Journal of Modern Applied Statistical Methods

Volume 5 | Issue 2 Article 15

11-1-2005

# Simulation of Non-normal Autocorrelated Variables

H.E.T. Holgersson
*Jönköping International Business School, Sweden*, thomas.holgersson@jibs.hj.se

Journal of Modern Applied Statistical Methods
November, 2006, Vol. 5, No. 2, 408- 416

1538 – 9472/06/$95.00

# Simulation of Non-normal Autocorrelated Variables

H.E.T. Holgersson
Department of Economics and Statistics
Jönköping International Business School

All statistical methods rely on assumptions to some extent. Two assumptions frequently met in statistical analyses are those of normal distribution and independence. When examining robustness properties of such assumptions by Monte Carlo simulations it is therefore crucial that the possible effects of autocorrelation and non-normality are not confounded so that their separate effects may be investigated. This article presents a number of non-normal variables with non-confounded autocorrelation, thus allowing the analyst to specify autocorrelation or shape properties while keeping the other effect fixed.

Key words: Autocorrelation, non-normality, confounding, robustness.

## Introduction

All statistical methods rely on assumptions to some extent. These assumptions, for example, may be that some moments are finite or that the variance is homogenous at all data points. Other assumptions involve normal distribution or independence. If some of the assumptions are violated then the expected properties of the method may no longer hold. For example, a statistical hypothesis test that requires independence of the data may seriously over reject under the null hypothesis if the data possess autocorrelation.

It is therefore important to investigate the robustness properties of statistical methods before they are applied to real data. Although modern computers are developing at a rapid pace, it has become increasingly popular to perform robustness studies of such assumptions by Monte Carlo simulations. However, when examining robustness to autocorrelation and non-normality, some technical problems arise.

Because autocorrelation usually is generated by a recursive sequence of random numbers, the central limit theorem will force the autocorrelated variable to be more normal when compared to the variable used to generate the sequence. For example, imagine the problem of investigating the robustness of a non-normality test to autocorrelation. If a skewed variable is used to recursively generate a skewed and autocorrelated variable, then this new variable will be more symmetric than the original one and will be more symmetric the larger the autocorrelation is. Thus, the simulation study will not reveal the separate effects of autocorrelation and non-normality as was intended.

Several such examples are to be found in the literature. For example, Shukur (2000) examined the robustness of an autocorrelation test to non-normality by generating a first order autoregressive process with non-normal disturbances and Bai and Ng (2005) applied the first order autoregressive process with non-normal disturbances to investigate the robustness of a non-normality test to autocorrelation. Such effects of confounding can be avoided by using alternative methods of generating the variables.

H. E. T. Holgersson is Associate Professor at Jönköping International Business School with research and teaching. His research has mainly been focused on assessing distributional properties. Research also includes multivariate analysis and financial statistics.

A number of variables are proposed that are non-normal but autocorrelated and allow for separate control of the shape properties (skewness/kurtosis) and autocorrelation. All proposed variables are easy to generate in standard software packages.

Non-Normal And Autocorrelated Variables

The question of how to simulate random variables with given distributional properties have been given a great deal of attention (see Johnson (1987) for a general description). One of the most frequently used methods is the so-called inverse method. Assume the problem consists of generating a variate $X$ whose distribution is specified by $F$ and that $F$ is strictly increasing with inverse function $F^{-1}$. Then, the inverse method consists of first generating a uniformly distributed variate $U$ and then calculating the variable of interest by $X = F^{-1}(U)$. This method is fast and simple because all statistical software provides facilities for generating uniformly distributed variates.

Unfortunately, the issue becomes much more complicated when generating autocorrelated variables with given distribution because the mapping $U \mapsto X$ usually will change the autocorrelation pattern of $U$ drastically. Furthermore, it is not an easy task to generate autocorrelated variates which are uniformly distributed. This suggests that the inverse method is not very useful for the problem of generating autocorrelated variates with given marginal distribution, and so other methods are usually applied for that purpose. In particular, such variates are frequently generated by finite order ARMA models, often the AR(1) process, with given distribution of the disturbances. Unfortunately, this method will not result in the distribution aimed at. This can be seen from the following: Consider the linear process defined by

$$Y_t - \mu_Y = \sum_{i=0}^{\infty} \psi_i \delta_{t-i} \tag{2.1}$$

where $\delta_t$ are some zero mean independently identically distributed variables and the $\psi_i$´s are constants such that $\sum_{i=0}^{\infty} |\psi_i| < \infty$. Without loss of generality, assume that $\mu_Y = 0$ and $Var(\delta_t) = 0 < \sigma_\delta^2 < \infty$. The sequence (2.1) is known as the infinite order moving average process, also referred to as a linear process, and encompasses all stationary variables according to Wold´s decomposition theorem. It is seen directly from (2.1) that because the right hand is the sum of a (possible infinite) number of variables, The left hand side will in general be at least as normally distributed as the $\delta_t$ due to the central limit theorem. For example, the AR(1) process defined by

$$Y_t = \phi Y_{t-1} + \delta_t \tag{2.2}$$

is a special case of (2.1) where $\psi_i = \phi^i$. Then it follows that the distribution of $Y_t$ will be more normal when compared to $\delta_t$ for $\phi \neq 0$. Now, it is not generally true that $|\psi_i| > 0 \ \forall_i$. An example is the finite order MA($q$) model.

However, it is readily seen that already the skewness of an MA(1) process is closer to that of a normal distribution when compared to the skewness of $\delta$. In fact, it may be shown that for a process defined by

$$Y_t = \delta_t + \theta \delta_{t-1} \tag{2.3}$$

the skewness of $Y_t$ is given by

$$\beta_{1,Y} := E\left[Y - E[Y]\right]^3 \Big/ \left(E\left[Y - E[Y]\right]^2\right)^{3/2}$$
$$= \beta_{1,\delta} \frac{\left(1+\theta^3\right)}{\left(1+\theta^2\right)^{3/2}}$$

where $\beta_{1,\delta}$ is the skewness coefficient of $\delta$ (see Appendix). Hence, if for example $\beta_{1,\delta} = 5$ and $\theta = -0.7$ the skewness becomes

$\beta_{1,Y} = 5\left(1-0.7^3\right)\Big/\left(1+0.7^2\right)^{3/2} \approx 1.8$, which is less than half that of $\beta_{1,\delta}$. Thus, if one wishes to investigate the power properties of an autocorrelation test when applied to non-normal data, and applies model (2.3) for different values of $\theta$, then the effect of non-normality will be confounded with the power properties because the skewness is a direct function of the autocorrelation.

In light of the above discussion, one may wonder how autocorrelated and non-normal variables should be generated then. The fact that autocorrelation in general will smooth out non-normality suggests that autocorrelated normally distributed variables should be generated at a first step, and the non-normality should be imposed in the second step. Furthermore, the transformation to non-normality should result in a simple relation between the autocorrelation of the original variable and that of the transformed, thus allowing for total control of the autocorrelation pattern. In the following, that principle will be applied in a series of theorems that describe the generation of the variables and its properties.

Theorem 1.

Let $Y_t = \phi Y_{t-1} + \delta_t$ where $\delta_t \overset{iid}{\sim} N(0,1)$, $|\phi| < 1$, and define $\tilde{Y}_t := Y_t \Big/ \sqrt{\sigma_Y^2}$ where $\sigma_Y^2 = V(Y) = 1\Big/\left(1-\phi^2\right)$. Then $\tilde{Y}_t^2 \sim \chi^2_{(1)}$ independently of the value of $\phi$, and the autocorrelations of $\tilde{Y}_t^2$ is given by $\rho_{\tilde{Y}^2}(k) = \phi^{2k}$.

Proof of Theorem 1:

The variance of the AR(1) process is well known to be given by $\sigma_Y^2 = 1\Big/\left(1-\phi^2\right)$ and hence $\tilde{Y}_t \sim N(0,1)$ and the chi-square distribution of $\tilde{Y}_t^2$ follows directly. The autocovariances $\gamma_{\tilde{Y}^2}(k)$ of $\tilde{Y}^2$ are given by:

$$
\begin{aligned}
\gamma_{\tilde{Y}^2}(k) &= E\left[\tilde{Y}_t^2 \tilde{Y}_{t-k}^2\right] - E\left[\tilde{Y}_t^2\right] E\left[\tilde{Y}_{t-k}^2\right] \\
&= \sigma_Y^{-2\cdot 2} E\left[Y_t^2 Y_{t-k}^2\right] - 1\cdot 1 \\
&= \sigma_Y^{-2\cdot 2} E\left[\left(\phi Y_{t-1} + \varepsilon_t\right)^2 Y_{t-k}^2\right] - 1 \\
&= \sigma_Y^{-2\cdot 2} E\left[\left(\phi^2 Y_{t-1}^2 + 2\phi Y_{t-1}\varepsilon_t + \varepsilon_t^2\right) Y_{t-k}^2\right] - 1 \\
&= E\begin{bmatrix} \phi^2 \sigma_Y^{-2} Y_{t-1}^2 \sigma_Y^{-2} Y_{t-k}^2 \\ +2\phi\sigma_Y^{-2} Y_{t-1}\varepsilon_t \sigma_Y^{-2} Y_{t-k}^2 + \sigma_Y^{-2}\varepsilon_t^2 \sigma_Y^{-2} Y_{t-k}^2 \end{bmatrix} - 1 \\
&= \phi^2 E\left[\tilde{Y}_{t-1}^2 \tilde{Y}_{t-k}^2\right] \\
&\quad + \left\{-\phi^2 E\left[\tilde{Y}_{t-1}^2\right] E\left[\tilde{Y}_{t-k}^2\right] + \phi^2 E\left[\tilde{Y}_{t-1}^2\right] E\left[\tilde{Y}_{t-k}^2\right]\right\} \\
&\quad + 0 + \sigma_Y^{-2} - 1 \\
&= \phi^2 \left(E\left[\tilde{Y}_{t-1}^2 \tilde{Y}_{t-k}^2\right] - E\left[\tilde{Y}_{t-1}^2\right] E\left[\tilde{Y}_{t-k}^2\right]\right) \\
&\quad + \phi^2 E\left[\tilde{Y}_{t-1}^2\right] E\left[\tilde{Y}_{t-k}^2\right] + \sigma_Y^{-2} - 1 \\
&= \phi^2 \gamma_{\tilde{Y}^2}(k-1) + \phi^2 \cdot 1 \cdot 1 + \frac{1}{\left(1\Big/\left(1-\phi^2\right)\right)} - 1 \\
&= \phi^2 \gamma_{\tilde{Y}^2}(k-1).
\end{aligned}
$$

Furthermore,

$$
\begin{aligned}
\gamma_{\tilde{Y}^2}(0) &= E\left[\tilde{Y}_t^2 \tilde{Y}_t^2\right] - E\left[\tilde{Y}_t^2\right] E\left[\tilde{Y}_t^2\right] \\
&= 3 - 1\cdot 1 \\
&= 2
\end{aligned}
$$

By using recursion, it follows that that $\gamma_{\tilde{Y}^2}(k) = 2\phi^{2\cdot k}$. The autocorrelation function of $\tilde{Y}_t^2$ is thus determined by $\rho_{\tilde{Y}^2}(k) = \gamma_{\tilde{Y}^2}(k) \big/ \gamma_{\tilde{Y}^2}(0) = \phi^{2k}$ and Theorem 1 follows. In other words, the autocorrelation behaves like that of an AR(1) process with autoregressive parameter $\phi^2$ while the distribution is $\chi^2_{(1)}$. Also note that the shape property is independent of the autoregressive parameter.

Now, this variable is highly skewed and may not be appropriate in situations where near-normal distributions are required. The next theorem proposes a general $\chi^2_{(r)}$ distributed variable (which limits a normal distribution as $r$ increases) with non-confounded autocorrelation:

Theorem 2.

Let $Z_t := \sum_{j=1}^{r} \tilde{Y}_{jt}^2$ where $\tilde{Y}_{jt}, j = 1,2,...,r$ are mutually independent variables defined as in Theorem 1 with common autocorrelation parameter $\phi$. Then the skewness and kurtosis of $Z_t$ are given by $\beta_{1,Z} = \sqrt{8/r}$ and $\beta_{2,Z} = 3 + 12/r$ respectively, independently of $\phi$, and the autocorrelation of $Z_t$ is given by $\rho_Z(k) = \phi^{2k}$, independently of $r$.

Proof of Theorem 2:

The chi-square distribution follows directly from the fact that a sum of $r$ independent $\chi^2_{(1)}$ variates is distributed as $\chi^2_{(r)}$. The skewness and kurtosis of such variates are well known and can be found in Johnson et al (1994). The autocovariance is obtained by using the property $E\left[\tilde{Y}_t^2 \tilde{Y}_{t-k}^2\right] - 1 = \gamma_{\tilde{Y}^2}(k)$ (given in the derivation of Theorem 1):

$$
\begin{aligned}
\gamma_Z(k) &:= E[Z_t Z_{t-k}] - E[Z_t]E[Z_{t-k}] \\
&= E\left[\sum_{i=1}^{r} \tilde{Y}_{i,t}^2 \sum_{i=1}^{r} \tilde{Y}_{i,t-k}^2\right] - E\left[\sum_{i=1}^{r} \tilde{Y}_{i,t}\right] E\left[\sum_{i=1}^{r} \tilde{Y}_{i,t-k}\right] \\
&= E\left[\sum_{i=1}^{r} \tilde{Y}_{i,t}^2 \tilde{Y}_{i,t-k}^2 + \sum_{i\neq j}^{r} \tilde{Y}_{i,t}^2 \tilde{Y}_{j,t-k}^2\right] - r \cdot r \\
&= \sum_{i=1}^{r} E\left[\tilde{Y}_{i,t}^2 \tilde{Y}_{i,t-k}^2\right] + \sum_{i\neq j}^{r} E\left[\tilde{Y}_{i,t}^2\right] E\left[\tilde{Y}_{j,t-k}^2\right] \\
&= rE\left[\tilde{Y}_t^2 \tilde{Y}_{t-k}^2\right] + r(r-1) - r^2 \\
&= r\left(E\left[\tilde{Y}_t^2 \tilde{Y}_{t-k}^2\right] - 1\right) \\
&= r\gamma_{\tilde{Y}^2}(k) \\
&= r2\phi^{2k}.
\end{aligned}
$$

In particular, $\gamma_Z(0) = r2\phi^0$ and so $\rho_Z(k) = \phi^{2k}$ as was to be shown. In other words, the shape-part of the distribution of $Z_t$ is $\chi^2_{(r)}$ and the autocorrelation-part of the distribution is $\rho_Z(k) = \phi^{2k}$, and none of the effects is confounded to the other. Thus, the skewness and kurtosis can be determined over an arbitrary (though discrete) range of values independently of the autocorrelation. The next theorem proposes a symmetric non-normal variable with non-confounded autocorrelation:

Theorem 3.

Let $Z_t$ be defined as in Theorem 2. Also, let $c_t = \phi c_{t-1} + \varepsilon_t$ where $\varepsilon_t \overset{iid}{\sim} N(0,1)$, and define $\tilde{c}_t := c_t \Big/ \sqrt{1/(1-\phi^2)}$ and $W_t := \tilde{c}_t (Z_t - r)$. Then the skewness and kurtosis of $W_t$ are given by $\beta_{1,W} = 0$ and $\beta_{2,W} = 3(3 + 12/r)$ independently of $\phi$ while the autocorrelations are given by $\phi^{3k}$, independently of $r$.

Proof of Theorem 3:

Firstly, note that $\tilde{c}_t$ and $(Z_t - r)$ are mutually independent and that $E[W_t] = 0$. Hence, the skewness and kurtosis of $W_t$ are given by:

$$
\begin{aligned}
\beta_{1,W} &= E[W_t]^3 \Big/ E[W_t]^{3/2} \\
&= E\left[\tilde{c}_t^3\right] E\left[(Z-r)^3\right] \Big/ E[W_t]^{3/2} \\
&= 0,
\end{aligned}
$$

and

$$
\begin{aligned}
\beta_{W,2} &= E\left[W_t^4\right] \Big/ \left(E\left[W_t^2\right]\right)^2 \\
&= E\left[\tilde{c}_t^4\right] E\left[Z_t^4\right] \Big/ \left(E\left[\tilde{c}_t^2\right] E\left[Z_t^2\right]\right)^2 \\
&= \left\{E\left[\tilde{c}_t^4\right] \Big/ \left(E\left[\tilde{c}_t^2\right]\right)^2\right\} \left\{E\left[Z_t^4\right] \Big/ \left(E\left[Z_t^2\right]\right)^2\right\} \\
&= 3\{3 + 12/r\}.
\end{aligned}
$$

Hence, the $W$ variable is symmetric with kurtosis determined by $r$ with range $9 < \beta_{2,W} \leq 45$. Furthermore, by using the results of the proofs of Theorem 2 and Theorem 3, the autocovariances may be obtained:

$$\begin{aligned}
\gamma_W(k) &= E[W_t W_{t-1}] - 0 \\
&= E[\tilde{c}_t Z_t \tilde{c}_{t-k} Z_{t-k}] \\
&= E[\tilde{c}_t \tilde{c}_{t-k}] E[(Z_t - r)(Z_{t-k} - r)] \\
&= \phi^k E\{(Z_t Z_{t-k}) - rZ_t - rZ_{t-1} + r^2\} \\
&= \phi^k \{E(Z_t Z_{t-k}) - r^2\} \\
&= \phi^k \{rE[\tilde{Y}_t^2 \tilde{Y}_{t-k}^2] + r(r-1) - r^2\} \\
&= \phi^k r \{E[\tilde{Y}_t^2 \tilde{Y}_{t-k}^2] - 1\} \\
&= \phi^k r \gamma_{\tilde{Y}^2}(k) \\
&= \phi^k 2r\phi^{2k}.
\end{aligned}$$

Hence, $\gamma_W(k) = 2r\phi^{3k}$ and in particular $\gamma_W(0) = 2r$. Thus, the autocorrelations are given by $\rho_W(k) = 2r\phi^{3k}/2r = \phi^{3k}$ independently of $r$ as was to be shown. Thus, Theorem 3 provides a symmetric but non-normal variable where the autocorrelations can be identical to those of an AR(1) process by putting the autocorrelation parameter of the original variable equal to $\phi^{1/3}$.

In general, the variables presented in Theorems 1-3 share the property that they all have autocorrelations that decay slowly in the sense that they are non zero at all lags if $\phi \neq 0$. In many instances it is of interest to generate autocorrelations that are zero above a certain lag. Therefore, some short memory processes of MA(1) type will also be proposed. These are presented below:

Theorem 4.

Let $Y_t = \delta_t - \theta\delta_{t-1}$ where $\delta_t \overset{iid}{\sim} N(0,1)$ and define $\tilde{Y}_t := Y_t / \sqrt{\sigma_Y^2}$ where $\sigma_Y^2 = 1 + \theta^2$. Then, $\tilde{Y}_t^2 \sim \chi_{(1)}^2$ independently of $\phi$ and the autocorrelations are given by $\rho_{\tilde{Y}_t^2}(1) = \theta^2 / (1+\theta^2)^2$ $k = 1$ and $\rho_{\tilde{Y}_t^2}(k) = 0$ $k > 1$ independently of $r$.

Proof of Theorem 4.

The chi-square distribution follows trivially as $\tilde{Y}$ is a standard normally distributed variate. The second order moments are given by

$$\begin{aligned}
\gamma_{\tilde{Y}^2}(0) &= V(\tilde{Y}_t^2) = 2. \\
\gamma_{\tilde{Y}^2}(1) &= E[\tilde{Y}_t^2 \tilde{Y}_{t-1}^2] - E[\tilde{Y}_t^2] E[\tilde{Y}_{t-1}^2] \\
&= \left(1/(1+\theta^2)^2\right) E[Y_t^2 Y_{t-1}^2] - 1 \cdot 1 \\
&= \left(1/(1+\theta^2)^2\right) E\left[(\delta_t - \theta\delta_{t-1})^2 (\delta_{t-1} - \theta\delta_{t-2})^2\right] - 1 \\
&= \left(1/(1+\theta^2)^2\right) \left\{ \begin{matrix} E[\delta_t^2 \delta_{t-1}^2] + E[\delta_t^2 \theta^2 \delta_{t-2}^2] \\ + E[\theta^2 \delta_{t-1}^4] + E[\theta^4 \delta_{t-1}^2 \delta_{t-2}^2] \end{matrix} \right\} - 1 \\
&= (1 + 4\theta^2 + \theta^4) / (1+\theta^2)^2 - 1 \\
&= 2\theta^2 / (1+\theta^2)^2.
\end{aligned}$$

An analogous proof reveals that $\gamma_{\tilde{Y}^2}(k) = 0$ if $k > 1$. Hence the autocorrelations of $\tilde{Y}_t^2$ are given by $\rho_{\tilde{Y}_t^2}(1) = 2\theta^2 / 2(1+\theta^2)^2 = \theta^2 / (1+\theta^2)^2$ $k = 1$ and $\rho_{\tilde{Y}_t^2}(k) = 0$ $k > 1$ as was to be shown. Thus, the autocorrelations of $\tilde{Y}_t^2$ are those of an MA(1) process where the autocorrelation at lag 1 equals the root of $\theta^2 / (1+\theta^2)^2$ which, in turn, is bounded between 0 and 0.25 (the maximum being reached at $\theta = 1$). This variable may also be extended to an arbitrary $\chi_{(r)}^2$ distribution:

Theorem 5.

Let $Z_t := \sum_{j=1}^{r} \tilde{Y}_{jt}^2$ where $\tilde{Y}_{jt},\ j=1,2,...,r$ are mutually independent variables defined as in Theorem 4 with common autoregressive parameter $\theta$. Then the skewness and kurtosis of $Z_t$ are given by $\beta_{1,Z} = \sqrt{8/r}$ and $\beta_{2,Z} = 3 + 12/r$ independently of $\theta$, and the autocorrelations of $Z$ are given by $\rho_{\tilde{Y}_t^2}(1) = \theta^2 \big/ \left(1+\theta^2\right)^2 \ k=1$ and $\rho_{\tilde{Y}_t^2}(k) = 0 \ k>1$, independently of $r$.

Proof of Theorem 5.

The skewness and kurtosis are motivated in Theorem 2. Analogous to the proof of Theorem 2, $\gamma_Z(k) = r\left(E\left[\tilde{Y}_t^2 \tilde{Y}_{t-k}^2\right] - 1\right)$, and as the variance is given by $\gamma_{Z_t}(0) = Var\left[Z_t\right] = 2r$ it follows that the autocorrelations are given by

$$\rho_{\tilde{Y}_t^2}(1) = r\left\{2\theta^2 \big/ \left(1+\theta^2\right)^2\right\} \big/ 2r$$
$$= \theta^2 \big/ \left(1+\theta^2\right)^2 \ k=1$$

and $\rho_{\tilde{Y}_t^2}(k) = 0 \ k>1$ as was to be shown. The $Z$ variable may also be symmetrized according to the following:

Theorem 6.

Let $Z_t$ be defined as in Theorem 5. Also, let $c_t = \varepsilon_t - \theta\varepsilon_{t-1}$ where $\varepsilon_t \sim N(0,1)$, and define $\tilde{c}_t := c_t \big/ \sqrt{\left(1+\theta^2\right)}$ and $W_t := \tilde{c}_t\left(Z_t - r\right)$. Then, the skewness and kurtosis of $W_t$ are given by $\beta_{1,W} = 0$ and $\beta_{2,W} = 3\left(3 + 12/r\right)$ independently of $\theta$ and the autocorrelations of $W_t$ are given by $\rho_W(1) = -\theta^3 \big/ \left(1+\theta^2\right)^2,\ \rho_W(k) = 0 \ \ k>1$ independently of $r$.

Proof of Theorem 6.

The skewness and kurtosis are given in the proof of Theorem 3. The autocovariance of $W_t$ is

$$\begin{aligned}
\gamma_W(1) &= E\left[W_t W_{t-1}\right] \\
&= E\left[\tilde{c}_t \tilde{c}_{t-1}\right] E\left[\left(Z_t - r\right)\left(Z_{t-1} - r\right)\right] \\
&= -\theta\left[E\left(Z_t Z_{t-1}\right) - r^2\right] \\
&= -\theta\left(\left(Cov\left[Z_t Z_{t-1}\right] + r^2\right) - r^2\right) \\
&= -\theta r\left(E\left[\tilde{Y}_t^2 \tilde{Y}_{t-k}^2\right] - 1\right) \\
&= -\theta r \gamma_{\tilde{Y}^2}(k).
\end{aligned}$$

Hence,

$$\gamma_W(0) = 2r,\ \gamma_W(1) = -r\theta\left\{2\theta^2 \big/ \left(1+\theta^2\right)^2\right\},$$

$\gamma_W(k) = 0 \ k>1$ and the theorem follows. Hence, the autocorrelations behave like those of an MA(1) process with MA parameter determined by the root of $-\theta^3 \big/ \left(1+\theta^2\right)^2$ which is bounded in the interval $(-0.32, 0.32)$ with maximum at $\theta = \pm 1.73$.

The variables proposed above are all univariate. It will sometimes be of interest to generate multivariate variables with cross correlation between pairwise marginal variables. When imposing such cross correlation one wish to do that in a manner that does not alter the marginal distributions. This can be achieved by letting one or several variables used to form the marginal variables be identical (fixed) in all marginal variables. The next theorem describes such variables and its main properties:

Theorem 7.

Let $\tilde{Y}_{ijt},\ i=1,2,...,P,\ j=1,2,...,r$ be mutually independent variables defined as in Theorem 1 or Theorem 4, depending on whether the AR(1) or MA(1) process have been used to generate the autocorrelation, with common autocorrelation parameter $\phi$ (or $\theta$). Then define $Z_{i,t} := \sum_{j=1}^{h} \tilde{Y}_{1,j,t}^2 + \sum_{j=h+1}^{r} \tilde{Y}_{i,j,t}^2$, $i = 1,2,...,P$ and let $\mathbf{Z}_t = \left[ Z_{1,t}, ..., Z_{P,t} \right]$ be a random vector of the marginal variables $Z_{i,t}$. Then, the cross correlations between two marginal variables of $\mathbf{Z}_t$ are given by

$$Corr\left(Z_{i,t}, Z_{i',t}\right) = h/r, \quad i \neq i'.$$

Proof of Theorem 7.

Assume for the moment that $h = 1$. Then, on observing that each $Z_{i,t}$ is a chi square variate, the covariance becomes:

$$\begin{aligned}
&Cov\left(Z_{i,t}, Z_{i',t}\right) \\
&= E\left[Z_{i,t} Z_{i',t}\right] - E\left[Z_{i,t}\right] E\left[Z_{i',t}\right] \\
&= E\left[\left(\tilde{Y}_{1,1,t}^2 + \sum_{j=2}^{r} \tilde{Y}_{i,j,t}^2\right)\left(\tilde{Y}_{1,1,t}^2 + \sum_{j=2}^{r} \tilde{Y}_{i',j,t}^2\right)\right] - r^2 \\
&= E\begin{bmatrix} \left(\tilde{Y}_{1,1,t}^2\right)^2 + \tilde{Y}_{1,1,t}^2 \sum_{j=2}^{r} \tilde{Y}_{i',j,t}^2 \\ + \left(\tilde{Y}_{i,2,t}^2 \tilde{Y}_{1,1,t}^2 + \tilde{Y}_{i,2,t}^2 \sum_{j=2}^{r} \tilde{Y}_{i',j,t}^2\right) + \\ ... + \left(\tilde{Y}_{i,r,t}^2 \tilde{Y}_{1,1,t}^2 + \tilde{Y}_{i,r,t}^2 \sum_{j=2}^{r} \tilde{Y}_{i',j,t}^2\right) \end{bmatrix} - r^2 \\
&= 3 + 1 \cdot (r-1) + \left(1 \cdot 1 + 1(r-1)\right) + \\
&... + \left(1 \cdot 1 + 1(r-1)\right) - r^2 \\
&= 3 + (r-1) + r + ... + r - r^2 \\
&= 2.
\end{aligned}$$

Because the marginal $Z_{i,t}$ variables are distributed as $\chi_{(r)}^2$ it follows that the correlation is given by $Corr\left(Z_{i,t}, Z_{i',t}\right) = 2/2r = 1/r$. By applying an analogous proof for a general $0 \leq h \leq r$ it is seen that $Cov\left(Z_{i,t}, Z_{i',t}\right) = 2h$ and hence the cross correlation is given by $Corr\left(Z_{i,t}, Z_{i',t}\right) = 2h/2r = h/r$ which completes the proof.

In other words, Theorem 7 proposes a random vector with marginal variables of the kind described in Theorem 2 (or Theorem 5) though with cross correlations given by $h/r$. It is also possible to generate symmetric multivariate variables as shown in the next theorem:

Theorem 8.

Let $Z_t$ be defined as in Theorem 7 and let $W_t := \tilde{c}_t\left(Z_t - r\right)$ where $\tilde{c}_t$ is defined as in Theorem 3 (or Theorem 6). Then the cross correlation between $W_{i,t}$ and $W_{i',t}$ is given by $Corr\left(W_{i,t}, W_{i',t}\right) = h/r$ independently of $\phi$ (or $\theta$).

Proof of Theorem 8.

The cross covariance is given by

$$\begin{aligned}
&Cov\left(W_{i,t}, W_{i',t}\right) \\
&= E\left(W_{i,t}, W_{i',t}\right) - 0 \\
&= E\left(\tilde{c}_t\left(Z_{i,t} - r\right)\tilde{c}_t\left(Z_{i',t} - r\right)\right) \\
&= E\left[\tilde{c}_t^2\right] E\left(\left(Z_{i,t} - r\right)\left(Z_{i',t} - r\right)\right) \\
&= E\left(Z_{i,t} Z_{i',t}\right) - r^2 \\
&= \left(2h + r^2\right) - r^2 \\
&= 2h.
\end{aligned}$$

Finally,

$$\begin{aligned}
&Var\left(W_{i,t}\right) \\
&= E\left(W_{i,t}^2\right) - E^2\left(W_{i,t}\right) \\
&= E\left(\tilde{c}_t^2\left(Z_{i,t} - r\right)^2\right) - 0 \\
&= E\left(\tilde{c}_t^2\right) E\left(Z_{i,t} - r\right)^2 \\
&= 1 \cdot Var\left(Z_{i,t}\right) \\
&= 2r.
\end{aligned}$$

This completes the proof.

The variables proposed above allow for separate control of shape properties and autocorrelations. In particular, the variables formed by transformations of AR(1) processes (those of Theorems 1-3) behave like AR(1) processes and the domain of the autoregressive parameter remains $-1<\phi<1$. In other words these variables provide a tool for generating AR(1) processes with non-normal distributions. On the other hand, the domains of the autoregressive parameter (which is equal to the whole real line though usually kept in the span $-1<\theta<1$) of the MA(1) type variables (those of Theorems 4-6) do not remain when transformed to non-normality.

This might be a drawback in some instances though they do provide a tool for investigating the effect of non-normal short memory processes. In general, the proposed variables should cover most robustness problems met in practice. Further research in the matter could involve the development of generating non-normal processes with long memory of ARFIMA type with autocorrelations non-confounded with shape properties. Other relevant issues involve non- stationary processes with given shape properties. This, however, is beyond the scope of this article and is left for future studies.

## Conclusion

In this article, it is argued that care must be taken when simulating autocorrelated and non-normal variables so that the autocorrelation is not a function of the shape property and vice versa. Furthermore, a number of random variables specially designed for simulation studies concerning shape properties and autocorrelation are proposed. The variables involve univariate or multivariate distributions, which are symmetric or skewed and have short memory or long memory. Thus, they cover a fairly wide range of applications. Furthermore, all variables are easily generated in any standard statistical software packages that have facilities to generate AR or MA processes.

## References


Bai, J. & Ng, S. (2005). Tests for skewness, kurtosis, and normality for time series data. *Journal of Business & Economic Statistics, 23*(1), 49-60.

Johnson, M. E. (1987). *Multivariate statistical simulation*. Wiley.

Johnson, N. L., Kotz, S., & Balakrishnan, N. (1994). *Continous univariate distributions,* Vol. 1. Wiley.

Shukur, G. (2000), The robustness of the systemwise Breauch-Godfrey auto-correlation test for non-normal error terms. *Communications in Statistics, Simulation, & Computation* , 29(2), 419-448.


## Appendix

Let $\mu_{X,h} := E\left[X_t^h\right]$ $\mu_{\varepsilon,h} := E\left[\varepsilon_t^h\right]$ for $h=1,2,3$. Then, if $\varepsilon_t$ is a sequence of iid (zero mean) uniformly integrable variables, the following holds:

$$\begin{aligned}
\mu_{X,2} &= E\left[\left(\sum_{i=0}^{\infty}\theta_i\varepsilon_{t-i}\right)^2\right] \\
&= E\left[\sum_{i=0}^{\infty}\theta_i^2\varepsilon_{t-i}^2+\sum\sum_{i\neq j}\theta_i\,\varepsilon_{t-i}\theta_j\,\varepsilon_{t-j}\right] \\
&= \sum_{i=0}^{\infty}\theta_i^2 E\left[\varepsilon_{t-i}\right] \\
&= \mu_{\varepsilon,2}\sum_{i=0}^{\infty}\theta_i^2.
\end{aligned}$$

$$\begin{aligned}
\mu_{X,3} &= E\left[\left(\sum_{i=0}^{\infty}\theta_i\varepsilon_{t-i}\right)^3\right] \\
&= E\left[\sum_{i=0}^{\infty}\begin{matrix}\theta_i^3\varepsilon_{t-i}^3+3\sum\sum_{i\neq j}\theta_i^2\,\varepsilon_{t-i}^2\theta_j\varepsilon_{t-j}\\ +\sum\sum\sum_{i\neq j\neq k}\theta_i\,\varepsilon_{t-i}\theta_j\varepsilon_{t-j}\theta_k\varepsilon_{t-k}\end{matrix}\right] \\
&= \sum_{i=0}^{\infty}\theta_i^3 E\left[\varepsilon_{t-i}^3\right] \\
&= \mu_{\varepsilon,3}\sum_{i=0}^{\infty}\theta_i^3.
\end{aligned}$$

$$\mu_{X,4} = E\left[\left(\sum_{i=0}^{\infty}\theta_i\varepsilon_{t-i}\right)^4\right]$$
$$= E\left[\sum_{i=0}^{\infty}\theta_i^4\varepsilon_{t-i}^4 + 4\sum_{i\neq j}\sum\theta_i^3\,\varepsilon_{t-i}^3\theta_j\varepsilon_{t-j} + \right.$$
$$3\sum_{i\neq j}\sum\theta_i^2\,\varepsilon_{t-i}^2\theta_j^2\varepsilon_{t-j}^2 + 6\sum_{i\neq j\neq k}\sum\sum\theta_i^2\,\varepsilon_{t-i}^2\theta_j\varepsilon_{t-j}\theta_k\varepsilon_{t-k}$$
$$\left. +6\sum_{i\neq j\neq k\neq l}\sum\sum\sum\theta_i^2\,\varepsilon_{t-i}^2\theta_j\varepsilon_{t-j}\theta_k\varepsilon_{t-k}\theta_l\varepsilon_{t-l}\right]$$
$$= \sum_{i=0}^{\infty}\theta_i^4 E\left[\varepsilon_{t-i}^4\right]$$
$$+3\sum_{i\neq j}\sum\theta_i^2\,E\left[\varepsilon_{t-i}^2\right]\theta_j^2 E\left[\varepsilon_{t-j}^2\right]$$
$$= \mu_{\varepsilon,4}\sum_{i=0}^{\infty}\theta_i^4 + 3\,\mu_{\varepsilon,2}^2\sum_{i\neq j}\sum\theta_i^2\,\theta_j^2.$$

Hence, if an MA(1) process is determined by

$$X_t = \varepsilon_t + \theta\varepsilon_{t-1} \tag{A1}$$

it follows that the skewness is given by

$$\beta_{1,X} = \frac{\mu_{X,3}}{\mu_{X,2}^{3/2}}$$
$$= \frac{\mu_{\varepsilon,3}\sum_{i=0}^{\infty}\theta_i^3}{\left(\mu_{\varepsilon,2}\sum_{i=0}^{\infty}\theta_i^2\right)^{3/2}}.$$
$$= \frac{\mu_{\varepsilon,3}\left(1^3+\theta^3\right)}{\mu_{\varepsilon,2}^{3/2}\left(1^2+\theta^2\right)^{3/2}}$$
$$= \beta_{1,\varepsilon}\frac{\left(1+\theta^3\right)}{\left(1+\theta^2\right)^{3/2}} \tag{A2}$$

Because $\frac{\left(1+\theta^3\right)}{\left(1+\theta^2\right)^{3/2}}$ is bounded in the span $\left(-1,1\right)$ it follows that $\left|\beta_{1,X}\right| \leq \left|\beta_{1,\varepsilon}\right|$.

Furthermore, the kurtosis of the MA(1) process is given by

$$\beta_{2,X} = \frac{\mu_{X,4}}{\mu_{X,2}^2}$$
$$= \frac{\mu_{\varepsilon,4}\sum_{i=0}^{\infty}\theta_i^4 + 3\,\mu_{\varepsilon,2}^2\sum_{i\neq j}\sum\theta_i^2\,\theta_j^2}{\left(\mu_{\varepsilon,2}\sum_{i=0}^{\infty}\theta_i^2\right)^2}$$
$$= \frac{\mu_{\varepsilon,4}\sum_{i=0}^{\infty}\theta_i^4}{\left(\mu_{\varepsilon,2}\sum_{i=0}^{\infty}\theta_i^2\right)^2} + \frac{3\,\mu_{\varepsilon,2}^2\sum_{i\neq j}\sum\theta_i^2\,\theta_j^2}{\left(\mu_{\varepsilon,2}\sum_{i=0}^{\infty}\theta_i^2\right)^2}$$
$$= \frac{\mu_{\varepsilon,4}\left(1+\theta^4\right)}{\left(\mu_{\varepsilon,2}\left(1+\theta^2\right)\right)^2} + 3\mu_{\varepsilon,2}^2\frac{\left(1\cdot\theta^2+\theta^2\cdot 1\right)}{\left(\mu_{\varepsilon,2}\left(1+\theta^2\right)\right)^2}$$
$$= \frac{\mu_{\varepsilon,4}\left(1+\theta^4\right)}{\mu_{\varepsilon,2}^2\left(1+\theta^2\right)^2} + 3\mu_{\varepsilon,2}^2\frac{2\theta^2}{\mu_{\varepsilon,2}^2\left(1+\theta^2\right)^2}$$
$$= \beta_{2,\varepsilon}\frac{\left(1+\theta^4\right)}{\left(1+\theta^2\right)^2} + 3\frac{2\theta^2}{\left(1+\theta^2\right)^2}. \tag{A3}$$

Note that (A3) is 3 if $\beta_{2,\varepsilon} = 3$. Furthermore, it may be shown that the kurtosis of $\beta_{2,X}$ is always closer to the kurtosis of the normal distribution when compared to $\beta_{2,\varepsilon}$, i.e. $\left|\beta_{2,X}-3\right| \leq \left|\beta_{2,\varepsilon}-3\right|$.