
Generating Multivariate Random Associated Data

Part 2

This is the second example to generate multivariate random associated data. This example shows how to generate ordinal, categorical, data. It is a little more complex than generating continuous data in that the correlation matrix and the marginal distribution is required.

This example creates a 2-dimensional dataset. However, this can easily be extended to multiple variables. The correlation matrix R for this 2-dimensional example.

$$R = \begin{bmatrix} 1 & -0.6 \\ -0.6 & 1 \end{bmatrix} \quad (1)$$

The R code below will generate an ordinal dataset with a correlation matrix of

$$R = \begin{bmatrix} 1 & -0.5469243 \\ -0.5469243 & 1 \end{bmatrix} \quad (2)$$

Increasing the sample size will let the correlation coefficients converge on the target correlations.

```
library(GenOrd);
set.seed(1);
##Sets the marginals.
##The values are cumulative so for the first variable the first marginal will be .1, the second is .2.
marginal = list(c(0.1,0.3,0.6),c(0.4,0.7,0.9));

##Checks the lower and upper bounds of the correlation coefficients.
corrcheck(marginal);

##Sets the correlation coefficients
R = matrix(c(1,-0.6,-0.6,1),2,2); ##Correlation matrix

n = 500;

##Selects and ordinal sample with given correlation R and given marginals.
m = ordsample(n, marginal, R);

##compare it with the pre-defined R
cor(m);
```