datasim.R

Wolf

Sun May 08 15:57:22 2016

# load dependencies  
library(mirt)

## Loading required package: stats4

## Loading required package: lattice

# import item parameters  
ipar <- read.table("resources/ipar.dat", header=T, row.names=1,  
 stringsAsFactors=F)  
  
# import correlation matrix   
fcorr <- read.table("resources/fcorr.dat", header=T, row.names=1)  
  
# Cholesky decomposition of the facet correltations to obtain weights  
# for data simulation  
w <- t(chol(fcorr))  
  
# Specify the number of variables. Also serves as a sanity check,  
# should equal 10 (assuming 10 facets in the dataset).  
nvar <- dim(w)[1]  
  
# Specify the number of observations each data set should have  
nobs <- 500  
  
# How many datasets to make?  
nds <- 5000  
  
# set seed  
s = 1234  
  
## other parameters  
# identify slope parameters  
a <- ipar[,1] # item slopes  
  
# Match slope parameters to their respective facets  
tmp <- matrix(rep(0,1000), nrow=100, ncol=10)  
for(i in 1:10) {  
 if(!i==1) {  
 x <- ((i-1)\*10)+1  
 } else x <- i  
 tmp[x:(x+9),i] <- a[x:(x+9)]  
}  
a <- tmp  
  
# item thresholds  
d <- as.matrix(ipar[,2:7])  
  
#while() {  
 ### Make data sets  
 ## Generate a set of factor scores  
 # set seed  
 set.seed(s)  
  
 # Generate a matrix of random numbers  
 r <- matrix(rnorm(nvar\*nobs,0,1), nrow=nvar, ncol=nobs)  
  
 # Compute the matrix product of w and r  
 f <- w %\*% r  
  
 # transpose d  
 f <- t(f)  
  
 # Cast d to a data frame  
 f <- as.matrix(f)  
  
 ## Generate item level data  
 nd <- simdata(a, d, itemtype="graded", Theta=f, mins=1)  
 colnames(nd) <- rownames(ipar)  
   
 write.table(nd, "cyw/simd1.dat", row.names=F)  
   
 #increment the seed  
 s <- s + 1  
#}  
  
  
  
  
  
## insert careless respondents  
  
## export dataset to file  
  
## i++