Multivariate Normal Random Numbers

statgraphics 18®
centurion

Revised: 10/11/2017

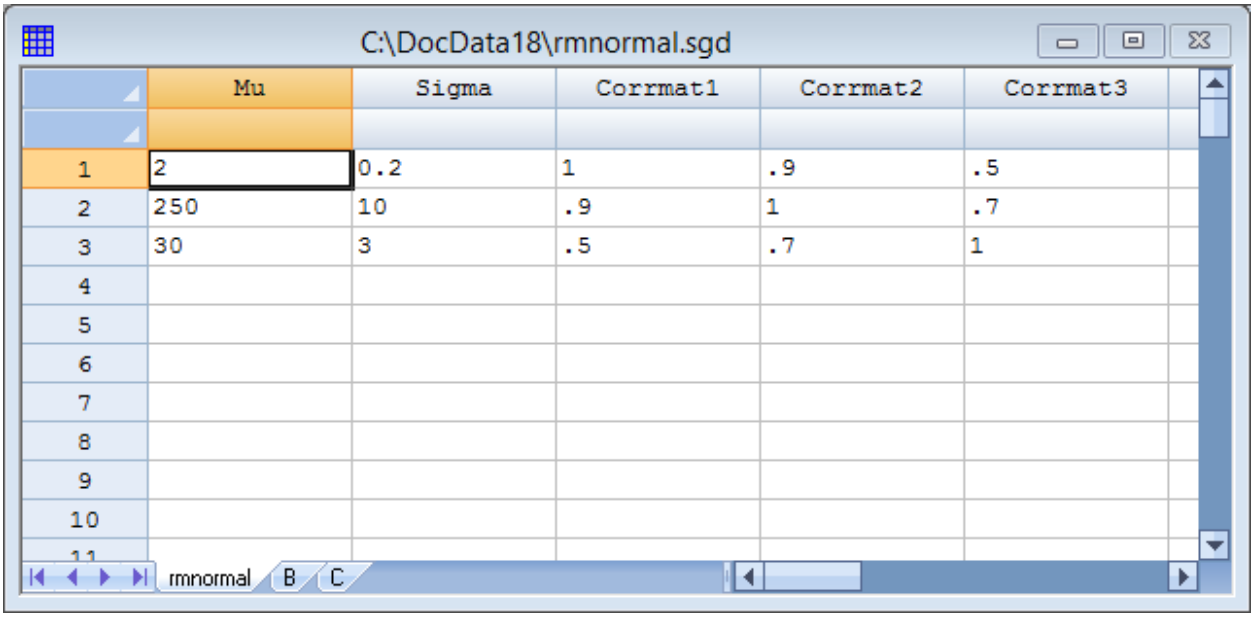
Summary

This procedure generates random numbers from a multivariate normal distribution involving up to 12 variables. The user inputs the variable means, standard deviations, and the correlation matrix. Random samples are generated which may be saved to the Statgraphics databook.

Sample StatFolio: *rmnormal.sgp*

Sample Data:

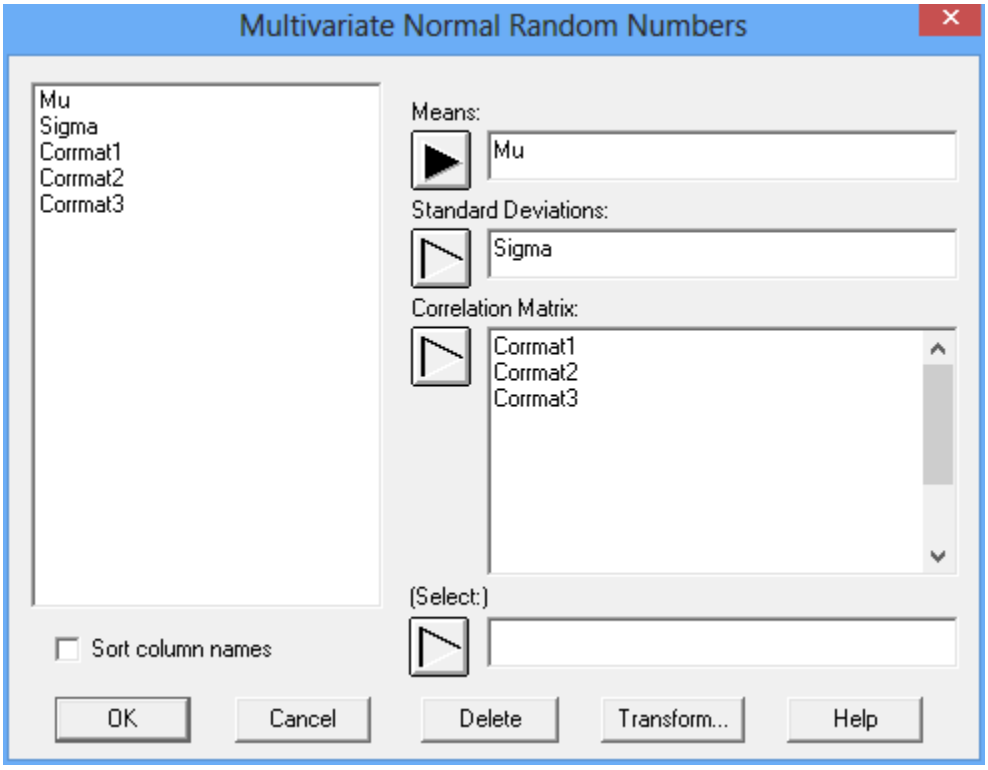
Suppose n = 200 random samples are desired from a trivariate normal distribution. The file *rmnormal.sgd* contains the desired means, standard deviations, and correlation matrix:



The first 3 rows correspond to the 3 variables for which random numbers are desired. The first 2 columns contain the means and standard deviations. The last 3 columns contain the correlation matrix.

Data Input

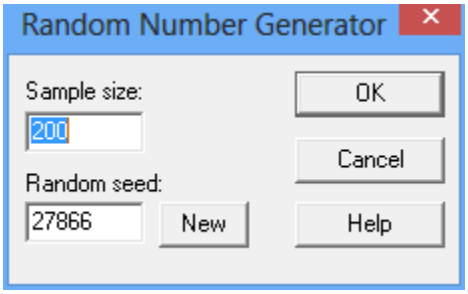
To generate random numbers for *m* variables, choose *Multivariate Normal Random Numbers* from the *Tools* menu. The data input dialog box is shown below:



- **Means:** numeric column containing the means for the *m* variables.
- **Standard Deviations:** numeric column containing the standard deviations for the *m* variables.
- **Correlation Matrix:** *m* numeric columns containing the correlation matrix.
- **Select:** subset selection.

Analysis Options

After specifying the variables, an *Analysis Options* dialog box is displayed:



- **Sample size**: *n*, the number of random observations desired.
- **Random seed**: an integer between 1 and 32767 which initializes the pseudo-random number generator. Each unique seed generates a different set of random numbers.
- **New**: push this button to generate a new seed.

Analysis Summary

The *Analysis Summary* shows input selections and the sample statistics for the generated random numbers:

Multivariate Normal Random Numbers

Variable	*Mean*	*Standard Deviation*	*Sample Mean*	*Sample Standard Deviation*
X1	2.0	0.2	1.9823	0.184928
X2	250.0	10.0	248.498	10.0603
X3	30.0	3.0	29.6177	3.15069

Correlations

	X1	X2	X3
X1	1.0	0.9	0.5
X2	0.9	1.0	0.7
X3	0.5	0.7	1.0

Sample Correlations

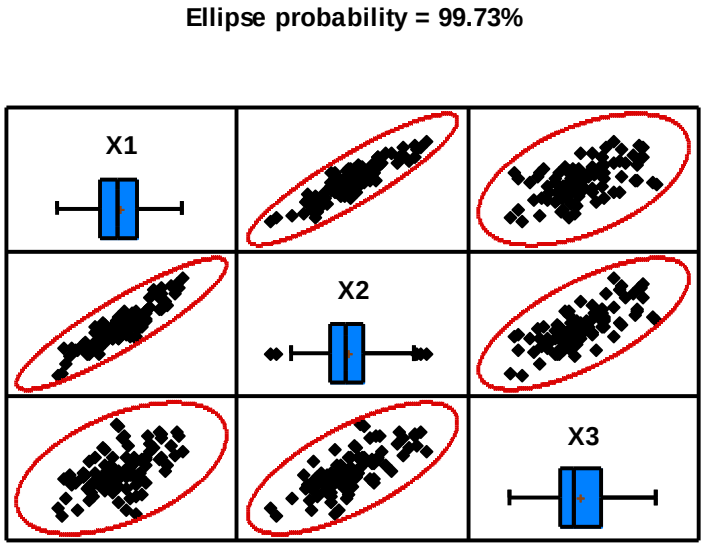
	X1	X2	X3
X1	1.0	0.903716	0.537067
X2	0.903716	1.0	0.732331
X3	0.537067	0.732331	1.0

Sample size: 100
Seed for random number generator: 17571

An error message will be displayed if the input means, standard deviations or correlation matrix is not valid.

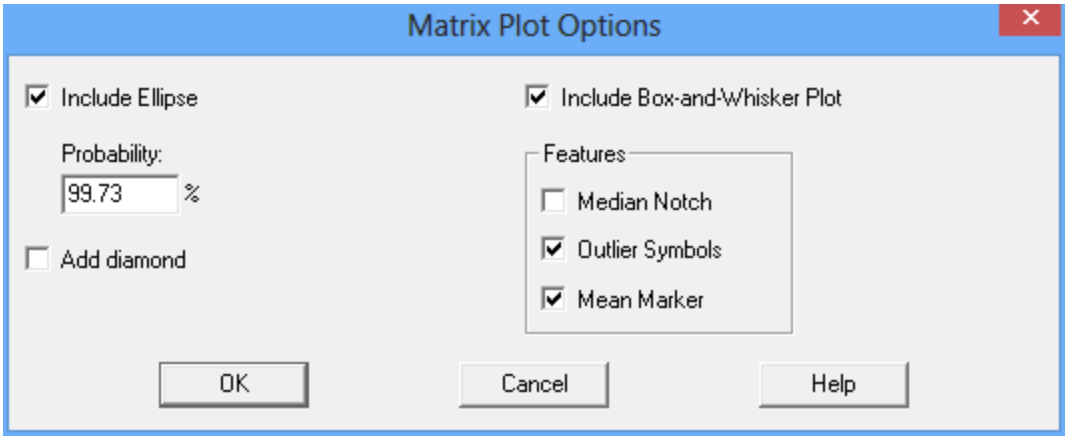
Matrix Plot

This pane displays the generated random numbers in the form of a matrix plot:



Each cell of the matrix is a scatter plot for 2 of the variables. If desired, a box and whisker plot may be added to the diagonal cells of the matrix. In addition, an elliptical region may be drawn containing a specified percentage of the probability for each marginal bivariate normal distribution.

Pane Options



- **Include Ellipse**: whether to include an ellipse for each pair of factors. Note that the ellipses are based on the input means, standard deviations and correlation matrix, not on the sample data.

- **Probability**: the probability included within the ellipse.

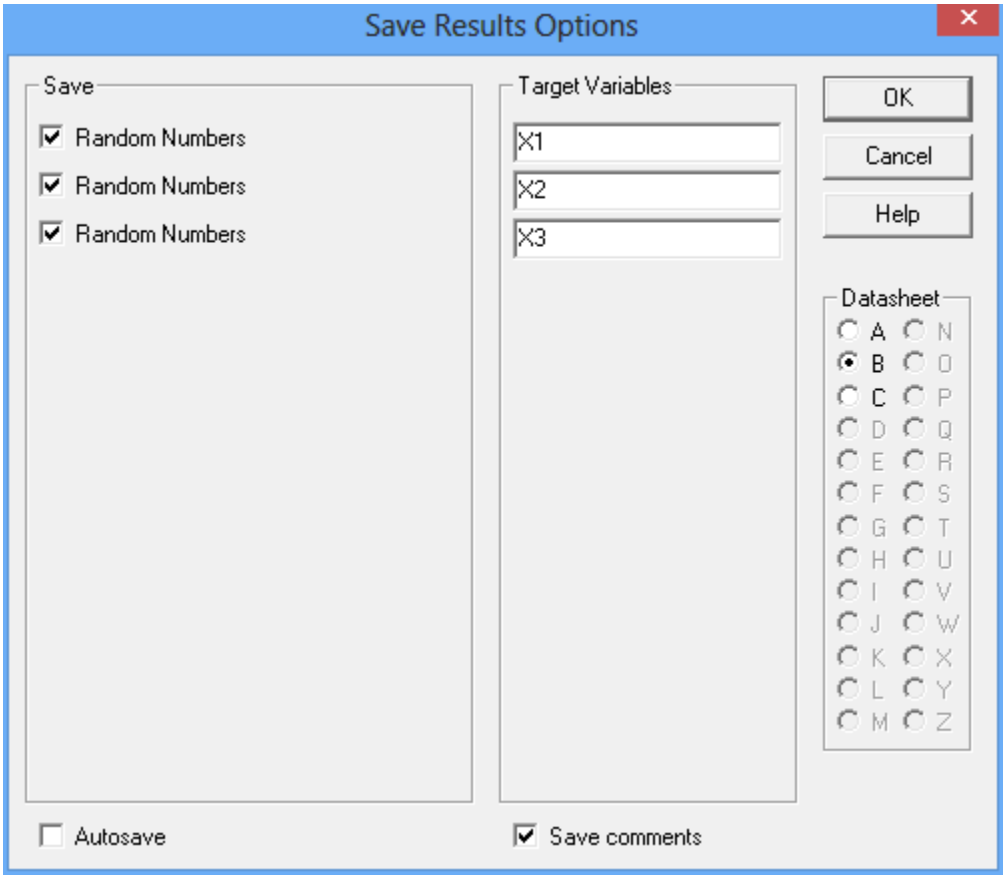
- **Include Box-and-Whisker Plot**: whether to include box-and-whisker plots in the display.

- **Features**: the box-and-whisker plots may include a notch to indicate a 95% confidence interval for the median, outlier symbols to indicate the presence of outside points, and/or a plus sign to indicate the location of the sample mean.

- **Add diamond:** if selected, a diamond will be added to the plot showing a 100(1-α)% confidence interval for the mean at the default system confidence level.

Save Results

To save the generated random numbers to the Databook, press the *Save Results* button on the analysis toolbar. This displays the dialog box shown below:



- **Save**: check all of the checkboxes.
- **Target variables**: select names for the columns in which the random numbers will be saved.
- **Datasheet**: select the sheet in which the data will be saved.
- **Autosave**: check this box if you wish the random numbers to be resaved automatically whenever a new set is generated.

Calculations

The multivariate random numbers are created by first generating *m* random numbers Z from a standard normal distribution. The desired random numbers are obtained from

$$X = \mu + AZ$$

where A is an *m* by *m* matrix such that $AA^T = \Sigma$ (the specified covariance matrix).

The ellipses are drawn using the marginal bivariate normal distribution for each pair of variables. The enclosed region satisfies the equation

$$(X - \mu)^T \Sigma^{-1} (X - \mu) \leq \chi_2^2(p)$$

where X is a bivariate random variable, μ is a 2 by 1 vector of means, Σ^{-1} is the inverse of the 2 by 2 covariance matrix for the 2 variables, and $X_2^2(p)$ is the value of the chi-square distribution with 2 degrees of freedom at which the cumulative distribution function equals *p*.