Computer Generation of Random Vectors from Continuous Multivariate Distributions

Paola Palmitesta  
Department of Quantitative Methods  
University of Siena

Corrado Provasi  
Department of Statistical Sciences  
University of Padova

Abstract

The evolution of statistical inference in the last years has been induced also by the development of new computational tools which have led both to the solution of classical statistical problems and to the implementation of new methods of analysis. In this framework, a relevant computational tool is given by simulation which leads to deal with methodological problems not solvable analytically. Simulation is frequently used for the generation of variates from known distributions with Monte Carlo methods for the study of the behavior of statistics for which the sampling distribution is unknown. Moreover, Monte Carlo algorithms have been developed which can be also used to approximate the stochastic integrals, as in bayesian statistical analysis and in optimization problems. The application of such algorithms needs the application of advanced computational tools and, particularly, computer algebra systems are surely one of the most suitable tools for the simulation of complex stochastic models.

In this paper we present the methods used for the implementation of some procedures which use Monte Carlo algorithms to obtain random vectors from multivariate continuous distributions in a computer algebra system, specifically Mathematica.

Keywords: Random Vector Generators; Continuous Multivariate Distributions; Monte Carlo Algorithms; Algebraic Software.

1 Introduction

The evolution of statistical inference in the last years has been induced also by the development of new computational tools which have led both to the solution of classical statistical problems and to the implementation of new methods of analysis. In literature many examples of the recent interrelation between numeric computation and statistics can be found, as in Thisted (1988), Fishman (1996), Tanner (1996) and Gentle (1998). In this framework, a relevant computational tool consists of simulation which leads to deal with methodological problems not solvable analytically.

Simulation is frequently used for the generation of variates from known distributions with Monte Carlo methods for the study of the behavior of statistics with unknown sampling distribution. Moreover, Monte Carlo algorithms have been developed which can be also used to approximate the stochastic integrals, as in bayesian statistical analysis and in optimization problems (for a review see Ripley, 1987).

The application of such algorithms needs the application of advanced computational tools and, particularly, computer algebraic systems are surely one of the most suitable tools for the simulation of complex stochastic models. In statistics, computer algebraic systems have been used in various contexts; in particular, among others, Heller (1991) shows how
computer algebraic systems can be used in a wide variety of statistical problems; Kendall (1988 and 1990) provides some procedures for the symbolic computation of expressions used in the analysis of the diffusion of euclidean forms; Silverman and Young (1987) use symbolic computation in order to study some bootstrap distributions and Venables (1985) in order to obtain maximum likelihood estimates in particular cases; Andrews and Stafford (1993) provide some procedures for the symbolic computation of asymptotic expansions in statistics; Provasi (1996) presents some symbolic procedures applicable to compute moments of order statistics of some frequently used random variables (see also Varian, 1992 and 1996, for applications in econometrics).

In this paper we consider the application of the algebraic software Mathematica (Wolfram, 1996) to obtain random vectors from continuous multivariate distributions with Monte Carlo methods. In general, the generation of multivariate distributions is not easily implemented, because the usual method based on the inverse of the cumulative distribution function used with univariate distributions can not be applied. This problem is stated and addressed in Section 2. The functions built in Mathematica to generate random vectors from specific distributions are presented and discussed in Section 3, while in Section 4 we present some computational algorithms to generate random vectors from multivariate distributions with given marginals and correlation matrix. Section 5 contains concluding remarks.

2 Methods for Generating Multivariate Distributions

Consider the random variable $X$ with cumulative distribution function (cdf) $F_X$ and assume that $F_X^{-1}(q) = \inf\{x : F_X(x) \geq q\}$, $0 < q < 1$. The traditional Monte Carlo method is based on the well-known result that, for every random variable $U$ uniformly distributed on $(0, 1]$, we have that $X$ and $F_X^{-1}(U)$ have the same cdf. As a consequence, a Monte Carlo simulation of a random variable $X$ can be done at first drawing a uniform random number $u$ from $U \sim \text{Uniform}(0, 1)$ and then inverting $u$ by means of $x = F_X^{-1}(u)$.

Likewise, a Monte Carlo simulation of $p$ random variables $(X_1, \ldots, X_p)$ with a dependence structure can start with $p$ uniform random variables $(U_1, \ldots, U_p)$ with the same dependence structure. However, in general the definition of a multivariate distribution with known marginals and some dependence structure is very complex and sometimes ambiguous, therefore the construction of Monte Carlo algorithms for multivariate distributions satisfying the usual assumptions could often be difficult.

In alternative, a method which converts the problem of the generation of a $p$-dimensional random vector $X = (X_1, \ldots, X_p)^\prime$ into the generation of $p$ univariate random variables can be applied when the probability density of $X$ can be divided in factors as follows:

$$f_{X_1,\ldots,X_p}(x_1,\ldots,x_p) = f_{X_1}(x_1)f_{X_2|X_1}(x_2|x_1)\cdots f_{X_p|X_1,\ldots,X_{p-1}}(x_p|x_1,\ldots,x_{p-1}).$$

Parrish (1990) used this method to generate random vectors from Pearson multivariate distributions.

When the computation of the conditional distribution of $X$ is difficult, a transformation of the vector can be considered. In this case, $X$ is represented as a function of independent univariate distributions. Notwithstanding the apparent easiness for applying this method, the search of a particular transformation to generate $X$ could be difficult when $X$ is specified by its probability density function $f_{X_1,\ldots,X_p}$.

Other approaches for generating multivariate distributions are based either on the extension to the multivariate case of the rejection method or on iterative methods. Among
them, we cite importance sampling, in which sampling points are concentrated in the most relevant regions for \( X \) rather than in the whole region, and methods based on Markov chains. For the latter methods, the idea is to simulate independent realizations forming an irreducible Markov chain, having as stationary distribution the distribution under study (for a review see Boswell et al., 1993).

Referring mainly on the transformation to independent forms, in the next two sections we present the methods we used in the construction of Mathematica functions, in order to generate random vectors from many multivariate continuous distributions (some of these methods can be found in extended form in Johnson (1987)).

### 3 Generating Specific Multivariate Distributions

As disclosed below, in this section we show the methods we used for the construction of the Mathematica functions which lead to the generation of random vectors from specified multivariate distributions. In general, these distributions are an extension to the multivariate case of known random variables which are often used in statistical applications with unidimensional samples.

The symbolic form of the Mathematica functions is

\[
\text{RandomArray}[\text{distribution}[\text{param}_1,\text{param}_2,\ldots],n]
\]

where \( \text{distribution} \) is the name of one the distributions in Tab. 1 with the expected parameters and \( n \) is the number of random vectors to be generated. These functions have been organized in the package MultiRand\(^1\).

**Multuniform Distribution.** The density function of the random vector \( \mathbf{U}^{(p)} = (U_1,\ldots, U_p) \)' uniformly distributed on the unit sphere in \( \mathbb{R}^p \) is given by

\[
f_{U_1,\ldots,U_p}(u_1,\ldots,u_p) = \frac{\Gamma(p/2)}{\pi^{p/2}} \left( 1 - \sum_{i=1}^{p} u_i^2 \right)^{-1}, \quad \sum_{i=1}^{p} u_i^2 < 1, \: u_i > 0, \: i = 1,\ldots,p,
\]

where \( \Gamma(\cdot) \) is the gamma function. A technique for the generation of random vectors from \( \mathbf{U}^{(p)} \) is based on the transformation (Muller, 1959)

\[
U_i = \frac{X_i}{(X_1 + \cdots + X_p)^{1/2}}, \quad i = 1,\ldots,p,
\]

where \( X_1,\ldots,X_p \) are independent standard normal variates (see Tashiro, 1977, for some others methods for the generation of random vectors from \( \mathbf{U}^{(p)} \)).

\(^1\)The package MultiRand.m can be downloaded as a zipped file at the following URL:


In MultiRand.m the generation of variates from univariate distributions is done using the package ContinuousDistribution.m.
<table>
<thead>
<tr>
<th>Distribution Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>MultinormalDistribution[μ, Σ]</td>
<td>multivariate normal distribution with mean vector $\mu$ and scale matrix $\Sigma$.</td>
</tr>
<tr>
<td>DirichletDistribution[α]</td>
<td>Dirichelet distribution with parameter vector $\alpha$.</td>
</tr>
<tr>
<td>MultivariateChiSquareDistribution[m, Σ]</td>
<td>multivariate chi square distribution with $m$ degrees of freedom and scale matrix $\Sigma$.</td>
</tr>
<tr>
<td>MultivariateFDistribution[m, k, Σ]</td>
<td>multivariate $F$ distribution with $(m, k)$ degrees of freedom and scale matrix $\Sigma$.</td>
</tr>
<tr>
<td>MultivariateTDistribution[m, μ, Σ]</td>
<td>multivariate $t$ distribution with $m$ degrees of freedom, mean vector $\mu$ and scale matrix $\Sigma$.</td>
</tr>
<tr>
<td>MultivariateJohnsonDistribution[μ, σ, a, b, R, option]</td>
<td>multivariate Johnson distribution with mean vector $\mu$, standard deviation vector $\sigma$, skewness indices vector $a$, kurtosis indices vector $b$ and with $R$ that is either a correlation matrix or a Spearman’s rank correlation matrix or Kendall’s rank correlation matrix according to the value given to the logical variable option.</td>
</tr>
<tr>
<td>MultivariateBetaDistribution[α, β]</td>
<td>multivariate beta distribution with parameter vectors $\alpha$ and $\beta$.</td>
</tr>
<tr>
<td>MultivariateGammaDistribution[θ, λ, γ]</td>
<td>multivariate gamma distribution with shape parameter vector $\theta$, scale parameter vector $\lambda$ and location parameter vector $\gamma$.</td>
</tr>
<tr>
<td>MultivariateExtremeValueDistribution[μ, σ, ρ]</td>
<td>multivariate extreme values distribution with position parameter vector $\mu$, scale parameter vector $\sigma$ and dependence parameter $\rho$.</td>
</tr>
<tr>
<td>MultivariateInverseGaussianDistribution[χ, ψ, α]</td>
<td>multivariate inverse gaussian distribution with parameter vectors $\chi$ and $\alpha$ and scalar parameter $\psi$.</td>
</tr>
<tr>
<td>MultivariateUniformDistribution[p, α]</td>
<td>$p$-dimensional uniform distribution with scalar parameter $\alpha$.</td>
</tr>
<tr>
<td>MultivariateBurrDistribution[c, d, k]</td>
<td>multivariate Burr distribution with parameter vectors $c$ and $d$ and scalar parameter $k$.</td>
</tr>
<tr>
<td>MultivariateParetoDistribution[θ, α]</td>
<td>multivariate Pareto distribution with location parameter vector $\theta$ and scalar parameter $\alpha$.</td>
</tr>
<tr>
<td>MultivariateLogisticDistribution[p, α]</td>
<td>p-dimensional logistic distribution with scalar parameter $\alpha$.</td>
</tr>
<tr>
<td>MultivariatePearsonTypeIIDistribution[v, μ, Σ]</td>
<td>multivariate Pearson Type II distribution with shape parameter $v$, mean vector $\mu$ and scale matrix $\Sigma$.</td>
</tr>
<tr>
<td>MultivariatePearsonTypeVIIIDistribution[v, μ, Σ]</td>
<td>multivariate Pearson Type VII distribution with shape parameter $v$, mean vector $\mu$ and scale matrix $\Sigma$.</td>
</tr>
<tr>
<td>MultivariatePowerExponentialDistribution[β, μ, Σ]</td>
<td>multivariate power exponential distribution with shape parameter $\beta$, mean vector $\mu$ and scale matrix $\Sigma$.</td>
</tr>
</tbody>
</table>
Dirichlet Distribution. Let \( Y_1, \ldots, Y_p, Y_{p+1} \) be independent standard gamma random variables with shape parameters \( \alpha_i > 0 \) for \( i = 1, \ldots, p, p + 1 \), and let

\[
X_i = \frac{Y_i}{Y_1 + \cdots + Y_p + Y_{p+1}}, \quad i = 1, \ldots, p.
\]

Then, the random vector \( \mathbf{X} = (X_1, \ldots, X_p)' \) has a Dirichlet distribution with density function given by (cf. Fang, Kotz and Ng, 1990, p. 17)

\[
f_{X_1, \ldots, X_p}(x_1, \ldots, x_p) = (B_p(\alpha))^{-1} \prod_{i=1}^{p+1} x_i^{\alpha_i - 1}, \quad \sum_{i=1}^{p+1} x_i = 1, \quad x_i > 0, \quad i = 1, \ldots, p, p + 1,
\]

where \( B_p(\alpha) = (\prod_{i=1}^{p+1} \Gamma(\alpha_i))/\Gamma(\alpha) \), \( \alpha = \sum_{i=1}^{p+1} \alpha_i \) and \( \alpha = (\alpha_1, \ldots, \alpha_p, \alpha_{p+1})' \). Hence, a technique for the generation of random vectors from a Dirichlet distribution with parameter vector \( \alpha \) can be based on the transformation (1) (see Narayan, 1990, for some others methods for the generation of random vectors from a Dirichlet distribution).

Multinormal and Related Distributions. Let the distribution of the random vector \( \mathbf{X} = (X_1, \ldots, X_p)' \) be a multivariate normal \( N_p(\mu, \Sigma) \) with mean vector \( \mu = (\mu_1, \ldots, \mu_p)' \) and covariance matrix \( \Sigma = (\sigma_{ij}) \). Moreover, let the lower triangular matrix \( \mathbf{A} \) be obtained by the Cholesky decomposition \( \mathbf{A} \mathbf{A}' = \Sigma \). Then, given the random vector \( \mathbf{Z} = (Z_1, \ldots, Z_p)' \) with independent standard normal components, the simulation of \( N_p(\mu, \Sigma) \) can be run using the transformation \( \mathbf{X} = \mu + \mathbf{AZ} \) (Scheuer and Stoller, 1962).

A multivariate distribution with chi-squared marginals can be obtained as a function of multinormal random vectors. In fact, let the independent random vectors \( \mathbf{X}_i = (X_{ij}, \ldots, X_{ip})', \quad j = 1, 2, \ldots, m \), identically distributed as multivariate normal \( N_p(0, \Sigma) \) and let \( Y_i = \sum_{j=1}^{m} X_{ij}^2/\sigma_{ii} \) for \( i = 1, \ldots, p \). Then, the joint distribution of \( (Y_1, \ldots, Y_p) \) is characterized by the scale matrix \( \Sigma \) and has chi-squared marginals with \( m \) degrees of freedom (cf. Mardia, Kent and Bibby, 1979, p. 67). Moreover, let \( S^2 \) be a chi-squared random variable with \( k \) degrees of freedom independent from \( Y_i \) and let \( F_i = k Y_i/(m S^2) \) for \( i = 1, \ldots, p \). The joint distribution of \( (F_1, \ldots, F_p) \) is called multivariate \( F \) distribution with \( (m, k) \) degrees of freedom and covariance matrix \( \Sigma \) of the accompanying multivariate normal distribution (Schuurmann, Krishnaiah and Chattopadhyay, 1975). We can note that the multivariate chi-squared distribution can be obtained by the main diagonal of the Wishart matrix generated with the Bartlett algorithm (Bartlett, 1933).

Also a random vector distributed as a multivariate \( t \) distribution can be written as a function of a multinormal random vector. As a matter of fact, if \( \mathbf{Z} \) is a multivariate normal \( N_p(0, \Sigma) \) and \( S^2 \) is a chi-squared random variable with \( m \) degrees of freedom independent from \( \mathbf{Z} \), then the joint distribution of \( (T_1, \ldots, T_p) \), where \( T_i = (\sqrt{\mathbf{Z}}/S) + \mu_i \) for \( i = 1, \ldots, p \), has a multivariate \( t \) distribution with \( m \) degrees of freedom, mean vector \( \mu = (\mu_1, \ldots, \mu_p)' \) and scale matrix \( \Sigma \). When \( \Sigma = \mathbf{I} \), where \( \mathbf{I} \) is the identity matrix, and \( m = 1 \), the multivariate Student’s \( t \) distribution is equal to the multivariate Cauchy distribution (see Johnson and Kotz, 1972, Cap. 37, for details on the various types of multivariate Student’s \( t \) distribution).

Multivariate Johnson Distribution. A continuous random variable \( X \) belongs to one of the distribution families of the Johnson translation system (1949a) if the transformation

\[
Z = \gamma + \delta g((X - \xi)/\lambda)
\]
is true, where \( Z \) indicates the standard normal distribution, \( \gamma \in \mathbb{R} \) and \( \delta > 0 \) are shape parameters, \( \xi \in \mathbb{R} \) is a location parameter, \( \lambda > 0 \) is a scale parameter, and \( g(\cdot) \) is one of the following transformations:

\[
g(y) = \begin{cases} 
\ln(y), & \text{for the } S_L \text{ (lognormal) family,} \\
\sinh^{-1}(y), & \text{for the } S_U \text{ (unbounded) family,} \\
\ln[y/(1 - y)], & \text{for the } S_B \text{ (bounded) family,} \\
y, & \text{for the } S_N \text{ (normal) family.}
\end{cases}
\]

(2)

An important characteristic of the Johnson system, useful in Monte Carlo studies, is that the family to which \( X \) belongs can be unambiguously identified by skewness and kurtosis, as in Hill, Hill and Holder (1974). Moreover, we can obtain random variables from \( X \) at first by the generation of variates from \( Z \) and then applying the inverse translation \( X = \xi + \lambda \ g^{-1}([Z - \gamma]/\delta] \),

where

\[
g^{-1}(z) = \begin{cases} 
e^z, & \text{for the } S_L \text{ (lognormal) family,} \\
\sinh(z), & \text{for the } S_U \text{ (unbounded) family,} \\
1/(1 + e^{-z}), & \text{for the } S_B \text{ (bounded) family,} \\
z, & \text{for the } S_N \text{ (normal) family.}
\end{cases}
\]

(3)

Johnson (1949b) has proposed a bivariate distribution based on the univariate translation system which can be easily extended to higher dimensions. Then, the random vector \( X = (X_1, \ldots, X_p)' \) with \( p \) components has a Johnson multivariate distribution if the transformation

\[
Z = \gamma + \delta \  g[\lambda^{-1}(X - \xi)] \sim N_p(0, \Sigma),
\]

(4)
is true, where \( \Sigma = \mathbf{R} = (\rho_{ij}) \) is a correlation matrix, \( g[(y_1, \ldots, y_p)'] = [g_1(y_1), \ldots, g_p(y_p)]' \) identifies the family to which the marginals of the translation belong, \( \gamma = (\gamma_1, \ldots, \gamma_p)' \), \( \delta = \text{diag}(\delta_1, \ldots, \delta_p) \), \( \xi = (\xi_1, \ldots, \xi_p)' \) and \( \lambda = \text{diag}(\lambda_1, \ldots, \lambda_p) \) are, respectively, the shape, location and scale parameters of the components \( X_i \) for \( i = 1, \ldots, p \). Therefore, having determined the families of the marginals of \( X \) on the basis of the first four moments, the generation of the random vectors is done generating \( Z \) from a multivariate normal distribution \( N_p(0, \Sigma) \) and then applying the inverse translation \( X = \xi + \lambda \ g^{-1}[\delta^{-1}(Z - \gamma)] \),

using the vectors of parameters previously determined and the vector of the inverse translation functions \( g^{-1}[(z_1, \ldots, z_p)'] = [g_1^{-1}(z_1), \ldots, g_p^{-1}(z_p)] \), where \( g_i^{-1}(\cdot) \) is defined by (3) for \( i = 1, \ldots, p \).

This method takes to the generation of random vectors with the same moments of the marginals, but not necessarily with the same correlation matrix, because the (4) is not linear. This suggests to measure the association among the components of \( X \) by means of the Spearman’s rank correlation matrix \( \mathbf{R}_S = (\hat{\rho}_{ij}) \) or the Kendall’s rank correlation matrix \( \mathbf{R}_T = (\tau_{ij}) \) whose elements, as known, are not only invariant regarding to non linear transformations but, in the normal case, are also connected to the correlation coefficient by the relations

\[
\rho_{ij} = 2 \sin \left( \frac{\pi}{6} \hat{\rho}_{ij} \right)
\]

and
\[ \rho_{ij} = \sin \left( \frac{\pi}{2} \tau_{ij} \right). \]

As a consequence, this method takes to the generation of random vectors from \( X \) with the same moments of the marginals and given matrix \( R_S \) or \( R_r \). In the next section we show an extension of the multivariate Johnson distribution which leads to the generation of random vectors with given correlation matrix.

**Multivariate Beta and Gamma Distributions.** Mathai and Moschopoulos (1991 and 1993) have introduced two multivariate distributions with beta and gamma marginals which can be easily indirectly simulated. Let \( B_1, \ldots, B_p \) be beta random variables with density function

\[ f_{B_i}(b_i) = \frac{\Gamma(\alpha_i + \beta_i)}{\Gamma(\alpha_i)\Gamma(\beta_i)} b_i^{\alpha_i-1}(1-b_i)^{\beta_i-1}, \]

with \( b_i \in (0, 1) \) parameters \( \alpha_i > 0 \) and \( \beta_i > 0 \) for \( i = 1, \ldots, p \), and let \( U \) be a uniform random variable with support on \((0, 1] \). Moreover, let \( U, B_1, \ldots, B_p \) be mutually independent and let

\[ X_i = B_i U^{1/(\alpha_i+\beta_i)}, \quad i = 1, \ldots, p. \] (5)

Then, the random vector \( X = (X_1, \ldots, X_p)' \) has components Beta\((\alpha_i, \beta_i + 1)\) for \( i = 1, \ldots, p \) and the following properties can be obtained directly from the definition:

\[ E(X_i) = \frac{\alpha_i}{(\alpha_i + \beta_i + 1)}, \]
\[ \text{Var}(X_i) = \frac{\alpha_i(\beta_i + 1)}{(\alpha_i + \beta_i + 1)^2((\alpha_i + \beta_i + 2))}, \]
\[ \text{Cov}(X_i, X_j) = \frac{\alpha_i \alpha_j}{(\alpha_i + \beta_i + 1)(\alpha_j + \beta_j + 1)[(\alpha_i + \beta_i + 1)(\alpha_j + \beta_j + 1) - 1]}, \quad i \neq j. \]

Now, consider the independent gamma random variables \( V_0, V_1, \ldots, V_p \) with density function

\[ f_{V_i}(v_i) = \frac{1}{\lambda_i^\theta_i \Gamma(\theta_i)} (v_i - \gamma_i)^{\theta_i-1} \exp \left[-(v_i - \gamma_i)/\lambda_i \right], \]

with \( v_i > \gamma_i \) and parameters \( \theta_i > 0, \lambda_i > 0 \) and \( \gamma_i \in \mathbb{R} \) for \( i = 0, 1, \ldots, p \), and let

\[ Y_i = \frac{\lambda_i}{\lambda_0} V_0 + V_i, \quad i = 1, \ldots, p. \] (6)

Then, the random vector \( Y = (Y_1, \ldots, Y_p)' \) has components Gamma\((\theta_0 + \theta_i, \lambda_i, (\gamma_0/\lambda_0)\lambda_i + \gamma_i)\) for \( i = 1, \ldots, p \), and, directly from the definition or from the moment generating function, the following properties can be obtained:

\[ E(Y_i) = (\theta_0 + \theta_i)\lambda_i + (\gamma_0/\lambda_0)\lambda_i + \gamma_i, \]
\[ \text{Var}(Y_i) = (\theta_0 + \theta_i)^2 \lambda_i^2, \]
\[ \text{Cov}(Y_i, Y_j) = \theta_0 \lambda_i \lambda_j, \quad i \neq j. \]

Therefore, the generation of random vectors from \( X \) or \( Y \) simply requires the generation of independent variates from the beta and uniform or gamma random variables and then the application of the transformations (5) or (6). Note that for both distributions the correlation structure is positive.
Multivariate Extreme Value Distribution. Suppose that the cdf of the random vector \( \mathbf{X} = (X_1, \ldots, X_p)' \) follows the logistic model

\[
F_{X_1, \ldots, X_p}(x_1, \ldots, x_p) = \exp \left\{- \left[ \sum_{i=1}^{p} \exp \left( -\frac{x_i - \mu_i}{\rho \sigma_i} \right) \right]^{\rho} \right\},
\]

where \( x_i \in \mathbb{R} \) and parameters \( \mu_i \in \mathbb{R} \) and \( \sigma_i > 0 \) for \( i = 1, \ldots, p \) and \( \rho \in [0, 1] \) measures the dependence among the marginals. The extremes \( \rho \to 1 \) and \( \rho \to 0 \) correspond, respectively, to the independence and complete dependence. Then, the components of \( \mathbf{X} \) have an extreme value distribution (or Gumbel distribution) with cdf

\[
F_{X_i}(x_i) = \exp \left[ - \exp \left( -\frac{x_i - \mu_i}{\sigma_i} \right) \right], \quad i = 1, \ldots, p.
\]

Shi (1995) has computed the density function of \( \mathbf{X} \) and he has proposed a transformation which allows its simulation. Let

\[
y_i = \exp \left( -\frac{x_i - \mu_i}{\sigma_i} \right), \quad i = 1, \ldots, p,
\]

\[
z = \left( \sum_{i=1}^{p} y_i^{1/\rho} \right)^{\rho}.
\]

Then, (7) can be written as

\[
F_{X_1, \ldots, X_p}(x_1, \ldots, x_p) = \exp \left\{- \left( \sum_{i=1}^{p} y_i^{1/\rho} \right)^{\rho} \right\} = e^{-z},
\]

from which we can derive the density function

\[
f_{X_1, \ldots, X_p}(x_1, \ldots, x_p) = \frac{\partial^p F_{X_1, \ldots, X_p}}{\partial x_1 \cdots \partial x_p} = \left( \prod_{i=1}^{p} \frac{y_i^{1/\rho}}{\sigma_i} \right) z^{-p/\rho} Q_p(z, \rho) e^{-z},
\]

where

\[
Q_p(z, \rho) = \left( \frac{p - 1}{\rho} - 1 + z \right) Q_{p-1}(z, \rho) - z \frac{\partial Q_{p-1}(z, \rho)}{\partial z}, \quad Q_1(z, \rho) = 1.
\]

Let

\[
\begin{align*}
T_1 &= \cos^2 \theta_1, \\
T_i &= \left( \prod_{j=1}^{i-1} \sin^2 \theta_j \right) \cos^2 \theta_i, \quad i = 2, \ldots, p - 1, \\
T_p &= \prod_{j=1}^{p-1} \sin^2 \theta_j, \quad \theta_j \in [0, \frac{\pi}{2}], \quad j = 1, \ldots, p - 1.
\end{align*}
\]

and consider the transformation

\[
y_i = z T_i^\rho, \quad i = 1, \ldots, p,
\]

or

\[
x_i = \mu_i - \sigma_i (\log z + \rho \log T_i), \quad i = 1, \ldots, p.
\]
Then, the density function of the random vector \((Z, \Theta_1, \Theta_2, \ldots, \Theta_{p-1})\) is given by (Shi, 1995)

\[
g_{Z, \theta_1, \ldots, \theta_{p-1}}(z, \theta_1, \ldots, \theta_{p-1}) = h_Z(z) \prod_{j=1}^{p-1} g_{\theta_j}(\theta_j),
\]

where

\[
h_Z(z) = \frac{\rho^{p-1}}{(p-1)!} Q_p(z, \rho) e^{-z}, \quad z > 0,
\]

\[
g_{\theta_j}(\theta_j) = 2(p-j) \cos \theta_j (\sin \theta_j)^{2(p-j)-1}, \quad \theta_j \in \left[0, \frac{\pi}{2}\right], \quad j = 1, \ldots, p-1,
\]

therefore the random variables \(Z, \Theta_1, \ldots, \Theta_{p-1}\) are independent among them.

Because \(Q_p(z, \rho)\) is a \(p-1\)-polynomial of \(z\), if we write the density function of \(Z\) as

\[
h_Z(z) = \sum_{j=1}^{p} q_{p,j} \phi(z, j),
\]

where

\[
\phi(z, k) = \frac{1}{\Gamma(k)} z^{k-1} e^{-z}, \quad z > 0,
\]

we note that \(Z\) has a mixture gamma distribution with shape parameter \(k\) and weights \(q_{p,j}\) depending on \(Q_p(z, \alpha)\). It is easy to obtain the following recurrence relations from the (8):

\[
q_{p,1} = \frac{\Gamma(p-\rho)}{\Gamma(p)\Gamma(1-\rho)},
\]

\[
(p-1)q_{p,j} = (p-j\rho)q_{p-1,j} + \rho(j-1)q_{p-1,j-1}, \quad j = 2, \ldots, p-1,
\]

\[
q_{p,p} = \rho^{p-1}.
\]

Now, the random vectors can be easily generated from the distribution (7) with marginals distributed as an extreme value distribution. As a matter of fact, applying the inverse transformation method, \(\sin^2 \Theta_j, \cos^2 \Theta_j\) can be obtained, respectively, from \(U^{1/(p-j)}\) and \(1 - U^{1/(p-j)}\), where \(U\) is the uniform distribution on \((0, 1]\), while the variates from the random variable \(Z\) with mixture gamma distribution can be generated using the composition method. Therefore, using the (9), the simulated random vectors from \(X\) are obtained.

**Multivariate Inverse Gaussian Distribution**  
Barndorff-Nielsen, Blæsild and Seshadri (1992) have proposed an extension to the multivariate case of the inverse gaussian distribution using an additive random effects model which can be easily simulated. Let \(X\) be a random variable InverseGaussian(\(\chi, \psi\)) with density function

\[
f_X(x) = \left[\frac{\chi}{2\pi}\right]^{1/2} e^{\sqrt{\chi} \psi} x^{-3/2} \exp \left[-\frac{1}{2} \left(\frac{\chi}{x} + \psi x\right)\right], \quad x > 0,
\]

with parameters \(\chi > 0\) and \(\psi > 0\). Now, let us suppose that the components of the random vector \(X = (X_1, \ldots, X_p)'\) are defined using the additive random effects model

\[
X_i = \alpha_i + V_i, \quad i = 1, \ldots, p,
\]

\(10\)
where $T \sim \text{InverseGaussian}(\chi, \psi)$ and $V_1, \ldots, V_p$, where $V_i \sim \text{InverseGaussian}(\chi_i, \psi/\alpha_i)$, are independent among them and $\alpha_1, \ldots, \alpha_p$ are positive parameters. Then, $X$ has InverseGaussian($\sqrt{\chi \alpha_i + \sqrt{\chi_i}^2}, \psi/\alpha_i$), $i = 1, \ldots, p$, marginals and from the moment generating function the following properties can be obtained:

$$E(X_i) = \sqrt{\frac{(\sqrt{\chi \alpha_i} + \sqrt{\chi_i})^2}{\psi/\alpha_i}},$$

$$\text{Var}(X_i) = \sqrt{\frac{(\sqrt{\chi \alpha_i} + \sqrt{\chi_i})^2}{(\psi/\alpha_i)^3}},$$

$$\text{Cov}(X_i, X_j) = \frac{\alpha_i \alpha_j \chi^2}{\sqrt{\chi \psi}}, \quad i \neq j.$$

Therefore, the generation of random vectors from $X$ simply requires at first the generation of independent variates from inverse gaussian random variables and then the application of the transformation (10). In this case too, note that the model used to build the multivariate distribution leads to a positive correlation structure, because it depends only on positive parameters.

The Cook-Johnson Family of Multivariate Uniform Distributions. Let $U = (U_1, \ldots, U_p)'$ be a p-dimensional uniform distribution with support on the hypercube $(0, 1]^p$ and with cdf

$$F_{U_1, \ldots, U_p}(u_1, \ldots, u_p) = \left\{ \sum_{i=1}^{p} u_i^{-1/\alpha} - p + 1 \right\}^{-\alpha}, \quad u_i \in (0, 1], \quad \alpha > 0. \quad (11)$$

Cook and Johnson (1981) have studied this family of distributions and have shown that

$$\lim_{\alpha \to 0} F_{U_1, \ldots, U_p}(u_1, \ldots, u_p) = \min [u_1, \ldots, u_p]$$

and

$$\lim_{\alpha \to \infty} F_{U_1, \ldots, U_p}(u_1, \ldots, u_p) = \prod_{i=1}^{p} u_i;$$

therefore the correlation among the components of $U$ approaches to its maximum when $\alpha \to 0$ and approaches to zero when $\alpha \to \infty$.

A simulation of $U$ can be run using the following algorithm: let $Y_1, \ldots, Y_p$ be i.i.d. standard exponential distributions and let the distribution of $V$ be a standard gamma with shape parameter $\alpha$. Then, the cumulative distribution function of the random variables

$$U_i = [1 + Y_i/V]^{-\alpha}, \quad i = 1, \ldots, p,$$

have a cdf given by (11).

For a set of arbitrary marginal distributions, $F_{X_1}, \ldots, F_{X_p}$, we can define the joint cdf as

$$F_{X_1, \ldots, X_p}(x_1, \ldots, x_p) = \left\{ \sum_{i=1}^{p} F_{X_i}(x_i)^{-1/\alpha} - p + 1 \right\}^{-\alpha} \quad (12)$$

If we assume that the random vector $X = (X_1, \ldots, X_p)'$ has a multivariate distribution with cdf given by the equation (12), then a simulation of it can be easily implemented inverting $(U_1, \ldots, U_p)$ with $(F_{X_1}^{-1}, \ldots, F_{X_p}^{-1})$. In this paper we have implemented the (11) also with the marginal distributions considered by Cook and Johnson (1981):
(i) Burr distribution:

\[ F_{X_i}(x_i) = 1 - (1 + d_i x_i^{c_i})^{-k}, \]

with \( x_i > 0 \) and parameters \( c_i > 0, d_i > 0 \) for \( i = 1, \ldots, p \), and \( k > 0 \).

(ii) Pareto distribution:

\[ F_{X_i}(x_i) = \left( \frac{\theta_i}{x_i} \right)^a, \]

with \( x_i > \theta_i \) and parameters \( \theta_i > 0 \) for \( i = 1, \ldots, p \), and \( \alpha > 0 \).

(iii) Logistic distribution:

\[ F_{X_i}(x_i) = \left[ 1 + \exp(-x_i) \right]^{-\alpha}, \]

with \( x_i \in \mathbb{R} \) and parameter \( \alpha > 0 \).

**Elliptically Contoured Distributions.** An interesting class of multivariate distributions including the Gaussian distribution as a special case is given by the class of elliptically contoured distributions (Johnson, 1987, ch. 6; see also Fang, Kotz and Ng, 1990). A random vector \( X = (X_1, \ldots, X_p)' \) is elliptically contoured with scale matrix \( \Sigma \) and location vector \( \mu \) if

\[ X = \mu + RAU(p), \tag{13} \]

where \( A \) is a lower triangular matrix obtained by the Cholesky decomposition of \( \Sigma \), \( U(p) \) is a random vector distributed as a uniform on the surface of the unit \( p \)-dimensional hypersphere as defined below and \( R \) is a random variable independent from \( U(p) \). The random variable \( R^2 \) has the distribution of \( (X - \mu)'\Sigma^{-1}(X - \mu) \) and in some cases of practical usage has a recognizable distribution which can be easily generated; therefore we can simulate an observation from \( X \) using its stochastic representation (13). These cases include the multivariate Pearson Type II and Pearson Type VII distributions which, according to Johnson (1987), are suitable for simulation studies.

The random vector \( X = (X_1, \ldots, X_p)' \) has a multivariate Pearson Type II distribution if the density function is given by

\[ f_{X_1, \ldots, X_p}(x_1, \ldots, x_p) = \frac{\Gamma(p/2 + v + 1)}{\Gamma(v + 1)\pi^{p/2}|\Sigma|^{-1/2}} \left[ 1 - (x - \mu)'\Sigma^{-1}(x - \mu) \right]^m, \]

where \( x = (x_1, \ldots, x_p)' \) and \( v > -1 \), \( p \) is the dimension of \( X \). The support of the distribution is restricted to the region \( (x - \mu)'\Sigma^{-1}(x - \mu) \leq 1 \).

The generation of random vectors from this distribution on the basis of the (13) is very easy, because the distribution of the quadratic form \( R^2 = (X - \mu)'\Sigma^{-1}(X - \mu) \) is beta with parameters \( p/2 \) and \( v + 1 \).

The random vector \( X = (X_1, \ldots, X_p)' \) has a multivariate Pearson Type VII distribution if the density function

\[ f_{X_1, \ldots, X_p}(x_1, \ldots, x_p) = \frac{\Gamma(v)}{\Gamma(v + p/2)\pi^{p/2}|\Sigma|^{-1/2}} \left[ 1 + (x - \mu)'\Sigma^{-1}(x - \mu) \right] \]

where \( x = (x_1, \ldots, x_p)' \), \( x_i \in \mathbb{R} \) for \( i = 1, \ldots, p \), and \( v > p/2 \), \( p \) is the dimension of \( X \).
In this case the distribution of $T = R^2$ has density function

$$f_T(t) = \frac{\Gamma(v)}{\Gamma(p/2)\Gamma(v-p/2)} t^{p/2-1} (1+t)^{-v}, \quad t > 0.$$ 

This density is a Pearson Type VI distribution and can be generated as $T = Y/(1 - Y)$, where $Y$ is beta with parameters $p/2$ e $v - p/2$ (Johnson and Kotz, 1972, Cap. 37, have proposed a different parametrization of this distribution referring it to the multivariate $t$ seen below).

Finally, Gómez, Gómez-Villegas and Marín (1998) have proposed a multivariate generalization of the exponential power distribution which is elliptically contoured. The random vector $X = (X_1, \ldots, X_p)'$ follows a multivariate exponential power distribution if has density function

$$f_{X_1,\ldots,X_p}(x_1,\ldots,x_p) = \frac{p\Gamma(p/2)}{\pi^{p/2}2^{1+p^2/2\beta}} |\Sigma|^{-1/2} \exp \left\{ -\frac{1}{2} \left[ (x - \mu)'\Sigma^{-1}(x - \mu) \right]^{\beta} \right\},$$

where $x = (x_1,\ldots,x_p)'$, $x_i \in \mathbb{R}$ for $i = 1,\ldots,p$, and $\beta > 0$ is a shape parameter, $p$ is the dimension of $X$. Gómez, Gómez-Villegas e Marín (1998) have proved that the (13) is the stochastic representation of $X$ when the distribution of $R$ is given by

$$f_R(r) = \frac{p}{\Gamma \left( 1 + \frac{p}{2\beta} \right) 2^{p^2/2\beta}} r^{p-1} \exp \left\{ -\frac{1}{2} r^{2\beta} \right\}, \quad r > 0.$$ 

The generation of random vectors from this distribution on the basis of the (13) is very easy, because the distribution of the random variable $T = R^{2\beta}$ is a gamma with shape parameter $p/(2\beta)$ and scale parameter 2.

4 Generating Random Vectors with Given Marginal Distributions and Given Correlation Matrix

The dependence structure of non elliptical multivariate distributions showed in the last section is ruled in general by the values of the parameters of the distributions. As a consequence, it is not easy to run Monte Carlo experiments with these distributions when it is necessary to define a priori their dependence structure as it happens, for example, when the robustness of statistical methods is investigated. In this section we show the methods we used to generate random vectors from non elliptical distributions with given marginals and correlation matrix with Mathematica (about the definition of this kind of vectors see, for example, Tiit, 1984).

4.1 Infinitely divisible distributions

Suppose that $X$ is a random variable with cdf $F_X$ and characteristic function $\varphi(t)$. If $[\varphi(t)]^{1/n}$ is the characteristic function of a random variable for each positive entire number $n$, then $X$ is infinitely divisible, that is $F_X$ is the distribution of a sum of $n$ independent random variables, where $[\varphi(t)]^{1/n}$ is the characteristic function of each of these distributions. For example, the gaussian, Poisson and gamma distributions are infinitely divisible random variables. Park and Shin (1998) have proposed an algorithm to generate random
vectors from multivariate distributions with marginals in the class of the random variables closed respect to the sum and given correlation matrix. Sim (1993) too has proposed an algorithm to generate Poisson and gamma random vectors with given marginals and covariance matrix. In this paper, on the contrary, we have followed the approach of Prékopa and Szántai (1978) to implement an algorithm using Mathematica which allows the generation of random vectors from a gamma distribution with given mean vector and covariance matrix. Differently from the method we followed, which always has a solution, the algorithms by Park and Shin (1998) and Sim (1993) do not always converge.

Let $$X = (X_1, \ldots, X_p)'$$ be a multivariate distribution with gamma marginals with density function

$$f_{X_i}(x_i) = \frac{1}{\lambda_i^\theta_i \Gamma(\theta_i)} x_i^{\theta_i-1} e^{-x_i/\lambda_i}, \quad x_i > 0,$$

with parameters $$\theta_i > 0$$ and $$\lambda_i > 0$$ for $$i = 1, \ldots, p$$. The aim is to generate random vectors from $$X$$ when the mean vector $$\mu = (\mu_1, \ldots, \mu_p)'$$ and the covariance matrix $$\Sigma = (\sigma_{ij})$$ are given with positive elements. Obviously, in this case also the correlation matrix with positive elements $$R = (\rho_{ij})$$ is defined and the parameters of $$X_i$$ can be derived from the corresponding mean and variance using the relations $$\theta_i = \mu_i^2/\sigma_{ii}$$ and $$\lambda_i = \sigma_{ii}/\mu_i$$ for $$i = 1, \ldots, p$$.

Now, consider the standard multivariate Gamma distribution $$Y = (Y_1, \ldots, Y_p)'$$ obtained using the transformation $$Y = B^{-1}X$$, where $$B = \text{diag}(\lambda_1, \ldots, \lambda_p)'$$. Then, we have that

$$E(Y_i) = \eta_i = \theta_i,$$

$$\text{Var}(Y_i) = \delta_{ii} = \theta_i, \quad i = 1, \ldots, p,$$

$$\text{Cov}(Y_i, Y_j) = \delta_{ij} = \rho_{ij} \sqrt{\theta_i \theta_j}, \quad \rho_{ij} \geq 0, \quad i, j = 1, \ldots, p, \quad i \neq j.$$

At this point we introduce the independent gamma random variables again in standard form $$Z_1, \ldots, Z_k$$ with shape parameters, respectively, $$\xi_1, \ldots, \xi_k$$, $$k \geq p$$, such that $$Y = T \xi$$, where $$Z = (Z_1, \ldots, Z_k)'$$ and $$T$$ is an incidence matrix $$(0, 1)$$ of dimension $$(p \times k)$$ and rank $$p$$. Then,

$$E(Y) = \eta = T\xi,$$

$$\text{Cov}(Y) = \Delta = T\Omega T',$$

with $$\eta = (\eta_1, \ldots, \eta_p)'$$, $$\Delta = (\delta_{ij})$$, $$\xi = (\xi_1, \ldots, \xi_k)'$$ and $$\Omega = \text{diag}(\xi_1, \ldots, \xi_k)$$. Therefore, the generation of random vectors from $$Y$$ to run stochastic simulations with Monte Carlo methods is very easy, because it is sufficient to generate independent Gamma random vectors with the usual methods and, then, multiply them by $$T$$. Therefore, the generation of random vectors from $$X$$ can be obtained by means of the transformation $$X = BY$$.

Since $$\mu$$ and $$\Sigma$$ are given, and consequently also $$\eta$$ and $$\Delta$$, the values assigned to $$\xi_1, \ldots, \xi_k$$ should satisfy the following conditions:

$$\begin{cases}
T\xi = \eta, \\
\bar{T}\xi = c, \\
\xi \geq 0,
\end{cases}$$
where \( c = \{ \delta_{11}, \delta_{12}, \ldots, \delta_{1p}, \delta_{22}, \ldots, \delta_{pp} \} \) and \( \tilde{T} \) is a matrix of dimension \( (p(p + 1)/2) \times k \) obtained multiplying the lines of \( T \) among them in the same order of the elements of \( c \). As a consequence of \( \eta_i = \delta_{ii} = \theta_i \) for \( i = 1, \ldots, p \), these conditions are also satisfied solving the following system:

\[
\begin{aligned}
\tilde{T}\xi &= c, \\
\text{subject to } \xi &\geq 0.
\end{aligned}
\]

The following remarks can be done.

(i) With the aim of finding a parameter vector which satisfies the constraints for a wide domain, it is convenient to use an incidence matrix with a number of elements different from zero at least equal to covariances. Prékopa and Szántai (1978) suggest to use an incidence matrix with a number of column vectors equal to the number of all the different combinations which can be chosen from \( p \) different elements (removing the column vector composed of all 0, the columns are \( 2^p - 1 \) ). Using a pattern based on a recursive relation, Ronning (1977) has obtained three different incidence matrices of dimension \( p \times (k(k - 1)/2) \).

(ii) The given conditions can determine \( \theta_1, \ldots, \theta_p \) unambiguously. If this is not the case, i.e. there are at least two vectors \( \xi \) satisfying the conditions, there are infinite solutions.

(iii) Since the solutions depend on \( T \), it is not always possible to satisfy them. However, it is possible to face the problem as a linear programming problem minimizing the absolute difference between the corresponding elements of \( \tilde{T}\xi \) and \( c \). Therefore, it is necessary to solve the following system (Prékopa and Szántai, 1978):

\[
\begin{aligned}
\text{Minimize } & \sum_{i=1}^{p(p+1)/2} u_i + \sum_{i=1}^{p(p+1)/2} v_i, \\
\text{subject to } & u - v + \tilde{T}\xi = c, \\
& u \geq 0, \ v \geq 0, \ \xi \geq 0.
\end{aligned}
\]

The function of \texttt{MultiRand} which allows the generation of \( n \) random vectors from a multivariate gamma with this method and with approximately mean vector \( \mu \), covariance matrix \( \Sigma \) and incidence matrix of Prékopa and Szántai (1978) is

\[
\text{RandomArray[MultivariateGammaTypeDistribution[\mu,\Sigma],n]}
\]

4.2 Normal Copula

One of the most used methods to deal with multivariate distributions is the function copula, which is a cdf \( C \) on the unit hypercube \((0, 1]^p\) with uniform marginals such that, given the cumulative distribution functions \( F_{X_1}, \ldots, F_{X_p} \) and \( C \) (cf. Nelsen, 1999),

\[
F_{X_1,\ldots,X_p}(x_1,\ldots,x_p) = C(F_{X_1}(x_1),\ldots,F_{X_p}(x_p)).
\]

In general, the modelling of different correlated distributions is easier if a normal copula is used, which is not only very flexible in the choice of correlation parameters, but also easier to be applied in Monte Carlo studies.
Assume that \( Z = (Z_1, \ldots, Z_p)' \) has a multivariate normal distribution with standard marginals \( Z_i \sim N(0,1) \) and correlation matrix \( R \) and indicate the joint cdf of \( Z \) with \( \mathcal{G}_{Z_1, \ldots, Z_p}(z_1, \ldots, z_p) \). Then, the cdf

\[
C(u_1, \ldots, u_p) = \mathcal{G}_{Z_1, \ldots, Z_p}(\Phi^{-1}(z_1), \ldots, \Phi^{-1}(z_p)),
\]

where \( \Phi(\cdot) \) indicates the cdf of the standard normal, defines a multivariate cdf known as normal copula. Therefore, for any set of cdf of given marginals \( F_{X_1}, \ldots, F_{X_p} \), the random variables

\[
X_1 = F_{X_1}^{-1}(\Phi(z_1)), \ldots, X_p = F_{X_p}^{-1}(\Phi(z_p)),
\]

have joint cdf

\[
\mathcal{F}_{X_1, \ldots, X_p}(x_1, \ldots, x_p) = \mathcal{G}_{Z_1, \ldots, Z_p}(\Phi^{-1}(F_{X_1}(z_1)), \ldots, \Phi^{-1}(F_{X_p}(z_p)))
\]

with cdf of the marginals given by \( F_{X_1}, \ldots, F_{X_p} \).

Also if the gaussian copula does not have a simple analytical expression, it is very easy to build Monte Carlo algorithms using it. Suppose that a set of correlated distributions \( (X_1, \ldots, X_p) \) is given with cdf of the marginals given by \( F_{X_1}, \ldots, F_{X_p} \). If we assume that the multivariate distribution \( (X_1, \ldots, X_p) \) can be described by the normal copula, then it is sufficient to draw a random vector from a multivariate normal with standard marginals and correlation matrix \( R \) with the usual method and then to apply the (14) to obtain a random vector from the desired distribution. However, the random vectors generated like that not necessarily have the same given correlation matrix \( R \), therefore it can be suitable to measure the association among the given marginals using either the Spearman’s rank correlation matrix \( R_S = (\tilde{\rho}_{ij}) \) or the Kendall’s rank correlation matrix \( R_T = (\tau_{ij}) \), whose elements, as seen before, are not only invariant regarding to non linear transformations but, in the normal case, are also connected to the correlation coefficient.

We have used the normal copula function to implement a procedure in \texttt{MultiRand} which generate random vectors from a multivariate distribution with univariate marginals following a generalized extreme value distribution (Joe, 1994). For this, a usual parametrization of the cdf is given by

\[
F_X(x) = \exp\left\{-\left(1 + \xi\frac{x - \mu}{\sigma}\right)^{-1/\xi}\right\}, \quad 1 + \xi\frac{x - \mu}{\sigma} \geq 0,
\]

where \( \mu \in \mathbb{R} \) and \( \sigma > 0 \) are, respectively, position and scale parameters and \( \xi \in \mathbb{R} \) is a shape parameter. The case \( \xi < 0 \) corresponds to the Weibull distribution, \( \xi = 0 \) to the Gumbel and \( \xi > 0 \) to the Fréchet distribution. Therefore, the function of \texttt{MultiRand} which allows the generation of \( n \) random vectors from this distribution with parameter vectors of the marginals \( \text{mu}, \sigma \) and \( \xi \) is given by

\[
\text{RandomArray}[\text{MultivariateGeneralizedExtremeValueDistribution}[^\text{mu},^\text{beta},^\text{sigma},^\text{R},^\text{option}],n]
\]

where \( R \) indicates either a correlation matrix or a Spearman’s rank correlation matrix or Kendall’s rank correlation matrix according to the value given to the logical variable \( \text{option} \). Note that an extreme value multivariate distribution defined like that allows to master, in Monte Carlo studies, the dependence structure among the marginals in a more flexible way respect to the extreme value multivariate distribution described in the last section.
4.3 Extension of the Multivariate Johnson Distribution

Stanfield et al. (1996) have obtained an extension of the multivariate Johnson distribution which allows the generation of random vectors with given correlation matrix and marginals with the first four moments given. Suppose that the random vector \( X = (X_1, \ldots, X_p)' \) has mean \( \mu_X = (\mu_{X_1}, \ldots, \mu_{X_p})' \), standard deviations matrix \( \sigma_X = \text{diag}(\sigma_{X_1}, \ldots, \sigma_{X_p}) \) and correlation matrix \( R_X \). Therefore, let the lower triangular matrix \( L_X = (l_{ij}) \) be obtained by the Cholesky decomposition \( L_X L_X' = R_X \).

If the multivariate distribution \( Y = (Y_1, \ldots, Y_p)' \) has independent standard Johnson components, so that \( \text{E}(Y_i) = 0 \) and \( \text{Var}(Y_i) = 1 \) for \( i = 1, \ldots, p \), then

\[
W = \mu_X + \sigma_X L_X Y
\]

(15)

has the same mean vector and covariance matrix of \( X \). Assume that the parameters \( \gamma_i, \delta_i, \lambda_i \) and \( \xi_i \) of \( Y_i \) are arranged in a way that the \( i \)-th component \( W_i \) of the random vector \( W \) has the same skewness and kurtosis of \( X_i \) for \( i = 1, \ldots, p \).

Let \( a_X \) and \( b_X \) be vectors whose \( i \)-th element is, respectively, the skewness and kurtosis of \( X_i \) and, in a similar way, let \( a_Y \) and \( b_Y \) be the vectors indicating the skewness and kurtosis of the components of \( Y \). Moreover, define the \( k \)-fold Hadamard product of \( L_X \) as \( L_X^{(k)} = (l_{ij}^k) \) for \( k = 3, 4 \), together with the auxiliary vector \( \Psi_X = (\psi_1, \ldots, \psi_p)' \), where

\[
\psi_i = 6 \sum_{j=1}^{p} \sum_{k=j+1}^{p} l_{ij}^2 l_{ik}^2, \quad i = 1, \ldots, p.
\]

Now, if the random vector \( W \) is generated according to the transformation (15) it is to show that the vectors \( a_W \) and \( b_W \) of the skewness and kurtosis indices describing the components of \( W \) have the following relation with the vectors \( a_Y \) and \( b_Y \) of the skewness and kurtosis describing the components of \( Y \):

\[
\begin{align*}
\{ & a_W = L_X^{(3)} a_Y, \\
& b_W = L_X^{(4)} b_Y + \Psi_X.
\end{align*}
\]

Therefore, the parameter vectors \( \gamma, \delta, \lambda \) and \( \xi \) of the Johnson distribution which determine the distribution of \( Y \) can be modified to satisfy the conditions

\[
\begin{align*}
\{ & a_Y = [L_X^{(3)}]^{-1} a_X, \\
& b_Y = [L_X^{(4)}]^{-1} (b_X - \Psi_X).
\end{align*}
\]

Then, the transformed random vector \( W \) has in this case the same correlation matrix of \( X \) and also skewness and kurtosis of \( W_i \) are same of those of \( X_i \) for \( i = 1, \ldots, p \).

The function of \texttt{MultiRand} which allows to obtain \( n \) random vectors from the extension of a Johnson multivariate distribution with mean vector \( \mu \), standard deviation vector \( \sigma \), skewness vector \( a \) and kurtosis vector \( b \) and correlation matrix \( R \) is

\[
\text{RandomArray}[\text{ExtensionMultivariateJohnsonDistribution}[\mu, \sigma, a, b, R], n]
\]
5 Concluding remarks

As stressed in the introduction, the study of the behavior of sampling statistics is usually done by means of the simulation of known distributions. In this framework, it is necessary that algorithms allowing the simulation of multivariate distribution are available and, to this purpose, we have prepared a package in an algebraic software, specifically Mathematica, which allows the generation of various continuous multivariate distributions.

The distributions taken into consideration can be divided into two groups. A first group refers to specific distributions for which it is possible to write analytically the density function or the moment generating function and, as a consequence, to derive the properties of the marginals. Such marginals, in general, are an extension to the multivariate case of known random variables, which are used very often in statistical applications with unidimensional samples. On the contrary, a second group refers to distributions with a given dependence structure of the components and are defined in indirect form. We can say that the high quantity of multivariate distributions, which can be generated using MultiRand, could be sufficient to run very complex Monte Carlo experiments.

In the future we plan to implement in Mathematica other multivariate distributions, too, for example, the multivariate Liouville distributions (cf. Fang, Kotz e Ng, 1990) and the extension to the multivariate case of stable distributions (cf. Nolan, 1998).

6 References


