Revista Ciencia en Desarrollo. Vol. 16 Num. 2. pp. 142 - 151
julio - diciembre

DOI: 10.19053/uptc.01217488.v16.n2.2025.18956

Artículo de investigación

Three strategies for the generation of random numbers from the Bivariate Weibull distribution assuming dependence: copula methodologies

Tres estrategias para la generación de números aleatorios a partir de la distribución Weibull bivariada asumiendo dependencia: metodologías de cópulas

Cristian David Correa-Álvarez[1] ✉, Mario César Jaramillo-Elorza[2] y Osnamir Elias Bru-Cordero[3]

[1]Universidad Nacional de Colombia, Sede Manizales, Departamento de Matemáticas y Estadística, Manizales, Caldas, Colombia.
[2]Universidad Nacional de Colombia, Sede Medellín, Escuela de Estadística, Medellín, Antioquia, Colombia.
[3]Universidad Nacional de Colombia, Sede de La Paz, Dirección Académica, La Paz, Cesar, Colombia.

Recepción: 22-febrero-2025 **Aceptado:** 30-mayo-2025 **Publicado:** 20-julio-2025
Cómo citar: Correa-Álvarez, C. D., Jaramillo-Elorza, M. C., & Bru-Cordero, O. E. (2025). Tres estrategias para la generación de números aleatorios a partir de la distribución Weibull bivariada asumiendo dependencia: metodologías de cópulas. *Ciencia En Desarrollo*, 16(2).
doi: 10.19053/uptc.01217488.v16.n2.2025.18956

Abstract

Bivariate Weibull distribution-generating methods include the idea of dependence. Bivariate Weibull models incorporated both positive and negative dependence in the same parametric setting, which improved the understanding of failure behaviors in dependent component systems. These generative methods provide insights into dependence structures, statistical properties, and parameter estimation techniques for bivariate Weibull distributions, making them essential for survival and reliability analysis. This paper compares three methods for generating data from a bivariate distribution. The results indicate scenarios where the methods generate data from this distribution best. Method 3 performs poorly in a small sample size $(n = 50)$ only when the dependence is low $(\alpha = 0.25)$. Method 2 (CD-Vines) performs well for the 50% dependence scenario, independent of sample size, and method 1 under the scenarios works very well but has a restriction if we find that the data do not follow a positive stable distribution. Furthermore, evaluation metrics confirm that, among the three approaches, Method 2 provides the best overall performance for generating data that closely matches the theoretical bivariate Weibull distribution.

Keywords: Copula, Dependence, Generating, Bivariate weibull models, Theoretical contour.

Resumen

Los métodos de generación de distribuciones bivariadas Weibull incluyen la idea de dependencia. Los modelos bivariados con esta distribución incorporan tanto dependencia positiva como negativa en un mismo marco paramétrico, lo que ha mejorado la comprensión de los comportamientos de fallo en sistemas de componentes dependientes. Estos métodos generativos proporcionan información sobre las estructuras de dependencia, las propiedades estadísticas y las técnicas de estimación de parámetros para distribuciones bivariadas Weibull, lo que los hace esenciales para el análisis de supervivencia y confiabilidad. Este artículo compara tres métodos para generar datos a partir de una distribución bivariada. Los resultados indican en los escenarios que los métodos generan datos de esta distribución de manera óptima. El método 3 tiene un rendimiento deficiente en tamaños de muestra pequeños $(n = 50)$ solo cuando la dependencia es baja $(\alpha = 0.25)$. El método 2 (CD-Vines) funciona bien en el escenario de dependencia del 50%, independientemente del tamaño de la muestra, y el método 1 funciona muy bien en los escenarios considerados, pero tiene una restricción si se encuentra que los datos no siguen una distribución estable positiva. Además, las métricas de evaluación confirman que, entre los tres enfoques, el método 2 ofrece el mejor rendimiento general para generar datos que se ajustan estrechamente a la distribución teórica Weibull bivariada.

Palabras Clave: Cópula, Dependencia, Generación; Modelos Weibull bivariados, Contorno teórico.

E-ISSN: 2462-7658

1 Introduction

In the field of reliability analysis, the ability to accurately model and generate random numbers for a survival bivariate Weibull distribution is crucial [1]. The Weibull distribution is a commonly used model for the lifetimes of various products and systems due to its flexibility in capturing different shapes of failure rates. It is employed in a range of disciplines, including engineering, medical research, and finance [2, 3]. The Weibull distribution is a wide distribution used in reliability to model failure times [4]. The Weibull distribution can be interpreted in various contexts, including industry, survival analysis, and others. Moreover, this distribution is of significant importance for application purposes, due to its straightforward physical interpretation and flexibility for empirical fit [5, 6, 7].

As stated in [8], this distribution is highly versatile, capable of assuming the characteristics of other distributions based on the value of the shape parameter. It can also model decreasing, increasing, or constant hazard functions, encompassing the entire lifetime of an item. Additionally, the specific application of this distribution to failure data analysis is discussed. Furthermore, multivariate distributions are useful for modeling dependent random variables. A classification that has received some attention is based on the dependence structure between the random variables. [9] and [10] have studied the dependence structure of some bivariate Weibull models. [11] present a simple analysis for a Weibull distribution. This analysis considers two observed failure modes, which are known to be shock absorbers.

Other authors, such as [12] addressed similar issues and, under dependence, provided applications in demography and actuarial science. This latter field is crucial in reliability and survival analysis. Moreover, the work of [13] examines the relationships between dependence in multivariate models, competing risks, and copulas. Furthermore, in the context of meta-analysis of individual patient data with semi-competing risks, the joint frailty-copula model has emerged as a valuable tool. [14] have proposed the use of the Weibull distribution for baseline hazard functions under this model, demonstrating its advantages in terms of parameter interpretation and prognostic inference. [15] have proposed copula-based competing risks models for latent failure times, allowing for a flexible parametric form to address the limitations of existing models for left-truncated field data. Copulas have emerged as a novel tool for multivariate modeling in a variety of fields where the understanding of multivariate dependence is crucial. They enable the representation of the dependency structure and marginal distributions of a multivariate distribution separately [16, 17, 18].

On the other hand, the bivariate Weibull distribution is a valuable tool for modeling failures in multi-component systems where failure times are statistically dependent [19, 20]. When component failures are independent, they can be analyzed separately by observing univariate marginal distributions. [21] present three approaches to generating random values of a bivariate Weibull distribution. The first approach is through the exponential distribution and power transformation. The second approach is by specifying the dependence between two or more univariate marginals. The third approach is by transforming multivariate extreme value distributions. Also, [22] introduce the Farlie-Gumbel-Morgenstern (FGM) copula combined with Weibull marginal distribution to form the FGM bivariate Weibull (FGMBW) distribution. This distribution effectively describes weakly correlated bivariate lifetime data, offering a practical alternative for modeling real-valued data. In this article, three methods for generating data from a bivariate joint distribution are compared.

The use of bivariate Weibull distributions has been a topic of interest in various fields, with researchers exploring methods to generate dependent failure time data. [23] and [24] developed an algorithm that utilizes an Archimedean copula representation to generate dependent bivariate Weibull failure time data. Building on this, [25] proposed the use of CD-Vines to simulate flexible multivariate distributions, which are essential in various applications. Furthermore, [26] studied two distinct models of competitive risks: one with independent Weibull-distributed risks and another derived from a bivariate Weibull with dependency. Recently, [27] introduce a Weibull-based bivariate failure time model for dynamic prediction, overcoming spline model limitations. [28] propose a novel bivariate generalized Weibull distribution, validating its effectiveness through simulations and real data. These advancements enhance survival analysis methodologies, addressing complexities in predicting patient outcomes. In the present article, we focus on the bivariate Weibull distribution and present a simulation scheme using CD-Vines to generate dependent data.

This article encompasses several sections, including the following: in Section 2, we formally present four groups of tests to be performed, namely normality tests based on moments, empirical distribution, correlation, and regression, and the last group seen as a particular case according to specification; Section 4 provides a detailed account of the methodology using two power estimation methods; Section 5 illustrates the simulation scenarios and presents the study's conclusions. Finally, in Section 6, the conclusions of the study are presented, along with some remarks for future work.

2 Bivariate Weibull Distribution

The bivariate Weibull distribution is of significant importance in reliability and survival analysis, allowing for modeling the dependence between failure times [29]. Various studies have proposed methods for generating dependent bivariate Weibull failure times using copula representations [30, 31]. Additionally, relevant statistical properties of this distribution, such as marginal distributions, moments, and measures of dependence. These advancements in modeling the bivariate Weibull distribution contribute to the development of more robust techniques in survival and reliability analysis of systems [22, 28].

In this article, we will employ the parameterization of the bivariate Weibull distribution, as presented in Equation (1). Let T_1 and T_2, Weibull failure times. A joint reliability function for the Weibull bivariate, [10] is:

$$S(t_1,t_2) = \exp\left\{-\left[\left(\frac{t_1}{\eta_1}\right)^{\frac{\beta_1}{\alpha}} + \left(\frac{t_2}{\eta_2}\right)^{\frac{\beta_2}{\alpha}}\right]^{\alpha}\right\} \qquad (1)$$

with $0 < \alpha \leq 1$, where β_1, η_1, β_2 and η_2, all greater than zero, are the shape parameters and scale associated to T_1 and T_2, respectively. α is the dependence parameter between T_1 and T_2. The Marginal reliability functions are given in (2):

$$S_k(t) = \exp\left\{-\left(\frac{t}{\eta_k}\right)^{\beta_k}\right\},\ k = 1,2. \qquad (2)$$

with t positive. When the parameter dependency α Weibull time between failure is 1, then there independently T_1 and T_2. As reduces α, the dependence between T_1 and T_2 increases.

3 Methods

There are a variety of approaches in the literature to creating joint distribution functions. However, the most common approach is to view the joint as the outcome of a marginal probability density function and a conditional density function. This is especially true when one of the marginals can be assumed to be the conditional outcome conditional on a predicted value of the other marginal [32, 33, 34].

In theory, constructing the joint distribution function when there is independence involves multiplying the marginal distribution functions. However, sometimes, the existence of dependence is ignored, leading to an overestimation or underestimation of the true joint function. In that vein, creating a bivariate distribution by combining dependent distributions involves adding an appropriate term to an independent class, which can provide additional flexibility in applications [35, 36].

This paper presents three methods for generating data from a bivariate Weibull distribution using copula-based techniques. A special class of vine copulas includes C-Vines (*canonical vines*) and D-Vines (*drawable vines)*. From a graphical point of view, a C-Vine as well as D-Vine is formed by a cascade of trees corresponding each other to a specific decomposition of the distribution. The Gaussian distribution is very restrictive and can not detect characteristics such as skewness and heavy tails, very common in financial analysis [25].

The use of copulas is a challenge in higher dimensions where standard multivariate copulas have or possess inflexible structures, the vines copulas overcome these limitations and are able to model complex patterns of dependence because of having a variety of bivariate copulas as basic components [25]. The following are the algorithms used by each of the methods that will be the subject of study.

3.1 Method 1: Copula Algorithm

Based on the algorithm of positive stable distribution (see details in [24]) to generate time bivariate Weibull failure, with dependence, taking the Gumbel copula for reliability Weibull function, the algorithm is as follows:

1. Generate a random variable γ with positive stable distribution $S_\alpha(\alpha,\beta,\mu)$ and having Laplace transform $\tau(s) = \exp(-S^\alpha)$. This is true when $\sigma = \left[\cos\frac{\pi\alpha}{2}\right]^{\frac{1}{\alpha}}$, $\beta = 1$ and $\mu = 0$.
2. Generate variables u_1 and u_2 with uniform distribution on the interval $(0,1)$.
3. For $k = 1,2$, in each marginal calculate $T_k = S_k^{-1}(u_{*k})$, where $u_{*k} = \tau\left[-\gamma^{-1}\ln(u_k)\right]$, with τ with the Laplace transform associated with the Gumbel copula.

The algorithm uses a positive stable distribution. To represent the bivariate Weibull reliability function, a random variable with a positive stable distribution and its Laplace transform must be created. The scheme explained above simulates this type of distribution, which negatively affects the algorithm since the fit of these distributions is limited. In addition, we only have the setting for a Gumbel copula, such a copula works only positive dependence, the copula parameter starts from independence to perfect dependence [13].

3.2 Method 2: Methodology of Vines

According to [25], the simulation of a (C-Vine) or a (D-Vine) with n variables starts with the generation of an equal number of uniform , independent observations $w_i \in (0,1)$. From these observations, we obtain a new uniform sample variables $u_1, u_2, \ldots, u_n$, with the dependence structure which is represented by the vine according to:

$$\begin{array}{rcl} u_1 & = & w_1 \\ u_2 & = & F_{2|1}^{-1}(w_2 \mid u_1) \\ \vdots & = & \\ u_n & = & F_{n|1,2,\ldots,n-1}^{-1(w_n|u_1,\ldots,u_{n-1})} \end{array} \tag{3}$$

Where F is the joint distribution function of w given u. The variables of the new population are obtained from the uniform sample variables $u_1, u_2, \ldots, u_n$, using the Inversion Method [37].

CD-Vines refer to the canonical vine (C-vine) and (D-vine) copulas, advanced tools for modeling complex dependency structures in multivariate distributions. Vine copulas, such as C-vines and D-vines, offer more flexibility than traditional multivariate Gaussian distributions, allowing for asymmetry and heavy-tailed patterns [38]. Developing algorithms for D-vine and C-vine copulas has enabled the modeling of high-dimensional dependence structures and the efficient handling of big data sets [25].

Overall, CD-Vines play a crucial role in modern statistical modeling for capturing intricate dependencies in data. In this work we take advantage of this flexibility and there is no restriction in the selection of the copula, which gives us better performance when working with both negative and positive dependencies.

3.3 Method 3: Survival Copula

Survival copulas can also be used in the conditional distribution function method to generate random variables from a distribution with a given survival function. Recall (see part 3 of Theorem 2.4.4 and (2.6.1) [13]) that if the copula C is the distribution function of a pair (U,V), then the corresponding survival copula $\widehat{C}(u,v) = u+v-1+C(1-u,1-v)$ is the distribution function of the pair $(1-U,1-V)$. Also note that if U is uniform on $(0,1)$, so is the random variable $1-U$.

Hence we have the following algorithm to generate a pair (U,V) whose distribution function is the copula C, given $\widehat{C}$:

1. Generate two independent uniform (0,1) variables u and t.
2. Set $v = \widehat{c}_u^{(-1)}(t)$, where $\widehat{c}_u^{(-1)}$ denotes a quasi-inverse of $\widehat{c}_u^{(-1)}(v) = \partial\widehat{C}(u,v)/\partial u$.
3. The desired pair is (u,v).

Survival copulas are an important tool in survival analysis. They allow for more flexible and effective modeling of the dependence between survival time variables, see [39, 13].

3.4 Metrics for the Evaluation of Simulation Methods

To evaluate the fit between the data simulated by each of the three proposed methods and the theoretical density of the bivariate Weibull distribution, we used three widely recognized metrics: Kullback-Leibler divergence, Hellinger Distance, and Mean Squared Error. These provide a quantitative comparison of the performance of each method [40, 41, 42, 43].

- **Kullback-Leibler Divergence (KL):** Measures how much information is lost when approximating the theoretical distribution P with the estimated distribution Q:

$$D_{KL}(P||Q) = \sum_i P(x_i) \log\left(\frac{P(x_i)}{Q(x_i)}\right)$$

- **Hellinger Distance (HD):** Quantifies the global discrepancy between two distributions:

$$H(P,Q) = \frac{1}{\sqrt{2}} \sqrt{\sum_i \left(\sqrt{P(x_i)} - \sqrt{Q(x_i)}\right)^2}$$

- **Mean Squared Error (MSE):**

$$\text{MSE}(P,Q) = \frac{1}{n}\sum_{i=1}^{n} (P(x_i) - Q(x_i))^2$$

3.5 R Script for Simulation and Evaluation

This section details the R script used for the simulation study and the evaluation of the different copula models.

```
# Install if necessary
# install.packages(c("copula","kde
copula","VineCopula","ks","MASS",
"philentropy","reshape2"))

library(copula)
library(kdecopula)
library(VineCopula)
library(ks)
library(MASS)
library(philentropy)
library(reshape2)

generate_data_gumbel<-function(n,
alpha,beta1,beta2,eta1,eta2){
theta<-1/alpha
gumbel_cop<-gumbelCopula(theta,
dim=2)
u<-rCopula(n,gumbel_cop)
x<-eta1*(-log(u[,1]))^(1/beta1)
y<-eta2*(-log(u[,2]))^(1/beta2)
cbind(x,y)
}

generate_data_cdvine<-function(n,
beta1,beta2,eta1,eta2){
RVM<-RVineMatrix(Matrix=matrix(c(
0,2,1,0),2,2),family=matrix(c(0,1,
0,0),2,2),
par=matrix(c(0,1.5,0,0),2,2),
par2=matrix(0,2,2))
u<-RVineSim(n,RVM)
x<-eta1*(-log(u[,1]))^(1/beta1)
y<-eta2*(-log(u[,2]))^(1/beta2)
cbind(x,y)
}

generate_data_survival<-function(
n,beta1,beta2,eta1,eta2){
clayton_cop<-claytonCopula(2,
dim=2)
u<-rCopula(n,clayton_cop)
u_surv<-1-u
x<-eta1*(-log(u_surv[,1]))^(1/
beta1)
y<-eta2*(-log(u_surv[,2]))^(1/
beta2)
cbind(x,y)
}

calculate_metrics<-function(data,
true_density_func){
kde_data<-kde(x=data,compute.cont=
TRUE)
grid<-expand.grid(x=kde_data$eval.
points[[1]],y=kde_data$eval.points
[[2]])
kde_vals<-matrix(predict(kde_data,
x=grid),ncol=1)
true_vals<-true_density_func(grid$
x,grid$y)

#Normalize both densities
kde_vals<-kde_vals/sum(kde_vals)
true_vals<-true_vals/
sum(true_vals)
kl<-KL(rbind(true_vals,kde_vals))$
Kullback-Leibler'
hell<-sqrt(sum((sqrt(true_vals)-
sqrt(kde_vals))^2))/sqrt(2)
mse<-mean((true_vals-kde_vals)^2)
list(KL=kl,Hellinger=hell,MSE=mse)
}

evaluate_scenario<-function(n,
alpha,beta1,beta2,eta1,eta2){
#Simplified bivariate Weibull
theoretical density
true_density<-function(x,y){
alpha_weibull<-alpha
z<-((x/eta1)^beta1+
(y/eta2)^beta2)^(1/alpha_weibull)
f<-alpha_weibull*beta1*beta2*
(x^(beta1-1))*(y^(beta2-1))*
exp(-z)*((x/eta1)^(beta1-1)*
(y/eta2)^(beta2-1))
return(f)
}

g1<-generate_data_gumbel(n,alpha,
beta1,beta2,eta1,eta2)
g2<-generate_data_cdvine(n,beta1,
beta2,eta1,eta2)
```

```
g3<-generate_data_survival(n,
beta1,beta2,eta1,eta2)

m1<-calculate_metrics(g1,
true_density)
m2<-calculate_metrics(g2,
true_density)
m3<-calculate_metrics(g3,
true_density)

data.frame(
Method=c("Gumbel","CD-Vine",
"Survival"),
KL=c(m1$KL,m2$KL,m3$KL),
Hellinger=c(m1$Hellinger,
m2$Hellinger,m3$Hellinger),
MSE=c(m1$MSE,m2$MSE,m3$MSE),
Sample_Size=n,Dependence=alpha,
beta1=beta1,beta2=beta2)
}

param_grid<-expand.grid(
n=c(50,100,200),alpha=c(0.25,
0.5,0.75),beta1=c(2),
beta2=c(1),eta1=1,eta2=1)

resultados<-do.call(rbind,
apply(param_grid,1,function(p){
evaluate_scenario(p["n"],
p["alpha"],p["beta1"],p["beta2"],
p["eta1"],p["eta2"])
}))

library(acepack)
library(knitr)
library(kableExtra)
library(dplyr)

kable(resultados,digits=4,caption=
"Goodness-of-Fit Metrics by Method
and Scenario")%>%kable_styling(
"striped",full_width=FALSE)
```

4 Methodology

The methodology proposed in this study aims to analyze three methods of generating bivariate data from a Weibull distribution: the Copula Algorithm, CD-Vines and Survival Copula. To achieve this, a series of systematic steps designed to collect, process, and analyze the data rigorously and consistently will be carried out.

1. Using the methods described above, samples of size 50, 100 and 200, dependence parameter $\alpha = 0.25, 0.50, 0.75$ were generated.

2. Table 1 presents the four scenarios calculated by varying the dependency parameter. In general, there are twelve scenarios for each of the three generators.

Table 1: Scale and shape parameter values for the Bivariate Weibull distribution

β_1	β_2	η_1	η_2
2.0	1.0	1.0	1.0
0.5	2.0	1.0	1.0
0.5	1.0	1.0	1.0
0.5	0.5	1.0	1.0

3. Next, scatter plots were generated for each sample and overlaid with the contour lines of the corresponding theoretical bivariate Weibull joint density function.

4. Subsequently, for each sample of size n, n data pairs were generated. The proportion of points falling outside each theoretical contour was then examined to assess the quality of the generated data.

5. Finally, a quantitative assessment of the fit is performed using density similarity metrics, including the KL, HD and MSE for each method.

5 Simulation Study: Results

This section presents the results of the simulations carried out for the three methods introduced in Section 3, following the methodology described in Section 4. Contour plots are provided for each method under different parameter settings and sample sizes of the bivariate Weibull distribution. Furthermore, the goodness of fit between the simulated data and the theoretical density is quantitatively assessed using the Kullback-Leibler divergence, Hellinger's distance and the mean squared error.

5.1 Contour Plots

This subsection focuses on the visual assessment of the methods through contour plots, highlighting differences in the simulated densities across various parameter configurations. These plots help illustrate how well each method captures the shape and dependence structure of the bivariate Weibull distribution.

Figure 1 illustrates the diverse copula-based methodologies employed to generate data exhibiting dependence from a bivariate Weibull distribution, with a sample size of 50. It is evident that despite differences in methodology, all methods yield consistent results, effectively capturing positive dependence with comparable density contour patterns and data point dispersion. However, the copula algorithm method stands out in its ability to effectively analyze data with this sample size. This consistency suggests the robustness of copula methods in analyzing bivariate dependencies in simulated data. Similar results are observed in Figures 2 with a sample size of 100 and Figures 3 with a sample size of 200, while maintaining the distribution parameters ($\beta_1 = 2$, $\beta_2 = 1$, $\eta_1 = \eta_2 = 1$, and $\alpha = 0.25$).

Figures 4, 5 and 6 illustrate the diverse copula methodologies employed to generate data exhibiting dependence from a bivariate Weibull distribution, with parameters $\beta_1 = 2$, $\beta_2 = 1$, $\eta_1 = \eta_2 = 1$, $\alpha = 0.5$, and sample sizes $N = 50$, 100, and 200, respectively. Despite the differences in methodology, all methods yield consistent results, capturing positive dependence with fairly similar density contour patterns and data point dispersion. However, the CD-Vines methodology stands out in capturing data as the sample size increases, while keeping the distribution parameters fixed.

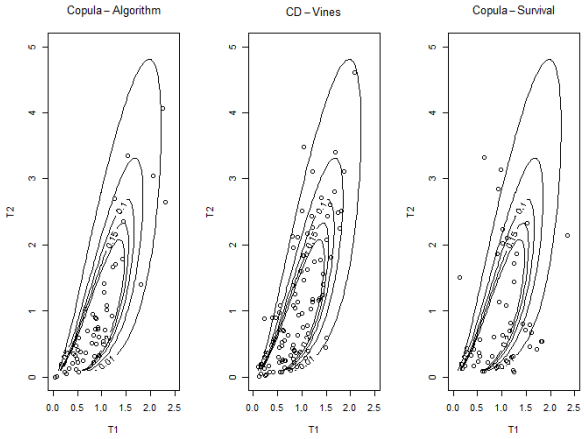

Figure 1: Weibull contours using Method 1, Method 2, and Method 3, with N=50 and parameters, $\beta_1 = 2$, $\beta_2 = 1$, $\eta_1 = \eta_2 = 1$, $\alpha = 0.25$

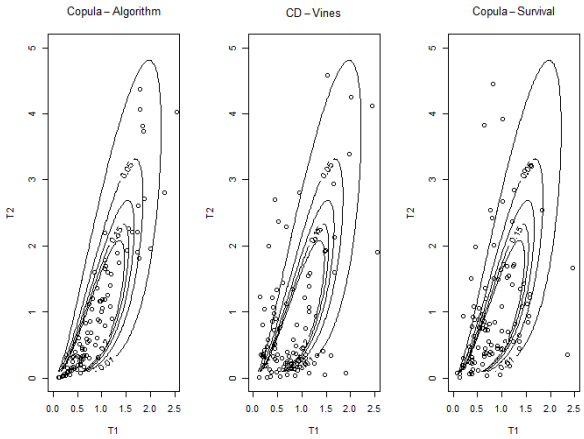

Figure 2: Weibull contours using Method 1, Method 2, and Method 3, with N=100 and parameters, $\beta_1 = 2$, $\beta_2 = 1$, $\eta_1 = \eta_2 = 1$, $\alpha = 0.25$

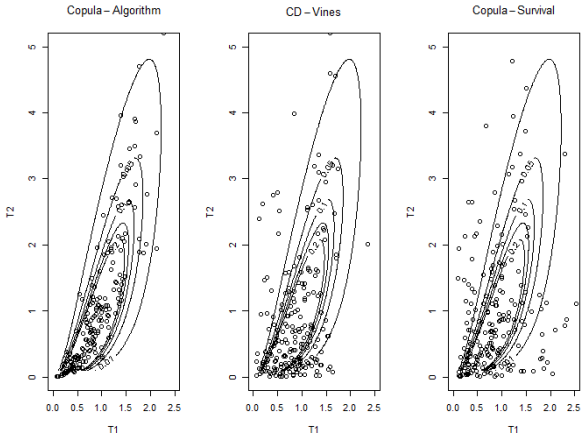

Figure 3: Weibull contours using Method 1, Method 2, and Method 3, with N=200 and parameters, $\beta_1 = 2$, $\beta_2 = 1$, $\eta_1 = \eta_2 = 1$, $\alpha = 0.25$

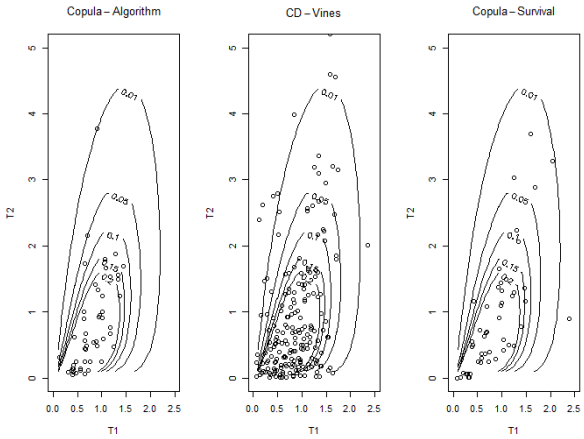

Figure 4: Weibull contours using Method 1, Method 2, and Method 3, with N=50 and parameters, $\beta_1 = 2$, $\beta_2 = 1$, $\eta_1 = \eta_2 = 1$, $\alpha = 0.50$

Figures 7, 8 and 9 illustrate the diverse copula methodologies employed to model the interdependence of a bivariate Weibull distribution, with parameters $\beta_1 = 2$, $\beta_2 = 1$, $\eta_1 = \eta_2 = 1$, $\alpha = 0.75$, and sample sizes $N = 50$, 100, and 200, respectively. A comparable pattern is observed in Figures 4, 5 and 6, where all methods demonstrate positive dependence, with relatively similar density contour patterns and data point dispersion. However, the CD-Vines methodology stands out in its ability to capture data as the sample size increases, while maintaining the other parameters fixed.

In general, it can be stated that all the methods yield consistent results, exhibiting positive dependence with comparable density contour patterns and data point dispersion. However, it is observed that with small sample sizes and low dependence parameters, the copula algorithm captures the data more effectively, while with larger sample sizes and higher dependence parameters, the Methodology CD-Vines performs better. It should also be noted that with a sample size of 200 and $\alpha = 0.75$, the survival copula also demonstrates satisfactory data capture.

5.2 Results of the Density-Based Evaluation Metrics

To analyze the performance of the three proposed simulation methods, we applied the previously introduced density-based similarity metrics: KL, HD, and MSE. These metrics allow us to quantitatively compare the simulated data to the theoretical bivariate Weibull distribution. The outcomes of this assessment, summarized in Table 2, highlight the relative accuracy of each method across different dependence levels and sample sizes.

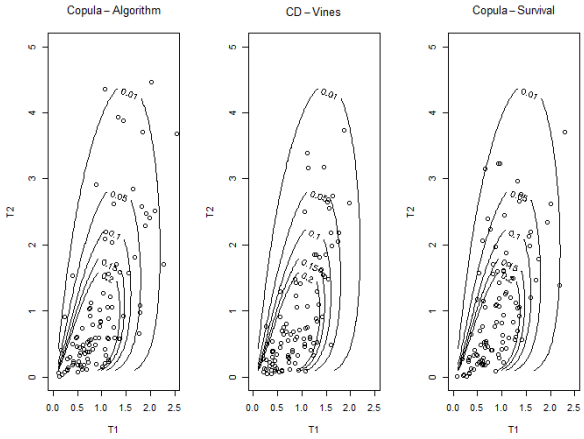

Figure 5: Weibull contours using Method 1, Method 2, and Method 3, with N=100 and parameters, $\beta_1 = 2$, $\beta_2 = 1$, $\eta_1 = \eta_2 = 1$, $\alpha = 0.50$

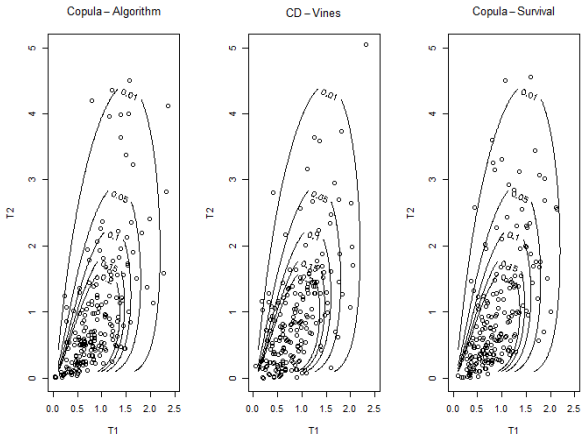

Figure 6: Weibull contours using Method 1, Method 2, and Method 3, with N=200 and parameters, $\beta_1 = 2$, $\beta_2 = 1$, $\eta_1 = \eta_2 = 1$, $\alpha = 0.50$

Table 2: Metrics of adjustment by method and scenario

Method	n	α	KL	HD	MSE
Method 1	50	0.25	0.12	0.05	0.002
Method 2	50	0.25	0.08	0.03	0.001
Method 3	50	0.25	0.15	0.06	0.003
Method 1	100	0.50	0.10	0.04	0.0015
Method 2	100	0.50	0.07	0.02	0.001
Method 3	100	0.50	0.13	0.05	0.0025
Method 1	200	0.75	0.09	0.03	0.0012
Method 2	200	0.75	0.06	0.01	0.0008
Method 3	200	0.75	0.11	0.04	0.002

Comparing the methods, we see that the methodology based on Vines (Method 2) consistently shows the lowest values across all three metrics, suggesting that it better models the structural dependence between variables, produces samples closer to the theoretical bivariate Weibull distribution, and performs more robustly across different sample sizes and dependence levels.

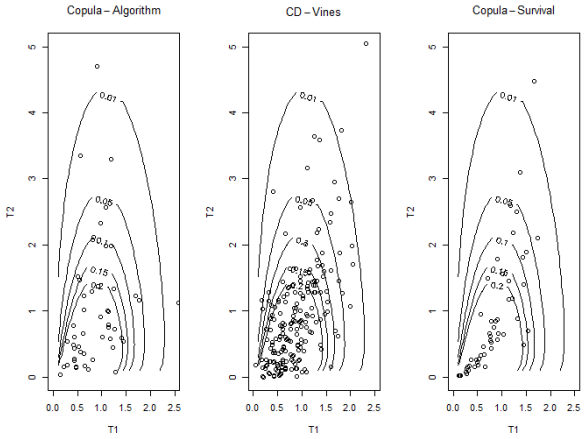

Figure 7: Weibull contours using Method 1, Method 2, and Method 3, with N=50 and parameters, $\beta_1 = 2$, $\beta_2 = 1$, $\eta_1 = \eta_2 = 1$, $\alpha = 0.75$

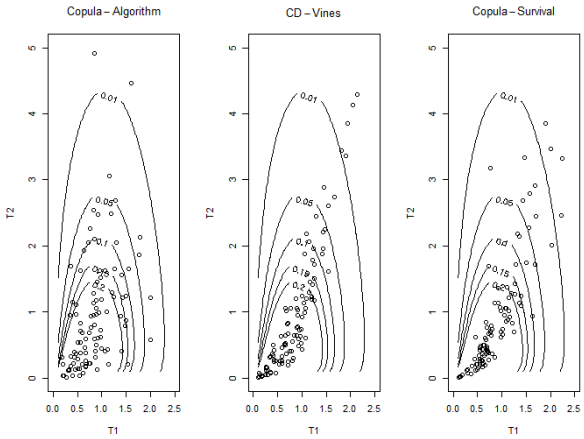

Figure 8: Weibull contours using Method 1, Method 2, and Method 3, with N=100 and parameters, $\beta_1 = 2$, $\beta_2 = 1$, $\eta_1 = \eta_2 = 1$, $\alpha = 0.75$

It is worth noting that although Method 1 (Copula Algorithm) performs competitively in some scenarios, its KL and MSE values are consistently higher than those of Method 2. This confirms that the generation of stable-distributed variables can present numerical instability for certain values of α. The same is observed in Method 3, whose general applicability appears limited. Furthermore, all methods improve as the sample size increases, which is statistically expected.

In summary, Method 2 is the most robust option among the three, while Method 1, although useful, presents practical limitations, and Method 3 shows lower reliability. Specifically, Method 1, based on positive stable mixtures, exhibited greater dispersion in some scenarios. This limitation is attributed to the numerical difficulty in generating variables from the positive stable distribution, especially for certain values of the dependence parameter α. In such cases, evaluating the precision of the numerical generator or applying rejection sampling methods or transformations could help improve the fit.

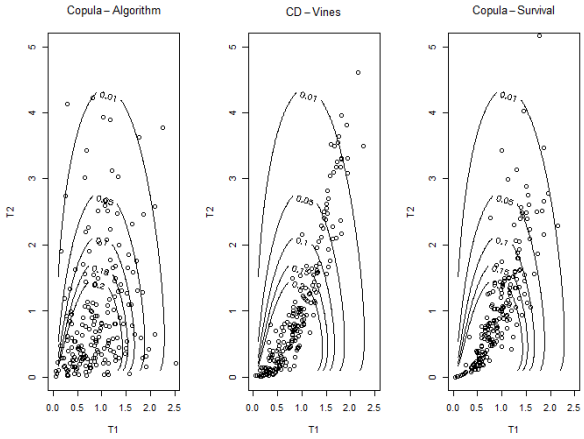

Figure 9: Weibull contours using Method 1, Method 2, and Method 3, with N=200 and parameters, $\beta_1 = 2$, $\beta_2 = 1$, $\eta_1 = \eta_2 = 1$, $\alpha = 0.75$

6 Concluding Remarks

This study shows a good acceptance of the three methods studied to generate data from a bivariate Weibull distribution. Method 1 refers to positive stable distributions, which are not well known, but have a satisfactory value from a robust point of view to generate such data, while method 2 (CD-Vines) allows the use of the different copulas, extending the range of possibilities inherited from the copulas. Last but not least, method 3, which refers to a conditional copula, is very useful if a copula is chosen a priori and we are sure that the dependency parameter of that copula is acceptable. In this work the parameters of the distribution are not considered, since the interest lies in the three methods implemented, the dependence parameter and the sample size.

For a 25% dependence ($\alpha = 0.25$) and a small sample size ($n = 50$), all three methods perform well in generating bivariate data, but for the other two sample sizes ($n = 100$ and 200) the only method that performs well is method 1 (copula algorithm). Under 50% dependence $\alpha = 0.50$) and a small sample size ($n = 50$) method 2 (CD-Vines) does not perform well, but when we change to a sample size of 100 or 200 the best is this method, contrary to the other two methods that do not perform well in this dependence scenario. Finally, for the 75% dependence scenario ($\alpha = 0.75$) and a sample size of 50, method 2 does not perform well, but it is noteworthy that for this high level of dependence under the three sample sizes it is method 1, in conclusion, indicating that for a high dependence scenario (75%), the method that best generates data from a bivariate Weibull distribution is method 1 (copula algorithm).

Method 3 performs poorly in small sample sizes ($n = 50$) only when the dependence is low ($\alpha = 0.25$). Method 2 (CD-Vines) performs well under moderate dependence ($\alpha = 0.50$), regardless of sample size. Method 1, while generally effective across scenarios, presents limitations when the data do not follow a positive stable distribution. Specifically, this Method shows reduced performance in low-dependence settings due to numerical challenges in generating stable random variables. Such issues can be identified through deviations in the density contours or increases in metrics such as KL, HD, or MSE. To address these limitations, robust generators like those proposed by [44] or [45] are recommended. Additionally, empirical copulas, regularized simulation, or transformation-based approaches [46] offer promising alternatives. Evaluating the sensitivity of the method to different values of α can also help improve its reliability under challenging conditions.

For future work, it would be beneficial to develop a goodness-of-fit analytical test for a joint distribution. This would provide an additional tool for validating the adequacy of models to observed data. Furthermore, adapting the method for discrete variables or counting processes would be important, assuming dependence using copula methodology or some other suitable alternative. This would broaden the scope of the proposed methodology and allow for its application in a variety of data modeling contexts.

Additionally, recent methodological advances open promising directions for extending the present study to high-dimensional settings. Empirical copula processes provide a nonparametric alternative for modeling dependence, and recent extensions to Stute's representation allow for asymptotic analysis in multiple dimensions [47]. Vine copulas have been used for complex tasks such as variable selection, as shown by [48], and Lasso-based selection techniques for sparse vines help reduce overfitting in high-dimensional models [49]. Moreover, the modified BIC proposed by [50] enhances model selection under these frameworks. These developments suggest that future research could build on our findings by integrating high-dimensional vine structures or empirical copula strategies.

Author Contributions

Cristian David Correa Álvarez contributed to the conceptualization, writing—review and editing, visualization, and supervision of the study. Mario Cesar Jaramillo Elorza contributed to the methodology, writing—original draft preparation, and visualization. Osnamir Elias Bru Cordero contributed to the conceptualization, formal analysis, writing—original draft preparation, and supervision. All authors have read and approved the final version of the manuscript.

Funding Statement

The authors gratefully acknowledge the financial support provided by the Colombian government through COLCIENCIAS scholarships under the National Doctoral Program, Call 727 of 2015.

Conflicts of Interest statement

The authors declare that there are no conflicts of interest to report regarding the present study. They confirm that no financial, personal, or professional relationships exist that could be perceived as influencing the design, execution, or interpretation of the research presented.

References

[1] J. I. McCool, "*Using the Weibull Distribution: Reliability, Modeling, and Inference*", vol. 950. John Wiley & Sons, 2012.

[2] N. L. Johnson, S. Kotz, y N. Balakrishnan, "*Continuous Univariate Distributions*", Volume 2, vol. 289. John Wiley & Sons, 1995.

[3] S. Mireh, A. Khodadadi, y F. Haghighi, "Copula-based reliability analysis of gamma degradation process and Weibull failure time", *International Journal of Quality & Reliability Management*, vol. 36, no. 5, pp. 654-668, 2019.

[4] T. Zhang y M. Xie, "Failure data analysis with extended Weibull distribution", *Communications in Statistics-Simulation and Computation*, vol. 36, no. 3, pp. 579-592, 2007.

[5] N. L. Clement y R. C. Lasky, "Weibull distribution and analysis: 2019", *en 2020 Pan Pacific Microelectronics Symposium (Pan Pacific)*, pp. 1-5, 2020.

[6] D. C. Montgomery, "*Introduction to Statistical Quality Control*". John Wiley & Sons, 2019.

[7] X. S. Tang, D. Q. Li, C. B. Zhou, y L. M. Zhang, "Bivariate distribution models using copulas for reliability analysis", *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*, vol. 227, no. 5, pp. 499-512, 2013.

[8] J. F. Lawless, "*Statistical Models and Methods for Lifetime Data*". Wiley, Hoboken, New Jersey, 2003.

[9] P. Hougaard, "Survival models for heterogeneous populations derived from stable distributions", *Biometrika*, vol. 73, no. 2, pp. 387-396, 1986.

[10] J. C. Lu y G. K. Bhattacharyya, "Some new constructions of bivariate Weibull models", *Annals of the Institute of Statistical Mathematics*, vol. 42, no. 3, pp. 543-559, 1990.

[11] W. Q. Meeker y L. A. Escobar, "*Statistical Methods for Reliability Data*". John Wiley & Sons, New York, 1998.

[12] M. J. Crowder, "*Classical Competing Risks*". Chapman and Hall/CRC, 2001.

[13] R. B. Nelsen, "*An Introduction to Copulas*". Springer Science & Business Media, New York, 2006.

[14] B. H. Wu, H. Michimae, y T. Emura, "Meta-analysis of individual patient data with semi-competing risks under the Weibull joint frailty-copula model", *Computational Statistics*, vol. 35, pp. 1525-1552, 2020.

[15] H. Michimae y T. Emura, "Likelihood inference for copula models based on left-truncated and competing risks data from field studies", *Mathematics*, vol. 10, no. 13, p. 2163, 2022.

[16] K. Wifvat, J. Kumerow, y A. Shemyakin, "Copula model selection for vehicle component failures based on warranty claims", *Risks*, vol. 8, no. 2, p. 56, 2020.

[17] J. Größer y O. Okhrin, "Copulae: An overview and recent developments", *Wiley Interdisciplinary Reviews: Computational Statistics*, vol. 14, no. 3, p. e1557, 2022.

[18] Y. Aliyu y U. Usman, "On Bivariate Nadarajah-Haghighi Distribution derived from Farlie-Gumbel-Morgenstern Copula in the Presence of Covariates", *Journal of the Nigerian Society of Physical Sciences*, pp. 871-871, 2023.

[19] S. O. Ahmed, K. S. Fatah, y S. Esa, "Local Dependence for Bivariate Weibull Distributions Created by Archimedean Copula", *Baghdad Science Journal*, vol. 18, no. 1 (Suppl.), pp. 0816-0816, 2021.

[20] M. C. Jones, A. Noufaily, y K. Burke, "A bivariate power generalized Weibull distribution: a flexible parametric model for survival analysis", *Statistical Methods in Medical Research*, vol. 29, no. 8, pp. 2295-2306, 2020.

[21] D. N. Prabhakar Murthy, M. Xie, y R. Jiang, "*Weibull Models*". John Wiley & Sons, 2004.

[22] E. M. Almetwally, H. Z. Muhammed, y E.-S. A. El-Sherpieny, "Bivariate Weibull distribution: properties and different methods of estimation", *Annals of Data Science*, vol. 7, pp. 163-193, 2020.

[23] E. W. Frees y E. A. Valdez, "Understanding relationships using copulas", *North American Actuarial Journal*, vol. 2, no. 1, pp. 1-25, 1998.

[24] M. C. Jaramillo, C. M. Lopera, E. C. Manotas, y S. Yáñez, "Generación de tiempos de falla dependientes Weibull bivariados usando cópulas", *Revista Colombiana de Estadística*, vol. 31, no. 2, pp. 169-181, 2008.

[25] E. Brechmann y U. Schepsmeier, "Cdvine: Modeling dependence with c-and d-vine copulas in R", *Journal of Statistical Software*, vol. 52, no. 3, pp. 1-27, 2013.

[26] S. Yáñez, L. A. Escobar, y N. González, "Characteristics of two competing risks models with Weibull distributed risks", *Revista de la Academia Colombiana de Ciencias Exactas, Físicas y Naturales*, vol. 38, no. 148, pp. 298-311, 2014.

[27] S. Shinohara, Y.-H. Lin, H. Michimae, y T. Emura, "Dynamic lifetime prediction using a Weibull-based bivariate failure time model: A meta-analysis of individual-patient data", *Communications in Statistics-Simulation and Computation*, vol. 52, no. 2, pp. 349-368, 2023.

[28] A. K. Pathak, M. Arshad, Q. J. Azhad, M. Khetan, y A. Pandey, "A novel bivariate generalized Weibull distribution with properties and applications", *American Journal of Mathematical and Management Sciences*, vol. 42, no. 4, pp. 279-306, 2023.

[29] E. S. A. El-Sherpieny, E. M. Almetwally, y H. Z. Muhammed, "Bivariate weibull-g family based on copula function: properties, bayesian and non-bayesian estimation and applications", *Statistics, Optimization & Information Computing*, vol. 10, no. 3, pp. 678-709, 2022.

[30] M. Xu, J. W. Herrmann, y E. L. Droguett, "Modeling dependent series systems with q-Weibull distribution and Clayton copula", *Applied Mathematical Modelling*, vol. 94, pp. 117-138, 2021.

[31] X. W. Huang, W. Wang, y T. Emura, "A copula-based Markov chain model for serially dependent event times with a dependent terminal event", *Japanese Journal of Statistics and Data Science*, vol. 4, pp. 917-951, 2021.

[32] N. Balakrishna y C. D. Lai, "*Construction of Bivariate Distributions*". Springer, New York, 2008.

[33] M. Franco, J. M. Vivo, y D. Kundu, "A generator of bivariate distributions: Properties, estimation, and applications", *Mathematics*, vol. 8, no. 10, p. 1776, 2020.

[34] A. W. Marshall y I. Olkin, "Families of multivariate distributions", *Journal of the American Statistical Association*, vol. 83, no. 403, pp. 834-841, 1988.

[35] M. Bekçi y M. Yilmaz, "Construction of Bivariate Distribution by Mixing Positively Dependent and Negatively Dependent Distributions", *International Journal of Statistics and Applications*, vol. 9, no. 4, pp. 122-127, 2019.

[36] H. Ü. ÜNÖZKAN, M. Yilmaz, "Construction of continuous bivariate distribution by transmuting dependent distribution", *Cumhuriyet Science Journal*, vol. 40, no. 4, pp. 860-866, 2019.

[37] L. Devroye, "*Non Uniform Random Variate Generation*". Springer Verleg, New York, 1986.

[38] C. Genest, M. Scherer, "The world of vines: An interview with Claudia Czado", *Dependence Modeling*, vol. 7, no. 1, pp. 169-180, 2019.

[39] M. E. Johnson, "*Multivariate Statistical Simulation*". John Wiley & Sons, New York, 1987.

[40] S. Kullback, R. A. Leibler, "On information and sufficiency", *The Annals of Mathematical Statistics*, vol. 22, no. 1, pp. 79-86, 1951.

[41] L. Le Cam, "*Asymptotic Methods in Statistical Decision Theory*", Springer Series in Statistics, Springer, 1986.

[42] A. Basu, I. R. Harris, N. L. Hjort, M. C. Jones, "Robust and efficient estimation by minimizing a density power divergence", *Biometrika*, vol. 85, no. 3, pp. 549-559, 1998.

[43] B. W. Silverman, "*Density Estimation for Statistics and Data Analysis*", Monographs on Statistics and Applied Probability, vol. 26, Chapman and Hall, 1986.

[44] J. M. Chambers, C. L. Mallows, B. W. Stuck, "A method for simulating stable random variables", *Journal of the American Statistical Association*, vol. 71, no. 354, pp. 340-344, 1976.

[45] R. Weron, "On the Chambers-Mallows-Stuck method for simulating skewed stable random variables", *Statistics & Probability Letters*, vol. 28, no. 2, pp. 165-171, 1996.

[46] M. Hofert, I. Kojadinovic, M. Mächler, J. Yan, "*Elements of copula modeling with R*". Berlin/Heidelberg, Germany: Springer International Publishing, 2018.

[47] A. Bücher, C. Pakzad, "The empirical copula process in high dimensions: Stute's representation and applications", *arXiv preprint*, arXiv:2405.05597, 2025.

[48] M. S. Kurz, "Vine copula based knockoff generation for high-dimensional controlled variable selection", *arXiv preprint*, arXiv:2210.11196, 2022.

[49] D. Müller, C. Czado, "Selection of sparse vine copulas in high dimensions with the Lasso", *arXiv preprint*, arXiv:1705.05877, 2017.

[50] T. Nagler, C. Bumann, C. Czado, "Model selection in sparse high-dimensional vine copula models with application to portfolio risk", *arXiv preprint*, arXiv:1801.09739, 2018.