Contributions to the class of beta-generated distributions

I.J.H. Visagie[1], D.J. de Waal[2], S.L. Makgai[2], and A. Bekker[2]

[1]*Subject Group Statistics, North-West University, South Africa,*
jaco.visagie@nwu.ac.za
[2]*Statistics Department, University of Pretoria, South Africa*

Abstract

The beta generator technique, introduced by Eugene et al. (2002), entails constructing a univariate distribution function as a composite function of two distribution functions. The success of this technique in the univariate setting has prompted research into the possibility of generalisation to the bivariate case. Such a generalisation, using copulas, can be found in Samanthi and Sepanski (2019).

In this paper, we construct bivariate distribution functions by passing a bivariate distribution function as an argument to the univariate beta distribution function. The class of distributions obtained is identical to that studied in Samanthi and Sepanski (2019); however, the elementary elements of the two classes differ (i.e., some distributions are simple to construct using one of the techniques considered and difficult to construct using the other). This paper provides a rigorous derivation of the parameter space of the beta-generated distributions, as well as a result relating to the dependence structure of the marginals. Finally, a practical example is included demonstrating the use of a beta-generated distribution in the modelling of observed losses in the energy market.

Keywords: Baseline distribution; Beta distribution; Bivariate normal distribution; Generator; Independence; Parameter space.

1 Introduction

The following technique for the construction of flexible univariate distributions is presented in Eugene et al. (2002). Let F denote the distribution function of a $beta\,(a,b)$ random variable with $a>0$ and $b>0$; i.e.,

$$F\left(x\right)=\frac{1}{B(a,b)}\int_0^x y^{a-1}\left(1-y\right)^{b-1}\mathrm{d}y,\qquad x\in\left(0,1\right),$$

where $B(a,b)=\int_0^1 x^{a-1}(1-x)^{b-1}\mathrm{d}x$ (see Gradshteyn and Ryzhik (2007)). If G denotes a distribution function, then

$$H(x)=F(G(x)) \tag{1}$$

is a distribution function with the same support as that of G. G is generally referred to as the baseline distribution, while F is known as the generator. Although various distributions have been used as generators, the original, and most commonly used, choice of F is the beta distribution function. We restrict our attention to this choice of F for the remainder of this paper.

Note that, by its construction, G is a special case of H, obtained when $a=b=1$. A distribution obtained using the construction in (1) is more flexible than the baseline distribution used for the construction; the beta-generator technique introduces two shape parameters, denoted by a and b, into the new distribution. For more details on the distributions that have been constructed using this technique, as well as their applications and properties, the interested reader is referred to the following papers. In the context of reliability, Nadarajah and Kots (2004) and Nadarajah and Kots (2006) respectively introduce the beta-Gumbel and the beta-exponential distribution. The beta generalised-exponential distribution is introduced in Barreto-Souza et al. (2010), while Nadarajah and Gupta (2004) and Akinsete et al. (2008) respectively introduce and study the beta-Fréchet and the beta-Pareto distributions.

The Kumaraswamy distribution is a beta-type distribution with certain advantages in terms of tractability, such as a simple closed form expression for the distribution function; see Jones (2009). This distribution has also been employed as a generator when constructing distribution using the technique of Eugene et al. (2002). For example, see Cordeiro et al. (2010) for the construction of the Kumaraswamy-Weibull distribution.

Alexander et al. (2010) introduces the class of generalised beta-generated distributions. These distributions have three shape parameters in the generator distribution as opposed to the two provided by the beta distribution. Zografos and Balakrishnan (2009) considers beta-generated distributions and also demonstrates the use of the generalised gamma distribution as a generator (as an alternative of the beta distribution). A recent review of the construction of distributions via compounding can be found in Tahir and Cordeiro (2016).

The popularity of the construction technique in (1) for the univariate case has prompted research into its use for the construction of higher dimensional distributions. Samanthi and Sepanski (2019) generalised this construction technique by proposing that the univariate baseline distribution be replaced by a bivariate copula function. If G_1 and G_2

denote two univariate distribution functions, $C : \mathbb{R}^2 \to \mathbb{R}$ denotes a copula function and f is the density function of a *beta* (a, b) random variable, then

$$H(x_1, x_2) = \int_0^{C(G_1(x_1), G_1(x_2))} f(y)\mathrm{d}y \tag{2}$$

is a bivariate distribution function if, and only if, $1 \leq a < \infty$ and $0 < b \leq 1$. Note that, in the bivariate case, the use of the beta-generator technique requires additional restrictions on the values of a and b, which is not the case when using univariate distributions.

There is ongoing interest in the construction of bivariate and multivariate distributions, using various techniques, as is witnessed by the following papers. Jones (2004) introduces families of distributions arising from order statistics. Arnold et al. (2006) introduces multivariate distributions based on the distributions in Jones (2004). The beta-generated class of distributions is also closely related to the distribution of order statistics; see Eugene et al. (2002).

In addition to considering a detailed treatment of the distributions proposed in Zografos and Balakrishnan (2009), Nadarajah et al. (2015) proposes bivariate generalisation of the mentioned distributions. Sarabia et al. (2014) proposes three bivariate and multivariate distributions with beta-generated marginals as well as providing a discussion on bivariate distributions related to the beta-generator technique. This discussion includes the work contained in Arnold et al. (1999), Arnold et al. (2001) as well as Sarabia and Gómez-Déniz (2008), all of which relate to the construction of distributions based on marginal and conditional distributions. The discussion also makes reference to the distributions proposed in Jones and Larsen (2004) and Arnold et al. (2006). While both of these papers propose multivariate versions of the beta-generated distributions, the first is only suitable for modelling data above the diagonal and the second does not, in general, have beta-generated marginals.

The current paper proposes a construction technique which is similar to the construction provided in Samanthi and Sepanski (2019); i.e., we replace the copula function in the upperbound of the integral in (2) with a bivariate distribution function. Although the class of distributions obtained using the newly proposed method is identical to the class of distributions obtained using (2), certain distributions are much simpler to construct using one method or the other. This is explained in more detail in Section 2.

The remainder of the paper is structured as follows. In Section 2, we propose an alternative technique for the construction of the class of beta-generated distributions which does not require any knowledge of the theory of copulas. This section contains a rigorous derivation of the parameter space of the proposed class of distributions. Samanthi and Sepanski (2019) also provides a derivation of this parameter space; however, the mentioned derivation relies heavily on the theory of copulas, whereas an elementary proof is here. Samanthi and Sepanski (2019) provides a sufficient condition for the independence of random variables with distribution function given by (2). In Section 3, we prove that the mentioned condition is also necessary for independence. Some comments regarding the link between the value of b and the dependence structure are also included in this section. Although the aim of this paper is not to provide a detailed look at the computational aspects associated with the beta-generated class of distributions, a numerical example demonstrating the use of these distributions in the modelling of observed losses in the energy market is provided in Section 4. Finally, Section 5 provides some conclusions as well as directions for future research.

2 Construction technique

We begin this section by considering the parameter space of the bivariate beta-generated distributions. In order to proceed, we need to consider the concept of two-increasingness of a of a bivariate function. A bivariate function, $\gamma : \Omega \to \mathbb{R}$, with $\Omega = \Omega_1 \otimes \Omega_2$ is said to be two-increasing if, for every x_1, x_2 and positive δ_1, δ_2 such that $x_1, x_1 + \delta_1 \in \Omega_1$ and $x_2, x_2 + \delta_2 \in \Omega_2$, we have that

$$\gamma(x_1 + \delta_1, x_2 + \delta_2) + \gamma(x_1, x_2) \geq \gamma(x_1 + \delta_1, x_2) + \gamma(x_1, x_2 + \delta_2). \tag{3}$$

It is well-known that a bivariate distribution function is required to be two-increasing. Note that, if γ is a distribution function of the two random variables X_1 and X_2, then the difference between the left and right hand sides of (3) is the following probability; $P(x_1 \leq X_1 \leq x_1 + \delta_1, x_2 \leq X_2 \leq x_2 + \delta_2)$.

Theorem 1. *Let*

$$H(x_1, x_2) = F(G(x_1, x_2)), \tag{4}$$

where F is a (non-degenerate) beta(a, b) distribution function and G is a continuous bivariate distribution function. H is two-increasing and therefore a bivariate distribution function if, and only if, $a \in [1, \infty)$ and $b \in (0, 1]$.

A proof of this theorem can be found in the Appendix.

We are thankful to an anonymous referee for pointing out a possible relaxation of the assumptions required to prove the above theorem. Although a formal proof of the statement below is still lacking, we believe that the following conjecture may hold.

Conjecture 1 *Let $H(x_1, x_2) = F(G(x_1, x_2))$, where F is a distribution function and G is a continuous bivariate distribution function. H is two-increasing if, and only if, the derivative of F is convex.*

The support of the random variables with distribution functions H and G are identical. As was the case when using univariate distribution functions, G is a special case of H, obtained when $a = b = 1$ and H is a generalisation of G containing two shape parameters; a and b. The marginals of H are univariate beta-generated distributions (the proof of this statement is trivial and it is omitted here). As a result, the mathematical properties of the marginals can be found in the statistical literature mentioned in the previous section.

It can be shown that the density function corresponding to the distribution function in (4) is given by

$$h(x_1, x_2) = f(G(x_1, x_2))g(x_1, x_2) + f'(G(x_1, x_2))\frac{\partial}{\partial x_1} G(x_1, x_2)\frac{\partial}{\partial x_2} G(x_1, x_2), \qquad (5)$$

where f' denotes the derivative of the beta density function.

As a specific example, consider the bivariate beta-normal distribution. Let $\mu_1, \mu_2 \in \mathbb{R}$, $\sigma_1, \sigma_2 > 0$, $\rho \in [-1, 1]$, $z_1 = (x_1 - \mu_1)/\sigma_1$ and $z_1 = (x_2 - \mu_2)/\sigma_2$. In this case, the bivariate normal distribution can be expressed as

$$G(y_1, y_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \int_{-\infty}^{y_1} \int_{-\infty}^{y_2} \exp\left(-\frac{1}{2(1-\rho^2)}(z_1^2 + z_2^2 - 2\rho z_1 z_2)\right) \mathrm{d}x_2 \mathrm{d}x_1.$$

Below we refer to a distribution as standard bivariate beta-normal if $\mu_1 = \mu_2 = 0$ and $\sigma_1 = \sigma_2 = 1$. Using G above as the baseline distribution, we obtain the bivariate beta-normal distribution function

$$H(x_1, x_2) = F(G(x_1, x_2)) = \frac{B(\Phi_\rho(z_1, z_2; a, b))}{B(a, b)},$$

with $\Phi_\rho(\cdot, \cdot)$ denoting the standard normal distribution function with correlation ρ and $B(x; a, b) = \int_0^x t^{a-1}(1-t)^{b-1}$ the upper incomplete beta integral. The bivariate beta-normal density function is

$$\begin{aligned} h(x_1, x_2) &= \frac{(\Phi_\rho(z_1, z_2))^{a-1}}{B(a, b)} \left[(\phi_\rho(z_1, z_2))^{a-1}(1 - \Phi_\rho(z_1, z_2)\phi_\rho(z_1, z_2))^{b-1}\right. \\ &+ \left.\frac{1}{\sigma_1\sigma_2}\left(\frac{a-1}{\Phi_\rho(z_1, z_2)} + \frac{1-b}{1-\Phi_\rho(z_1, z_2)}\right)(1 - \Phi_\rho(z_1, z_2))^{b-1}\,\phi_\rho(z_1, z_2)\right], \end{aligned}$$

with $\phi_\rho(\cdot,\cdot)$ denoting the standard normal density function.

Let $\Phi(\cdot)$ and $\phi(\cdot)$ respectively denote the standard normal distribution and density functions in the univariate case. The marginal of h evaluated in x_1 can be expressed as

$$h(x_1,\infty) = \frac{1}{\sigma_1} f(\Phi(z_1))\phi(z_1) = \frac{\Phi(z_1)^{a-1}(1-\Phi(z_1))^{b-1}\phi(z_1)}{\sigma_1 B(a,b)}, \qquad (6)$$

with a similar result holding for x_2. The equation in (6) can be recognised as the density function of the beta-normal distribution studied in Eugene et al. (2002). Although the moments of the beta-normal distribution are generally not available in closed form, certain parameter combinations allow closed form expressions, see Eugene et al. (2002).

In order to evaluate the effect of the parameters a and b on the shape of the bivariate beta-normal distribution we include several graphs showing densities and contour plots. All calculations are performed using MATLAB (2019). In each case we set $\mu_1 = \mu_2 = 0$ and $\sigma_1 = \sigma_2 = 1$ while the values of a, b and ρ are varied. Figures 1, 2 and 3 show the resulting densities and contour plots. While it is not possible to draw exact conclusions regarding the exact effect that the values of a and b have on the moments of the distribution or on the dependence structure from the graphs, we can gain some insight. For instance increasing the value of a reduces the variance of the marginal distributions, while reducing the value of b increases the correlation as well as the tail dependence between the marginals.

In the univariate case, the beta-exponential distribution has been showed to be bimodal for certain parameter sets, see Famoye and Eugene (2004). Note that the corresponding bivariate distribution is unimodal for all parameter sets. This difference can be attributed to the difference between the parameter spaces of these distributions; Famoye and Eugene (2004) reports that the maximum value of a which results in bimodality in the univariate case is approximately 0.214, while the minimum value for a in the bivariate case is 1.

As was mentioned above, the class of distributions obtained using the techniques described in Samanthi and Sepanski (2019) and the current paper are identical. This can easily be shown to be the case using Sklar's theorem, see Sklar (1959), which asserts in the bivariate case that, for every distribution there exists a corresponding copula and for every copula there exists a corresponding distribution. However, it is worth noting that even though the two classes contain the same distributions, the elementary elements in each class differ. For example, if we use the Dirichlet distribution as the baseline, then constructing the beta-Dirichlet distribution is conceptually simple when using the

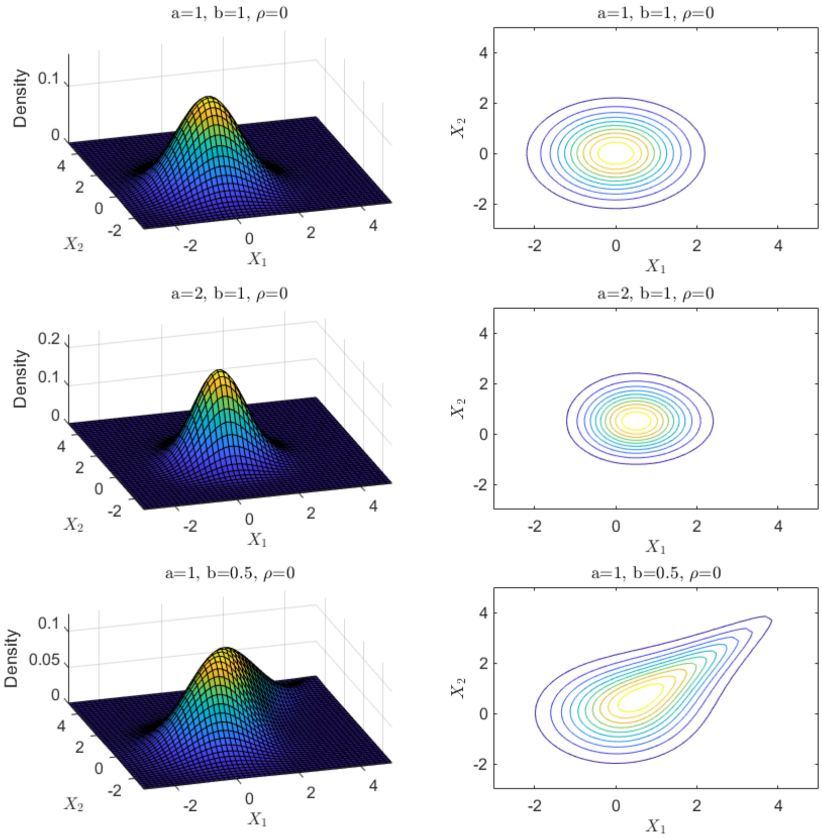

Figure 1. Density functions and contour plots of the standard bivariate beta-normal with $\rho = 0$.

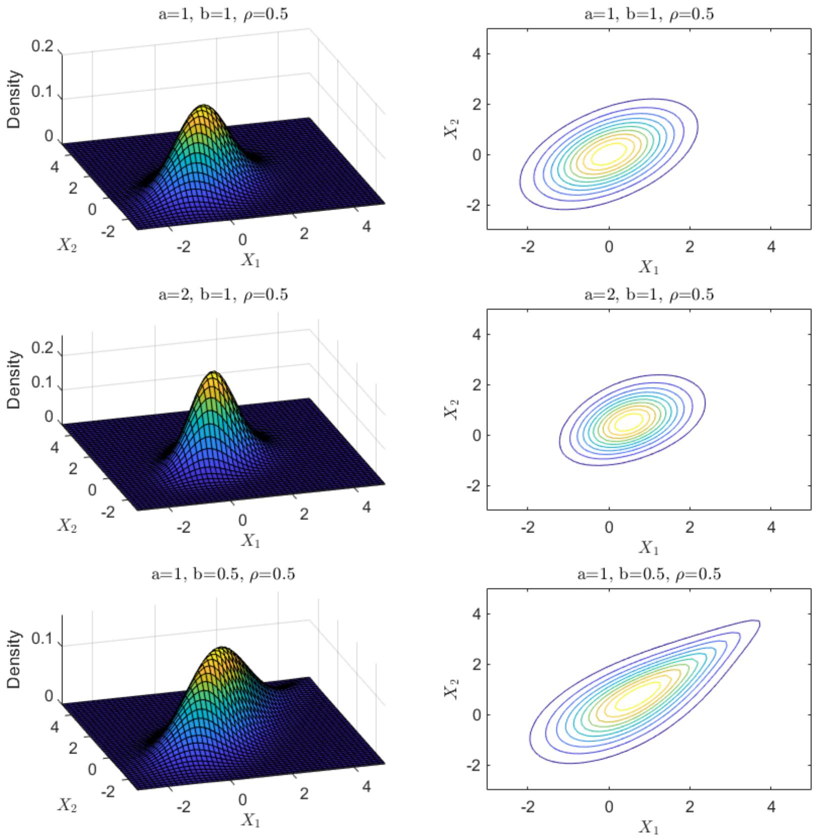

Figure 2. Density functions and contour plots of the standard bivariate beta-normal with $\rho = 0.5$.

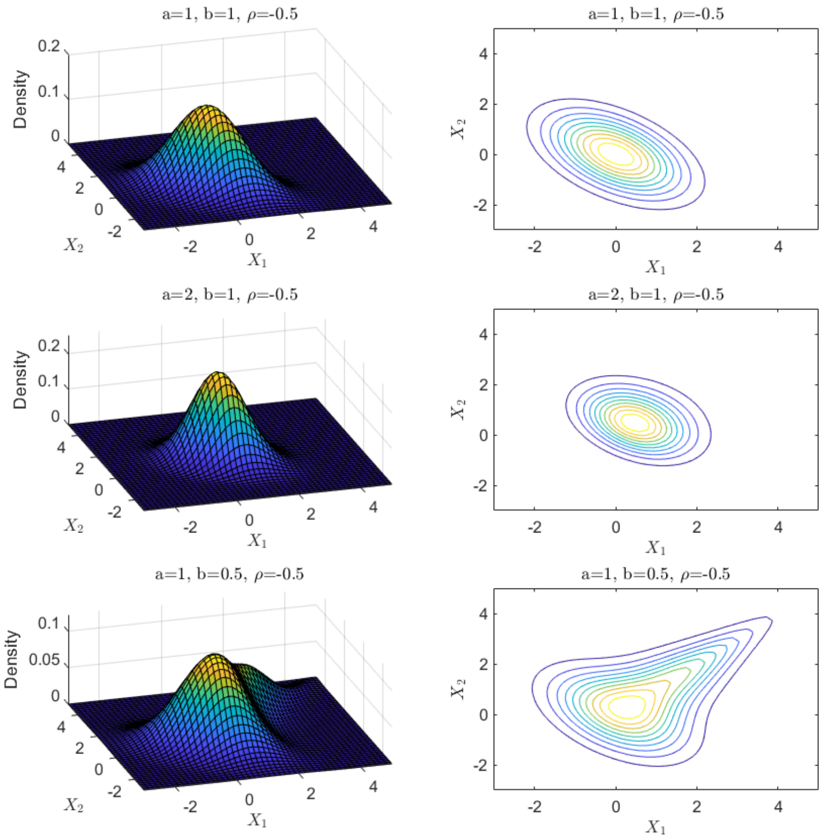

Figure 3. Density functions and contour plots of the standard bivariate beta-normal with $\rho = -0.5$.

technique proposed in the current paper. I.e., let the baseline density be Dirichlet;

$$g(x_1,x_2)=\frac{\Gamma(\theta_0+\theta_1+\theta_2)}{\Gamma(\theta_1)\Gamma(\theta_2)\Gamma(\theta_3)}x_1^{\theta_1-1}x_2^{\theta_2-1}(1-x_1-x_2)^{\theta_0-1},$$

with $\theta_0,\theta_1,\theta_2>0$ and $0\leq x_1+x_2\leq 1$, see Kotz et al. (2000). The beta-Dirichlet distribution can be constructed by setting

$$H(x_1,x_2)=F\left(\int_0^{x_1}\int_0^{min(x_2,1-x_1)}g(y_1,y_2)\mathrm{d}y_2\mathrm{d}y_1\right). \tag{7}$$

Although the calculation of the distribution function in (7) requires numerical integration, the implementation of the methodology is straightforward. The same distribution can, theoretically, be constructed using the copula-based approach of Samanthi and Sepanski (2019). However, this requires that we use a copula with the same dependence structure as that of the Dirichlet distribution. By Sklar's theorem, we know that such a copula exists. However this copula is not known in closed form. Although both construction methods mentioned may be used in order to construct the beta-Dirichlet distribution, the method advocated for in the current paper proves simpler to understand and implement in the case of this specific distribution.

3 Dependence structure

The application of the beta-generator technique alters the dependence structure of the distribution. Samanthi and Sepanski (2019) investigates the dependence structure of a random variable with distribution function H in the case where the baseline distribution has independent marginals. The mentioned paper shows that, in this case, $b=1$ is a sufficient condition in order for H to possess independent marginals. Below, we show that $b=1$ is also a necessary condition for independence to hold. In order to prove this result, we shall make use of the following result.

Theorem 2. *If F is a continuous distribution function with support $(0,1)$, then $F(x_1x_2)=F(x_1)F(x_2)$ if, and only if, $F(x)=x^t$, for some $t>0$.*

For a proof, see Engel (1998), pages 274 and 275. Using the result in Theorem 2, we can now prove the following theorem.

Theorem 3. *If $G(x_1,x_2)=G(x_1,\infty)\,G(\infty,x_2)$, then $b=1$ is a necessary and sufficient condition for $H(x_1,x_2)=F(G(x_1,x_2))$ to have independent marginals.*

Proof. X_1 and X_2 are independent if, and only if, for all (x_1, x_2) in the support of H,

$$\begin{aligned} H(x_1, x_2) &= H(x_1, \infty)H(\infty, x_2) \\ \iff \quad & F(G(x_1, x_2)) = F(G(x_1, \infty))F(G(\infty, x_2)) \\ \iff \quad & F(x) = x^t, \end{aligned} \tag{8}$$

for all $x \in (0, 1)$ and some $t > 0$, by Theorem 1. Note that (8) holds if, and only if, the density, f, has the following form;

$$f(x) = tx^{t-1}, \tag{9}$$

for some $t > 0$.

If $b \neq 1$, then f is clearly not in the form given in (9). On the other hand, if $b = 1$, then

$$\begin{aligned} f(x) &= \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1}(1-x)^{b-1} \\ &= \frac{\Gamma(a+1)}{\Gamma(a)} x^{a-1} \\ &= ax^{a-1}, \end{aligned}$$

which is of the form given in (9). □

The result above shows that, in the case where G possesses independent marginals, the value of a does not influence the dependence structure of the beta-generated distribution. This is especially interesting since we have seen in Figures 1, 2 and 3 above that the value of a alters the shape of both marginal distributions.

Although we are unable to provide a mathematical proof, we conjecture the following results relating to the dependence structure of beta-generated distributions based on numerical results obtained.

Conjecture 2 *Let ρ_1 and ρ_2 respectively denote the correlation between the marginal distributions of the generator distribution $G(x_1, x_2)$ and those of the beta-generated distribution $H(x_1, x_2) = F(G(x_1, x_2))$ where $F(x)$ is the beta distribution function with parameters $a \geq 1$ and $0 < b \leq 0$. In this case, $\rho_1 \leq \rho_2$, with equality holding only for $b = 1$. Furthermore,*

$$\lim_{b \downarrow 0} \rho_2 = 1,$$

for every G and for every value of $a \geq 1$.

4 Practical application

As was mentioned above, the aim of this section is not to study the computational aspects of the class of distributions discussed above (we plan to discuss these aspects in a forthcoming publication). As a result, we adopt a somewhat unorthodox approach below. Instead of considering observed data and then showing that the proposed models fit these data better than standard models, we opt to use a dataset that can readily be modelled using standard techniques. We then demonstrate that the distributions proposed above can be used in order to model this dataset without introducing an inordinate amount of complexity and computational difficulties.

Below we use the beta-bivariate normal distribution, discussed in Section 2, to model the joint distribution of the daily log-losses of oil (denoted by X_1 below) and gas prices (denoted by X_2 below) as measured from 13 January 2005 until 13 May 2005. The means of X_1 and X_2 are calculated to be -0.0001 and -0.0003 respectively, while the standard deviations are 0.0204 and 0.0200. The correlation between X_1 and X_2 is 0.4831. For a more detailed discussion of the data, see Bustince et al. (2013).

Financial log-losses are often assumed to be normally distributed; see Cont (2001) for a study of the empirical properties of financial losses. We investigate the normality of the observed log-losses using a graphical procedure before turning our attention to a formal goodness-of-fit test. Figures 4 and 5 show kernel density estimates of X_1 and X_2 respectively. These estimates are obtained using Matlab's *ksdensity.m* function. The figures also show the fitted normal densities superimposed using dashed lines. In both figures, the deviation between the kernel density estimate and the fitted normal density is small enough that we conclude that the normal distribution is an appropriate model.

Turning our attention to a formal goodness-of-fit test, we consider the well-known Kolmogorov-Smirnov test (also known as the Lilliefors test) for normality using Matlab's *lillietest.m* function. For the log-losses associated with the oil and gas prices, we observe p-values of 7.1% and 11.7% respectively when testing the null hypothesis of normality. As a result, we reject the assumption of normality for neither of the marginal distributions at a 5% significance level. As a result, we conclude that the univariate normal distribution is an appropriate model for both marginal distributions. We now fit the bivariate beta-normal distribution to the data in order to see if this distribution is able to accurately model not only the marginals of the distribution, but also the dependence structure.

The beta-bivariate normal distribution has a total of seven parameters. In order to speed up computation we use a method similar to that explained in Visagie (2018). That is, we use Nelder-Mead optimisation, which requires starting values; see Nelder and

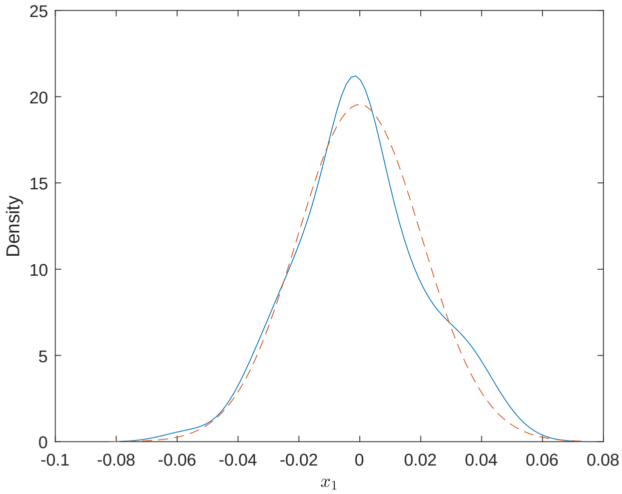

Figure 4. Kernel density estimate (solid line) of log-losses of oil prices with fitted normal density (dashed line) superimposed.

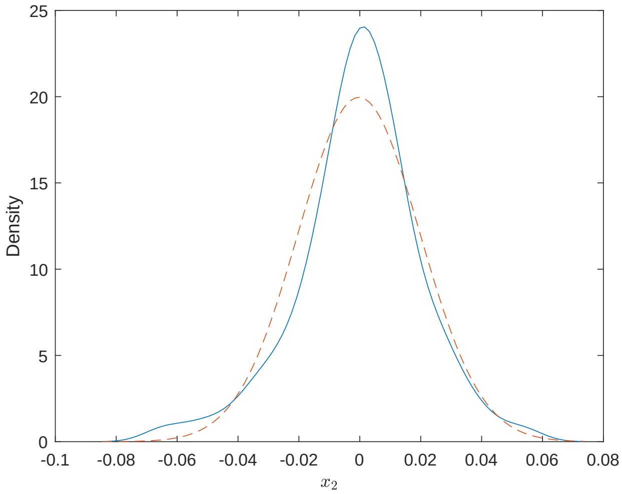

Figure 5. Kernel density estimate (solid line) of log-losses of gas prices with fitted normal density (dashed line) superimposed.

Mead (1965). In Visagie (2018), starting values are obtained by randomly generating parameter values in the specified subset of the parameter space (chosen such that the modeller believes that there is a high probability that this subset contains the true parameter values) and calculating the log-likelihood function value. This procedure is repeated a large number of times, and the parameter values associated with the largest log-likelihood function value is chosen as the starting values used in the Nelder-Mead optimisation.

In the current context, the method explained above would require the generation of seven random numbers, one for each parameter in the model. Heuristically, in order for the generated parameter set to be "good starting values", each of the seven generated parameter values should be "close to the true value". The large number of parameters in the model may cause the number of starting values required to be excessive. In order to remedy this situation, we generate random starting values only for a and b. Given these (fixed) values, we fit a univariate beta-normal distribution to each of the marginals using maximum likelihood estimation (and the *fminsearch.m* function in Matlab). This accounts for all of the parameters in the model, except the covariance parameter in the bivariate normal baseline distribution. This parameter is then also estimated by maximising the log-likelihood function in the subsequent step. The entire process is repeated a large number of times and the log-likelihood is calculated for each pair of (a, b) values generated. A parameter set with a large log-likelihood value is then used as the starting values in the final optimisation process in which all seven parameters are estimated (by maximising the joint log-likelihood).

It should be noted that users of *R*, see R Core Team (2019), may opt to use another approach when fitting this distribution. They may use the *Newdistns* package, see Nadarajah and Rocha (2016b), in order to fit the marginal distributions first and then use these parameters as starting values for the joint estimation procedure. More details on the *Newdistns* package can be found in Nadarajah and Rocha (2016a).

Using the Matlab procedure detailed above, we obtain the following set of starting values:

$$(\widehat{\mu}_1, \widehat{\mu}_2, \widehat{\sigma}_1, \widehat{\sigma}_2, \widehat{\rho}, \widehat{a}, \widehat{b}) = (-0.0083, -0.0084, 0.0003, 0.0003, 0.3343, 1.0711, 0.6675).$$

Using these starting values in the optimisation procedure, we obtain the following parameter estimates:

$$(\widehat{\mu}_1, \widehat{\mu}_2, \widehat{\sigma}_1, \widehat{\sigma}_2, \widehat{\rho}, \widehat{a}, \widehat{b}) = (-0.0025, -0.0030, 0.0005, 0.0004, 0.5254, 1.0701, 1.0000).$$

In this case the values of the log-likelihood functions achieved by the fitted bivariate normal distribution and the fitted bivariate beta-normal distribution are very similar.

In order to visually judge the fit of the beta-bivariate normal distribution to the observed values of X_1 and X_2, consider Figure 6. This figure contains contour a plot of the fitted density with the observed log-losses superimposed. The figure seems to indicate that the model fits the data well. Formal goodness-of-fit testing procedure for the beta-bivariate normal distribution have yet to be developed.

5 Conclusion

This paper proposes a technique for the construction of the bivariate beta-generated class of distributions. An elementary proof of the parameter space this class of distributions is included. Although the class of distributions constructed using the newly proposed technique is identical to the class constructed in Samanthi and Sepanski (2019), some distributions are simpler to construct using the newly proposed technique. Furthermore, the current paper does not require any knowledge of copula theory, which may make the results presented more easily accessible to certain readers.

A sufficient condition for the independence of the marginals of a beta-generated distribution (under the independence copula) is presented in Samanthi and Sepanski (2019); the condition being that the second shape parameter of the beta distribution equals 1. The current paper proves that this condition is, in fact, also necessary for independence to hold in this case. We also visually inspect the effect that the shape parameters have on the dependence structure of the distribution through the use of density and contour plots.

A practical example is included in which the observed log-losses of oil and gas prices are modelled. Using both graphical and formal techniques, we observe that the normal distribution is a realistic model for the marginals. We then fit a beta-bivariate normal distribution to the data and we see that the performance of this distribution (as measured by the likelihood function) is similar to that of the bivariate normal distribution. Visual inspection indicates that the bivariate beta-normal model fits the observed data well. However, it should be noted that formal goodness-of-fit testing for this distribution has not been addressed in the literature. This, together with the two conjectures stated in the paper, constitutes possible avenues for further research.

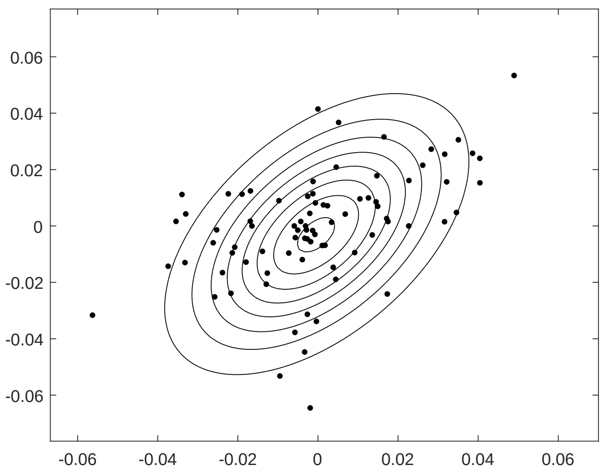

Figure 6. Observed data together with a contour plot of the fitted density.

6 Appendix

The proof for Theorem 1 can be found below.

Proof. Let Ω denote the (two dimensional) domain of G, and let $\Omega = \Omega_1 \otimes \Omega_2$, where Ω_1 and Ω_2 denote the domains of the marginal distributions of G respectively. Let x_1, x_2 and $\delta_1, \delta_2 > 0$ be arbitrary points such that $x_1, x_1 + \delta_1 \in \Omega_1$ and $x_2, x_2 + \delta_2 \in \Omega_2$.

H is two-increasing if, and only if,

$$\begin{aligned}
& H\ \ (x_1+\delta_1, x_2+\delta_2) - H\left(x_1+\delta_1, x_2\right) \geq H\left(x_1, x_2+\delta_2\right) - H\left(x_1, x_2\right) \\
\iff & F(G(x_1+\delta_1, x_2+\delta_2)) - F(G(x_1+\delta_1, x_2)) \geq F(G(x_1, x_2+\delta_2)) - F(G(x_1, x_2)) \\
\iff & \int_{G(x_1+\delta_1, x_2)}^{G(x_1+\delta_1, x_2+\delta_2)} f(t)\mathrm{d}t \geq \int_{G(x_1, x_2)}^{G(x_1, x_2+\delta_2)} f(t)\mathrm{d}t. \qquad (10)
\end{aligned}$$

Below, it is shown that $a \in [1, \infty)$ and $b \in (0, 1]$ are both sufficient and necessary for (10) to hold.

Sufficiency:

Let $a \in [1, \infty)$ and $b \in (0, 1]$. Note that f is non-decreasing on $[0, 1]$.

Consider the case where

$$G(x_1 + \delta_1, x_2) \geq G(x_1, x_2 + \delta_2). \qquad (11)$$

In this case

$$\begin{aligned}
\int_{G(x_1+\delta_1, x_2)}^{G(x_1+\delta_1, x_2+\delta_2)} f(t)\mathrm{d}t \quad & \geq \quad \left[G\left(x_1+\delta_1, x_2+\delta_2\right) - G\left(x_1+\delta_1, x_2\right)\right] f(G\left(x_1+\delta_1, x_2\right)) \\
& \geq \quad \left[G\left(x_1, x_2+\delta_2\right) - G\left(x_1, x_2\right)\right] f(G\left(x_1+\delta_1, x_2\right)) \\
& \geq \quad \left[G\left(x_1, x_2+\delta_2\right) - G\left(x_1, x_2\right)\right] f(G\left(x_1, x_2+\delta_2\right)) \\
& \geq \quad \int_{G(x_1, x_2)}^{G(x_1, x_2+\delta_2)} f(t)\mathrm{d}t,
\end{aligned}$$

where the first inequality follows from the non-decreasingness of f, and the second, from the two-increasingness of G. The third inequality is due to the non-decreasingness of f as well as the assumption made in (11), and the final inequality is again due to the non-decreasingness of f. This shows that (10) holds.

Next, we consider the case where

$$G(x_1 + \delta_1, x_2) < G(x_1, x_2 + \delta_2). \qquad (12)$$

In this case

$$
\begin{aligned}
&\int_{G(x_1+\delta_1,x_2)}^{G(x_1+\delta_1,x_2+\delta_2)} f(t)\mathrm{d}t \\
&\quad= \int_{G(x_1+\delta_1,x_2)}^{G(x_1,x_2+\delta_2)} f(t)\mathrm{d}t + \int_{G(x_1,x_2+\delta_2)}^{G(x_1+\delta_1,x_2+\delta_2)} f(t)\mathrm{d}t \\
&\quad\geq \int_{G(x_1+\delta_1,x_2)}^{G(x_1,x_2+\delta_2)} f(t)\mathrm{d}t + \left[G\left(x_1+\delta_1,x_2+\delta_2\right) - G\left(x_1,x_2+\delta_2\right)\right] f(G\left(x_1,x_2+\delta_2\right)) \\
&\quad\geq \int_{G(x_1+\delta_1,x_2)}^{G(x_1,x_2+\delta_2)} f(t)\mathrm{d}t + \left[G\left(x_1+\delta_1,x_2\right) - G\left(x_1,x_2\right)\right] f(G\left(x_1,x_2+\delta_2\right)) \\
&\quad\geq \int_{G(x_1+\delta_1,x_2)}^{G(x_1,x_2+\delta_2)} f(t)\mathrm{d}t + \left[G\left(x_1+\delta_1,x_2\right) - G\left(x_1,x_2\right)\right] f(G\left(x_1+\delta_1,x_2\right)) \\
&\quad\geq \int_{G(x_1,x_2)}^{G(x_1+\delta_1,x_2)} f(t)\mathrm{d}t + \int_{G(x_1+\delta_1,x_2)}^{G(x_1,x_2+\delta_2)} f(t)\mathrm{d}t \\
&\quad\geq \int_{G(x_1,x_2)}^{G(x_1,x_2+\delta_2)} f(t)\mathrm{d}t,
\end{aligned}
$$

where the equality follows from the continuity of f and the first inequality is due to the non-decreasingness of f. The second inequality follows from the two-increasingness of G. The third inequality is justified by the by the non-decreasingness of f as well as the assumption made in (12), while the penultimate inequality is due to the non-decreasingness of f. This shows that (10) holds.

Since both the assumptions in (11) and (12) cannot be false simultaneously, H is two-increasing if $a \in [1,\infty)$ and $b \in (0,1]$.

Necessity:

Note that $G(x_1,x_2)$ is continuous and non-decreasing in both x_1 and x_2. As a result, there exists a $k \in \Omega \subset \mathbb{R}^2$ such that $G(k) = c$, for all $c \in (0,1)$.

In order to demonstrate that (10) does not hold, it suffices to show that

$$
\int_{c_1}^{c_2} f(t)\mathrm{d}t \geq \int_{c_3}^{c_4} f(t)\mathrm{d}t \iff \frac{\int_{c_1}^{c_2} f(t)\mathrm{d}t}{\int_{c_3}^{c_4} f(t)\mathrm{d}t} > 1, \tag{13}
$$

for some $c_4 - c_3 \geq c_2 - c_1$.

Let $a \in (0,1)$. There exists a $k \in (0,1)$, such that $f(t)$ is strictly decreasing on $(0,k)$ since F is non-degenerate. Choose $c_2 \in (0,k)$. Note that $c_2 > 0$; therefore, $f(c_2) < \infty$. From the non-degenerateness of F, we have that $\lim_{c\downarrow 0} f(c) = \infty$ and that f is continuous on $(0,1)$. As a result, there exists a $c_1 \in (0, c_2)$ such that $f(c_1) = 4f(c_2)$.

Let $\epsilon \leq \frac{c_2 - c_1}{2}$ be some arbitrary positive number, then

$$c_1 < c_1 + \frac{\epsilon}{2} < c_2 - \epsilon < c_2.$$

Furthermore,

$$\int_{c_1}^{c_1+\epsilon/2} f(t)\mathrm{d}t \geq \frac{\epsilon}{2} f\left(c_1 + \frac{\epsilon}{2}\right),$$

from the decreasingness of f on $[c_1, c_1 + \frac{\epsilon}{2}]$. We also have that

$$\int_{c_2-\epsilon}^{c_2} f(t)\mathrm{d}t \leq \epsilon f\left(c_2 - \epsilon\right),$$

from the decreasingness of f on $[c_2 - \epsilon, c_2]$. Let

$$\phi\left(\epsilon\right) = \frac{\int_{c_1}^{c_1+\frac{\epsilon}{2}} f(t)\mathrm{d}t}{\int_{c_2-\epsilon}^{c_2} f(t)\mathrm{d}t} \geq \frac{f(c_1 + \frac{\epsilon}{2})}{2f(c_2 - \epsilon)}. \tag{14}$$

Note that $f(t)$ is continuous on $(0,1)$. As a result, $\int_{c_1}^{c_1+\frac{\epsilon}{2}} f(t)\mathrm{d}t$ and $\int_{c_2-\epsilon}^{c_2} f(t)\mathrm{d}t$ are continuous in ϵ. From the non-degenerateness of F, we have that $f(t) > 0$ for all $t \in (0,1)$ and that $\int_{y_1}^{y_2} f(t)\mathrm{d}t > 0$ for all $y_1 < y_2$ such that $y_1, y_2 \in [0,1]$.

Since $\int_{c_1}^{c_1+\frac{\epsilon}{2}} f(t)\mathrm{d}t$ and $\int_{c_2-\epsilon}^{c_2} f(t)\mathrm{d}t$ are both continuous in ϵ and strictly positive, we have that $\phi(\epsilon)$ is continuous in ϵ.

Consider the limit where ϵ approaches 0 from above.

$$\begin{aligned}
\lim_{\epsilon \downarrow 0} \phi(\epsilon) &= \lim_{\epsilon \downarrow 0} \frac{\int_{c_1}^{c_1+\epsilon/2} f(t)\mathrm{d}t}{\int_{c_2-\epsilon}^{c_2} f(t)\mathrm{d}t} \\
&\geq \lim_{\epsilon \downarrow 0} \frac{f(c_1 + \frac{\epsilon}{2})}{2f(c_2 - \epsilon)} \\
&= \frac{f(c_1)}{2f(c_2)} \\
&= 2,
\end{aligned}$$

where the inequality follows from (14). Together with the continuity of $\phi(\epsilon)$, this shows that there exists an $\tilde{\epsilon}$ such that

$$\frac{\int_{c_1}^{c_1+\tilde{\epsilon}/2} f(t)\mathrm{d}t}{\int_{c_2-\tilde{\epsilon}}^{c_2} f(t)\mathrm{d}t} \geq 2,$$

which demonstrates (13). Therefore, if $a \in (0,1)$, then H is not two-increasing.

Let $b \in (1,\infty)$. There exists an $m \in (0,1)$, such that $f(t)$ is strictly decreasing on $(m,1)$ since F is non-degenerate. Choose $c_1 \in (m,1)$. Note that $c_1 < 1$; therefore, $f(c_1) > 0$. Since F is non-degenerate, we have that $\lim_{g\uparrow 1} f(c) = 0$ and that f is continuous on $(0,1)$. As a result, there exists a $c_2 \in (c_1,1)$ such that $f(c_2) = \frac{1}{4}f(c_1)$.

Let $\epsilon \leq \frac{c_2-c_1}{2}$ be some positive number, then $c_1 < c_1 + \frac{\epsilon}{2} < c_2 - \epsilon < c_2$. Furthermore,

$$\int_{c_1}^{c_1+\frac{\epsilon}{2}} f(t)\mathrm{d}t \geq \frac{\epsilon}{2} f\left(c_1 + \frac{\epsilon}{2}\right),$$

since f is decreasing on $\left[c_1, c_1 + \frac{\epsilon}{2}\right]$. We also have that

$$\int_{c_2-\epsilon}^{c_2} f(t)\mathrm{d}t \leq \frac{\epsilon}{2} f\left(c_1 - \epsilon\right),$$

since f is decreasing on $[c_2-\epsilon, c_2]$.

Let

$$\phi\left(\epsilon\right) = \frac{\int_{c_1}^{c_1+\frac{\epsilon}{2}} f(t)\mathrm{d}t}{\int_{c_2-\epsilon}^{c_2} f(t)\mathrm{d}t} \geq \frac{f(c_1+\frac{\epsilon}{2})}{2f(c_2-\epsilon)}. \tag{15}$$

Note that $f(t)$ is continuous on $(0,1)$. As a result, $\int_{c_1}^{c_1+\frac{\epsilon}{2}} f(t)\mathrm{d}t$ and $\int_{c_2-\epsilon}^{c_2} f(t)\mathrm{d}t$ are continuous in ϵ. Since F is non-degenerate, we have that $f(t) > 0$ for all $t \in (0,1)$ and that $\int_{y_1}^{y_2} f(t)\mathrm{d}t > 0$ for all $y_1 < y_2$ such that $y_1, y_2 \in [0,1]$.

Since $\int_{c_1}^{c_1+\frac{\epsilon}{2}} f(t)\mathrm{d}t$ and $\int_{c_2-\epsilon}^{c_2} f(t)\mathrm{d}t$ are both continuous in ϵ and strictly positive, we have that $\phi(\epsilon)$ is continuous in ϵ.

Consider the limit where ϵ approaches 0 from above.

$$\begin{aligned}
\lim_{\epsilon\downarrow 0}\phi(\epsilon) &= \lim_{\epsilon\downarrow 0}\frac{\int_{c_1}^{c_1+\epsilon/2} f(t)\mathrm{d}t}{\int_{c_2-\epsilon}^{c_2} f(t)\mathrm{d}t} \\
&\geq \lim_{\epsilon\downarrow 0}\frac{f(c_1+\frac{\epsilon}{2}}{2f(c_2-\epsilon)} \\
&= \frac{f(c_1)}{2f(c_2)} \\
&= 2,
\end{aligned}$$

where the inequality follows from (15). Together with the continuity of $\phi(\epsilon)$, this shows that there exists a $\tilde{\epsilon}$ such that

$$\frac{\int_{c_1}^{c_1+\tilde{\epsilon}/2} f(t)\mathrm{d}t}{\int_{c_2-\tilde{\epsilon}}^{c_2} f(t)\mathrm{d}t} \geq 2,$$

which demonstrates (13). Therefore, if $b \in (1, \infty)$, then H is not two-increasing.

□

References

Akinsete, A., Famoye, F., and Lee, C. (2008). The beta-Pareto distribution. *Statistics*, 42(6):547–563.

Alexander, C., Cordeiro, G. M., and Ortega, E. M. M. (2010). Generalized beta-generated distributions. *Comput. Statist. Data. Anal.*, 56(6):1880–1897.

Arnold, B. C., Castillo, E., and Sarabia, J. M. (1999). *Conditional Specification of Statistical Models*. Springer Verlag, New York.

Arnold, B. C., Castillo, E., and Sarabia, J. M. (2001). Conditionally specified distributions: an introduction (with discussion). *Stat. Sci.*, 16:249–274.

Arnold, B. C., Castillo, E., and Sarabia, J. M. (2006). Families of multivariate distributions involving the Rosenblatt construction. *J. Amer. Statist. Assoc.*, 101:1652–1662.

Barreto-Souza, W., Santos, A. H. S., and Cordeiro, G. M. (2010). The beta generalized-exponential distribution. *J. Stat. Comput. Simul.*, 80(2):159–172.

Bustince, H., Fernandez, J., Mesiar, R., and Calvo, T. (2013). *Aggregation Functions in Theory and in Practise: Advances in Intelligent Systems and Computing*. Springer, Berlin, Heidelberg.

Cont, R. (2001). Empirical properties of asset returns: stylized facts and statistical issues. *J. Comput. Finance*, 1:223–236.

Cordeiro, G. M., Ortega, E. M., and Nadarajah, S. (2010). The Kumaraswamy Weibull distribution with application to failure data. *J. Franklin Inst.*, 347(8):1399–1429.

Engel, A. (1998). *Problem-Solving Strategies: Problem books in Mathematics*. Springer Verlag, New York.

Eugene, N., Lee, C., and Famoye, F. (2002). Beta-normal distribution and its applications. *Commun. Stat. - Theory and Methods*, 31(4):497–512.

Famoye, F.and Lee, C. and Eugene, N. (2004). Beta-normal distribution: bimodality properties and application. *J. Mod. Appl. Stat. Methods*, 3(1):85–103.

Gradshteyn, I. S. and Ryzhik, I. M. (2007). *Table of Integrals, Series and Products.* Academic Press, San Diego.

Jones, M. (2009). Kumaraswamy's distribution: a beta-type distribution with some tractability advantages. *Stat. Methodol.*, 6(1):70–81.

Jones, M. and Larsen, P. V. (2004). Multivariate distributions with support above the diagonal. *Biometrika*, 91:975–986.

Jones, M. C. (2004). Families of distributions arising from distributions of order statistics. *TEST*, 13:1–43.

Kotz, S., Balakrishnan, N., and Johnson, N. L. (2000). *Continuous Multivariate Distributions: Models and Applications, Volume 1, Second edition.* John Wiley & Sons, Inc., New York.

MATLAB (Release 2019). *The MathWorks, Inc., Natick, Massachusetts, United States.*

Nadarajah, S., Cordeiro, G. M., and Ortega, E. M. M. (2015). The Zografos-Balakrishnan-G family of distributions: mathematical properties and applications. *Commun. Stat. - Theory and Methods*, 44:186–215.

Nadarajah, S. and Gupta, A. (2004). The beta-Fréchet distribution. *Far East J. Theor. Stat.*, 14:15–24.

Nadarajah, S. and Kots, S. (2004). The beta-Gumbel distribution. *Math. Probab. Eng.*, 10:323–332.

Nadarajah, S. and Kots, S. (2006). The beta exponential distribution. *Reliab. Eng. Syst. Safety*, 91:689–697.

Nadarajah, S. and Rocha, R. (2016a). Newdistns: An R package for new families of distributions. *J. Stat. Softw*, 69(1-32).

Nadarajah, S. and Rocha, R. (2016b). *Newdistns: Computes PDF, CDF, Quantile and Random Numbers, Measures of Inference for 19 General Families of Distributions.* R package version 2.1.

Nelder, J. A. and Mead, R. (1965). A simplex method for function minimization. *Comput. J.*, 7:308–313.

R Core Team (2019). *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria.

Samanthi, R. G. M. and Sepanski, J. (2019). A bivariate extension of the beta generated distribution derived from copulas. *Commun. Stat. - Theory and Methods*, 48(5):1043–1059.

Sarabia, J. M. and Gómez-Déniz, E. (2008). Construction of multivariate distributions: a review of some recent results (with discussion). *SORT*, 32:3–36.

Sarabia, J. M., Prieto, F., and Jordá, V. (2014). Bivariate beta-generated distributions with applications to well-being data. *J. Stat. Dist. Applic.*, 1(15):1–35.

Sklar, A. (1959). Fonctions de répartition à n dimensions et leurs marges. *Publ. Inst. Statist. Univ. Paris*, 8:229–231.

Tahir, M. H. and Cordeiro, G. M. (2016). Compounding of distributions: a survey and new generalized classes. *J. Stat. Dist. Applic.*, 3(16):1–35.

Visagie, I. J. H. (2018). On parameter estimation in multi-parameter distributions. *Stat., Optim. Inf. Comput.*, 6:452–467.

Zografos, K. and Balakrishnan, N. (2009). On families of beta and generalised gamma-generated distributions and associated inference. *Stat. Methodol.*, 6:344–362.