International Journal of Engineering Inventions
e-ISSN: 2278-7461, p-ISSN: 2319-6491
Volume 6, Issue 11 [November. 2017] PP: 50-55

Bivariate Weibull Chi-squaremodel based on Gaussian copula

MervetKhalifahAbd Elaal[1,2]
[1]*King Abdulaziz University Department of Statistics, Jeddah, Saudi Arabia*
[2]*Al-Azhar University, Department of Statistics, Cairo, Egypt*

Abstract: *The Weibull distribution is widely used as a lifetime distribution in many fields such as social scienceandreliabilityengineering. The aim of this paper is to introduce anew bivariateWeibullChi-squaredistribution based on Gaussian copula that is a popular used in various applications like econometrics and finance. We explainthe goodness of fit test for copula and use both parametric and semi-parametric methodsto estimate the model parameters. Finally, Simulationissuggested to illustrate methods of inference and examine the satisfactory performance of the proposed distribution.*

Key words: *Weibull Chi-square distribution; Bivariate Weibull Chi square distribution; Maximum likelihood method; copula; Parametric and semiparametric methods*

Date of Submission: 04-12-2017 Date of acceptance: 16-12-2017

I. Introduction

Recently, there has been an increased interest in defining new generators for univariate continuous families of distributions by introducing one or more additional shape parameter(s) to the baseline distribution. For instance, Cordeiro,et al. [4],Bourguignon et al. [3] proposed a generator of distributions called the Weibull-Gclass, Nadarajahet al.[8], and among others. The class of Weibull G distributions (WG) has received an increasing amount of attention in recent years. Many studiesconducted based on the properties and inferences of Weibull G distributions with a consideration to their applications. In this paper, we introduce a bivariate Weibull Chi-square distribution in the dependence structure and illustrate its applicability.

The(WG) probability density function (PDF) has the following

$$f(t,\alpha,\beta,\delta)=\frac{\alpha}{\beta^{\alpha}}\frac{g(t,\delta)}{1-G(t,\delta)}\left\{-\frac{\log[1-G(t,\delta)]}{\beta}\right\}^{\alpha-1} exp\left\{-\left[-\frac{\log[1-G(t,\delta)]}{\beta}\right]^{\alpha}\right\},\quad t\geq 0 \qquad (1)$$

where $G(t,\delta)$ and $g(t,\delta)$, are Cdfand PDFof any baseline distribution depends on a parameter vector δ, t is in the range of g(t, δ), $\beta > 0$ is the scale parameter and $\alpha > 0$ is the shape parameter. The (WG) distributionfunction (Cdf)is given by

$$F(t,\alpha,\beta,\delta)=1-exp\left\{\left[-\frac{\log[1-G(t,\delta)]}{\beta}\right]^{\alpha}\right\},\qquad t\geq 0 \qquad (2)$$

Various Class Weibull G distributions have been discussed such as Weibull Pareto distribution by Alzaatreh, et al. [2].Copulas are a general tool to construct multivariate distributions and measure the dependence structure between random variables. The paper of Abd elaal [1] provided several methods of constructing bivariate distributions with copula functions.The main aim of this article is to introduce bivariate Weibull Chi square (BWCH) model based on the most used copula function named Gaussian copula with a suitable organization. The paper is organized as follows. Section 2presents the bivariateWeibull Chi-square (BWCH) model based on Gaussian copula function. The maximum likelihood estimates (MLEs) for the model parameters are demonstrated in Section 3. In Section 4, the flexibility of the model is explained. Finally, the performance of the suggested model using a simulation data is discussed Section 5.

II. Bivariate Weibull Chi-square distribution based on Gaussian copula

Suppose that $g(t)$ is Chi-square distribution. We have g(t;r) $=\frac{2^{-\frac{r}{2}}}{\Gamma\left(\frac{r}{2}\right)}t^{\frac{r}{2}-1}\exp\left(-\frac{t}{2}\right),\ t,\alpha,\beta,r>0$, and G(t;r) is $\frac{1-\Gamma\left(t,\frac{r}{2}\right)}{\Gamma\left(\frac{r}{2}\right)}$, where $\Gamma\left(t,\frac{r}{2}\right)$ *is incomplete gamma* andthen the Weibull- Chi-square (WCH) distribution PDF and cdf distribution are given by respectively

$$f(t,\alpha,\beta,r)=\frac{\alpha\ \frac{2^{-\frac{r}{2}}}{\Gamma(\frac{r}{2})}t^{\frac{r}{2}-1}\exp\left(-\frac{t}{2}\right)}{\beta^{\alpha}\left[1-\frac{1-\Gamma(t,\frac{r}{2})}{\Gamma(\frac{r}{2})}\right]}\left\{-\frac{log\left[1-\frac{1-\Gamma(t,\frac{r}{2})}{\Gamma(\frac{r}{2})}\right]}{\beta}\right\}^{\alpha-1}exp\left\{-\left[-\frac{log\left[1-\frac{1-\Gamma(t,\frac{r}{2})}{\Gamma(\frac{r}{2})}\right]}{\beta}\right]^{\alpha}\right\},$$

$$0<t<r<\infty,\qquad \alpha,\beta,r>0. \tag{3}$$

And

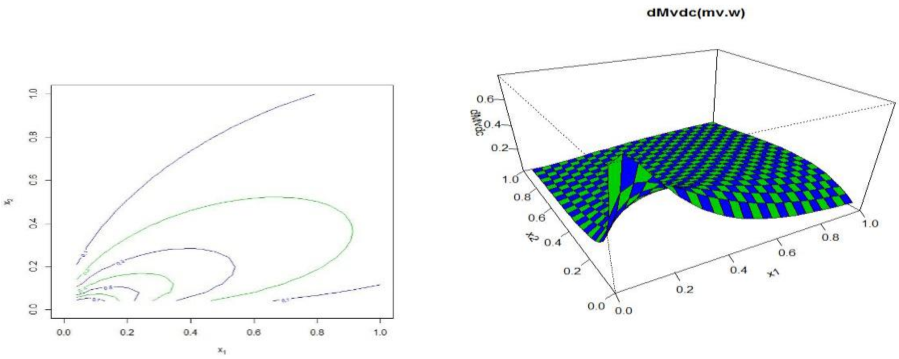
$$F(t,\alpha,\beta,r)=1-exp\left\{-\left[-\frac{log\left[1-\frac{1-\Gamma(t,\frac{r}{2})}{\Gamma(\frac{r}{2})}\right]}{\beta}\right]^{\alpha}\right\},\ \ t,\alpha,\beta,r>0,\ \ (4)$$

The density of the WCH distribution can be right-skewed. This fact implies that the WCHand BWCHdistributions can be very useful to fit different data sets with various shapes. Now, let T_1 and T_2 are following Weibull-Chi-square (WCH)distribution then the bivariate Weibull-Chi-square(BWCH) distribution which defined as the joint PDF of T_1 and T_2 based on Gaussiancopula becomes

$$f\ \left(t_1,t_2,\underline{\alpha},\underline{\beta},\underline{r}\right)$$

$$=\prod_{j=1}^{2}\frac{\alpha_j\frac{2^{-\frac{r_j}{2}}}{\Gamma(\frac{r_j}{2})}t^{\frac{r_j}{2}-1}\exp\left(-\frac{t_j}{2}\right)}{\beta_j{}^{\alpha_j}\left[1-\frac{1-\Gamma(t_j,\frac{r_j}{2})}{\Gamma(\frac{r_j}{2})}\right]}\left\{-\frac{log\left[1-\frac{1-\Gamma(t_j,\frac{r_j}{2})}{\Gamma(\frac{r_j}{2})}\right]}{\beta_j}\right\}^{\alpha_j-1}exp\left\{-\left[-\frac{log\left[1-\frac{1-\Gamma(t_j,\frac{r_j}{2})}{\Gamma(\frac{r_j}{2})}\right]}{\beta_j}\right]^{\alpha_j}\right\}$$

$$\left\{\frac{1}{\sqrt{1-\rho^2}}(\exp[\frac{-\rho}{2(1-\rho^2)}\{\rho(z_1^2+\ z_2^2)-2z_1\ z_2\ \}])\right\},t_j,\alpha_j,\beta_j,r_j>0 \tag{5}$$



(a) (b)

Fig.3 BWCH based on Gaussian copula: (a) Contour plot and (b) PDF curve

III. Parameter Estimation

In this section, we provide the estimation of the unknown parameters of BWCH distribution. There are two approaches to fitting copula models; parametric and semi-parametric methods.

3.1 Parametric methods of estimation

There are two approaches to fitting BWCH models. One approach is to estimate the marginal and copula parameters separately. The second approach is to obtain the estimation of the marginal and copula parameter from the pseudo-observations separately names modified ML.

3.1.1 Maximum likelihood estimation (ML)

The log-likelihood function expressed as

$log\,L=\sum_{i=1}^{n}[log\,f_1(t_{1i})+log\,f_2(t_{2i})+log\,c(F_1(t_{1i}),F_2(t_{2i}))]$(6)

The estimation of BWCH distribution parameters obtained by ML in two-steps. The first step is estimating the parameters of marginal distribution F_1 and F_2 by MLE separately as

$$\log L_j\ =\sum_{i=1}^{n}\log f_j\left(t_{ji}\right)\quad,\quad j=1,2. \tag{7}$$

Then, estimating copula parameters by maximizing the copula density is;

$\log L=\sum_{i=1}^{n}log\,c\left(F_1(t_{1i}),F_2(t_{2i})\right)$(8)

By considering the first step with (WCH) distribution, the parameters of each marginal distribution are estimated using MLE method. Now, if$t_1, \dots, t_n$ is a random sample from WE(α_j,β_j,r_j), then the log-likelihood function $L(\alpha_j, \beta_j, r_j)$becomes

$$\log L_j(t_j, \alpha_j, \beta_j, r_j) = nlog(\alpha_j) + \sum_{i=1}^{n} log \left[\frac{2^{-\frac{r_j}{2}}}{\Gamma(\frac{r_j}{2})} t^{\frac{r_j}{2}-1} \exp\left(-\frac{t_{ji}}{2}\right)\right] - n\alpha_j log(\beta_j) - \sum_{i=1}^{n} log \left[1 - \frac{1-\Gamma(t_{ji}, \frac{r_j}{2})}{\Gamma(\frac{r_j}{2})}\right] +$$

nαj−nlog−log 1−1−Γ *(tjirj2)*Γ(*rj2)βj*−*i=1n*− *log 1*−1−Γ *(tji,rj2)*Γ(*rj2)βjαj*.(9)

$$\frac{\partial \log L_j(t_j, \alpha_j, \beta_j, r_j)}{\partial \alpha_j} = \frac{n}{\alpha_j} - n\, log(\beta_j) + n\, log\left\{-\frac{log\left[1 - \frac{1-\Gamma(t_{ji}, \frac{r_j}{2})}{\Gamma(\frac{r_j}{2})}\right]}{\beta_j}\right\} = 0 \quad (10)$$

$$\frac{\partial \log L_j(t_{ji}, \alpha_j, \beta_j, r_j)}{\partial \beta_j} = \frac{-n\alpha_j}{\beta_j} + \frac{n(\alpha_j - 1)^2}{\beta_j} + \alpha_j \sum_{i=1}^{n} \left[-\frac{log\left[1 - \frac{1-\Gamma(t_{ji}, \frac{r_j}{2})}{\Gamma(\frac{r_j}{2})}\right]}{\beta_j}\right]^{\alpha_j} \frac{1}{\beta_j^{\alpha_j+1}} = 0\,. \ (11)$$

$$\frac{\partial \log L_j(t_j, \alpha_j, \beta_j, r_j)}{\partial r_j} = \frac{\partial \sum_{i=1}^{n} log\left[\frac{2^{-\frac{r_j}{2}}}{\Gamma(\frac{r_j}{2})} t^{\frac{r_j}{2}-1} \exp\left(-\frac{t_{ji}}{2}\right)\right]}{\partial r_j} - \frac{\partial \sum_{i=1}^{n} log\left[1 - \frac{1-\Gamma\left(t_{ji}, \frac{r_j}{2}\right)}{\Gamma\left(\frac{r_j}{2}\right)}\right]}{\partial r_j} + \frac{(n\alpha_j - n)\partial\, log\left[\frac{-log\left[1 - \frac{1-\Gamma\left(t_{ji}, \frac{r_j}{2}\right)}{\Gamma\left(\frac{r_j}{2}\right)}\right]}{\beta_j}\right]}{\partial r_j} + \frac{\partial \sum_{i=1}^{n}\left[\frac{-log\left[1 - \frac{1-\Gamma(t_{ji}, \frac{r_j}{2})}{\Gamma(\frac{r_j}{2})}\right]}{\beta_j}\right]^{\alpha_j}}{\partial r_j} = 0 \quad (12)$$

The solution of the system of nonlinear equations (10),(11) and (12)gives the MLE of α_j, β_j,and r_j.Then copula density will estimated as given,

$$\log L(\theta) = \sum_{i=1}^{n} log\, c\left(\widehat{F_1}(t_{1i}), \hat{F}_2(t_{2i})\right) (13)$$

Where $\widehat{F_1}(t_1)$ and , $\hat{F}_2(t_2)$ denote the ML estimates of the parameters from first step. The solution of the nonlinear equation (13) gives the MLE of θ.

3.1.2 Modified maximum likelihood estimation (MML)

We will propose a modified ML method to obtain the model parameters of BWCH as follows. Firstly, the parameter estimation of marginal distribution F_1 and $F_2$by MLE separately computed as $\log L_j = \sum_{i=1}^{n} \log f_j(t_{ji})$ $, j = 1,2$.

The solution of the system of nonlinear equations (10),(11) and (12)gives the MLE of α_j, β_j,and r_j.Secondly, estimate copula parameters by maximizing the copula density as $\log L(\theta) = \sum_{i=1}^{n} \log[c_\theta(\widehat{U}_i, \widehat{V}_i)]$(14)

Where $\widehat{U}_i$, $\hat{V}_i$are pseudo-observations computed from $\hat{U}_i = \frac{R_{1i}}{n+1} = \frac{n}{n+1}\hat{F}_1(t_{1i})$, $\hat{V}_i = \frac{R_{2i}}{n+1} = \frac{n}{n+1}\hat{F}_1(t_{2i})$, $R_{1i}, R_{2i}$are respectively the ranks of t_{1i}, t_{2i}.It is important to respect that the margins Cdf.s estimated parametrically from the first step.

3.2 Semi-parametric methods of estimation

This section presents the semiparametric methods to estimate the copula model parameter.

Methods-of-moments

Following Kojadinovic and Yan [7], let c be a bivariate random sample from Cdf $C_\theta[F_1(t_1), F_2(t_2)]$, where F_1 and F_2 are continuous Cdf.s and C_θ is an absolutely continuous copula such that $\theta \in \mathcal{O}$, where $\mathcal{O}$ is an open subset of R^2. Furthermore, let $R_1, \dots, R_n$ are the vectors of ranks associated with $t_1, \dots, t_n$ unless otherwise stated. In what follows, all vectors are row vectors. Method-of-moments approaches are based on the inversion of a consistent estimator of a moment of the copula C_θ. The two best-known moments, Spearman's rho and Kendall's tau, are respectively given by

$$\rho(\theta) = 12\int_{[0,1]^2} \mathrm{u\, v dC_\theta(u,v)} - 3, \quad (14)$$

and $$\tau(\theta) = 4\int_{[0,1]^2} \mathrm{C_\theta(u,v)dC_\theta(u,v)} - 1. (15)$$

Consistent estimators of these two moments can be expressed as

$$\rho_n = \frac{12}{n(n+1)(n-1)}\sum_{i=1}^{n} R_{i,1}R_{i,2} - 3\frac{n+1}{n-1}, \quad (16)$$

And $$\tau_n = \frac{4}{n(n-1)}\sum_{i=1}^{n} 1[t_{i,1} \le t_{j,1}]1[t_{i,2} \le t_{j,2}] - 1 (17)$$

When the functions ρ and τ are one-to-one, consistent estimators of θ is given by

$$\theta_{n,\rho} = \rho^{-1}(\rho_n), \theta_{n,\tau} = \tau^{-1}(\tau_n).$$

It called inversion of Kendall's (itau) and inversion of Spearman's rho (irho) respectively. For more information, see Kojadinovic and Yan [7].

IV. Goodness of fit tests for copula

The idea of this test is to compare the empirical copula with the parametric estimator derived under the null hypothesis seeDobrić and Schmid [5]. That is, test if C is well-represented by a specific copula C_θ

$$H_0: C = C_\theta Vs. \quad H_1: C \ne C_\theta$$

Two approaches are commonly used in the literature to test the goodness of fit of a copula see Genest, et al.[6]. The goodness of fit tests based on the empirical process

$$\mathbb{C}_n(u,v) = \sqrt{n}\{C_n(u,v) - C_{\theta_n}(u,v)\},$$

where$C_n(u,v)$ is the empirical copula of the data of T_1 and T_2

$$C_n(u,v) = \frac{1}{n}\sum_{i=1}^{n} 1(U_{i,n} \le u, V_{i,n} \le v), \quad u,v \in [0,1],$$

$U_{i,n}, V_{i,n}$arepseudo observations from C calculated from data as follows. $U_{i,n} = \frac{R_{1i}}{n+1}$, $V_{i,n} = \frac{R_{2i}}{n+1}$, R_{1i} , $R_{2i}$are respectively the ranks of t_{1i}, t_{2i}.Here $C_n(u,v)$ is a consistent estimator and θ_n is an estimator of θ obtained using the pseudo observations. According to Genest et al.[6], and Kojadinovic et al., [7], the test statistics is the Cramer-von Miss and is defined as$S_n = \sum_{i=1}^{n}\{C_n(U_{i,n}, V_{i,n}) - C_{\theta_n}(U_{i,n}, V_{i,n})\}^2$

V. Simulation Data

In this section, a new bivariate proposed BWCH model based on Gaussian copula is presented. The correlation measures Kendall's tau and Spearman's rho of two variables with BWCH distribution are obtained and used to provide the values of copula parameter. Considering the following values of marginal and copula parametersof BWCH distribution based on Gaussian copulawith different sizes of sample (n = 30, 50, 100, and 150), where Gaussian copula parameter$\theta_G = 0.8$.The estimations of parameters for the model by Gaussian copula and the corresponding bias, mean squared errors and relative mean squared errors based on 1000 replications are reported in Table 1, 2, and 3.
To sum up, we observe the follows.

Table 1. The estimates, the bias, the mean squared errors and the relative mean squared errors of parameters by simulation study for BWCH distribution based on Gaussian copula

Sample Size		Estimates ,bias,0.5 mean square errors and relative mean square errors of Parameter						
		$\alpha_1 = 0.7$	$b_1 = 0.8$	$r_1 = 4$	$\alpha_2 = 0.8$	$b_2 = 0.7$	$r_2 = 3$	$\theta_G = 0.8$
n=30	ML	1.300 0.600 1.904 2.719	0.899 0.099 1.159 1.449	4.249 0.249 8.333 2.083	1.158 0.358 1.297 1.622	1.052 0.352 2.060 2.943	3.204 5.532 5.532 1.844	0.511 0.289 1.134 0.385
	MML	1.300 0.600 1.904 2.719	0.899 0.099 1.159 1.449	4.249 0.249 8.333 2.083	1.158 0.358 1.297 1.622	1.052 0.352 2.060 2.943	3.204 5.532 5.532 1.844	0.763 0.037 0.067 0.083
n=50	ML	1.22 0.422 1.107	0.811 0.0107 0.150	4.300 0.300 6.659	1.004 0.204 0.725	0.943 0.243 0.254	3.200 4.176 4.176	0.547 0.253 0.064

		1.582	0.188	1.665	0.906	0.363	1.392	0.080
	MML	1.22	0.811	4.300	1.004	0.943	3.200	0.771
		0.422	0.0107	0.300	0.204	0.243	4.176	0.029
		1.107	0.150	6.659	0.725	0.254	4.176	0.001
		1.582	0.188	1.665	0.906	0.363	1.392	0.001
n=100	ML	0.964	0.747	4.274	0.866	0.870	3.194	0.669
		0.264	0.053	0.274	0.068	0.170	2.831	0.131
		0.521	0.067	4.412	0.370	0.127	2.831	0.017
		0.745	0.084	1.103	0.463	0.182	0.944	0.021
	MML	0.964	0.747	4.274	0.866	0.870	3.194	0.781
		0.264	0.053	0.274	0.068	0.170	2.831	0.019
		0.521	0.067	4.412	0.370	0.127	2.831	0.000
		0.745	0.084	1.103	0.463	0.182	0.944	0.000
n=150	ML	0.926	0.735	4.155	0.807	0.841	3.166	0.670
		0.226	0.065	0.155	0.007	0.141	2.0253	0.130
		0.447	0.622	2.912	0.884	1.108	2.025	0.229
		1.852	0.777	2.679	1.105	0.430	2.484	0.286
	MML	0.926	0.735	4.155	0.807	0.841	3.166	0.804
		0.226	0.065	0.155	0.007	0.141	2.0253	0.004
		0.447	0.622	2.912	0.884	1.108	2.025	0.008
		1.852	0.777	2.679	1.105	0.430	2.484	0.011

Table 2. The estimates, the bias, the mean squared errors and the relative mean squared errors of correlation parameter by simulation study for BWCH distribution based on Gaussian copula

Sample Size	$\theta_G = 0.8$				
	Estimates	bias	MSE	RMSE	Method Estimation
n=30	0.511	0.289	1.134	0.385	ML
	0.763	0.037	0.067	0.083	MML
	0.796	0.004	0.010	0.013	Itau
	0.797	0.003	0.022	0.007	IRho
n=50	0.547	0.253	0.064	0.080	ML
	0.771	0.029	0.001	0.001	MML
	0.755	0.045	0.002	0.003	Itau
	0.762	0.038	0.001	0.002	IRho
n=100	0.669	0.131	0.017	0.021	ML
	0.781	0.019	0.000	0.000	MML
	0.777	0.023	0.001	0.001	Itau
	0.776	0.024	0.001	0.001	IRho
n=150	0.670	0.130	0.229	0.286	ML
	0.804	0.004	0.008	0.011	MML
	0.806	0.006	0.021	0.026	Itau
	0.804	0.004	0.012	0.015	IRho

1. As expected, most results improve with increasing in sample size.

2. For most selected values of $\alpha_1, b_1, r_1, \alpha_2, b_2, r_2$ and θ_G the bias, MSE and RMSE of the estimates $\hat{\alpha}_1, \hat{b}_1, \hat{r}_1, \hat{\alpha}_2, \hat{b}_2, \hat{r}_2$ and $\hat{\theta}_G$ become smaller as the sample size increased.

3. the efficient estimators of marginal parameters of the model differ according to the parameters. It seems that ML estimates$\hat{\alpha}_1, \hat{b}_1, \hat{r}_1, \hat{\alpha}_2, \hat{b}_2, \hat{r}_2$ and of the model are the same corresponding MML estimates.

4. For copula parameter, the MML provided efficient most estimates for the model with the marginals and Gaussian, copula parameters compared to ML, Itau, and Irho.

Now, to check if the selected parmetric copula function is suitable for the marginals, goodness of fit test statistics using selected copula function for the marginals is preformed. The results in Table (3) show a non signficant p-value obtained using parmetric bootstrap for Gaussian copula function which indicate that selected parmetric copula function provide approporiate fit to the marginals. In addition, estimate of the copula parmeter based on ML, MML, Itau, and Irho methods for the Gaussian copula.This estimates are used as intial value when fitting this copula model using WCH marginals.

Table 3. Goodness of fit test statistics with their p-values and estimate of the copula parameter for selected copula functions.

model	statistic	p-value	Estimate of copula parameter θ	Method estimation
BWCH	0.0235	0.3272	0.7949	Ml
	0.0235	0.3422	0.7949	MML
	0.0270	0.2792	0.7548	Itau
	0.0261	0.3651	0.7625	Irho

Acknowledgments

The author would like to thank the Editor-in-Chief, and the referee for constructive comments, which greatly improved the paper.

References

[1]. Abd Elaal, M. K. (2017).Bivariate beta exponential distributions based on copulas. InternationalOrganization of Scientific ResearchJournal of Mathematics,13(3), 7–19.doi: 10.9790/5728-1303010719.

[2]. Alzaatreh A, Famoye F, Lee C (2013).Weibull-Pareto Distribution and Its Applications Communications in Statistics–Theory and Methods,42(9),1673–1691. doi:10.1080/ 03610926.2011.599002.

[3]. Bourguignon, M., Silva, R.B. and Cordeiro, G.M. (2014).The Weibull–G family of probability distributions, Journal of Data Science 12, 53–68.

[4]. Cordeiro, G.M. and de Castro, M. (2011).A new family of generalized distributions, Journal of Statistical Computation and Simulation 81, 883–893.

[5]. Dobrić, J., and Schmid, F. (2007).A goodness of fit test for copulas based on Rosenblatt's transformation.Computational Statistics and Data Analysis, 51(9), 4633–4642. http://doi.org/10.1016/j.csda.2006.08.012

[6]. Genest, C., Rémillard, B., & Beaudoin, D. (2009). Goodness-of-fit tests for copulas: A review and a power study. Insurance: Mathematics and Economics, 44(2), 199–213.

[7]. Kojadinovic, I., and Yan, J. (2010). Comparison of three semiparametric methods for estimating dependence parameters in copula models, Insurance: Mathematics and Economics,47(1), 52–63.

[8]. Nadarajah, S., and Rocha, R. (2016).Newdistns: An R Package for New Families of Distributions, Journal of Statistical Software, 69(10), 1-32, doi:10.18637/jss.v069.i10.

MervetKhalifahAbd Elaal "Bivariate Weibull Chi-squaremodel based on Gaussian copula." International Journal of Engineering Inventions (IJEI), vol. 6, no. 11, 2017, pp. 50-55.