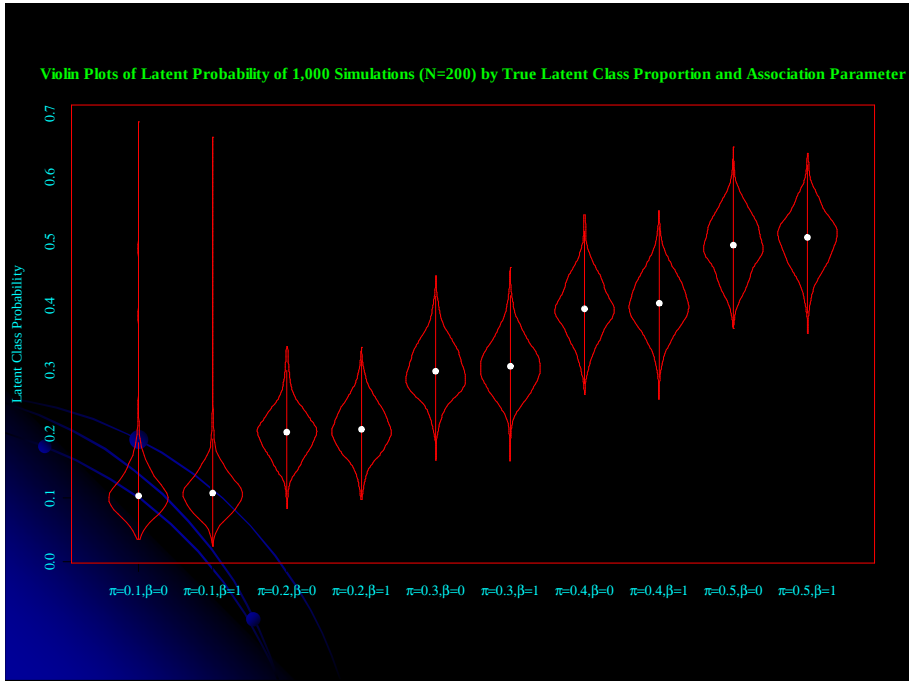
- Assuming for the data the model structure:

$$\log m_{xabcder^*} = u + u_x^X + u_a^A + u_b^B + u_c^C + u_d^D + u_e^E + \sum_j u_j^{R^*} + u_{xa}^{XA} + u_{xb}^{XB} + u_{xc}^{XC} + u_{xd}^{XD} + u_{xe}^{XE} + \beta_x \left(V_{r^*} - \bar{V}\right)$$

- Varying
  - presence of non-naïve MNAR association
  - true latent class proportions
  - sample size
- Using the model to estimate






# Conclusion

- R is the most flexible in facilitating simulation using existing specialty software
- The collection of necessary output can only be ad hoc
- The procedure can apply to other popular statistical software packages such as MPlus, which can also run in dos mode