A bivariate Pareto type I models

Mervat k. Abd Elaal 1,2,3 *, Hind M. Alzahrani 1

1 Department of Statistics, Faculty of Science, King Abdulaziz University, Jeddah, Saudi Arabia
2 Faculty of Commerce (Girls Campus), Al-Azhar University, Nasr city, Cairo, Egypt
3 Department of Statistics, Faculty of Science, King Abdulaziz University, P.O.Box 42805, Jeddah 21551, Saudi Arabia
*Corresponding author E-mail: mkabdeelaal@kau.edu.sa

Abstract

In this paper two new bivariate Pareto Type I distributions are introduced. The first distribution is based on copula, and the second distribution is based on mixture of and copula. Maximum likelihood and Bayesian estimations are used to estimate the parameters of the proposed distribution. A Monte Carlo Simulation study is carried out to study the behavior of the proposed distributions. A real data set is analyzed to illustrate the performance and flexibility of the proposed distributions.

Keywords: Bivariate Pareto Type I; Gaussian Copula; Maximum Likelihood Estimation; Bayesian Estimation.

1. Introduction

The Pareto distribution was first introduced by Vilfredo Pareto as a model for the distribution of income. It is used in a wide range of fields such as insurance, business, engineering, survival analysis, reliability and life testing, see for example Davis and Feldstein [1], Cox and Oakes [2], Cohen and Whitten [3] and Bhattacharya, [4]. The probability density function (Pdf) of the Pareto type I (P) distribution is given by

\[ f(t) = \alpha \beta \langle \frac{1}{t} \rangle^{\alpha+1}, t > \beta. \]  

The cumulative density function (Cdf) is given by

\[ F(T) = 1 - \left( \frac{t}{\beta} \right)^{\alpha}, t > \beta. \]  

(2)

The survivor function (SF) is given by

\[ S(T) = \left( \frac{t}{\beta} \right)^{-\alpha}, t > \beta. \]

The hazard rate function (HRF) is given by:

\[ h(T) = \frac{\alpha}{\beta}, t > \beta. \]  

(3)

The cumulative hazard rate function (CHRF) is given by:

\[ H(T) = -\alpha(\ln(\beta) - \ln(t)), t > \beta. \]  

(4)

Howlader [5] studied Bayesian prediction and estimation from Pareto distribution of the first kind. Bayesian estimators of the scale parameter of Pareto type I model have been obtained by direct method and Lindley’s approach, see Setiya, Kumar, and Pande [6], for more details see Mahmoud, Sultan, and Moshtreef [7].

A copula is a statistical method that approaches the joint distribution in terms of the marginal distributions and then links the marginal distribution functions together. A copula function captures the dependence relationships amongst the different random variables. This approach provides a general structure of modeling multivariate distributions. Sklar [8] has introduced this method in the context of probabilistic metric spaces. This approach has been formalized by Clemen and Winkler [9]. Copula have become a standard tool with many applications for examples, multi-asset pricing, credit portfolio modeling, risk management, see Longin and Solnik [10], Li [11], Patton [12], Joe [13], Lopez-Paz et al [14], and Board et al [15]. Adham and Walker [16] applied the M mixture representation of the Gompertz distribution in order to motivate a new family of distributions which extends naturally to the multivariate case using copula. In addition, they found out that the mixing idea and the use of copula method allowed full dependency structures and was easy to analyse. Many researchers have used copula to propose new bivariate and multivariate distributions. Kundu and Dey [17] studied the maximum likelihood estimators of the unknown parameters for the Marshall-Olkin bivariate Weibull distribution using EM algorithm. Diakarya [18] studied the properties of Archimedean copulas of stochastic processes and proposed analytical expressions of the survival copulas of Archimedean processes. Sankaran and Kundu [19] discussed several other new properties for bivariate Pareto model such as the maximum likelihood estimator by using two stage estimator and analyzed two data sets for the bivariate Pareto Type II distribution. Achcar et al [20] introduced Bayesian analysis for a bivariate generalized exponential distribution with censored data from Copula functions and using MCMC methods to simulate samples. Dou et al [21] used order statistics to construct multivariate distributions with fixed marginals of the Bernstein copula in terms of a finite mixture distribution.

The main aim of this article is to establish new bivariate Pareto type I distribution because of the important role of multivariate and bivariate Pareto type I in income’s analysis and ability to fit some upper tail of multivariate socio-economic and income’s data, see Mandelbroth [22] and Yeh [23]. The rest of the paper is organized as follows: in Section 2, we introduce bivariate Pareto Type I distribution based on Gaussian copula and bivariate Pareto Type I distribution based on mixture and Gaussian copula parameters.
estimation of the proposed new bivariate Pareto Type I distributions is performed using maximum likelihood and Bayesian methods in Section 3. In Section 4, Monte Carlo simulation study and analyses of real data are conducted to show the usefulness and flexibility of the proposed distributions. Finally, some concluding remarks are presented in Section 5.

2. Bivariate Pareto type I distributions

In this section Bivariate Pareto Type I (BPI) distribution based on Gaussian copula and BPI distribution based on mixture and Gaussian copula are constructed.

2.1. Construction of BPI distribution based on Gaussian copula

The simplest method to construct BPI with Gaussian copula is by using the inversion method for univariate distribution. Therefore, the joint Cdf is given by

\[ F(T_1, T_2) = C[F(T_1), F(T_2)] \]

Where \( T_1 \) and \( T_2 \) are identical independent distribution (i.i.d) from \( PI(\alpha_j, \beta_j) \).

Then, the joint Pdf of \( T_1 \) and \( T_2 \) is given by

\[ f(T_1, T_2) = C'[F(T_1), F(T_2)] f(T_1) f(T_2) \]

Therefore, the joint Pdf of \( T_1 \) and \( T_2 \) can be rewritten as

\[ f(T_1, T_2) = \left( \frac{\alpha_1 \beta_1}{\alpha_1 + \beta_1} \right) \left( \frac{\alpha_2 \beta_2}{\alpha_2 + \beta_2} \right) C_G \] (6)

For more explanation, see Joe [24], and Flores [25]. Observing that \( \rho \) is a parameter associated to the dependence between the random variable \( T_1 \) and \( T_2 \) which related to Kendall’s rank correlation and the Spearman’s rank correlation given by (7) and (8) respectively.

\[ \rho = 12 \int_{-1}^{1} \int_{-1}^{1} uv dC(u,v) - 3 \] (7)

\[ \tau = 1 - 4 \int_{-1}^{1} \int_{-1}^{1} \frac{\partial C}{\partial u} (u,v) \frac{\partial C}{\partial v} (u,v) dudv. \] (8)

Graphical representation of the Pdf, Cdf, and contours plots of the BPI distributions based on Gaussian copula for two different values of the copula parameter \( \rho \) are shown in Figure (1).

Fig.1: Pdf, Cdf and Contours of BPI Distribution Based on Gaussian Copula For \( \alpha_1 = 1.5 , \alpha_2 = 2, \beta_1 = .01, \beta_2 = .03, \rho = c(0.5, .70, 0.85) \)
2.2. Construction of BPI distribution based on mixture and Gaussian copula

Let $M$ denote the PDF for a random variable $T$ on $(0, \infty)$ which has a mixture representation. If $(T_1, T_2)$ is a two-dimensional random vector conditionally independent given $(U_1, U_2)$, where $U_1$ and $U_2$ have bivariate gamma distribution, then the joint PDF can be written in the form of compound distribution given by

$$f(T_1, T_2) = \int_{H(T_1)}^{\infty} \int_{H(T_2)}^{\infty} \prod_{j=1}^{2} f(T_j|U_j) f(T_j) f(U_j) \times C_G dU_1 dU_2 \left( \right) \tag{9}$$

For $j=1, 2$,

$$f(T_j|U_j) = \frac{\alpha_j}{U_j} \left[ \left( \ln(\beta_j) - \ln(T_j) \right) \right] \tag{10}$$

Therefore, the function can be rewritten as:

$$f(T_j|U_j) = U_j e^{-\alpha_j / U_j} \geq 0, j=1, 2. \tag{11}$$

For more details, see Walker and Stephens [26], Adham and Silva and Lopes [27]. The joint PDF of BPI distribution based on M mixture representation with Gaussian copula given by (9), can be rewritten as

$$f(T_1, T_2) = \int_{H(T_1)}^{\infty} \int_{H(T_2)}^{\infty} \prod_{j=1}^{2} \left[ \frac{\alpha_j}{U_j} e^{-\alpha_j / U_j} \right] C_G dU_1 dU_2. \tag{13}$$

$$j = 1, 2.$$

3. Estimation

The estimation of the parameters for BPI distribution based on Gaussian copula and BPI distribution based on mixture and Gaussian copula using maximum likelihood (ML) and Bayesian methods will be preformed.

3.1. Parameters estimation for BPI distribution based on Gaussian copula

3.1.1. Maximum likelihood estimation

$$l(\theta|T_1, T_2) =$$

$$n \ln(\alpha_1) + n \alpha_1 \ln(\beta_1) - (\alpha_1 + 1) \sum_{i=1}^{n} [\ln(t_{1i})] + n \ln(\alpha_2) + n \alpha_2 \ln(\beta_2) - (\alpha_2 + 1) \sum_{i=1}^{n} [\ln(t_{2i})] + \sum_{i=1}^{n} [\ln(C_{Gauss}(v_{1i}, v_{2i})]. \tag{14}$$

If $(T_1, T_2) = ((t_{11}, t_{21}), \ldots, (t_{1n}, t_{2n})$) are i.i.d sample of size $n$ from BPI distribution given in (6), then the log-likelihood function can be written as where $C_G$ given in (5), and $\beta_j, j = 1, 2$ is fixed, we don’t need to differential of estimate it, and $\theta = (\beta_1, \alpha_1, \beta_2, \alpha_2, \rho)$. The ML estimate of the unknown parameters can be obtained by maximizing (14) with respect to the unknown parameters $\beta_1, \alpha_1, \beta_2, \alpha_2$ and $\rho$, such that $\beta_1, \alpha_1, \beta_2, \alpha_2 > 0$, and $-1 < \rho < 1$. That is, differentiating (14) with respect to $\beta_1, \alpha_1, \beta_2, \alpha_2$, and $\rho$ and equating it to zero, the first partial derivatives are given by:

$$\frac{\partial l}{\partial \beta_1} = 0 \Rightarrow \hat{\alpha}_1 = \frac{n \sum_{i=1}^{n} \ln(\frac{\beta_1}{t_{1i}})}{\sum_{i=1}^{n} \ln(\beta_1)} \tag{15}$$

$$\frac{\partial l}{\partial \beta_2} = 0 \Rightarrow \hat{\alpha}_2 = \frac{n \sum_{i=1}^{n} \ln(\frac{\beta_2}{t_{2i}})}{\sum_{i=1}^{n} \ln(\beta_2)}$$

The ML estimate of the unknown parameters can be obtained by solving the system of nonlinear equations in (15) numerically. Sampling information matrix and approximate confidence interval.

Approximate confidence interval of the parameters $\theta$ can be obtained based on the asymptotic distribution of the ML estimates of $\theta$ when $\theta > 0$. Using the large sample and under appropriate regularity conditions, the ML estimates for the parameters $\hat{\theta}$ have approximately multivariate normal distribution with mean $\theta$ and asymptotic variance-covariance matrix $I^{-1}(\theta)$. See (Algorithm 1 in Appendix).

Then the $100(1-\gamma)\%$ approximate confidence interval for the parameters $\beta_j, \alpha_1, \beta_2, \alpha_2$, and $\rho$ are:

$$\hat{\beta}_j = \min\left(t_{1i}, \frac{\hat{\alpha}_j}{\sqrt{\hat{\rho}}} \right) \pm \frac{z_{\gamma/2}}{\sqrt{\text{var}(\hat{\beta}_j)}}, j = 1, 2 \text{ and } \sqrt{\hat{\rho}} \pm \frac{z_{\gamma/2}}{\sqrt{\text{var}(\hat{\rho})}}$$.

Where, $z_{\gamma/2}$ is the upper $(\gamma/2)$ the percentile of the standard normal distribution.

3.1.2. Bayesian estimation

Let $(T_1, T_2)$ be a bivariate random samples from BPI distribution given by (6) and assuming non informative independent priors for the parameters such that.

$$\pi(\beta_j) \propto \frac{1}{\beta_j} \pi(\alpha_j) \propto \frac{1}{\alpha_j}, j = 1, 2, (\rho) = \frac{1}{2} \tag{16}$$

Therefore, the joint posterior distribution can be written as

$$\pi(\alpha_1, \beta_1, \alpha_2, \beta_2, \rho, T_1, T_2) = \pi(\alpha_1) \pi(\beta_1) \pi(\alpha_2) \pi(\beta_2) \pi(\rho) \pi(T_1, T_2)$$

That is, simulate samples from the conditional distributions

$$\pi(\alpha_1|\beta_1, \beta_2, \alpha_2, \rho, T_1, T_2),$$

$$\pi(\alpha_2|\beta_1, \alpha_1, \beta_2, \rho, T_1, T_2),$$

$$\pi(\rho|\beta_1, \alpha_1, \beta_2, \alpha_2, T_1, T_2)$$

By using Metropolis-Hastings algorithm, since the conditional distributions in this case are not identified as known distributions, see Achcar at al [20].

3.2. Parameters estimation for BPI distribution based on mixture and Gaussian copula

3.2.1. Maximum likelihood estimation

Suppose that, $(T_1, T_2) = ((t_{11}, t_{21}), \ldots, (t_{1n}, t_{2n}))$ is a random samples from BPI distribution given in (13), and $(U_1, U_2) =$
\((u_{11}, u_{21}), ..., (u_{n_1}, u_{n_2})\) is a random samples from bivariate gamma distribution. The likelihood function is given by:

\[
L(\theta | T_1, T_2, U_1, U_2) = \prod_{i=1}^{n_1} f(T_1, T_2) = \prod_{i=1}^{n_1} \alpha_{ij} e^{-u_{ij}C_{\text{Gauss}}(u_{v_1}, v_2)}
\]

\[
l \left( u_{ji} > -\alpha_j (ln(\beta_j) - ln(\theta_{ji})) \right)
\]

Where \(C_{\text{Gauss}}(u_{v_1}, v_2)\) is given by (5), and \(v_j = F(U_j), j = 1, 2,\) and \(\theta = (\beta_1, \alpha_1, \beta_2, \alpha_2, \rho)\).

The likelihood function can be rewritten as

\[
L(\theta | T_1, T_2, U_1, U_2) = \prod_{i=1}^{n_1} \alpha_{ij} e^{-\sum_{i=1}^{n_1} \sum_{j=1}^{n} C_{\text{Gauss}}(u_{v_1}, v_2)}
\]

(18)

\[
l \left( \alpha_j < \min \left( \frac{u_{ji}}{-\ln(\theta_{ji})} \right) \right)
\]

The log-likelihood function can be written as

\[
\pi(\alpha_j, \beta_j, \rho | T_1, T_2) \propto \prod_{i=1}^{n_1} f(t_{1i}, t_{2i}; \beta_1, \alpha_1, \beta_2, \alpha_2, \rho) \prod_j^{\alpha_j} \pi(\alpha_j) \pi(\rho), j = 1, 2.
\]

(17)

\[
l(\theta | T_1, T_2, U_1, U_2) = \]

\[
n \ln \alpha_1 - \sum_{i=1}^{n_1} u_{1i} - \sum_{i=1}^{n_1} \ln t_{1i}
\]

\[
+ n \ln \alpha_2 - \sum_{i=1}^{n_1} u_{2i} - \sum_{i=1}^{n_1} \ln t_{2i}
\]

\[
+ \sum_{i=1}^{n_1} \ln C_{\text{Gauss}}(v_{1}, v_2)
\]

\[
l \left( \alpha_j < \min \left( \frac{u_{ji}}{-\ln(\theta_{ji})} \right) \right), j = 1, 2
\]

(19)

The ML estimates of the unknown parameters can be obtained by maximizing (19) with respect to the unknown parameters \(\beta_1, \alpha_1, \beta_2, \alpha_2, \) and \(\rho\).

That is, ML estimates can be obtained by solving numerical the five dimensional optimization problems. The first derivative are given by

\[
\frac{\partial l}{\partial \alpha_j} = 0 \Rightarrow \hat{\alpha}_j = \frac{n}{\sum_{i=1}^{n} \ln(\theta_{ji})}, j = 1, 2,
\]

\[
\frac{\partial l}{\partial \rho} = 0 \Rightarrow \hat{\rho} = \frac{\sum_{j=1}^{\alpha_j} \sum_{i=1}^{n_1} \frac{u_{ji}^{\frac{\alpha_j}{\alpha_j - 1}}}{\alpha_j - 1}}{n - j = -j}
\]

Sampling information matrix and approximate confidence interval is obtained by Algorithm 1 in Appendix.

### 3.2.2 Bayesian estimation

If we have a bivariate random sample \((T_1, T_2) = ((t_{11}, t_{21}), ..., (t_{1n}, t_{2n}))\), \(n=1, 2, ..., n\), from BPI distribution, then the corresponding latent variables \((U_1, U_2) = ((u_{11}, u_{21}), ..., (u_{1n}, u_{2n}))\), \(n=1, 2, ..., n\), is generated from gamma(2,1), where \(U_j \sim \text{gamma}(2,1), j = 1, 2\). The Gibbs sampler procedure is used to obtain Bayesian estimates of the parameters \((\beta_1, \alpha_1, \beta_2, \alpha_2, \rho)\) of the BPI distribution based on mixture and Gaussian copula. Assuming non-informative prior distribution of the parameters as in (17). Therefore, the joint posterior distribution can be written as

\[
\pi(\theta_i | T_1, T_2, U_1, U_2) \propto \prod_{i=1}^{j-1} \pi(\beta_j) \pi(\alpha_j) \pi(\rho) L(\beta_j, \alpha_j, \rho | T_1, T_2, U_1, U_2)
\]

\(j = 1, 2, \ldots, n\)

Where \(L(\theta_i | T_1, T_2, U_1, U_2)\) is given by (18).

Now, the final conditional distributions of the Gibbs sampler can be obtained by the following:

1) Sample \(U_j\) from \(\pi(U_j | \beta_j, \alpha_j, \rho, T_1, T_2)\) given in (18).

2) 

Sample \(\alpha_i\) from the mixture of the joint distribution from (17) for \(j = 1, 2\).

\[
\pi(\alpha_i | T_1, T_2, U_1, U_2) \propto \alpha_i^{n_i - 1} l \left( \alpha_i < \min \left( \frac{u_{ji}}{-\ln(\theta_{ji})} \right) \right), j = 1, 2, i = 1, 2 \ldots, n.
\]

If \(d_j = \min \left( \frac{u_{ji}}{-\ln(\theta_{ji})} \right)\), then the final conditional distribution of \(\alpha_i\) is given by

\[
\pi(\alpha_i | T_1, T_2, U_1, U_2) \propto \frac{\alpha_i^{n_i - 1}}{d_j} l(\alpha_i < d_j), j = 1, 2.
\]

The cumulative distribution is given by

\[
F(\alpha_i | T_1, T_2, U_1, U_2) = \int_{0}^{\alpha} \frac{\alpha^{n_i - 1}}{d_j} d\alpha_i
\]

\[
= \int_{0}^{\alpha} \frac{\alpha^{n_i - 1}}{d_j} d\alpha_i = \frac{\alpha^{n_i}}{d_j}.
\]

Let \(\delta = F(\alpha_i | T_1, T_2, U_1, U_2) \Rightarrow \delta = \frac{\alpha_i^{n_i}}{d_j}.
\]

By using inverse method to sample \(\alpha_i\)

\[
\alpha_i = d_j \delta^{\frac{1}{n_i}}
\]

(22)

Where \(\delta\) is Uniform \((0, 1)\).

3) Finally, sample \(\rho\) from

\[
\pi(\rho | \beta_j, \alpha_j, T_1, T_2, U_1, U_2) \propto \left(1 - \rho^2\right)^{\frac{n}{2}} e^{-\frac{1}{2} \sum_{i=1}^{n} \frac{u_{ji}^{\frac{\alpha_j}{\alpha_j - 1}}}{\alpha_j - 1}} \frac{\nu_1^{\frac{\nu_1}{2}}}{\Gamma\left(\frac{\nu_1}{2}\right)} \left(\frac{\nu_2}{\nu_1 + \nu_2} + \frac{u_{ji}^{\frac{\alpha_j}{\alpha_j - 1}}}{\alpha_j - 1}\right)^{-\frac{\nu_2}{2}}
\]

\(i = 1, 2 \ldots n\).

Metropolis hastings is used to obtain estimate of \(\rho\). See Abd Elaal et al [28].
4. Simulation study

<table>
<thead>
<tr>
<th>Sample size</th>
<th>Parameters</th>
<th>MLE</th>
<th>Bayesian estimation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( \alpha_1 )</td>
<td>Mean</td>
<td>RMSE</td>
</tr>
<tr>
<td>n=35</td>
<td>( \hat{\alpha}_1 )</td>
<td>1.2505</td>
<td>0.2534</td>
</tr>
<tr>
<td></td>
<td>( \hat{\alpha}_2 )</td>
<td>1.1535</td>
<td>0.2593</td>
</tr>
<tr>
<td></td>
<td>( \hat{\beta} )</td>
<td>0.8076</td>
<td>0.0350</td>
</tr>
<tr>
<td></td>
<td>( \hat{\rho} )</td>
<td>1.2408</td>
<td>0.0776</td>
</tr>
<tr>
<td>n=50</td>
<td>( \hat{\alpha}_1 )</td>
<td>1.1315</td>
<td>0.0696</td>
</tr>
<tr>
<td></td>
<td>( \hat{\alpha}_2 )</td>
<td>0.8009</td>
<td>0.0557</td>
</tr>
<tr>
<td></td>
<td>( \hat{\beta} )</td>
<td>1.2242</td>
<td>0.0430</td>
</tr>
<tr>
<td></td>
<td>( \hat{\rho} )</td>
<td>1.2042</td>
<td>0.0154</td>
</tr>
<tr>
<td>n=100</td>
<td>( \hat{\alpha}_1 )</td>
<td>1.1254</td>
<td>0.0453</td>
</tr>
<tr>
<td></td>
<td>( \hat{\alpha}_2 )</td>
<td>0.8012</td>
<td>0.0040</td>
</tr>
<tr>
<td></td>
<td>( \hat{\beta} )</td>
<td>1.1028</td>
<td>0.0132</td>
</tr>
<tr>
<td></td>
<td>( \hat{\rho} )</td>
<td>0.8018</td>
<td>0.0039</td>
</tr>
<tr>
<td>n=150</td>
<td>( \hat{\alpha}_1 )</td>
<td>1.1254</td>
<td>0.0453</td>
</tr>
<tr>
<td></td>
<td>( \hat{\alpha}_2 )</td>
<td>0.8012</td>
<td>0.0040</td>
</tr>
<tr>
<td></td>
<td>( \hat{\beta} )</td>
<td>1.1028</td>
<td>0.0132</td>
</tr>
<tr>
<td></td>
<td>( \hat{\rho} )</td>
<td>0.8018</td>
<td>0.0039</td>
</tr>
</tbody>
</table>

4.1. Simulation study of BPI distribution based on Gaussian copula

A Monte Carlo simulation study performed to investigate and compare the ML and Bayesian estimates of parameters \( \alpha_1, \alpha_2, \rho \), while \( \beta_1 \) and \( \beta_2 \) are fixed. Different sample sizes, \( n=35, 50, 100, 150 \), were considered using different values of the parameters, with \( \beta_1 \) and \( \beta_2 \) set to the minimum fixed values and the copula parameter taking the values \( \rho = (0.70, 0.80, 0.85) \). The BPI distribution is fitted to the data and the ML and Bayesian estimate of the parameters of BPI distribution based on Gaussian copula are obtained. Then, the average estimates along with their relative mean square error (RMSE) over 1000 replication are calculated. The results are reported in Tables 1 and 2.

<table>
<thead>
<tr>
<th>Sample size</th>
<th>Parameters</th>
<th>MLE</th>
<th>Bayesian estimation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( \hat{\alpha}_1 )</td>
<td>Mean</td>
<td>RMSE</td>
</tr>
<tr>
<td>n=35</td>
<td>( \hat{\alpha}_1 )</td>
<td>1.2708</td>
<td>0.2987</td>
</tr>
<tr>
<td></td>
<td>( \hat{\alpha}_2 )</td>
<td>1.2277</td>
<td>0.2792</td>
</tr>
<tr>
<td></td>
<td>( \hat{\beta} )</td>
<td>0.7162</td>
<td>0.0669</td>
</tr>
<tr>
<td></td>
<td>( \hat{\rho} )</td>
<td>1.2669</td>
<td>0.1058</td>
</tr>
<tr>
<td>n=50</td>
<td>( \hat{\alpha}_1 )</td>
<td>1.1525</td>
<td>0.0937</td>
</tr>
<tr>
<td></td>
<td>( \hat{\alpha}_2 )</td>
<td>0.7056</td>
<td>0.0229</td>
</tr>
<tr>
<td></td>
<td>( \hat{\beta} )</td>
<td>1.2315</td>
<td>0.0478</td>
</tr>
<tr>
<td></td>
<td>( \hat{\rho} )</td>
<td>1.1290</td>
<td>0.0475</td>
</tr>
<tr>
<td>n=100</td>
<td>( \hat{\alpha}_1 )</td>
<td>1.2164</td>
<td>0.0238</td>
</tr>
<tr>
<td></td>
<td>( \hat{\alpha}_2 )</td>
<td>1.1173</td>
<td>0.0259</td>
</tr>
<tr>
<td></td>
<td>( \hat{\beta} )</td>
<td>0.6998</td>
<td>0.0032</td>
</tr>
<tr>
<td>n=150</td>
<td>( \hat{\alpha}_1 )</td>
<td>1.2708</td>
<td>0.2987</td>
</tr>
<tr>
<td></td>
<td>( \hat{\alpha}_2 )</td>
<td>1.2277</td>
<td>0.2792</td>
</tr>
<tr>
<td></td>
<td>( \hat{\beta} )</td>
<td>0.7162</td>
<td>0.0669</td>
</tr>
<tr>
<td></td>
<td>( \hat{\rho} )</td>
<td>1.2669</td>
<td>0.1058</td>
</tr>
</tbody>
</table>

It can be seen from Tables 1 and 2 that, for all selected values of \( \alpha_1, \alpha_2, \rho \), the RMSE of the estimates \( \hat{\alpha}_1, \hat{\alpha}_2, \hat{\beta}, \hat{\rho} \) become smaller as the sample size increases. In addition, it can be seen that we have better estimates and smaller RMSE when the copula parameter \( \rho = 0.80 \). Moreover, the Bayesian method gave better and more accurate estimates for the parameters than the ML method especially with small samples size.

4.2. Simulation study of BPI distribution based on mixture and Gaussian copula

A Monte Carlo simulation study is performed to investigate and compare ML and Bayesian estimates of the parameters of BPI distribution based on Mixture and Gaussian copula. The comparison and performances of the estimates are studied mainly with respect to their RMSE. These are illustrated in Tables 3 and 4 using different sample sizes \( n=10, 25, 50, 100 \) and the different values of the parameters, with \( \beta_1 \) and \( \beta_2 \) set to minimum fixed values and copula parameter \( \rho = (0.70, 0.80) \). For each sample of generated data, the BPI distribution is fitted and the ML and Bayesian estimate of the parameters of BPI distribution based on mixture and Gaussian copula are obtained. Then, the average estimates along with their relative mean square error (RMSE) over 1000 replication are calculated.

<table>
<thead>
<tr>
<th>Sample size</th>
<th>Parameters</th>
<th>MLE</th>
<th>Bayesian estimation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( \hat{\alpha}_1 )</td>
<td>Mean</td>
<td>RMSE</td>
</tr>
<tr>
<td>n=10</td>
<td>( \hat{\alpha}_1 )</td>
<td>1.1557</td>
<td>0.0913</td>
</tr>
<tr>
<td></td>
<td>( \hat{\alpha}_2 )</td>
<td>0.7003</td>
<td>0.0287</td>
</tr>
<tr>
<td></td>
<td>( \hat{\beta} )</td>
<td>1.2431</td>
<td>0.0843</td>
</tr>
<tr>
<td>n=25</td>
<td>( \hat{\alpha}_1 )</td>
<td>1.1413</td>
<td>0.0832</td>
</tr>
<tr>
<td></td>
<td>( \hat{\alpha}_2 )</td>
<td>0.7001</td>
<td>0.0109</td>
</tr>
<tr>
<td></td>
<td>( \hat{\beta} )</td>
<td>1.2296</td>
<td>0.0453</td>
</tr>
<tr>
<td>n=50</td>
<td>( \hat{\alpha}_1 )</td>
<td>1.1296</td>
<td>0.0477</td>
</tr>
<tr>
<td></td>
<td>( \hat{\alpha}_2 )</td>
<td>0.7004</td>
<td>0.0059</td>
</tr>
<tr>
<td></td>
<td>( \hat{\beta} )</td>
<td>1.2179</td>
<td>0.0251</td>
</tr>
<tr>
<td>n=100</td>
<td>( \hat{\alpha}_1 )</td>
<td>1.1176</td>
<td>0.0260</td>
</tr>
<tr>
<td></td>
<td>( \hat{\alpha}_2 )</td>
<td>0.9680</td>
<td>0.0066</td>
</tr>
</tbody>
</table>

The results in Tables 3 and 4 indicate that for all selected values of \( \alpha_1, \alpha_2, \rho \), the RMSE of the estimates \( \hat{\alpha}_1, \hat{\alpha}_2, \hat{\beta}, \hat{\rho} \) become smaller as the sample size increases. The copula parameters \( \alpha_1 = 1.2, \alpha_2 = 1.1, \rho = 0.80 \) provides better estimate of the parameters copula to \( \rho = 0.7 \). Also, the Bayesian method provides better and more accurate
estimates for the parameters compared to the ML method especially with small samples sizes.

4.3. Models comparison

The performance of the two proposed BPI distributional models are compared based on RMSE. In addition, Akaike’s Information Criterion (AIC) and Bayesian Information Criterion (BIC) are calculated. The results are reported in Table 5 indicate that the BPI distribution based on mixture and Gaussian copula have lower RMSE, AIC, and BIC values compared to BPI distribution based on Gaussian copula. Therefore, we conclude that BPI distribution based on mixture and Gaussian copula is more flexible compared to BPI based on Gaussian copula.

4.5. Data analysis

This data set represents the two different measurements of stiffness, ‘Shock’ and ‘Vibration’ of each of 30 boards. Here T1 represents the first measurement (Shock) involving a shock wave down the board and T2 represents the second measurement (Vibration) is specified while vibrating the board. The data set was originally from William Galligan, and it has been reported in Johnson et al [28], and illustrated in Table 6. The PI distribution is fitted to the marginals.

Table 6: Summary for the Estimation and the Test for Comparisons Two Models

<table>
<thead>
<tr>
<th>Model</th>
<th>Method</th>
<th>$d_1$</th>
<th>$d_2$</th>
<th>$\beta$</th>
<th>AIC</th>
<th>BIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>BPI based on Gaussian copula</td>
<td>MLE</td>
<td>2.817</td>
<td>2.281</td>
<td>0.975</td>
<td>162.9042</td>
<td>167.1078</td>
</tr>
<tr>
<td>BPI based on Gaussian copula and Mixture</td>
<td>Bayesian</td>
<td>2.767</td>
<td>2.272</td>
<td>0.970</td>
<td>162.9042</td>
<td>167.1078</td>
</tr>
<tr>
<td>BPI based on Gaussian copula and Mixture</td>
<td>MLE</td>
<td>2.706</td>
<td>2.226</td>
<td>0.898</td>
<td>136.5605</td>
<td>140.7641</td>
</tr>
</tbody>
</table>

Table 7: Two Different Stiffness Measurements of 30 Boards

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1889</td>
<td>1651</td>
<td>2</td>
<td>2403</td>
<td>2048</td>
<td>3</td>
<td>2119</td>
<td>1700</td>
<td>37</td>
<td>1367</td>
<td>1195</td>
</tr>
<tr>
<td>4</td>
<td>1645</td>
<td>1627</td>
<td>5</td>
<td>1976</td>
<td>1916</td>
<td>6</td>
<td>1712</td>
<td>1713</td>
<td>94</td>
<td>1312</td>
<td>1168</td>
</tr>
<tr>
<td>7</td>
<td>1943</td>
<td>1685</td>
<td>8</td>
<td>104</td>
<td>1820</td>
<td>9</td>
<td>2983</td>
<td>2794</td>
<td>29</td>
<td>2132</td>
<td>1864</td>
</tr>
<tr>
<td>10</td>
<td>1745</td>
<td>1600</td>
<td>11</td>
<td>1710</td>
<td>1591</td>
<td>12</td>
<td>2046</td>
<td>1907</td>
<td>44</td>
<td>1831</td>
<td>1573</td>
</tr>
<tr>
<td>13</td>
<td>1840</td>
<td>1841</td>
<td>14</td>
<td>1867</td>
<td>1685</td>
<td>15</td>
<td>1859</td>
<td>1649</td>
<td>66</td>
<td>1723</td>
<td>1492</td>
</tr>
<tr>
<td>16</td>
<td>1954</td>
<td>2149</td>
<td>17</td>
<td>1325</td>
<td>1170</td>
<td>18</td>
<td>1419</td>
<td>1371</td>
<td>88</td>
<td>1573</td>
<td>1296</td>
</tr>
<tr>
<td>19</td>
<td>1828</td>
<td>1634</td>
<td>20</td>
<td>1725</td>
<td>1594</td>
<td>21</td>
<td>2276</td>
<td>2189</td>
<td>101</td>
<td>1753</td>
<td>1464</td>
</tr>
<tr>
<td>22</td>
<td>1899</td>
<td>1614</td>
<td>23</td>
<td>1633</td>
<td>1513</td>
<td>24</td>
<td>2061</td>
<td>1867</td>
<td>124</td>
<td>1652</td>
<td>1353</td>
</tr>
<tr>
<td>25</td>
<td>1856</td>
<td>1493</td>
<td>26</td>
<td>1777</td>
<td>1412</td>
<td>27</td>
<td>2168</td>
<td>1896</td>
<td>148</td>
<td>1573</td>
<td>1326</td>
</tr>
<tr>
<td>28</td>
<td>1655</td>
<td>1675</td>
<td>29</td>
<td>2326</td>
<td>2301</td>
<td>30</td>
<td>1490</td>
<td>1382</td>
<td>172</td>
<td>1753</td>
<td>1464</td>
</tr>
</tbody>
</table>

Table 7 shows the Kolmogorov-Smirnov test along with associated p-values for the two marginals.

Table 8: The K-S Test for the Data

<table>
<thead>
<tr>
<th>Sample</th>
<th>p-value</th>
<th>K-S</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T_1$(Shock)</td>
<td>0.06705</td>
<td>0.2318</td>
</tr>
<tr>
<td>$T_2$(Vibration)</td>
<td>0.05274</td>
<td>0.2462</td>
</tr>
</tbody>
</table>

The BPI distribution based on the Gaussian copula and the BPI distribution based on mixture and Gaussian copula are fitted and the results are shown in Table 8. The AIC and BIC values in Table 8 indicate that the BPI distribution based on mixture and Gaussian copula provides better fit for the data compared to BPI model based on Gaussian copula.

5. Summary remarks

In this article, we proposed two new bivariate distributions the first one is BPI distribution based on Gaussian copula and the second one is BPI distribution based on mixture and Gaussian copula. Parameter estimates of the proposed new BPI distributions are obtained using ML and Bayesian methods. Monte Carlo simulation study and analyses of real data are conducted to show the usefulness of the proposed distributions. We can conclude that the BPI distribution based on mixture and Gaussian copula is more flexible and performed better than the BPI distribution based on Gaussian copula.

References

Appendix

Algorithm 1: fisher information matrix

\[
I^{-1}(\theta) = \begin{pmatrix}
\frac{\partial^2 l}{\partial \theta_1^2} & \frac{\partial^2 l}{\partial \theta_1 \partial \theta_2} & \cdots & \frac{\partial^2 l}{\partial \theta_1 \partial \theta_n} \\
\frac{\partial^2 l}{\partial \theta_2 \partial \theta_1} & \frac{\partial^2 l}{\partial \theta_2^2} & \cdots & \frac{\partial^2 l}{\partial \theta_2 \partial \theta_n} \\
\vdots & \vdots & \ddots & \vdots \\
\frac{\partial^2 l}{\partial \theta_n \partial \theta_1} & \frac{\partial^2 l}{\partial \theta_n \partial \theta_2} & \cdots & \frac{\partial^2 l}{\partial \theta_n^2}
\end{pmatrix}^{-1}
\]

Where \( l \) is the asymptotic Fisher information matrix. The second partial derivatives will be simplified as follows:

1. \( l_{11} = -E \left[ \frac{\partial^2 l}{\partial \theta_1^2} \right] = \frac{n}{\alpha^2}, l_{12} = l_{21} = -E \left[ \frac{\partial^2 l}{\partial \theta_1 \partial \theta_2} \right] = 0, \]
   \( l_{13} = l_{31} = -E \left[ \frac{\partial^2 l}{\partial \theta_1 \partial \theta_3} \right] = \frac{n}{\beta^2} \sum_{i=1}^{n} \ln(t_{i1}) - \frac{r_1}{\beta_1}, \)
   \( i = 1, \)
   \( l_{33} = -E \left[ \frac{\partial^2 l}{\partial \theta_3^2} \right], \)

2. \( l_{14} = l_{41} = -E \left[ \frac{\partial^2 l}{\partial \theta_1 \partial \theta_4} \right] = \frac{n}{\gamma^2} \sum_{i=1}^{n} \ln(t_{i1}-\frac{r_2}{\beta_2}) - \frac{r_2}{\beta_2}, \)
   \( l_{44} = -E \left[ \frac{\partial^2 l}{\partial \theta_4^2} \right], \)

3. \( l_{12} = l_{21} = -E \left[ \frac{\partial^2 l}{\partial \theta_1 \partial \theta_2} \right] = \frac{n}{\gamma^2} \sum_{i=1}^{n} \ln(t_{i1}-\frac{r_2}{\beta_2}) - \frac{r_2}{\beta_2}, \)
   \( l_{22} = -E \left[ \frac{\partial^2 l}{\partial \theta_2^2} \right] = \frac{n}{\beta^2}, \)

4. \( l_{13} = l_{31} = -E \left[ \frac{\partial^2 l}{\partial \theta_1 \partial \theta_3} \right] = \frac{n}{\gamma^2} \sum_{i=1}^{n} \ln(t_{i1}-\frac{r_2}{\beta_2}) - \frac{r_2}{\beta_2}, \)
   \( l_{33} = -E \left[ \frac{\partial^2 l}{\partial \theta_3^2} \right], \)

Algorithm 2:

1. Introduce a non-negative latent variable \( \tau \), such that
   \[ \pi(U_j | \tau) \propto \pi(U_j | \tau) \cdot \pi(\tau) \cdot e^{-u_{ij} - \frac{\tau}{2} \cdot \frac{1}{\tau}} \]
   \[ 1 \left( \tau < e^{-\frac{2\gamma}{\beta}} \right) \cdot l(\mu_j > -\alpha \ln \left( \frac{\beta_j}{\alpha_j} \right)) + 1, j = 1, 2, \ldots, n. \]

2. Choose the initial values of \( U_j \) to be
   \[ u_{ij} = -\alpha \ln \left( \frac{\beta_j}{\alpha_j} \right) + 1, j = 1, 2, \ldots, n. \]

3. Sample \( \tau \) from Uniform \( (0, e^{-\frac{2\gamma}{\beta}}) \)

\[ \pi(\tau) = \frac{1}{e^{\frac{2\gamma}{\beta}}} \left( \tau < e^{-\frac{2\gamma}{\beta}} \right) \]

\[ \begin{align*}
x_i & < -2 \ln(\tau) \\
\text{Where} \quad x_i & = y_{ij}^2 + y_{jj}^2 - 2y_{ij}y_{jj} - 2 \ln(\tau) (1 - \rho^2) = y_{ij} + y_{jj} - 2\rho y_{ij}y_{jj} \\
& = y_{ij}^2 - 2y_{ij}(y_{jj}(\rho)) + y_{jj}^2 - y_{jj}^2(\rho^2) - 2\ln(\tau) (1 - \rho^2) = (y_{ij} - \rho y_{jj})^2 + y_{jj}^2(\rho^2) \\
& = y_{ij} - \rho y_{jj} = \pm \sqrt{2(1 - \rho^2) \left[ \ln(\tau) + \frac{y_{jj}^2}{2} \right]} \\
\text{Let} \quad q_{ij} & = \sqrt{2(1 - \rho^2) \left[ \ln(\tau) + \frac{y_{jj}^2}{2} \right]} \\
y_{ij} & = \rho y_{jj} \pm q_{ij} \]

let \( \delta_{ij} = \rho y_{jj} - q_{ij}, \delta_{2i} = \rho y_{jj} + q_{ij} \)

\[ u_{ij} > -\alpha \ln \left( \frac{\beta_j}{\alpha_j} \right). \]

Where

\[ y_{ij} \sim N(0, (1 - \rho^2)(1)) \]

Then

Let

\[ A_{ij} = \max \left[ -\alpha \ln \left( \frac{\beta_j}{\alpha_j} \right), F_{U_{ij}}^{-1}[\mu_1(\delta_{1i})], B_{ij} = F_{U_{ij}}^{-1}[\mu_2(\delta_{2i})] \right] \]

Then
\[(A_{ji} < u_{ji} < B_{ji})\]

4) Sample \(u_{ji}\) from \(f(u_{ji} | \tau)\)

\[\pi(u_{ji} | \tau) \propto e^{-u_{ji}} I(A_{ji} < u_{ji} < B_{ji})\]

\[k \int_{A_{ji}}^{B_{ji}} \pi(u_{ji} | \tau) \, du_{ji} = 1\]

\[k^{-1} = \int_{A_{ji}}^{B_{ji}} \pi(u_{ji} | \tau) \, du_{ji} = \int_{A_{ji}}^{B_{ji}} e^{-u_{ji}} \, du_{ji} = e^{-A_{ji}} - e^{-B_{ji}}\]

\[k = (e^{-A_{ji}} - e^{-B_{ji}})^{-1}\]

\[\pi(u_{ji} | \tau) = (e^{-A_{ji}} - e^{-B_{ji}})^{-1} e^{-u_{ji}}\]

\[I(A_{ji} < u_{ji} < B_{ji})\]

Find the Cdf of \(U_j\)

\[F(U_j) = \int_{A_{ji}}^{B_{ji}} \pi(u_{ji} | \tau) \, du_{ji} = \int_{A_{ji}}^{B_{ji}} (e^{-A_{ji}} - e^{-B_{ji}})^{-1} e^{-u_{ji}} \, du_{ji} =\]

\[(e^{-A_{ji}} - e^{-B_{ji}})^{-1} \times (e^{-A_{ji}} - e^{-u_{ji}}), j = 1, 2, \ldots, n\]

\(\pi(U_j | \tau)\) is a double truncated distribution that can be sampled by using the inverse distribution function method. Then, for \(v \sim Uniform(0, 1)\)

- Generating \(V \sim Uniform(0, 1)\)

\[V = F(u_{ji}) = \frac{(e^{-A_{ji}} - e^{-u_{ji}})}{(e^{-A_{ji}} - e^{-B_{ji}})}\]

\[(e^{-A_{ji}} - e^{-B_{ji}}) V = (e^{-A_{ji}} - e^{-u_{ji}})\]

\[e^{-u_{ji}} = e^{-A_{ji}} - (e^{-A_{ji}} - e^{-B_{ji}}) V\]

\[u_{ji} = -\ln[e^{-A_{ji}} - (e^{-A_{ji}} - e^{-B_{ji}}) V], j=1, 2, i=1, 2\ldots n.\]