mathematics

MDPI

Article

Power Families of Bivariate Proportional Hazard Models

Guillermo Martínez-Flórez [1,†], Carlos Barrera-Causil [2,*,†] and Artur J. Lemonte [3,†]

[1] Departamento de Matemáticas y Estadística, Facultad de Ciencias Básicas, Universidad de Córdoba, Montería 230002, Colombia
[2] Grupo de Investigación Davinci, Facultad de Ciencias Exactas y Aplicadas, Instituto Tecnológico Metropolitano, Medellín 050034, Colombia
[3] Departamento de Estatística, Universidade Federal do Rio Grande do Norte, Natal 59077-000, RN, Brazil
* Correspondence: carlosbarrera@itm.edu.co
† These authors contributed equally to this work.

Abstract: In this paper, we propose a general class of bivariate proportional hazard distributions, which is based on the family of asymmetric proportional hazard distributions and the bivariate Pareto copula. Distributional properties of the bivariate proportional hazard distribution are derived. We specialize the bivariate proportional hazard family of distributions to the normal case, and so we introduce the bivariate proportional hazard normal distribution. Parameter estimation by the maximum likelihood method of the bivariate proportional hazard normal distribution is then discussed. Finally, an application of the new bivariate distribution to real data is considered for illustrative purposes.

Keywords: asymmetric distribution; bivariate distribution; copula; two-stage estimation

MSC: 60E05

check for updates

Citation: Martínez-Flórez, G.; Barrera-Causil, C.; Lemonte, A.J. Power Families of Bivariate Proportional Hazard Models. *Mathematics* **2022**, *10*, 4410. https://doi.org/10.3390/math10234410

Academic Editor: Leonid V. Bogachev

Received: 14 October 2022
Accepted: 18 November 2022
Published: 23 November 2022

1. Introduction

The proportional hazard distribution was presented as an extension of the distribution of the minimum of a random sample replacing $n \in \mathbb{Z}$ with $b > 0$ in a random variable Z with probability density function (PDF) f, and cumulative distribution function (CDF) F. So, this distribution could be considered as a distribution of fractional order statistics ([1]). Here, we start defining the proportional hazard (PHF) distribution, which was studied by [2]; see also [3] where inference under progressively type II right-censored sampling for the PHF distribution is considered. Let F be a continuous CDF with PDF $f = dF$, and hazard function $h = f/(1-F)$. We say that Z has a PHF distribution associated with F and f, if its PDF is of the form

$$\varphi_F(z) = \beta f(z)\{1 - F(z)\}^{\beta-1}, \quad z \in \mathbb{R}, \tag{1}$$

where $\beta > 0$ is the shape parameter. We use the notation $Z \sim \mathrm{PHF}(\beta)$ to refer to this distribution. The CDF of the PHF distribution is given by

$$\mathbb{F}(z) = 1 - \{1 - F(z)\}^{\beta}, \quad z \in \mathbb{R}. \tag{2}$$

The hazard function of this model is $h_F(z) = \beta f(z)/\{1 - F(z)\}$; that is, the proportional hazard function of the PDF f. From Equation (2), we can use the inversion method for generating a random variable with PHF distribution; that is, if $U \sim \mathcal{U}(0,1)$, i.e., a uniform distribution on the interval $(0,1)$, then the random variable $X = \mu + \sigma F^{-1}(1 - (1-U)^{1/\beta})$ is distributed according to the PHF distribution with parameter vector $\boldsymbol{\theta} = (\mu, \sigma, \beta)'$, where $F^{-1}(\cdot)$ denotes the inverse of $F(\cdot)$. The distributional properties of the PHF distribution, as well as inferential procedures and information matrix for the model parameters are studied by [2].

The location-scale extension of model (1) is obtained using the linear transformation $X = \mu + \sigma Z$, where $\mu \in \mathbb{R}$ is a location parameter, $\sigma > 0$ is a scale parameter, and $Z \sim \text{PHF}(\beta)$. The PDF of the random variable X is given by

$$\varphi(x) = \frac{\beta}{\sigma} f\left(\frac{x-\mu}{\sigma}\right)\left(1 - F\left(\frac{x-\mu}{\sigma}\right)\right)^{\beta-1}, \quad x \in \mathbb{R}. \tag{3}$$

If a random variable X follows the model (3), it is denoted by $X \sim \text{PHF}(\mu, \sigma, \beta)$. If $f = \phi$ and $F = \Phi$, where ϕ and Φ detone the PDF and CDF of a standard normal distribution, we write $X \sim \text{PHN}(\mu, \sigma, \beta)$, and so we obtain the proportional hazard normal (PHN) distribution. Note that $\text{PHN}(\mu, \sigma, \beta = 1) \equiv \mathcal{N}(\mu, \sigma)$, and so the normal distribution is a special case of the PHN distribution. Hence, it is more flexible than the normal model in terms of skewness and kurtosis, due to the inclusion of the shape parameter β. We have that the Kumaraswamy-G (KwG) distribution, proposed by [4], is defined by the PDF

$$\varphi_G(z) = abg(z)G^{a-1}(z)\{1 - G^a(z)\}^{b-1}, \quad z \in \mathbb{R}, \tag{4}$$

where G is an arbitrary continuous CDF, and $g = dG$. Also, $a > 0$ and $b > 0$ are additional shape parameters to that of G. If Z is a random variable with PDF (4), we write $Z \sim \text{KwG}(a, b)$. It is worth stressing that when $a = 1$, we obtain the $\text{PHF}(b)$ distribution, i.e., $\text{KwG}(a = 1, b) \equiv \text{PHF}(b)$, and so this kind of distribution arises as a possible solution when we have asymmetric data.

The construction of multivariate distributions from copula theory can be carried out using the Sklar's theorem. The way to obtain a multivariate distribution from the Sklar's theorem is as follows: Let $X_1, \ldots, X_p$ be p random variables with continuous CDFs $F_{X_1}(x_1), \ldots, F_{X_p}(x_p)$, respectively. Then, according to Sklar's theorem, $F_{X_1,\ldots,X_p}(x_1, \ldots, x_p)$ has a unique copula representation:

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

Moreover, it is well known that many dependence properties of a multivariate distribution depend only on the corresponding copula. Therefore, many dependence properties of a multivariate distribution can be obtained by studying the corresponding copula. By using Clayton copula, Ref [5] proposed a bivariate extension of the power-normal distribution. Specifically, a bivariate random variable (X_1, X_2) has a bivariate power-normal distribution, denoted by $\text{BPN}(\delta, \beta_1, \beta_2)$, if for $\beta_1 > 0$ and $\beta_2 > 0$, (X_1, X_2) has the following joint CDF

$$F_{X_1,X_2}(x_1, x_2) = (\{\Phi(x_1)\}^{-\beta_1/\delta} + \{\Phi(x_2)\}^{-\beta_2/\delta} - 1)^{-\delta}, \quad (x_1, x_2) \in \mathbb{R}^2,$$

where $\delta > 0$ is the parameter that controls the dependence in the Clayton copula. The Clayton copula is usually attributed to [6], but it has also been studied by [7]. For $(u, v) \in [0, 1] \times [0, 1]$, it is defined as

$$C_\delta(u, v) = (u^{-1/\delta} + v^{-1/\delta} - 1)^{-\delta}, \quad \delta > 0.$$

The first systematic study of this model was conducted by [8], who interpreted the parameter δ as a measure of the dependence between u and v. Therefore, the independence between the variables is obtained when δ tends to zero. Obviously there are other forms of constructing multivariate distributions, and so new multitariate distributions have been proposed in the statistical literature. To mention a few, but not limited to, we refer the reader to the works by [9–12], among others.

It is worth emphasizing that the univariate PHF family of distributions has received significant attention over the last years in the statistical literature, mainly due to its flexibility in considering different models in its construction given by f and F; see expression (1). On the other hand, multivariate extensions of this univariate family of distributions has been little explored. This paper fills this gap and provides a bivariate extension of this univariate distribution. The bivariate PHF distribution is quite simple and, hence, may be widely applied in analyzing bivariate real data in practice. Additionally, distributional properties of the bivariate PHF distribution are investigated in details. An important claim to introduce the bivariate PHF distribution relies on the fact that the practitioners will have new bivariate model to use in bivariate settings. Additionally, the formulae related with the new bivariate model are manageable and with the use of modern computer resources and its numerical capabilities, the bivariate PHF distribution may prove to be an useful addition to the arsenal of applied statisticians. We hope that the bivariate distribution introduced in this paper may serve as an alternative bivariate model to some well-known bivariate model available in the statistical literature. We also hope that the bivariate PHF distribution may work better (at least in terms of model fitting) than some bivariate distributions available in the literature in certain practical situations, although it cannot always be guaranteed. Here, the bivariate PHF family of distributions we introduce is obtained from a multivariate distribution which was constructed from Pareto copula (descried in the next section) coupled with PHF marginals. It is worth stressing that many types of copulas are available o construct multivariate distributions, namely: Clayton copula, Frank copula, Joe copula, and Gumbel copula, among others. Perhaps, the most common copula is the Clayton copula. In this paper, instead, we shall consider the Pareto copula to introduce the new bivariate model due to its simplicity and interesting properties (descried in the next section). In addition, the Pareto copula yields a simple form of the likelihood function, so it makes the estimation of the model parameters easy to deal with. In short, it is evident that few works about bivariate generalizations of the PHF distribution have been published in the statistical literature and, hence, a new, simple and tractable bivariate PHF extension of the univariate PHF distribution through Pareto copula is welcome.

The paper is organized as follows. Section 2 presents the Pareto copula and discusses with details the structural properties derived from it. The bivariate proportional hazard distribution is proposed in Section 3. Distributional properties of this bivariate family of distributions are also derived in this section. In Section 4, we study the bivariate PHN distribution, where the univariate normal distribution is taken into account. Location-scale extension, as well as parameter estimation for the bivariate PHN distribution are also discussed in this section. An empirical appliction of the bivariate PHN distribution that considers real data is provided in Section 5 for illustrative purposes. Finally, Section 6 concludes the paper.

2. Pareto Copula

The Pareto distribution, or Pareto type I distribution, has been extensively used in economic literature and in reliability theory. Its CDF is given by

$$G_X(x) = 1 - \left(\frac{x}{\sigma}\right)^{-\alpha}, \quad x > 0,$$

where $\sigma > 0$ and $\alpha > 0$. Another type of Pareto distribution is the Pareto type II distribution, whose CDF takes the form

$$F_X(x) = 1 - \left(1 + \frac{x}{\sigma}\right)^{-\alpha}, \quad x > 0.$$

Its survival function is given by

$$S_X(x) = \left(1 + \frac{x}{\sigma}\right)^{-\alpha}, \quad x > 0.$$

The bivariate Pareto distribution of the random vector $X = (X_1, X_2)$ has joint CDF in the form

$$\begin{aligned} F_{X_1,X_2}(x_1, x_2) &= F_{X_1}(x_1) + F_{X_2}(x_2) - 1 \\ &\quad + [\{S_{X_1}(x_1)\}^{-1/\alpha} + \{S_{X_2}(x_2)\}^{-1/\alpha} - 1]^{-\alpha}, \end{aligned}$$

where $(x_1, x_2) \in \mathbb{R}^2_+$. Then, using the Sklar's theorem for continuous marginal distributions, the bivariate Pareto copula is given by

$$\begin{aligned} C(u_1, u_2) &= F_{X_1,X_2}(F^{-1}_{X_1}(u_1), F^{-1}_{X_2}(u_2)) \\ &= u_1 + u_2 - 1 + [(1 - u_1)^{-1/\alpha} + (1 - u_2)^{-1/\alpha} - 1]^{-\alpha}. \end{aligned}$$

Next, we shall consider some properties of the bivariate Pareto copula:

1. The joint PDF of the Pareto copula is

$$c(u_1, u_2) = \frac{\partial^2 C(u_1, u_2)}{\partial u_1 \partial u_2} = \frac{\alpha + 1}{\alpha} \frac{[(1 - u_1)(1 - u_2)]^{-\frac{\alpha+1}{\alpha}}}{[(1 - u_1)^{-1/\alpha} + (1 - u_2)^{-1/\alpha} - 1]^{\alpha+2}}.$$

2. The joint survival function of the Pareto copula is

$$S(u_1, u_2) = P(U_1 \geq u_1, U_2 \geq u_2) = [(1 - u_1)^{-1/\alpha} + (1 - u_2)^{-1/\alpha} - 1]^{-\alpha}.$$

3. For all $u_2 \in [0, 1]$, the bivariate Pareto copula is a concave function on u_1 for fixed u_2. This result follows from

$$\frac{\partial^2 C(u_1, u_2)}{\partial u_1^2} = \frac{\alpha + 1}{\alpha} \frac{(1 - u_1)^{-\frac{2\alpha+1}{\alpha}}}{[(1 - u_1)^{-1/\alpha} + (1 - u_2)^{-1/\alpha} - 1]^{\alpha+2}} \geq 0.$$

4. The bivariate dependency measures for continuous variables, usually used in copulas, are Kendall's tau correlation coefficient (τ), Spearman's rho (ρ_s), and the medial correlation coefficient. The first two coefficients are invariant to increasing transformations. The Kendall's tau measures the difference between the probability of two concordant random pairs and the probability of two discordant random pairs. For a continuous bivariate distribution function F, τ is defined by

$$\tau = 4 \int F dF - 1.$$

Spearman's rho correlation coefficient measures the correlation between the two CDFs, $F_{X_1}(x_1)$ and $F_{X_2}(x_2)$. Then, given that F_{X_1} and F_{X_2} are continuous uniform random variables on $(0, 1)$ with mean and variance $1/2$ and $1/12$, respectively, it can be shown that

$$\rho_s = 12 \int_0^1 \int_0^1 C(u_1, u_2) du_1 du_2 - 3.$$

If M_{X_1} and M_{X_2} denote the medians of X_1 and X_2, respectively, the medial correlation coefficient of X_1 and X_2, say $M_{X_1X_2}$, is given by (see [13])

$$M_{X_1X_2} = 4C\left(\frac{1}{2}, \frac{1}{2}\right).$$

For the bivariate Pareto copula, we have that

$$\tau = -\frac{2\alpha}{2\alpha+1},$$
$$\rho_s = 12\int_{-1}^{0}\int_{-1}^{0}\left[u_1^{-1/\alpha}+u_2^{-1/\alpha}-1\right]^{-\alpha}du_1du_2-3, \tag{5}$$
$$M_{X_1X_2} = 4\left(2^{(\alpha+1)/\alpha}-1\right)^{-\alpha}.$$

5. A non-negative function b is totally positive of order 2 (TP_2) if for all $x_1 < x_2$ and $y_1 < y_2$, with $x, y \in \mathbb{R}$, if is verified that

$$b(x_1,y_1)b(x_2,y_2) \geq b(x_2,y_1)b(x_1,y_2).$$

This is a condition of positive dependence when it is performed on the PDF, and means that two pairs with components matching high-low and low-low are more likely than two pairs with high-low and low-high components. For all $u_1, u_2 \in [0,1]$, the bivariate Pareto copula is a non-negative function of order 2 (or TP_2).

6. The bivariate tail dependence concept is related to the amount of dependency on the tail of the bivariate distribution, in the upper or lower quadrant. The λ symbol is used to determine a tail dependence parameter. If a bivariate copula C, with survival copula $\bar{C}$, is such that

$$\lambda_U = \lim_{u\to 1}\frac{\bar{C}(u,u)}{1-u}$$

exists, then C has an upper tail dependence if $\lambda_U \in (0,1]$ and has no upper tail dependence if $\lambda_U = 0$. It can be shown that

$$\lambda_U = \lim_{v\to 1}\frac{1-2v+C(v,v)}{1-v},$$

where v refers to the component u_2. Similarly, if a bivariate copula C is such that

$$\lambda_L = \lim_{u\to 0}\frac{C(u,u)}{u}$$

exists, then C has lower tail dependence if $\lambda_L \in (0,1]$ and has no lower tail dependence if $\lambda_L = 0$. The reasoning behind these definitions is that

$$\lambda_U = \lim_{u\to 1}\Pr[U_1 > u \mid U_1 > u], \quad \lambda_L = \lim_{u\to 0}\Pr[U_1 \leq u \mid U_1 \leq u].$$

For the bivariate Pareto copula, we have that $\lambda_u = \infty$ and $\lambda_L = 0$. Then, the Pareto copula has a lower tail dependence. That is, for a value u arbitrarily close to zero, there is a positive probability that one of the variables u_1 or u_2 takes values smaller than u, given that the other is smaller than u.

3. Bivariate Proportional Hazard Distribution

In what follows, the joint CDF of (X_1, X_2), $F_{X_1,X_2}(x_1,x_2)$, is constructed from the bivariate Pareto copula, with marginal distributions $X_1 \sim \mathrm{PHF}(\alpha_1)$ and $X_2 \sim \mathrm{PHF}(\alpha_2)$, where $\alpha_1 > 0$ and $\alpha_2 > 0$. Then, from the Sklar's theorem, the joint CDF of (X_1, X_2) is given by

$$F_{X_1,X_2}(x_1,x_2) = C(F_{X_1}(x_1), F_{X_2}(x_2)).$$

We have the following definition.

Definition 1. *A bivariate random vector* (X_1, X_2) *has a bivariate proportional hazard distribution, if its joint CDF is given by*

$$\begin{aligned} F_{X_1,X_2}(x_1, x_2) = 1 &- (1 - F_{X_1}(x_1))^{\alpha_1} - (1 - F_{X_2}(x_2))^{\alpha_2} \\ &+ [(1 - F_{X_1}(x_1))^{-\frac{\alpha_1}{\alpha}} + (1 - F_{X_2}(x_2))^{-\frac{\alpha_2}{\alpha}} - 1]^{-\alpha}, \end{aligned}$$

where $(x_1, x_2) \in \mathbb{R}^2$, $\alpha_1 > 0$, $\alpha_2 > 0$ *and* $\alpha > 0$.

Remark 1. *We denote the bivariate proportional hazard distribution in Definition 1 by* $(X_1, X_2) \sim \mathrm{BPHF}(\alpha_1, \alpha_2, \alpha)$.

Remark 2. *Let* $(X_1, X_2) \sim \mathrm{BPHF}(\alpha_1, \alpha_2, \alpha)$. *Then, the joint PDF of* (X_1, X_2) *has the form*

$$f_{X_1,X_2}(x_1, x_2) = \frac{(\alpha + 1)\alpha_1\alpha_2}{\alpha} \frac{f_{X_1}(x_1)[1 - F_{X_1}(x_1)]^{-\frac{\alpha_1}{\alpha}-1} f_{X_2}(x_2)[1 - F_{X_2}(x_2)]^{-\frac{\alpha_2}{\alpha}-1}}{[(1 - F_{X_1}(x_1))^{-\frac{\alpha_1}{\alpha}} + (1 - F_{X_2}(x_2))^{-\frac{\alpha_2}{\alpha}} - 1]^{\alpha+2}}.$$

We have the following propositions.

Proposition 1. *Let* $(X_1, X_2) \sim \mathrm{BPHF}(\alpha_1, \alpha_2, \alpha)$. *We have that*

(i) $X_j \sim \mathrm{PHF}(\alpha_j)$, *for* $j = 1, 2$.

(ii) *The CDF of* X_1, *given* $X_2 = x_2$, *is*

$$F_{X_1|X_2}(x_1 \mid X_2 = x_2) = \frac{[1 - F_{X_2}(x_2)]^{-\alpha_2(\frac{\alpha+1}{\alpha})}}{[(1 - F_{X_1}(x_1))^{-\frac{\alpha_1}{\alpha}} + (1 - F_{X_2}(x_2))^{-\frac{\alpha_2}{\alpha}} - 1]^{\alpha+1}}.$$

(iii) *The PDF of* X_1, *given* $X_2 = x_2$, *is*

$$f_{X_1|X_2}(x_1 \mid X_2 = x_2) = \frac{(\alpha + 1)\alpha_1}{\alpha} \frac{f_{X_1}(x_1)[1 - F_{X_1}(x_1)]^{-\frac{\alpha_1}{\alpha}-1}[1 - F_{X_2}(x_2)]^{-\alpha_2(\frac{\alpha-1}{\alpha})}}{[(1 - F_{X_1}(x_1))^{-\frac{\alpha_1}{\alpha}} + (1 - F_{X_2}(x_2))^{-\frac{\alpha_2}{\alpha}} - 1]^{\alpha+2}}. \tag{6}$$

(iv) *The joint survival function of* (X_1, X_2) *is*

$$\begin{aligned} S_{X_1,X_2}(x_1, x_2) &= P(X_1 \geq x_1, X_2 \geq x_2) \\ &= F_{X_1}(x_1) + F_{X_2}(x_2) - 1 + C(1 - F_{X_1}(x_1), 1 - F_{X_2}(x_2)) \\ &= [(1 - F_{X_1}(x_1))^{-\frac{\alpha_1}{\alpha}} + (1 - F_{X_2}(x_2))^{-\frac{\alpha_2}{\alpha}} - 1]^{-\alpha}. \end{aligned}$$

Proposition 2. *Let* $(X_1, X_2) \sim \mathrm{BPHF}(\alpha_1, \alpha_2, \alpha)$. *We have that* X_1 *is stochastically decreasing in* X_2, *and vice versa, for any value of* α_1, α_2, *and* α.

Proof. Since Pareto copula is a concave function on u_1 for fixed u_2, the result holds. □

The bivariate hazard rate of X_1 and X_2 was defined by [14] and it is given by

$$h(x_1, x_2) = \left(-\frac{\partial}{\partial x_1}, -\frac{\partial}{\partial x_2}\right) \log[\Pr(X_1 > x_1, X_2 > x_2)] = (h_1(x_1, x_2), h_2(x_1, x_2)).$$

Proposition 3. *Let* $(X_1, X_2) \sim \mathrm{BPHF}(\alpha_1, \alpha_2, \alpha)$. *We have that*

(i) *For fixed* x_2, $h_1(x_1, x_2)$ *is an increasing function of* x_1.

(ii) *For fixed* x_1, $h_2(x_1, x_2)$ *is an increasing function of* x_2.

Proof. This result can be verified as follows. Let $u_1 = (1 - F_{X_1}(x_1))^{-\frac{\alpha_1}{\alpha}}$ and $k_2 = (1 - F_{X_2}(x_2))^{-\frac{\alpha_2}{\alpha}} - 1$. It can be shown that

$$\frac{\partial h_1(x_1, x_2)}{\partial x_1} = \frac{\partial Q_1(x_1, x_2)}{\partial u_1}\frac{\partial u_1}{\partial x_1} < 0,$$

where $Q_1(x_1, x_2) = -\partial \log(u_1 + k)^{-\alpha}/\partial u_1$ and, hence, the result follows. □

Proposition 4. *Let* $(X_1, X_2) \sim \mathrm{BPHF}(\alpha_1, \alpha_2, \alpha)$. *We have that the Kendall's tau correlation coefficient, Spearman's rho and the medial correlation coefficient, are given by the expressions provided in* (5).

Proof. Kendall's tau and Spearman's rho correlation coefficients are invariant to increasing transformations, and so the resut follows. Note that is enough to take the transformation $u_1 = 1 - (1 - F_{X_1}(x_1))^{-\alpha_1}$, and $u_2 = 1 - (1 - F_{X_2}(x_2))^{-\alpha_2}$, which leads to the PDF Pareto copula. Since the medial coefficient is defined directly from the Pareto copula, the result is obtained directly. □

Remark 3. *The upper and lower limits for* ρ_s *are*

$$-\frac{4\alpha(\alpha+1) - 1/2}{(2\alpha+1)^2} \le \rho_s \le \frac{-2\alpha + 1/2}{2\alpha + 1}.$$

Proof. Since the Pareto copula has a lower tail dependence, the result holds. □

Proposition 5. *The bivariate distribution* $\mathrm{BPHF}(\alpha_1, \alpha_2, \alpha)$ *has a lower tail dependence.*

4. Bivariate Proportional Hazard Normal Distribution

When $F_{X_1} = F_{X_2} = \Phi$, i.e., the standard normal CDF, we obtain the bivariate PHN distribution, denoted by $\mathrm{BPHN}(\alpha_1, \alpha_2, \alpha)$. The joint PDF of $(X_1, X_2) \sim \mathrm{BPHN}(\alpha_1, \alpha_2, \alpha)$ is given by

$$f_{X_1,X_2}(x_1, x_2) = \frac{(\alpha+1)\alpha_1\alpha_2}{\alpha}\frac{\phi(x_1)[1-\Phi(x_1)]^{-\frac{\alpha_1}{\alpha}-1}\phi(x_2)[1-\Phi(x_2)]^{-\frac{\alpha_2}{\alpha}-1}}{[(1-\Phi(x_1))^{-\frac{\alpha_1}{\alpha}} + (1-\Phi(x_2))^{-\frac{\alpha_2}{\alpha}} - 1]^{\alpha+2}},$$

where $(x_1, x_2) \in \mathbb{R}^2$, $\alpha_1 > 0$, $\alpha_2 > 0$ and $\alpha > 0$. The joint CDF takes the form

$$\begin{aligned} F_{X_1,X_2}(x_1, x_2) = {} & 1 - (1-\Phi(x_1))^{\alpha_1} - (1-\Phi(x_2))^{\alpha_2} \\ & + [(1-\Phi(x_1))^{-\frac{\alpha_1}{\alpha}} + (1-\Phi(x_2))^{-\frac{\alpha_2}{\alpha}} - 1]^{-\alpha}. \end{aligned}$$

The marginal distributions of $(X_1, X_2) \sim \mathrm{BPHN}(\alpha_1, \alpha_2, \alpha)$ are $X_j \sim \mathrm{PHN}(\alpha_j)$ for $j = 1, 2$. The PDF of X_1, given $X_2 = x_2$, is given by

$$f(X_1 \mid X_2 = x_2) = \frac{(\alpha+1)\alpha_1}{\alpha}\frac{\phi(x_1)[1-\Phi(x_1)]^{-\frac{\alpha_1}{\alpha}-1}[1-\Phi(x_2)]^{-\alpha_2(\frac{\alpha-1}{\alpha})}}{[(1-\Phi(x_1))^{-\frac{\alpha_1}{\alpha}} + (1-\Phi(x_2))^{-\frac{\alpha_2}{\alpha}} - 1]^{\alpha+2}},$$

and the CDF of X_1, given $X_2 = x_2$, is

$$F(X_1 \mid X_2 = x_2) = \frac{[1-\Phi(x_2)]^{-\alpha_2(\frac{\alpha+1}{\alpha})}}{[(1-\Phi(x_1))^{-\frac{\alpha_1}{\alpha}} + (1-\Phi(x_2))^{-\frac{\alpha_2}{\alpha}} - 1]^{\alpha+1}}.$$

The joint survival function of $(X_1, X_2) \sim \mathrm{BPHN}(\alpha_1, \alpha_2, \alpha)$ has the form

$$S_{X_1,X_2}(x_1, x_2) = [(1-\Phi(x_1))^{-\frac{\alpha_1}{\alpha}} + (1-\Phi(x_1))^{-\frac{\alpha_2}{\alpha}} - 1]^{-\alpha}.$$

4.1. Location-Scale Extension

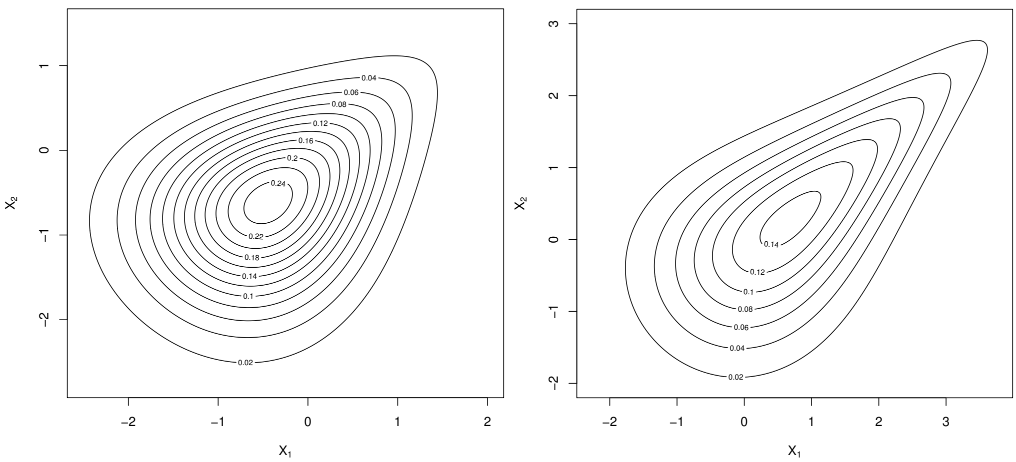
The family of BPHN distributions with location-scale parameters is defined as the joint distribution of $X_1 = \mu_1 + \sigma_1 Z_1$ and $X_2 = \mu_2 + \sigma_2 Z_2$, where $Z_j \sim \text{PHN}(\alpha_j)$, $\mu_j \in \mathbb{R}$ and $\sigma_j > 0$ for $j = 1, 2$. The corresponding joint PDF is given by

$$f_{X_1,X_2}(x_1, x_2) = \frac{(\alpha+1)\alpha_1\alpha_2}{\sigma_1\sigma_2\alpha} \frac{\phi(z_1)[1-\Phi(z_1)]^{-\frac{\alpha_1}{\alpha}-1}\phi(z_2)[1-\Phi(z_2)]^{-\frac{\alpha_2}{\alpha}-1}}{[(1-\Phi(z_1))^{-\frac{\alpha_1}{\alpha}} + (1-\Phi(z_2))^{-\frac{\alpha_2}{\alpha}} - 1]^{\alpha+2}},$$

where μ_j is the location parameter, σ_j is the scale parameter, and

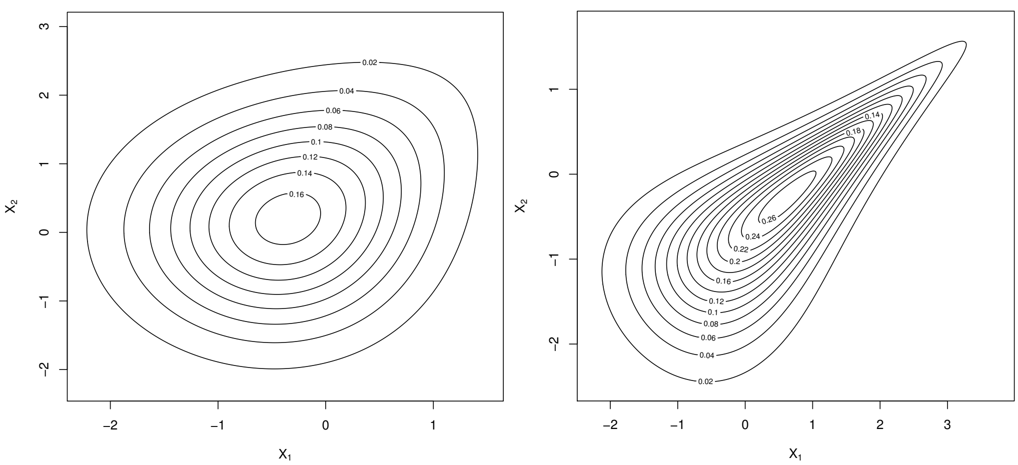
$$z_j = \frac{x_j - \mu_j}{\sigma_j}, \quad j = 1, 2.$$

We use the notation $\text{BPHN}(\mu_1, \sigma_1, \alpha_1, \mu_2, \sigma_2, \alpha_2, \alpha)$ to denote this location-scale extension. Figures 1 and 2 display contour plots of the BPHN distribution for some parameter values.



(**a**) (**b**)

Figure 1. Contour plots: (**a**) BPHN(0,1,1.75,0,1,2.25,2), and (**b**) BPHN(0,1,0.5,0,1,0.75,0.75).



(**a**) (**b**)

Figure 2. Contour plots: (**a**) BPHN(0,1,1.5,0,1,0.75,3.5), and (**b**) BPHN(0,1,0.75,0,1,2,0.5).

4.2. Parameter Estimation

In what follows, we consider the issue of estimating the parameters of the BPHN distribution. Given the dependency structure induced by Pareto copula, we use the maximum likelihood (ML) method, and the two-stage method (see [15]) to estimate the parameter vector $\boldsymbol{\theta} = (\mu_1, \sigma_1, \alpha_1, \mu_2, \sigma_2, \alpha_2, \alpha)'$. Initially, as in [5], we perform the following reparameterizations: $\alpha_1 = \alpha\gamma_1$, and $\alpha_2 = \alpha\gamma_2$, so that the interest is focused on the parameter vector $\boldsymbol{\theta}_1 = (\mu_1, \sigma_1, \gamma_1, \mu_2, \sigma_2, \gamma_2, \alpha)'$. Thus, for a sample of size n from the $\mathrm{BPHN}(\mu_1, \sigma_1, \gamma_1, \mu_2, \sigma_2, \gamma_2, \alpha)$ distribution, say $(x_{11}, x_{12}), \ldots, (x_{n1}, x_{n2})$, the log-likelihood function is given by

$$\begin{aligned}
\ell(\boldsymbol{\theta}_1) = n & \left[\log(\alpha+1) + \log(\alpha) + \sum_{j=1}^{2} \log(\gamma_j) - \sum_{j=1}^{2} \log(\sigma_j) \right] \\
& + \sum_{i=1}^{n} \log(\phi(z_{i1})) + \sum_{i=1}^{n} \log(\phi(z_{i2})) \\
& - \gamma_1 \sum_{i=1}^{n} \log[1-\Phi(z_{i1})] - \gamma_2 \sum_{i=1}^{n} \log[1-\Phi(z_{i2})] \\
& - (\alpha+2) \sum_{i=1}^{n} \log[(1-\Phi(z_{i1}))^{-\gamma_1} + (1-\Phi(z_{i2}))^{-\gamma_2} - 1],
\end{aligned}$$

where

$$z_{ij} = \frac{x_{ij} - \mu_j}{\sigma_j}, \quad j = 1, 2, \quad i = 1, 2, \ldots, n.$$

The ML estimate $\widehat{\boldsymbol{\theta}}_1 = (\widehat{\mu}_1, \widehat{\sigma}_1, \widehat{\gamma}_1, \widehat{\mu}_2, \widehat{\sigma}_2, \widehat{\gamma}_2, \widehat{\alpha})'$ of $\boldsymbol{\theta}_1 = (\mu_1, \sigma_1, \gamma_1, \mu_2, \sigma_2, \gamma_2, \alpha)'$ can be obtained from the direct maximization of the log-likelihood function $\ell(\boldsymbol{\theta}_1)$ with respect to the model parameters. Using the R programming language, specifically applying the `optim` function, the maximization of the log-likelihood function can be performed. Also, using the invariance property of the ML estimator, we can easily obtain the ML estimates $\widehat{\alpha}_1$ and $\widehat{\alpha}_2$ of α_1 and α_2, respectively.

The score functions are

$$\begin{aligned}
\frac{\partial \ell(\boldsymbol{\theta}_1)}{\partial \mu_j} = & \frac{1}{\sigma_j} \sum_{i=1}^{n} z_{ij} - \frac{\gamma_j}{\sigma_j} \sum_{i=1}^{n} \frac{\phi(z_{ij})}{1-\Phi(z_{ij})} \\
& + \frac{(\alpha+2)\gamma_j}{\sigma_j} \sum_{i=1}^{n} \frac{\phi(z_{ij})(1-\Phi(z_{ij}))^{-\gamma_j - 1}}{(1-\Phi(z_{i1}))^{-\gamma_1} + (1-\Phi(z_{i2}))^{-\gamma_2} - 1},
\end{aligned}$$

$$\begin{aligned}
\frac{\partial \ell(\boldsymbol{\theta}_1)}{\partial \sigma_j} = & -\frac{n}{\sigma_j} + \frac{1}{\sigma_j} \sum_{i=1}^{n} z_{ij}^2 - \frac{\gamma_j}{\sigma_j} \sum_{i=1}^{n} \frac{z_{ij}\phi(z_{ij})}{1-\Phi(z_{ij})} \\
& + \frac{(\alpha+2)\gamma_j}{\sigma_j} \sum_{i=1}^{n} \frac{z_{ij}\phi(z_{ij})(1-\Phi(z_{ij}))^{-\gamma_j - 1}}{(1-\Phi(z_{i1}))^{-\gamma_1} + (1-\Phi(z_{i2}))^{-\gamma_2} - 1},
\end{aligned}$$

$$\begin{aligned}
\frac{\partial \ell(\boldsymbol{\theta}_1)}{\partial \gamma_j} = & \frac{n}{\gamma_j} - \sum_{i=1}^{n} \log(1-\Phi(z_{ij})) \\
& + (\alpha+2) \sum_{i=1}^{n} \frac{(1-\Phi(z_{ij}))^{-\gamma_j} \log(1-\Phi(z_{ij}))}{(1-\Phi(z_{i1}))^{-\gamma_1} + (1-\Phi(z_{i2}))^{-\gamma_2} - 1},
\end{aligned}$$

$$\frac{\partial \ell(\boldsymbol{\theta}_1)}{\partial \alpha} = \frac{n}{\alpha+1} + \frac{n}{\alpha} - \sum_{i=1}^{n} \log[(1-\Phi(z_{i1}))^{-\gamma_1} + (1-\Phi(z_{i2}))^{-\gamma_2} - 1],$$

where $j = 1,2$. The ML estimate $\widehat{\boldsymbol{\theta}}_1 = (\widehat{\mu}_1, \widehat{\sigma}_1, \widehat{\gamma}_1, \widehat{\mu}_2, \widehat{\sigma}_2, \widehat{\gamma}_2, \widehat{\alpha})'$ can also be obtained by solving simultaneously the nonlinear system of equations

$$\begin{cases} \dfrac{\partial \ell(\boldsymbol{\theta}_1)}{\partial \mu_j} = 0, \\ \dfrac{\partial \ell(\boldsymbol{\theta}_1)}{\partial \sigma_j} = 0, \\ \dfrac{\partial \ell(\boldsymbol{\theta}_1)}{\partial \gamma_j} = 0, \\ \dfrac{\partial \ell(\boldsymbol{\theta}_1)}{\partial \alpha} = 0. \end{cases}$$

Numerical approaches are required for solving the above nonlinear system of equations, which requires the specification of initial values. To obtain initial values for the model parameters to be used in the iterative process, we can use the two-stage estimation procedure proposed by [15], since this procedure was defined for multivariate copula-based models. The first stage considers ML estimation from univariate marginals, while the second stage considers ML estimation of the dependence parameter with the other parameters held fixed from the first stage:

First stage. Considering that $X_j \sim \text{PHN}(\mu_j, \sigma_j, \alpha_j)$ for $j = 1,2$. We have that the ML estimaties of μ_j, σ_j and α_j are obtained from the solutions of the nonlinear equations

$$\begin{cases} \dfrac{n}{\alpha_j} + \displaystyle\sum_{i=1}^{n} \log(1 - \Phi(z_{ij})) = 0, \\ \displaystyle\sum_{i=1}^{n} z_{ij} + (\alpha_j - 1) \sum_{i=1}^{n} W_{ij} = 0, \\ \displaystyle\sum_{i=1}^{n} z_{ij}^2 + (\alpha + 1) \sum_{i=1}^{n} z_{ij} W_{ij} = n, \end{cases}$$

where

$$W_{ij} = \frac{\phi(z_{ij})}{1 - \Phi(z_{ij})}, \quad j = 1,2, \quad i = 1,2,\ldots,n.$$

The solution of the previous system of equations for each j allows us to obtain the ML estimates $\widetilde{\mu}_j$, $\widetilde{\sigma}_j$ and $\widetilde{\alpha}_j$ of μ_j, σ_j and α_j, respectively, for $j = 1,2$.

Second stage. In this stage, we estimate the parameter α by replacing the unknown parameters with the ML estimates from the previous stage and then maximize the resulting log-likelihood function; that is, we need to maximize the function

$$L(\alpha) = \sum_{i=1} \log[C_\alpha((1 - \Phi(z_{i1}))^{-\hat{\alpha}_1}, (1 - \Phi(z_{i2}))^{-\hat{\alpha}_2}],$$

where

$$\log[C_\alpha(u,v)] = \log\left(\frac{\alpha+1}{\alpha}\right) - \left(1 + \frac{1}{\alpha}\right) \log((1-u)(1-v)) \\ - (\alpha+2) \log[(1-u)^{-\frac{1}{\alpha}} + (1-v)^{-\frac{1}{\alpha}} - 1],$$

so that we have the score function

$$\frac{\partial L(\alpha)}{\partial \alpha} = -\frac{1}{\alpha(\alpha+1)} + \frac{1}{\alpha^2} \log[(1-u)(1-v)] \\ - \log[(1-u)^{-\frac{1}{\alpha}} + (1-v)^{-\frac{1}{\alpha}} - 1].$$

The solution of the resulting equation $\partial L(\alpha)/\partial \alpha = 0$ lead us to the estimate $\widetilde{\alpha}$ of α.

We can also consider another alternative to obtain initial values for the model parameters by using the method proposed by [16], where we initially normalize each variable to obtain a model that depends only on α_1, α_2 and α; that is, we have the log-likelihood function

$$\ell(\gamma_1, \gamma_2, \alpha) = \alpha n \log(\alpha(\alpha+1)\gamma_1\gamma_2) - \sum_{i=1}^{n}\sum_{j=1}^{2} \gamma_j \log(1 - \Phi(z_{ij}))$$
$$- (\alpha+2)\sum_{i=1}^{n} \log((1-\Phi(z_{i1}))^{-\gamma_1} + (1-\Phi(z_{i2}))^{-\gamma_2} - 1).$$

After doing

$$\frac{\partial \ell(\gamma_1, \gamma_2, \alpha)}{\partial \alpha} = 0,$$

we obtain

$$\widetilde{\alpha}(\gamma_1, \gamma_2) = \frac{(2-A) + \sqrt{4+A^2}}{2A},$$

where

$$A := A(\gamma_1, \gamma_2) = \frac{1}{n}\sum_{i=1}^{n} \log((1-\Phi(z_{i1}))^{-\gamma_1} + (1-\Phi(z_{i2}))^{-\gamma_2} - 1).$$

Thus, the new profiled log-likelihood function $\ell_p(\widetilde{\alpha}(\gamma_1, \gamma_2), \gamma_1, \gamma_2)$ is reached. After obtaining the estimates $\widetilde{\gamma}_1$ and $\widetilde{\gamma}_2$ of γ_1 and γ_2, respectively, we then obtain an estimate for α as $\widetilde{\alpha}(\widetilde{\gamma}_1, \widetilde{\gamma}_2)$. We can take as initial values for μ_1 and μ_2 the sample means, and for σ_1 and σ_2 the sample standard deviations, or we can replace $\widetilde{\gamma}_1$, $\widetilde{\gamma_2}$ and $\widetilde{\alpha}$ in the original log-likelihood to obtain the initial values for μ_1, μ_2, σ_1 and σ_2.

When n is large, we have that $\widehat{\boldsymbol{\theta}}_1 \overset{a}{\sim} \mathcal{N}_7(\boldsymbol{\theta}_1, \boldsymbol{J}(\boldsymbol{\theta}_1)^{-1})$, where "$\overset{a}{\sim}$" means approximately distributed, $\boldsymbol{J}(\boldsymbol{\theta}_1) = -E(\boldsymbol{H}(\boldsymbol{\theta}_1))$ is the expected Fisher information matrix for $\boldsymbol{\theta}_1$, and $\boldsymbol{H}(\boldsymbol{\theta}_1) = \partial^2\ell(\boldsymbol{\theta}_1)/\partial\boldsymbol{\theta}_1\partial\boldsymbol{\theta}_1'$ is the Hessian matrix, whose elements are provided in the Appendix A. Using the transformation method, we have that $\widehat{\boldsymbol{\theta}} \overset{a}{\sim} \mathcal{N}_7(\boldsymbol{\theta}, \boldsymbol{K}(\boldsymbol{\theta})^{-1})$, where $\boldsymbol{K}(\boldsymbol{\theta}) = \boldsymbol{D}'\boldsymbol{J}(\boldsymbol{\theta}_1)\boldsymbol{D}$ with $\boldsymbol{K}(\boldsymbol{\theta})^{-1}$ denoting its inverse, and

$$\boldsymbol{D} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \alpha & 0 & 0 & 0 & \dfrac{\alpha_1}{\alpha} \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \alpha & \dfrac{\alpha_2}{\alpha} \\ 0 & 0 & -\dfrac{\alpha}{\alpha_1} & 0 & 0 & -\dfrac{\alpha}{\alpha_2} & 1 \end{pmatrix}.$$

The elements of $\boldsymbol{J}(\boldsymbol{\theta}_1)$ are obtained in terms of numerical integration. On the other hand, the observed Fisher information matrix given by $-\boldsymbol{H}(\widehat{\boldsymbol{\theta}}_1)$ can also be used for computing asymptotic standard errors for the ML estimates of the model parameters (see Appendix A). After computing $-\boldsymbol{H}(\widehat{\boldsymbol{\theta}}_1)$, we obtain the observed Fisher information matrix under $\boldsymbol{\theta}$-parametrization as $-\widehat{\boldsymbol{D}}'\boldsymbol{H}(\widehat{\boldsymbol{\theta}}_1)\widehat{\boldsymbol{D}}$, where $\widehat{\boldsymbol{D}} := \boldsymbol{D}(\widehat{\alpha}_1, \widehat{\alpha}_2, \widehat{\alpha})$. It is worth emphisizing that the observed information matrix can be computed numerically from standard maximization routines, which now provide the observed information matrix as part of their output. For example, one can use the R functions `optim` or `nlm` to compute the observed Fisher information matrix numerically.

4.3. Monte Carlo Simulation

In the following, we consider Monte Carlo simulation experiments to evaluate the performance of the ML estimation procedure discussed in the previous section to estimate the BPHN distribution parameters. It is worth emphasizing that is rather easy to generate random variates from the BPHN distribution. We consider the following steps to generate (X_1, X_2) from BPHN$(\mu_1, \sigma_1, \alpha_1, \mu_2, \sigma_2, \alpha_2, \alpha)$ distribution:

Step 1. Generate two independent random variates $U_1 \sim \mathcal{U}(0,1)$ and $U_2 \sim \mathcal{U}(0,1)$, where $\mathcal{U}(0,1)$ means a uniform distribution on $(0,1)$.

Step 2. Compute

$$X_2 = \mu_2 + \sigma_2 \Phi^{-1}(1 - (1 - U_2)^{1/\alpha_2}),$$

where $\Phi^{-1}(\cdot)$ is the standard normal quantile function.

Step 3. Define $B_1 = [1 - \Phi(X_2)]^{-\alpha_2(\alpha+1)/\alpha}$ and $B_2 = [1 - \Phi(X_2)]^{-\alpha_2/\alpha} - 1$. Let $B_3 = [(B_1/U_1)^{1/(\alpha+1)} - B_2]^{-\alpha/\alpha_1}$. For each value of X_2, compute

$$X_1 = \mu_1 + \sigma_1 \Phi^{-1}(1 - B_3).$$

Then, $(X_1, X_2) \sim$ BPHN$(\mu_1, \sigma_1, \alpha_1, \mu_2, \sigma_2, \alpha_2, \alpha)$.

The Monte Carlo experiments were performed using the R programming language, and we consider 10,000 Monte Carlo replications. The performance of the ML procedure to estimate the BPHN distribution parameters was evaluated on the basis of the following quantities for each sample size: the empirical mean, and the empirical standard deviation, which are computed from 10,000 replications. The sample sizes we consider are $n = 90$ and 150. Without loss of generality, we consider $\mu_1 = 1.5$, $\sigma_1 = \sigma_2 = 1.0$, $\alpha_1 = 0.5$, $\mu_2 = -0.5$, $\alpha_2 = 1.2$, and $\alpha = 0.2$, 0.4, 0.6, 0.8, 1.0, 1.2 and 1.5. The emprical mean is listed in Table 1, while the empirical standard deviation is listed in Table 2. In general, from Table 1, note that the ML estimates of μ_1, σ_1, α_1, μ_2, σ_2 and α_2 are close to their respective true values, while the parameter that controls the dependence in the Pareto copula, α, is overestimated, mainly as this parameter increases. As expected, the standard deviation of the ML estimates of all parameters decreases as the sample size increases (see Table 2). In short, the Monte Carlo simulation experiments reveal that the ML method can be used effectively to estimate the BPHN parameters.

Table 1. Parameter estimates; empirical mean.

n	α	μ_1	σ_1	α_1	μ_2	σ_2	α_2	α
90	0.2	1.024	0.876	0.495	−0.378	1.034	1.539	0.321
	0.4	1.192	0.903	0.566	−0.403	1.008	1.371	0.632
	0.6	1.364	0.950	0.665	−0.483	0.986	1.271	0.941
	0.8	1.376	0.948	0.632	−0.555	0.965	1.176	1.256
	0.8	1.419	0.959	0.651	−0.576	0.959	1.178	1.553
	1.0	1.406	0.953	0.619	−0.591	0.956	1.145	1.821
	1.5	1.426	0.962	0.613	−0.612	0.949	1.121	2.303
150	0.2	0.990	0.864	0.471	−0.339	1.053	1.599	0.318
	0.4	1.172	0.898	0.540	−0.395	1.016	1.362	0.625
	0.6	1.319	0.936	0.608	−0.487	0.991	1.248	0.931
	0.8	1.391	0.955	0.635	−0.550	0.974	1.167	1.224
	1.0	1.427	0.963	0.643	−0.543	0.978	1.182	1.523
	1.2	1.430	0.963	0.628	−0.552	0.976	1.178	1.805
	1.5	1.458	0.972	0.627	−0.601	0.960	1.125	2.192

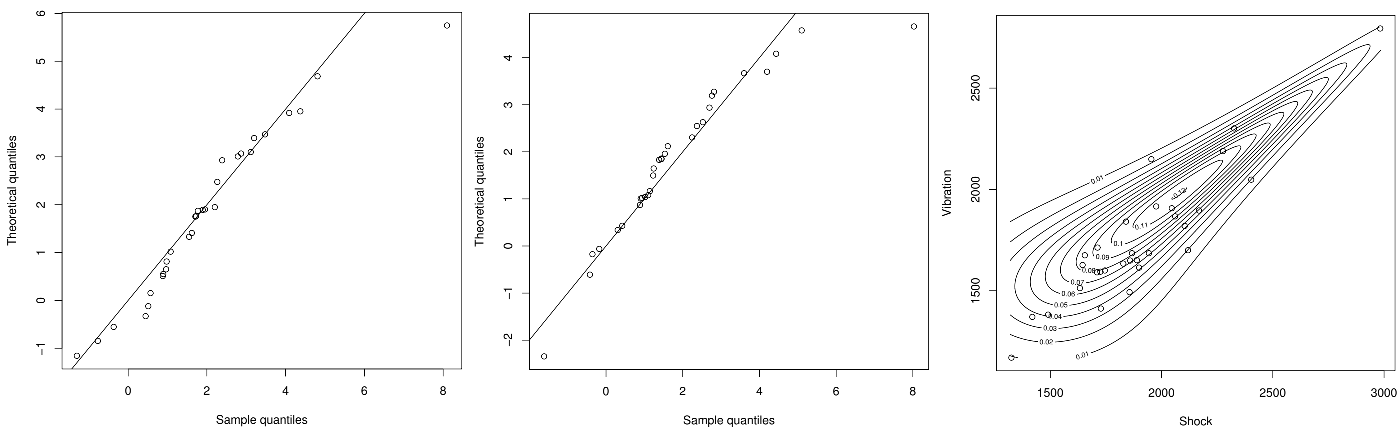
Table 2. Parameter estimates; empirical standard deviation.

n	α	μ_1	σ_1	α_1	μ_2	σ_2	α_2	α
90	0.2	0.251	0.117	0.158	0.195	0.081	0.326	0.037
	0.4	0.370	0.152	0.259	0.273	0.102	0.416	0.089
	0.6	0.460	0.175	0.389	0.314	0.117	0.452	0.153
	0.8	0.426	0.169	0.319	0.315	0.117	0.412	0.248
	1.0	0.480	0.178	0.372	0.387	0.137	0.533	0.348
	1.2	0.470	0.180	0.329	0.368	0.135	0.465	0.387
	1.5	0.463	0.181	0.326	0.364	0.124	0.483	0.620
150	0.2	0.236	0.103	0.147	0.193	0.072	0.326	0.028
	0.4	0.289	0.120	0.202	0.207	0.075	0.312	0.065
	0.6	0.348	0.137	0.255	0.263	0.094	0.359	0.113
	0.8	0.384	0.149	0.295	0.283	0.100	0.359	0.165
	1.0	0.420	0.156	0.317	0.293	0.101	0.378	0.223
	1.2	0.431	0.164	0.311	0.316	0.109	0.411	0.290
	1.5	0.432	0.161	0.311	0.352	0.119	0.428	0.368

5. Real Data Application

Next, we consider an application of the BPHN distribution to real data for illustrative purposes. The data set contains two different measures of stiffness of each of the 30 boards. The considered stiffness measures are the Shock (X_1) and Vibration (X_2) of each of the 30 boards. The data were reported by [17]. The first measurement involves sending a shock wave down the board, and the second measurement is determined while vibrating the board. All numerical computations provided in this were done by using the R program.

We assume that $(X_{1i}, X_{2i}) \sim \text{BPHN}(\mu_1, \sigma_1, \alpha_1, \mu_2, \sigma_2, \alpha_2, \alpha)$ for $i = 1, 2, \ldots, 30$. To estimate the BPHN model parameters, we initially estimate the marginal PDFs corresponding to each variable. We thus obtain the following estimates: $\widetilde{\mu}_1 = 1554.82$, $\widetilde{\sigma}_1 = 176.34$, $\widetilde{\alpha}_1 = 0.18$, $\widetilde{\mu}_2 = 1393.59$, $\widetilde{\sigma}_2 = 173.79$, and $\widetilde{\alpha}_2 = 0.17$. Figure 3a,b display the QQ-plots for the estimated marginal distributions. Note the goodness-of-fit of each estimated marginal distribution to the real data, which means that the BPHN distribution may be a good choice for modeling these data. Now, using the above estimates in the copula function, we obtain an estimate for the parameter α, that is, $\widetilde{\alpha} = 0.35$. Finally, taking the above estimates as initial values for the iterative process, we obtain the following ML estimates (standard errors in parentheses): $\widehat{\mu}_1 = 1529.91(184.47)$, $\widehat{\sigma}_1 = 277.53(48.13)$, $\widehat{\alpha}_1 = 0.28(0.17)$, $\widehat{\mu}_2 = 1428.44(87.82)$, $\widehat{\sigma}_2 = 184.49(33.33)$, $\widehat{\alpha}_2 = 0.16(0.05)$, and $\widehat{\alpha} = 0.27(0.12)$. Visual inspection of the contour plot in Figure 3c confirms a satisfactory fit of the BPHN distribution to the data.



(**a**) (**b**) (**c**)

Figure 3. QQ-plots and contour plot: (**a**) Shock, (**b**) Vibration, and (**c**) contour plot of the BPHN distribution.

Next, for the sake of comparison, we also consider the following bivariate distributions to fit the current real data: the bivariate skew-normal conditional (BVSNC) distribution introduced by [9], denoted by BVSNC$(\xi_1, \eta_1, \xi_2, \eta_2, \lambda)$, and the bivariate Birnbaum-Saunders (BVBS) distribution studied by [10], denoted by BVBS$(\alpha_1, \beta_1, \alpha_2, \beta_2, \lambda)$. The ML estimates (standard errors in parentheses) of the BVSNC parameters are $\widehat{\xi}_1 = 1927.51(26.65)$, $\widehat{\xi}_2 = 1745.35(30.38)$, $\widehat{\eta}_1 = 320.31(41.40)$, $\widehat{\eta}_2 = 313.23(40.43)$ and $\widehat{\lambda} = 10.88(5.01)$, whereas the ML estimates (standard errors in parentheses) of the BVBS parameters are $\widehat{\alpha}_1 = 0.1621(0.0210)$, $\widehat{\alpha}_2 = 0.1701(0.0219)$, $\widehat{\beta}_1 = 1917.22(22.26)$, $\widehat{\beta}_2 = 1734.76(29.32)$ and $\widehat{\lambda} = 10.29(4.86)$. To compare the bivariate distributions, we make use of the Akaike information criterion (AIC), and Bayesian information criterion (BIC). The smaller the values of AIC and BIC, the better the fitted distribution to the real data. Table 3 lists the AIC and BIC values for all bivariate distributions, which leads to the conclusion that the BPHN distribution is better than the other bivariate distributions to model the current bivariate real data.

Table 3. AIC and BIC values.

Measure	BVSNC	BVBS	BPHN
AIC	842.99	835.82	828.27
BIC	850.00	842.83	838.08

6. Concluding Remarks

The univariate proportional hazard distribution has found several applications in the literature and has many attractive properties. However, the extension of the univariate proportional hazard distribution in a multivariate setting has been so neglected in the statistical literature. In this paper, based on the Pareto copula, we have introduced a simple and tractable bivariate proportional hazard family of distributions that may be very useful to deal with bivariate data in practice. We also derive several distribution properties for this bivariate family of distributions. In addition, we particularize this general bivariate distribution to the case where the univariate normal distribution is considered, so that the bivariate proportional hazard normal distribution is obtained. The estimation of the bivariate proportional hazard normal model parameters is approached by the maximum likelihood method, and the observed and expected Fisher information matrixes are derived. Finally, an application to real data set is presented to illustrate the bivariate proportional hazard normal distribution in practice.

Author Contributions: Conceptualization, G.M.-F., A.J.L. and C.B.-C.; data curation, G.M.-F., A.J.L. and C.B.-C.; formal analysis, G.M.-F., A.J.L. and C.B.-C.; funding acquisition, C.B.-C.; investigation, G.M.-F. and A.J.L. and C.B.-C.; methodology, G.M.-F., A.J.L. and C.B.-C.; project administration, G.M.-F., A.J.L. and C.B.-C.; resources, G.M.-F., A.J.L. and C.B.-C.; supervision, G.M.-F., A.J.L. and C.B.-C.; visualization, C.B.-C.; writing—original draft, G.M.-F. and A.J.L. and C.B.-C.; writing—review and editing, A.J.L. and C.B.-C. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Instituto Tecnológico Metropolitano (ITM).

Data Availability Statement: Not applicable.

Acknowledgments: G.M.-F. acknowledges the support given by Universidad de Córdoba, Montería, Colombia. A.J.L. acknowledges the financial support of the Brazilian agency *Conselho Nacional de Desenvolvimento Científico e Tecnológico* (CNPq grant 304776/2019–0). C.B.-C. extends their sincere gratitude to the Instituto Tecnológico Metropolitano (ITM). The authors would like to thank the anonymous reviewers for their insightful comments and suggestions.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Observed Fisher Information Matrix

In the following, we calculate the elements of the Hessian matrix $\boldsymbol{H}(\boldsymbol{\theta}_1)$ and their respective expected values. For $j=1,2$ and $i=1,2,\ldots,n$, let

$$f_{ij}=\frac{\gamma_j}{\sigma_j}\phi(z_{ij})(1-\Phi(z_{ij}))^{-\gamma_j-1},\quad W_{ij}=\frac{\phi(z_{ij})}{1-\Phi(z_{ij})},$$

$$DN_i=(1-\Phi(z_{i1}))^{-\gamma_1}+(1-\Phi(z_{i2}))^{-\gamma_2}-1,$$
$$H_{ij}=(1-\Phi(z_{ij}))^{-\gamma_j}\log((1-\Phi(z_{ij}))).$$

We have the following derivatives ($j'=1,2$):

$$\frac{\partial^2\ell(\boldsymbol{\theta}_1)}{\partial\alpha^2}=-\frac{n}{(\alpha+1)^2}-\frac{n}{\alpha^2},\quad \frac{\partial^2\ell(\boldsymbol{\theta}_1)}{\partial\alpha\partial\gamma_j}=\sum_{i=1}^n\frac{H_{ij}}{DN_i},$$

$$\frac{\partial^2\ell(\boldsymbol{\theta}_1)}{\partial\alpha\partial\mu_j}=\sum_{i=1}^n\frac{f_{ij}}{DN_i},\quad \frac{\partial^2\ell(\boldsymbol{\theta}_1)}{\partial\alpha\partial\sigma_j}=\sum_{i=1}^n\frac{z_{ij}f_{ij}}{DN_i},$$

$$\frac{\partial^2\ell(\boldsymbol{\theta}_1)}{\partial\gamma_j^2}=-\frac{n}{\gamma_j^2}+(\alpha+2)\sum_{i=1}^n[H_{ij}-\log(1-\Phi(z_{ij}))]\frac{H_{ij}}{DN_i},$$

$$\frac{\partial^2\ell(\boldsymbol{\theta}_1)}{\partial\gamma_1\partial\gamma_2}=(\alpha+2)\sum_{i=1}^n\prod_{j=1}^2\frac{H_{ij}}{DN_i^2},$$

$$\begin{aligned}\frac{\partial^2\ell(\boldsymbol{\theta}_1)}{\partial\mu_j\partial\gamma_j}&=-\frac{1}{\sigma_j}\sum_{i=1}^n W_{ij}\\&\quad-(\alpha+2)\sum_{i=1}^n\frac{f_{ij}}{DN_i}\left[\log(1-\Phi(z_{ij}))-\frac{H_{ij}}{DN_i}-\frac{1}{\gamma_j}\right],\end{aligned}$$

$$\begin{aligned}\frac{\partial^2\ell(\boldsymbol{\theta}_1)}{\partial\sigma_j\partial\gamma_j}&=-\frac{1}{\sigma_j}\sum_{i=1}^n z_{ij}W_{ij}\\&\quad-(\alpha+2)\sum_{i=1}^n\frac{z_{ij}f_{ij}}{DN_i}\left[\log(1-\Phi(z_{ij}))-\frac{H_{ij}}{DN_i}-\frac{1}{\gamma_j}\right],\end{aligned}$$

$$\frac{\partial^2\ell(\boldsymbol{\theta}_1)}{\partial\mu_{j'}\partial\mu_j}=-(\alpha+2)\sum_{i=1}^n\frac{f_{ij}f_{ij'}}{DN_i^2},\quad j\neq j',$$

$$\frac{\partial^2\ell(\boldsymbol{\theta}_1)}{\partial\mu_{j'}\partial\gamma_j}=-(\alpha+2)\sum_{i=1}^n\frac{H_{ij}f_{ij'}}{DN_i^2},\quad j\neq j',$$

$$\frac{\partial^2\ell(\boldsymbol{\theta}_1)}{\partial\sigma_{j'}\partial\gamma_j}=-(\alpha+2)\sum_{i=1}^n\frac{z_{ij'}H_{ij}f_{ij'}}{DN_i^2},\quad j\neq j',$$

$$\begin{aligned}\frac{\partial^2\ell(\boldsymbol{\theta}_1)}{\partial\mu_j^2}&=-\frac{n}{\sigma_j^2}-\frac{\gamma_j}{\sigma_j^2}\sum_{i=1}^n[z_{ij}W_{ij}+W_{ij}^2]\\&\quad+\frac{\alpha+2}{\sigma_j}\sum_{i=1}^n\frac{f_{ij}}{DN_i}\left[z_{ij}-(\gamma_j+1)W_{ij}+\frac{\sigma_j}{\gamma_j}\frac{f_{ij}}{DN_i}\right],\end{aligned}$$

$$\frac{\partial^2\ell(\boldsymbol{\theta}_1)}{\partial\mu_j\partial\sigma_{j'}}=-(\alpha+2)\sum_{i=1}^n\frac{z_{ij'}f_{ij}f_{ij'}}{DN_i^2},\quad j\neq j',$$

$$\frac{\partial^2\ell(\boldsymbol{\theta}_1)}{\partial\mu_j\partial\sigma_j} = -\frac{2}{\sigma_j^2}\sum_{i=1}^n z_{ij} + \frac{\alpha+2}{\sigma_j}\sum_{i=1}^n \frac{f_{ij}}{DN_i}\left[z_{ij}^2 - 1 - (\gamma_j+1)z_{ij}W_{ij} + \frac{\sigma_j z_{ij} f_{ij}}{DN_i}\right] - \frac{\gamma_j}{\sigma_j}\sum_{i=1}^n [(z_{ij}^2 - z_{ij}W_{ij} - 1)W_{ij}],$$

$$\frac{\partial^2\ell(\boldsymbol{\theta}_1)}{\partial\sigma_j\partial\sigma_{j'}} = -(\alpha+2)\sum_{i=1}^n \frac{z_{ij}z_{ij'}f_{ij}f_{ij'}}{DN_i^2}, \quad j \neq j',$$

$$\frac{\partial^2\ell(\boldsymbol{\theta}_1)}{\partial\sigma_j^2} = \frac{n}{\sigma_j} - \frac{3}{\sigma_j^2}\sum_{i=1}^n z_{ij}^2 + \frac{\alpha+2}{\sigma_j}\sum_{i=1}^n \frac{z_{ij}f_{ij}}{DN_i}\left[z_{ij}^2 - 2 - \frac{(\gamma_j+1)}{\gamma_j}f_{ij} + \frac{z_{ij}f_{ij}}{DN_i}\right] - \frac{\gamma_j}{\sigma_j}\sum_{i=1}^n [(z_{ij}^2 + z_{ij}W_{ij} - 1)z_{ij}W_{ij}].$$

Let

$$a_{kj} = E\left[z^k\left(\frac{\phi(z)}{1-\Phi(z)}\right)^j\right],$$

and $g(u) = \Phi^{-1}(1 - u^{1/\gamma_j})$, where $\Phi^{-1}(\cdot)$ is the quantile function of the standard normal distribution. We have the expected values:

$$E\left(\frac{\partial^2\ell(\boldsymbol{\theta}_1)}{\partial\gamma_j^2}\right) = -\frac{n}{\gamma_j^2}\left[1 + \frac{2(\alpha+1)}{(\alpha+3)\alpha^2}\right],$$

$$E\left(\frac{\partial^2\ell(\boldsymbol{\theta}_1)}{\partial\alpha\partial\gamma_j}\right) = -\frac{n(\alpha+1)}{\alpha\gamma_j},$$

$$E\left(\frac{\partial^2\ell(\boldsymbol{\theta}_1)}{\partial\gamma_j\partial\gamma_{j'}}\right) = -n\alpha(\alpha+1)(\alpha+2)\int_1^\infty\int_1^0 \frac{uv\log(u)\log(v)}{(u+v-1)^{\alpha+4}}dudv, \quad j \neq j',$$

$$E\left(\frac{\partial^2\ell(\boldsymbol{\theta}_1)}{\partial\alpha^2}\right) = \frac{n}{\alpha^2} + \frac{n}{(1+\alpha)^2},$$

$$E\left(\frac{\partial^2\ell(\boldsymbol{\theta}_1)}{\partial\alpha\partial\mu_j}\right) = -\frac{n\alpha(\alpha+1)\gamma_j}{\sigma_j(\alpha+2)}\int_1^\infty \phi(g(u))u^{\frac{1}{\gamma_j}-\alpha-1}du,$$

$$E\left(\frac{\partial^2\ell(\boldsymbol{\theta}_1)}{\partial\alpha\partial\sigma_j}\right) = -\frac{n\alpha(\alpha+1)\gamma_j}{\sigma_j(\alpha+2)}\int_1^\infty g(u)\phi(g(u))u^{\frac{1}{\gamma_j}-\alpha-1}du,$$

$$E\left(\frac{\partial^2\ell(\boldsymbol{\theta}_1)}{\partial\mu_j\partial\gamma_j}\right) = -\frac{na_{01}}{\sigma_j} + \frac{n\alpha(\alpha+1)}{\sigma_j}\int_1^\infty \phi(g(u))\log(u)u^{\frac{1}{\gamma_j}-\alpha-1}du + \frac{n}{\sigma_j(\alpha+3)}\int_1^\infty \phi(g(u))\log(u)u^{\frac{1}{\gamma_j}-\alpha-1}du,$$

$$E\left(\frac{\partial^2\ell(\boldsymbol{\theta}_1)}{\partial\sigma_j\partial\gamma_j}\right) = -\frac{na_{11}}{\sigma_j} + \frac{n\alpha(\alpha+1)}{\sigma_j}\int_1^\infty g(u)\phi(g(u))\log(u)u^{\frac{1}{\gamma_j}-\alpha-1}du,$$

$$E\left(\frac{\partial^2 \ell(\boldsymbol{\theta}_1)}{\partial \mu_{j'} \partial \gamma_j}\right) = -\frac{n(\alpha+2)(\alpha+1)}{\sigma_j} \int_1^\infty \int_1^\infty \frac{u\phi(g(u)) \log(u) v^{\frac{1}{\gamma_j}+1}}{(u+v-1)^{\alpha+4}} dudv,$$

$$E\left(\frac{\partial^2 \ell(\boldsymbol{\theta}_1)}{\partial \sigma_{j'} \partial \gamma_j}\right) = -\frac{n\alpha(\alpha+1)(\alpha+2)}{\sigma_{j'}} \times \int_1^\infty \int_1^\infty \frac{ug(u)\phi(g(u)) \log(u) v^{\frac{1}{\gamma_j}+1}}{(u+v-1)^{\alpha+4}} dudv, \quad j \neq j',$$

$$E\left(\frac{\partial^2 \ell(\boldsymbol{\theta}_1)}{\partial \mu_j \partial \mu_{j'}}\right) = -\frac{n\alpha(\alpha+1)(\alpha+2)\gamma_j\gamma_{j'}}{\sqrt{2\pi}\sigma_j\sigma_{j'}} \times \int_1^\infty \int_1^\infty \frac{\phi(g(u)+g(v)) u^{\frac{1}{\gamma_j}+1}}{(u+v-1)^{\alpha+4}} dudv, \quad j \neq j',$$

$$E\left(\frac{\partial^2 \ell(\boldsymbol{\theta}_1)}{\partial \mu_j \partial \sigma_{j'}}\right) = -\frac{n\alpha(\alpha+1)(\alpha+2)\gamma_j\gamma_{j'}}{\sqrt{2\pi}\sigma_j\sigma_{j'}} \times \int_1^\infty \int_1^\infty \frac{g(v)\phi(g(u)+g(v)) u^{\frac{1}{\gamma_j}+1} v^{\frac{1}{\gamma_j}+1}}{(u+v-1)^{\alpha+2}} dudv, \quad j \neq j',$$

$$E\left(\frac{\partial^2 \ell(\boldsymbol{\theta}_1)}{\partial \sigma_j \partial \sigma_{j'}}\right) = -\frac{n\alpha(\alpha+1)(\alpha+2)\gamma_j\gamma_{j'}}{\sigma_j\sigma_{j'}} \times \int_1^\infty \int_1^\infty \frac{g(u)g(v)\phi(g(u)+g(v)) u^{\frac{1}{\gamma_j}+1} v^{\frac{1}{\gamma_j}+1}}{(u+v-1)^{\alpha+4}} dudv, \quad j \neq j',$$

$$E\left(\frac{\partial^2 \ell(\boldsymbol{\theta}_1)}{\partial \mu_j^2}\right) = -\frac{n}{\sigma_j^2} - \frac{n\gamma_j[a_{11}+a_{02}]}{\sigma_j^2} + \frac{n\alpha(\alpha+2)(\alpha+1)}{\sigma_j} \times \int_1^\infty \int_1^\infty \frac{\phi(g(u))}{(u+v-1)^{\alpha+3}} \left[g(u) - (\gamma_j+1)\phi(g(u)) u^{\frac{1}{\gamma_j}} + \frac{\phi(g(u)) u^{1+\frac{1}{\gamma_j}}}{(u+v-1)}\right] dudv,$$

$$
\begin{aligned}
E\left(\frac{\partial^2 \ell(\boldsymbol{\theta}_1)}{\partial \sigma_j^2}\right) = {} & \frac{n}{\sigma_j^2} - \frac{3na_{20}}{\sigma_j^2} - \frac{n\gamma_j}{\sigma_j^2}[a_{21} - a_{11} + a_{22}] \\
& - \frac{\alpha(\alpha+1)(\alpha+2)\gamma_j}{\sigma_j^2} \int_1^\infty \int_1^\infty \frac{g(u)\phi(g(u))u^{\frac{1}{\gamma_j}+1}}{(u+v-1)^{\alpha+3}} dudv \\
& + \frac{n\alpha(\alpha+1)(\alpha+2)}{\sigma_j} \times \\
& \int_1^\infty \int_1^\infty \left[\frac{\sqrt{2\pi}\gamma_j}{\sigma_j^2} \frac{\phi(g(u))u^{\frac{1}{\gamma_j}+1}}{(u+v-1)^{\alpha+3}} [-1 + (g(u))^2 + (\gamma_j+1)\sqrt{2\pi}\phi(g(u))] \right. \\
& \left. + \frac{\gamma_j^2 (g(u))^2 \phi(g(u)) u^{\frac{2}{\gamma_j}+2}}{\sqrt{2\pi}\sigma_j^2 (u+v-1)^2} \right] dudv.
\end{aligned}
$$

References

1. Durrans, S.R. Distributions of fractional order statistics in hydrology. *Water Resour. Res.* **1992**, *28*, 1649–1655. [CrossRef]
2. Martínez-Flórez, G.; Moreno-Arenas, G.; Vergara-Cardoso, S. Properties and Inference for Proportional Hazard Models. *Colomb. J. Stat.* **2013**, *36*, 95–114.
3. Wang, B.; Yu, K.; Jones, M. Inference under progressively type II right-censored sampling for certain lifetime distributions. *Technometrics* **2010**, *52*, 453–460. [CrossRef]
4. Cordeiro, G.M.; de Castro, M. A new family of generalized distributions. *J. Stat. Comput. Simul.* **2011**, *81*, 883–898. [CrossRef]
5. Kundu, D.; Gupta, R. Power normal distribution. *Statistics* **2013**, *47*, 110–125. [CrossRef]
6. Clayton, D.G. A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. *Biometrika* **1978**, *65*, 141–152. [CrossRef]
7. Kimeldorf, G.; Sampson, A.R. Uniform representations of bivariate distributions. *Commun. Stat. Theory Methods* **1975**, *4*, 617–627. [CrossRef]
8. Cook, R.D.; Johnson, M.E. A family of distributions for modelling non-elliptically symmetric multivariate data. *J. R. Stat. Soc. B* **1981**, *43*, 210–218. [CrossRef]
9. Arnold, B.C.; Castilho, E.; Sarabia, J.M. Conditionally specified multivariate skewed distributions. *Sankhya A* **2002**, *64*, 206–226.
10. Lemonte, A.J.; Martínez-Flórez, G.; Moreno-Arenas, G. Multivariate Birnbaum–Saunders distribution: Properties and associated inference. *J. Stat. Comput. Simul.* **2015**, *85*, 374–392. [CrossRef]
11. Martínez-Flórez, G.; Lemonte, A.J.; Salinas, H.S. Multivariate skew-power-normal distributions: Properties and associated inference. *Symmetry* **2019**, *11*, 1509. [CrossRef]
12. Zhang, L.; Xu, A.; An, L.; Li, M. Bayesian inference of system reliability for multicomponent stress-strength model under Marshall-Olkin Weibull distribution. *Systems* **2022**, *10*, 196. [CrossRef]
13. Nelsen, R.B. *An Introduction to Copulas*; Springer: New York, NY, USA, 2006.
14. Marshall, A.W. Some comments on the hazard gradient. *Stoch. Process. Their Appl.* **1975**, *3*, 293–300. [CrossRef]
15. Joe, H. Asymptotic efficiency of two-stage estimation method for copula-based models. *J. Multivar. Anal.* **2005**, *94*, 401–419. [CrossRef]
16. Gupta, D.; Gupta, R.C. Analyzing skewed data by power normal model. *Test* **2008**, *17*, 197–210. [CrossRef]
17. Johnson, R.A.; Wichern, D.W. *Applied Multivariate Statistical Analysis*; Pearson Prentice Hall: Upper Saddle River, NJ, USA, 2007.