Chilean Journal of Statistics
Vol. 4, No. 2, September 2013, 3–26

Distribution Theory
Research Paper

The beta Weibull Poisson distribution

Ana Percontini[1]*, Betsabé Blas[2], and Gauss M. Cordeiro[3]

[1]Departamento de Ciências Exatas, Universidade Estadual de Feira de Santana, Feira de Santana, Brazil,
[2]Departamento de Estatística, Universidade Federal de Pernambuco, Recife, Brazil,
[3]Departamento de Estatística, Universidade Federal de Pernambuco, Recife, Brazil

(Received: 20 November 2012 · Accepted in final form: 22 March 2013)

Abstract

Providing a wider distribution is always precious for statisticians. A new five-parameter distribution called the beta Weibull Poisson is proposed, which is obtained by compounding the Weibull Poisson and beta distributions. It generalizes several known lifetime models. We obtain some properties of the proposed distribution such as the survival and hazard rate functions, quantile function, ordinary and incomplete moments, order statistics and Rényi entropy. Estimation by maximum likelihood and inference for large samples are addressed. The potentiality of the new model is shown by means of a real data set. In fact, the proposed model can produce better fits than some well-known distributions.

Keywords: Beta distribution · Generating function · Lifetime data · Mean Deviation · Moment · Quantile function.

Mathematics Subject Classification: Primary 60E05 · Secondary 62P99.

1. Introduction

The Weibull distribution is a very popular model in reliability and it has been widely used for analyzing lifetime data. Several new models have been proposed that are either derived from or in some way are related to the Weibull distribution. When modelling monotone hazard rates, the Weibull distribution may be an initial choice because of its negatively and positively skewed density shapes. However, it does not provide a reasonable parametric fit for modelling phenomenon with non-monotone failure rates such as the bathtub shaped and the unimodal failure rates that are common in reliability and biological studies. An example of the bathtub-shaped failure rate is the human mortality experience with a high infant mortality rate which reduces rapidly to reach a low level. It then remains at that level for quite a few years before picking up again. Unimodal failure rates can be observed

*Corresponding author. Email: anappaixao@gmail.com

ISSN: 0718-7912 (print)/ISSN: 0718-7920 (online)

http://www.soche.cl/chjs

in course of a disease whose mortality reaches a peak after some finite period and then declines gradually.

The statistics literature is filled with hundreds of continuous univariate distributions. Recent developments focus on new techniques for building meaningful distributions. Several methods of introducing one or more parameters to generate new distributions have been studied in the statistical literature recently. Among these methods, the compounding of some discrete and important lifetime distributions has been in the vanguard of lifetime modeling. So, several families of distributions were proposed by compounding some useful lifetime and truncated discrete distributions.

In recent years, there has been a great interest among statisticians and applied researchers in constructing flexible distributions to facilitate better modeling of lifetime data. Several authors introduced more flexible distributions to model monotone or unimodal failure rates but they are not useful for modelling bathtub-shaped failure rates. Adamidis and Loukas (1998) proposed the exponential geometric (EG) distribution to model lifetime data with decreasing failure rate function and Gupta and Kundu (1999, 2001a,b) defined the generalized exponential (GE) (also called the exponentiated exponential) distribution. The last distribution has only increasing or decreasing failure rate function. Following the key idea of Adamidis and Loukas (1998), Kus (2007) introduced the exponential Poisson (EP) distribution which has a monotone failure rate. Lee et al. (2007) proposed a generalization of the Weibull distribution called the beta Weibull (BW) distribution. Barreto-Souza et al. (2010) studied a Weibull geometric (WG) distribution which extends the EG and Weibull distributions. In this paper, we propose a new compounding distribution, called the *beta Weibull Poisson* (BWP) distribution, by compounding the beta and Weibull Poisson (WP) distributions (Lu and Shi, 2012). The failure rate function of the WP distribution has various shapes. In fact, it can be increasing, decreasing, upside-down bathtub-shaped or unimodal.

The proposed generalization stems from a general class of distributions which is defined by the following cumulative distribution function (cdf)

$$F(x) = I_{G(x)}(a,b) = \frac{1}{B(a,b)} \int_0^{G(x)} w^{a-1}(1-w)^{b-1}dw, \qquad (1)$$

where $a > 0$ and $b > 0$ are two additional shape parameters to the parameters of the G-distribution, $B(a,b) = \Gamma(a)\Gamma(b)/\Gamma(a+b)$ is the beta function and $I_{G(x)}(a,b)$ denotes the incomplete beta function ratio evaluated at $G(x)$. The parameters a and b govern both the skewness and kurtosis of the generated distribution.

This class was proposed by Eugene et al. (2002) and has been widely used ever since. For example, Eugene et al. (2002) introduced the beta normal (BN) distribution, Nadarajah and Kotz (2004) defined the beta Gumbel (BGu) distribution and Nadarajah and Gupta (2004) proposed the beta Fréchet (BF) distribution. Another example is the beta exponential (BE) model studied by Nadarajah and Kotz (2006).

The probability density function (pdf) corresponding to (1) is given by

$$f(x) = \frac{g(x)}{B(a,b)} G(x)^{a-1}\{1 - G(x)\}^{b-1}, \qquad (2)$$

where $g(x) = dG(x)/dx$ is the baseline density function.

The paper is organized as follows. In Section 2, we define the BWP distribution and highlight some special cases. In Section 3, we demonstrate that the new density function is a linear combination of WP density functions. The proof is given in Appendix A. Also, we derive the survival and hazard rate functions, moments and moment generating function (mgf), order statistics and their moments and Rényi entropy. Maximum likelihood estimation of the model parameters and the observed information matrix are discussed in Section 4. In Section 5, we provide an application of the BWP model to the maintenance data with 46 observations reported on active repair times (hours) for an airborne communication transceiver. Concluding remarks are given in Section 6. Unless otherwise indicated, all results presented in the paper are new and original. It is expected that they could encourage further research of the new model.

2. The BWP distribution

We assume that Z has a truncated Poisson distribution with parameter $\lambda > 0$ and probability mass function given by

$$p(z;\lambda) = \mathrm{e}^{-\lambda}\lambda^{z}\,\Gamma^{-1}(z+1)(1-\mathrm{e}^{-\lambda})^{-1},\ z = 1,2,\ldots,$$

where $\Gamma(p) = \int_0^p x^{p-1}\mathrm{e}^{-x}dx$ (for $p > 0$) is the gamma function.

We define $\{W_i\}_{i=1}^{Z}$ to be independent and identically distributed random variable having the Weibull density function defined by

$$\pi(w;\alpha,\beta) = \alpha\beta w^{\alpha-1}\exp(-\beta w^{\alpha}), w > 0,$$

where $\alpha > 0$ is the shape parameter and $\beta > 0$ is the scale parameter.

We define $X = min\{W_1,\ldots,W_Z\}$, where the random variables Z and W's are assumed independent. The WP distribution of X has density function given by

$$g(x;\alpha,\beta,\lambda) = \dot{c}\,u\,x^{\alpha-1}\,\mathrm{e}^{\lambda u},\quad x > 0, \tag{3}$$

where $\dot{c} = \dot{c}(\alpha,\beta,\lambda) = \dfrac{\alpha\beta\lambda\mathrm{e}^{-\lambda}}{1-\mathrm{e}^{-\lambda}}$ and $u = \mathrm{e}^{-\beta x^{\alpha}}$.

The WP model is well-motivated for industrial applications and biological studies. As a first example, consider the time to relapse of cancer under the first-activation scheme. Suppose that the number, say Z, of carcinogenic cells for an individual left active after the initial treatment follows a truncated Poisson distribution and let W_i be the time spent for the ith carcinogenic cell to produce a detectable cancer mass, for $i \geq 1$. If $\{W_i\}_{i\geq 1}$ is a sequence of independent and identically distributed (iid) Weibull random variables independent of Z, then the time to relapse of cancer of a susceptible individual can be modeled by the WP distribution. Another example considers that the failure of a device occurs due to the presence of an unknown number, say Z, of initial defects of the same kind, which can be identifiable only after causing failure and are repaired perfectly. Define by W_i the time to the failure of the device due to the ith defect, for $i \geq 1$. If we assume that the W_i's are iid Weibull random variables independent of Z, which is a truncated Poisson random variable, then the time to the first failure is appropriately modeled by the WP distribution. For reliability studies, the proposed models for $X = \min\{W_i\}_{i=1}^{Z}$ and $T = \max\{W_i\}_{i=1}^{Z}$ can be used in serial and parallel systems with identical components, which appear in many industrial applications and biological organisms. The first activation scheme may be questioned by certain diseases. Consider that the number Z of latent factors

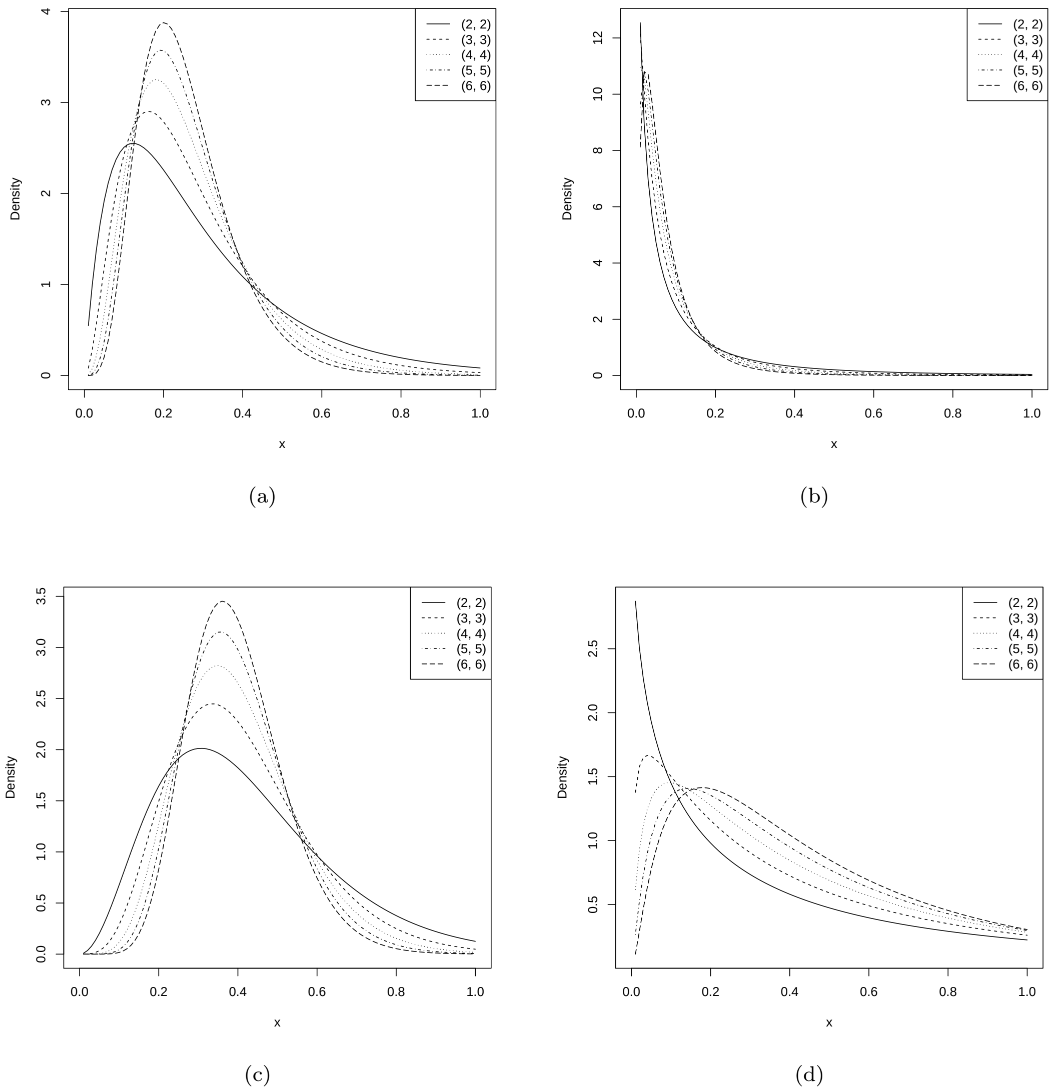

Figure 1. Plots of the BWP density function for: (a) $\alpha = 1$, $\beta = 2$ and $\lambda = 1$, (b) $\alpha = 0.5$, $\beta = 2$ and $\lambda = 1$, (c) $\alpha = 1.5$, $\beta = 2$ and $\lambda = 1$, (d) $\alpha = 0.5$, $\beta = 0.5$ and $\lambda = 2$.

that must all be activated by failure follows a truncated Poisson distribution and assume that W represents the time of resistance to a disease manifestation due to the ith latent factor has the Weibull distribution. In the last-activation scheme, the failure occurs after all Z factors have been activated. So, the WP distribution is able for modeling the time to the failure under last-activation scheme.

The cdf corresponding to (3) is

$$G(x) = \frac{\mathrm{e}^{\lambda u} - \mathrm{e}^{\lambda}}{1 - \mathrm{e}^{\lambda}}, \quad x > 0. \tag{4}$$

The BWP density function is obtained by inserting (3) and (4) in equation (2). It is given by

$$f(x) = c\, u\, x^{\alpha-1} \mathrm{e}^{\lambda u} (\mathrm{e}^{\lambda} - \mathrm{e}^{\lambda u})^{a-1} (\mathrm{e}^{\lambda u} - 1)^{b-1}, \tag{5}$$

where

$$c = \frac{\alpha\beta\lambda\, e^{-\lambda}(e^{\lambda} - 1)^{2-a-b}}{B(a,b)(1 - e^{-\lambda})}.$$

Hereafter, a random variable X having density function (5) is denoted by $X \sim \mathrm{BWP}(\alpha, \beta, \lambda, a, b)$.

The cumulative distribution of X is given by

$$F(x) = I_{G(x)}(a,b) = I_{(e^{\lambda u} - e^{\lambda})/(1-e^{\lambda})}(a,b). \tag{6}$$

We are motivated to study the BWP distributions because of the wide usage of the Weibull and the fact that the current generalization provides means of its continuous extension to still more complex situations. A second positive point of the current generalization is that the WP distribution is a basic exemplar of the proposed family. A third positive point is the the role played by the two beta generator parameters to the WP model. They can add more flexibility in the density function (5) by imposing more dispersion in the skewness and kurtosis of X and to control the tail weights.

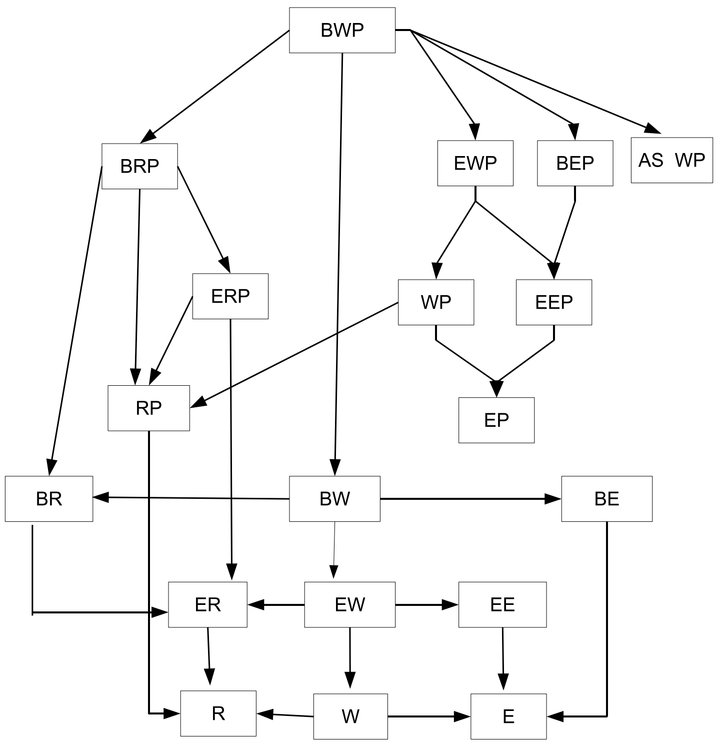

Figure 2. Relationships of the BWP sub-models.

The beta exponential Poisson (BEP) distribution is obtained from (5) when $\alpha = 1$. For $b = 1$, the exponentiated Weibull Poisson (EWP) distribution comes as a special model. In addition, for $\alpha = 1$, we obtain the exponentiated exponential Poisson (EEP) distribution.

On the other hand, if $\alpha = 2$, the beta Rayleigh Poisson (BRP) distribution is obtained. In addition, for $b = 1$, it follows the exponentiated Rayleigh Poisson (ERP) distribution. The beta Weibull (BW) distribution comes as the limiting distribution of the BWP distribution when $\lambda \to 0^+$. For $a = b = 1$, equation (5) becomes the WP density function. In addition, if $\alpha = 1$, we obtain the exponential Poisson (EP) distribution. The following distributions are new sub-models: the beta Rayleigh Poisson (BRP), exponentiated Weibull Poisson (EWP), beta exponential Poisson (BEP), exponentiated Rayleigh Poisson (ERP), beta Rayleigh (BR), Rayleigh Poisson (RP) and arc sine Weibull Poisson (ASWP) distributions (for more details, see Appendix B). Other sub-models are the beta exponential (BE), beta Weibull (BW), beta Rayleigh (BR), exponentiated Rayleigh (ER), exponentiated exponential (EE), exponentiated Weibull (EW), Rayleigh (R), Weibull (W) and exponential (E) distributions. Several special distributions of the BWP model are displayed in Figure 2.

3. Properties of the new distribution

3.1 Density function

We can derive a useful expansion for the BWP density function (see the proof in Appendix A) given by

$$f(x) = \sum_{r=0}^{\infty} \sum_{j=0}^{r} v_{r,j}\, g(x; \alpha, \beta, \lambda_{r,j}), \tag{7}$$

where $\lambda_{r,j} = \lambda(r - j + 1) > 0$ and

$$v_{r,j} = \frac{(-1)^j\,(r+1)\,v_r\,\mathrm{e}^{j\lambda}\,(1 - \mathrm{e}^{-\lambda_{r,j}})}{(r - j + 1)\,\mathrm{e}^{-\lambda_{r,j}}(1 - \mathrm{e}^{\lambda})^r(\mathrm{e}^{\lambda} - 1)} \binom{r}{j}.$$

Clearly, $\sum_{r=0}^{\infty} \sum_{j=0}^{r} v_{r,j} = 1$. Equation (7) reveals that the BWP density function is a linear combination of WP density functions. So, we can obtain some mathematical properties of the BWP distribution directly from those WP properties.

3.2 Cumulative function and quantiles

By integrating (7), the cdf $F(x)$ becomes

$$F(x) = \sum_{r=0}^{\infty} \sum_{j=0}^{r} v_{r,j}\, G(x; \alpha, \beta, \lambda_{r,j}). \tag{8}$$

Quantile functions are in widespread use in general statistics and often find representations in terms of lookup tables for key percentiles. For some baseline distributions with closed-form cdf, it is possible to obtain the quantile function in closed-form. However, for some other distributions, the solution is not possible. The quantile function, say $x = Q(z; \alpha, \beta, \lambda, a, b) = F^{-1}(z; \alpha, \beta, \lambda, a, b)$, of the BWP distribution follows by inverting (6) as

$$x = Q(z;\alpha,\beta,\lambda,a,b) = \left\{\log\left(\log[w + \mathrm{e}^{\lambda}(1-w)]^{\frac{1}{\lambda}}\right)^{-\frac{1}{\beta}}\right\}^{\frac{1}{\alpha}}, \tag{9}$$

where $w = Q_{a,b}(z)$ denotes the beta quantile function with parameters a and b.

Power series methods are at the heart of many aspects of applied mathematics and statistics. We can obtain the moments of the beta G distribution using a power series expansion for the quantile function $x = Q_G(u) = G^{-1}(u)$ of the baseline cdf $G(x)$ with easily computed non-linear recurrence equation for its coefficients.

When the function $Q(u)$ does not have a closed form expression, this function can usually be written in terms of a power series expansion of a transformed variable v, which is usually of the form $v = p(qu - t)^{\rho}$ for p, q, t and ρ known constants.

We can obtain a power series for $Q_{a,b}(z)$ in the Wolfram website given by

$$\begin{aligned} Q_{a,b}(z) = v &+ \frac{(b-1)}{(a+1)}v^2 + \frac{(b-1)(a^2+3ba-a+5b-4)}{2(a+1)^2(a+2)}v^3 \\ &+ \frac{v^4(b-1)}{3(a+1)^3(a+2)(a+3)}[a^4 + (6b-1)a^3 + (b+2)(8b-5)a^2 + \\ &(33b^2 - 30b + 4)a + b(31b-47) + 18] + O(v^5), \end{aligned} \tag{10}$$

where $v = [azB(a,b)]^{1/a}$ for $a > 0$.

The simulation of the BWP distribution is easy. If T is a random variable having a beta distribution with parameters a and b, then the random variable

$$X = \left\{\log\left(\log[W + \mathrm{e}^{\lambda}(1-W)]^{\frac{1}{\lambda}}\right)^{-\frac{1}{\beta}}\right\}^{\frac{1}{\alpha}}$$

follows the BWP distribution.

3.3 Survival and hazard rate functions

The BWP survival function is given by

$$S(x;\boldsymbol{\theta}) = 1 - F(x;\boldsymbol{\theta}) = 1 - I_{(\mathrm{e}^{\lambda u} - \mathrm{e}^{\lambda})/(1-\mathrm{e}^{\lambda})}(a,b),$$

where $\boldsymbol{\theta} = (\alpha,\beta,\lambda,a,b)$ is the vector of the model parameters. The failure rate function corresponding to (5) reduces to

$$h(x;\boldsymbol{\theta}) = \frac{f(x;\boldsymbol{\theta})}{S(x;\boldsymbol{\theta})} = \frac{c\,u\,x^{\alpha-1}\mathrm{e}^{\lambda u}(\mathrm{e}^{\lambda} - \mathrm{e}^{\lambda u})^{a-1}(\mathrm{e}^{\lambda u} - 1)^{b-1}}{\{1 - I_{(\mathrm{e}^{\lambda u} - \mathrm{e}^{\lambda})/(1-\mathrm{e}^{\lambda})}(a,b)\}}.$$

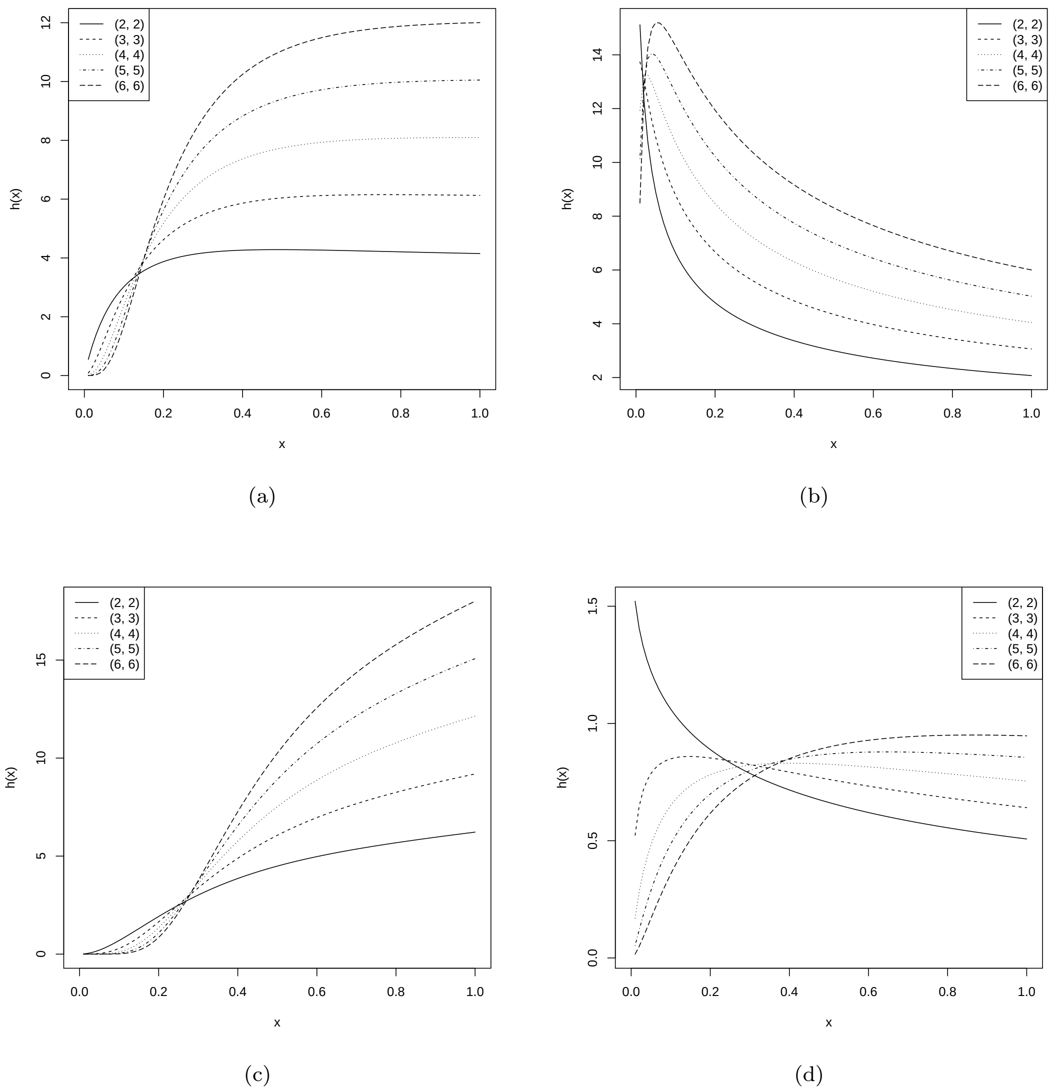

Figure 3. Plots of the BWP hazard rate function for (a) $\alpha = 1$, $\beta = 2$ and $\lambda = 1$; (b) $\alpha = 0.5$, $\beta = 2$ and $\lambda = 1$; (c) $\alpha = 1.5$, $\beta = 2$ and $\lambda = 1$; (d) $\alpha = 0.5$, $\beta = 0.5$ and $\lambda = 1$.

3.4 Moments

We hardly need to emphasize the necessity and importance of moments in any statistical analysis especially in applied work. Some of the most important features and characteristics of a distribution can be studied through moments (e.g., tendency, dispersion, skewness and kurtosis).

An expression for the mgf of X can be obtained from (7) using the WP generating function. Setting $y = \lambda_{r,j} \mathrm{e}^{-\beta x^{\alpha}}$ in the definition of the mgf, we can express it as

$$M_X(t) = \sum_{r=0}^{\infty} \sum_{j=0}^{r} v_{r,j} \, (\mathrm{e}^{\lambda_{r,j}} - 1)^{-1}$$
$$\times \int_0^{\lambda_{r,j}} \exp\{t(-\beta^{-1}[\log(y) - \log(\lambda_{r,j})])^{1/\alpha} + y\} dy.$$

Using the power series of the exponential function, after some simplification, we obtain

$$M_X(t) = \sum_{r,m,n=0}^{\infty} \sum_{j=0}^{r} q(r,m,n,j)\, J(\lambda_{r,j},m,n)\, t^n, \tag{11}$$

where

$$J(\lambda_{r,j},m,n) = \int_0^{\lambda_{r,j}} y^m \left(-\beta^{-1}[\log(y) - \log(\lambda_{r,j})]\right)^{\frac{n}{\alpha}} dy$$

and

$$q(r,m,n,j) = \frac{v_{r,j}}{(\mathrm{e}^{\lambda_{r,j}} - 1)\; m!\; n!}.$$

The last integral can be computed using the software Mathematica 8.0. Then,

$$M_X(t) = \sum_{r,m,n=0}^{\infty} \sum_{j=0}^{r} \varpi(r,m,n,j)\, \Gamma\left(\frac{\alpha+n}{\alpha}\right) t^n, \tag{12}$$

where

$$\varpi(r,m,n,j) = \beta^{-\frac{n}{\alpha}}\, \lambda_{r,j}^{m+1}\, (1+m)^{-\frac{\alpha+n}{\alpha}}\, q(r,m,n,j).$$

Equation (12) can be reduced to

$$M_X(t) = \sum_{n=0}^{\infty} \delta_n\, t^n, \tag{13}$$

where $\delta_n = \sum_{m,r=0}^{\infty} \sum_{j=0}^{r} \varpi(r,m,n,j)\, \Gamma\left(\frac{\alpha+n}{\alpha}\right)$, $n = 0, 1, \ldots$

Hence, the *nth* ordinary moment of X, say $\mu'_n = E(X^n)$, is simply given by $\mu'_n = n!\, \delta_n$. Further, the central moments (μ_n) and cumulants (κ_n) of X can be determined as

$$\mu_n = \sum_{s=0}^{n} (-1)^s \binom{n}{s} \mu_1'^{s}\, \mu'_{n-s} \qquad \text{and} \qquad \kappa_n = \mu'_n - \sum_{s=1}^{n-1} \binom{n-1}{s-1} \kappa_s\, \mu'_{n-s},$$

respectively, where $\kappa_1 = \mu'_1$. Then, $\kappa_2 = \mu'_2 - \mu_1'^2$, $\kappa_3 = \mu'_3 - 3\mu'_2\mu'_1 + 2\mu_1'^3$, $\kappa_4 = \mu'_4 - 4\mu'_3\mu'_1 - 3\mu_2'^2 + 12\mu'_2\mu_1'^2 - 6\mu_1'^4$, etc. The skewness $\gamma_1 = \kappa_3/\kappa_2^{3/2}$ and kurtosis $\gamma_2 = \kappa_4/\kappa_2^2$ follow from the second, third and fourth cumulants.

The nth descending factorial moment of X is

$$\mu'_{(n)} = E(X^{(n)}) = E\left[X(X-1) \times \cdots \times (X-n+1)\right] = \sum_{r=0}^{n} s(n,r)\, \mu'_r,$$

where

$$s(n,r)=\frac{1}{r!}\left[\frac{d^r}{dx^r}x^{(n)}\right]_{x=0}$$

is the Stirling number of the first kind which counts the number of ways to permute a list of n items into r cycles. So, we can obtain the factorial moments from the ordinary moments given before.

The incomplete moments of X can be expressed in terms of the incomplete moments of the WP distribution from equation (7). We obtain

$$\begin{aligned} m_n(y) = E(X^n|X<y) &= \sum_{r=0}^{\infty}\sum_{j=0}^{r} v_{r,j}\int_0^y x^n g(x;\alpha,\beta,\lambda_{r,j})dx \\ &= \sum_{r=0}^{\infty}\sum_{j=0}^{r} v_{r,j}\int_0^y x^n c\,u\,x^{\alpha-1}\mathrm{e}^{\lambda u}dx. \end{aligned} \tag{14}$$

Setting $z=\beta x^\alpha$ and integrating by parts, we can write

$$m_n(y)=\frac{\mathrm{e}^{-\lambda}y^n}{1-\mathrm{e}^{\lambda}}\sum_{r=0}^{\infty}\sum_{j=0}^{r} v_{r,j}\left\{n\sum_{s=0}^{\infty}\frac{\lambda^s}{s!}\left[\sum_{m=0}^{\infty}\frac{(-1)^m s^m(\beta y^\alpha)^m}{n+m\alpha}\right]-\mathrm{e}^{\lambda \mathrm{e}^{-\beta y^\alpha}}\right\}.$$

The sum in m converges to $(n+m\alpha)^{-1}\mathrm{e}^{-s\beta y^\alpha}$. Then, the *nth* incomplete moment of X becomes

$$m_n(y)=\sum_{r=0}^{\infty}\sum_{j=0}^{r} p_{r,j}\;y^n\left\{\sum_{s=0}^{\infty}\left[\frac{\lambda^s\mathrm{e}^{-s\beta y^\alpha}}{s!(n+m\alpha)}\right]-\frac{y^n\mathrm{e}^{-\lambda(1-\mathrm{e}^{-\beta y^\alpha})}}{1-\mathrm{e}^{\lambda}}\right\}, \tag{15}$$

where $p_{r,j}=\dfrac{n\,v_{r,j}\;\mathrm{e}^{-\lambda}}{1-\mathrm{e}^{\lambda}}$.

We can derive the mean deviations of X about the mean μ_1' and about the median M in terms of its first incomplete moment. They can be expressed as

$$\delta_1=2\big[\mu_1'\,F(\mu_1')-m_1(\mu_1')\big] \qquad \text{and} \qquad \delta_2=\mu_1'-2m_1(M), \tag{16}$$

where $\mu_1'=E(X)$ and $m_1(q)=\int_{-\infty}^{q}x\,f(x)\,dx$. The quantity $m_1(q)$ is obtained from (15) with $n=1$ and the measures δ_1 and δ_2 in (16) are immediately determined from these formulae with $n=1$ by setting $q=\mu_1^{'}$ and $q=M$, respectively. For a positive random variable X, the Bonferroni and Lorenz curves are defined as $B(\pi)=T_1(q)/[\pi\mu_1']$ and $L(\pi)=T_1(q)/\mu_1'$, respectively, where $q=F^{-1}(\pi)=Q(\pi)$ comes from the quantile function (9) for a given probability π.

The formulae derived along the paper can be easily handled in most symbolic computation software platforms such as Maple, Mathematica and Matlab. These platforms have currently the ability to deal with analytic expressions of formidable size and complexity. Established explicit expressions to calculate statistical measures can be more efficient than computing them directly by numerical integration. The infinity limit in the sums of these expressions can be substituted by a large positive integer such as 20 or 30 for most practical purposes.

3.5 Order statistics

Order statistics make their appearance in many areas of statistical theory and practice. Moments of order statistics play an important role in quality control and reliability, where some predictors are often based on moments of the order statistics. We derive an explicit expression for the density function of the ith order statistic $X_{i:n}$, say $f_{i:n}(x)$ (see Appendix C). For a beta-G model defined from the parent functions $g(x)$ and $G(x)$, $f_{i:n}(x)$ can be expressed as an infinite linear combination of WP density functions

$$f_{i:n}(x) = \sum_{l=0}^{\infty}\sum_{s=0}^{l} \gamma_{i:n}(l,s)\; g(x;\alpha,\beta,\lambda_{l,s}), \tag{17}$$

where $\lambda_{l,s} = \lambda(l-s+1)$ and

$$\gamma_{i:n}(l,s) = \sum_{j=0}^{n-i}\sum_{r,k=0}^{\infty} \frac{(-1)^{l+s+j+k}\, \mathrm{e}^{s\lambda} \binom{l}{s}\binom{n-i}{j}\binom{r+a-1}{k}\binom{k+b-1}{l}(1-\mathrm{e}^{-\lambda_{l,s}})\, c_{i+1-j,r}}{(l-s+1)(1-\mathrm{e}^{-\lambda})\,(1-\mathrm{e}^{\lambda})^{l}\, B(a,b)^{i+j} B(i,n-i+1)}.$$

An expression for the mgf of $X_{i:n}$ can be obtained from (17) using the WP generating function. Setting $y = \lambda_{l,s}\mathrm{e}^{-\beta x^{\alpha}}$ in the definition of the generating function, we obtain

$$M_{X_{i:n}}(t) = \sum_{l,m,n=0}^{\infty}\sum_{s=0}^{l} \varpi_i(l,m,n,s)\,\Gamma\left(\frac{\alpha+n}{\alpha}\right) t^{n}, \tag{18}$$

where

$$\varpi_i(l,m,n,s) = \frac{\beta^{-\frac{n}{\alpha}}\, \lambda_{l,s}^{m+1}\, (1+m)^{-\frac{\alpha+n}{\alpha}} \gamma_{i:n}(l,s)}{m!\; n!\,(\mathrm{e}^{\lambda_{l,s}}-1)}.$$

Equation (18) can be reduced to $M_{X_{i:n}}(t) = \sum_{n=0}^{\infty} \delta_{i:n}\, t^{n}$, where

$$\delta_{i:n} = \sum_{m,l=0}^{\infty}\sum_{s=0}^{l} \varpi_i(l,m,n,s)\,\Gamma\left(\frac{\alpha+n}{\alpha}\right), n = 0,1,\ldots$$

Hence, the *sth* ordinary moment of $X_{i:n}$ becomes $E(X_{i:n}^{s}) = s!\,\delta_{i:n}$.

3.6 Rényi entropy

The entropy of a random variable X with density function $f(x)$ is a measure of the uncertainty variation. The Rényi entropy is defined as

$$I_R(\rho) = (1-\rho)^{-1}\log\left\{\int f(x)^{\rho}dx\right\},$$

where $\rho > 0$ and $\rho \neq 1$. If a random variable X has the BWP distribution, we have

$$f(x)^{\rho} = \left[\frac{g(x;\theta)}{B(a,b)}\right]^{\rho} G(x)^{(a-1)\rho}\,[1-G(x)]^{(b-1)\rho}. \tag{19}$$

By expanding the binomial term, the following expansion holds for any real a,

$$G(x)^{(a-1)\rho+j} = \sum_{r=0}^{\infty} s_r[(a-1)\rho+j]\;G(x)^r,$$

where $s_r[(a-1)\rho+j] = \sum_{i=r}^{\infty}(-1)^{r+i}\binom{(a-1)\rho+j}{j}\binom{i}{r}$. Equation (19) can be rewritten as

$$f(x)^{\rho} = \left[\frac{g(x;\theta)}{B(a,b)}\right]^{\rho}\sum_{j=0}^{\infty}\sum_{r=0}^{\infty} q_{j,r}\;G(x)^r,$$

where $q_{j,r} = (-1)^j\binom{(b-1)\rho}{j}\;s_r[(a-1)\rho+j]$.

From equations (3) and (4), we obtain

$$f(x)^{\rho} = \left[\frac{c\,u\,x^{\alpha-1}\mathrm{e}^{\lambda u}}{B(a,b)}\right]^{\rho}\sum_{j=0}^{\infty}\sum_{r=0}^{\infty} q_{j,r}\left(\frac{\mathrm{e}^{\lambda u}-\mathrm{e}^{\lambda}}{1-\mathrm{e}^{\lambda}}\right)^{r}.$$

Then,

$$f(x)^{\rho} = \sum_{j=0}^{\infty}\sum_{r=0}^{\infty}\sum_{t=0}^{r} p_{j,r,t}\,u^{\rho}\,x^{\alpha(\rho-\frac{\rho}{\alpha})}\;\mathrm{e}^{\lambda(\rho+r-t)u}, \tag{20}$$

where $u = \mathrm{e}^{-\beta x^{\alpha}}$ and

$$p_{j,r,t} = \frac{q_{j,r}(-1)^t\binom{r}{t}\mathrm{e}^{\lambda t}c^{\rho}}{[B(a,b)]^{\rho}(1-\mathrm{e}^{\lambda})^r}.$$

Using the power series expansion $\mathrm{e}^{\lambda(\rho+r-t)u} = \sum_{s=0}^{\infty}\dfrac{[\lambda(\rho+r-t)]^s}{s!}\mathrm{e}^{-s\beta x^{\alpha}}$ in (20) and setting $y = \beta s x^{\alpha}$, the Rényi entropy reduces to

$$I_R(\rho) = (1-\rho)^{-1}\log\left\{\sum_{j=0}^{\infty}\phi_j(\rho)\,\Gamma\left(\rho+\frac{1-\rho}{\alpha}\right)\right\}, \tag{21}$$

where

$$\phi_j(\rho) = \sum_{r,s=0}^{\infty} \sum_{t=0}^{r} \frac{p_{j,r,t}\, \lambda^s (\rho + r - t)^s}{\alpha\, s!\, (\beta s)^{\frac{1-\rho}{\alpha} + \rho}}.$$

4. Maximum likelihood estimation

Let $x_1, \ldots, x_n$ be a random sample of size n from the BWP$(a, b, \alpha, \beta, \lambda)$ distribution. The log-likelihood function for the vector of parameters $\boldsymbol{\theta} = (a, b, \alpha, \beta, \lambda)^T$ can be expressed as

$$\begin{aligned} l(\boldsymbol{\theta}) = n\,[\log(\alpha\beta\lambda) - \lambda - \log[B(a,b)] - \log(1 - \mathrm{e}^{-\lambda}) - \log(1 - \mathrm{e}^{\lambda})^{a+b-2}] \\ + (\alpha - 1) \sum_{i=1}^{n} \log(x_i) - \beta \sum_{i=1}^{n} x_i^{\alpha} + \lambda \sum_{i=1}^{n} u_i \\ + (a - 1) \sum_{i=1}^{n} \log(\mathrm{e}^{\lambda u_i} - \mathrm{e}^{\lambda}) + (b - 1) \sum_{i=1}^{n} \log(1 - \mathrm{e}^{\lambda u_i}), \end{aligned}$$

where $u_i = \exp(-\beta x_i^{\alpha})$ is a transformed observation. The components of the score vector $U(\boldsymbol{\theta})$ are given by

$$\begin{aligned} U_{\alpha}(\boldsymbol{\theta}) &= \frac{n}{\alpha} + \sum_{i=1}^{n} \log(x_i) - \beta \sum_{i=1}^{n} x_i^{\alpha} \log(x_i) - \lambda\beta \sum_{i=1}^{n} u_i x_i^{\alpha} \log(x_i) \\ &\quad + \lambda\beta \sum_{i=1}^{n} u_i x_i^{\alpha} \mathrm{e}^{\lambda u_i} \log(x_i) \left(\frac{1-a}{\mathrm{e}^{\lambda u_i} - \mathrm{e}^{\lambda}} + \frac{b-1}{1 - \mathrm{e}^{\lambda u_i}} \right), \\ U_{\beta}(\boldsymbol{\theta}) &= \frac{n}{\beta} - \sum_{i=1}^{n} x_i^{\alpha} - \lambda \sum_{i=1}^{n} u_i x_i^{\alpha} + \lambda \sum_{i=1}^{n} u_i x_i^{\alpha} \mathrm{e}^{\lambda u_i} \\ &\quad \times \left(\frac{1-a}{\mathrm{e}^{\lambda u_i} - \mathrm{e}^{\lambda}} + \frac{b-1}{1 - \mathrm{e}^{\lambda u_i}} \right), \end{aligned}$$

$$
\begin{aligned}
U_\lambda(\boldsymbol{\theta}) &= \frac{n}{\lambda} - n + \frac{n\mathrm{e}^{-\lambda}}{1-\mathrm{e}^{-\lambda}} - \frac{n(a+b-2)\mathrm{e}^{\lambda}\log(1-\mathrm{e}^{\lambda})^{a+b-2}}{(1-\mathrm{e}^{\lambda})\log(1-\mathrm{e}^{\lambda})} \\
&\quad + \sum_{i=1}^{n} u_i + (a-1)\sum_{i=1}^{n}\frac{u_i\mathrm{e}^{\lambda u_i}-\mathrm{e}^{\lambda}}{\mathrm{e}^{\lambda u_i}-\mathrm{e}^{\lambda}} - (b-1) \\
&\quad \times \sum_{i=1}^{n}\frac{u_i\mathrm{e}^{\lambda u_i}}{1-\mathrm{e}^{\lambda u_i}}, \\
U_a(\boldsymbol{\theta}) &= -n\left[\psi(a)-\psi(a+b)\right] + n\log(\mathrm{e}^{\lambda}-1)^{a+b-2}\log[\log(1-\mathrm{e}^{\lambda})] \\
&\quad + \sum_{i=1}^{n}\log(\mathrm{e}^{\lambda u_i}-\mathrm{e}^{\lambda}), \\
U_b(\boldsymbol{\theta}) &= -n\left[\psi(b)-\psi(a+b)\right] + n\log(\mathrm{e}^{\lambda}-1)^{a+b-2}\log[\log(1-\mathrm{e}^{\lambda})] \\
&\quad + \sum_{i=1}^{n}\log(1-\mathrm{e}^{\lambda u_i}),
\end{aligned}
$$

where $\psi(\cdot)$ is the digamma function. The maximum likelihood estimates (MLEs) $\widehat{\boldsymbol{\theta}} = (\widehat{a}, \widehat{b}, \widehat{\alpha}, \widehat{\beta}, \widehat{\lambda})^T$ of $\boldsymbol{\theta} = (a, b, \alpha, \beta, \lambda)^T$ are the simultaneous solutions of the non-linear equations: $U_a(\boldsymbol{\theta}) = U_b(\boldsymbol{\theta}) = U_\alpha(\boldsymbol{\theta}) = U_\beta(\boldsymbol{\theta}) = U_\lambda(\boldsymbol{\theta}) = 0$. They can be solved numerically using iterative methods such as a Newton-Raphson type algorithm.

For interval estimation and hypothesis tests on the model parameters, we require the 5×5 observed information matrix $J = J(\boldsymbol{\theta})$ given in Appendix D. Under conditions that are fulfilled for parameters in the interior of the parameter space but not on the boundary, the asymptotic distribution of $\sqrt{n}(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta})$ is $N_5(0, I(\boldsymbol{\theta})^{-1})$, where $I(\boldsymbol{\theta})$ is the expected information matrix. In practice, we can replace $I(\boldsymbol{\theta})$ by the observed information matrix evaluated at $\widehat{\boldsymbol{\theta}}$, say $J(\widehat{\boldsymbol{\theta}})$. We can construct approximate confidence regions for the parameters based on the multivariate normal $N_5(0, J(\widehat{\boldsymbol{\theta}})^{-1})$ distribution.

Further, the likelihood ratio (LR) statistic can be used for comparing this distribution with some of its sub-models. We can compute the maximum values of the unrestricted and restricted log-likelihoods to construct the LR statistics for testing some sub-models of the BWP distribution. For example, the test of $H_0 : a = b = 1$ versus $H_1 : H_0$ *is not true* is equivalent to compare the BWP and WP distributions and the LR statistic becomes $w = 2\{l(\hat{a}, \hat{b}, \hat{\alpha}, \hat{\beta}, \hat{\lambda}) - l(1, 1, \tilde{\alpha}, \tilde{\beta}, \tilde{\lambda})\}$, where $\hat{a}$, $\hat{b}$, $\hat{\alpha}$, $\hat{\beta}$ and $\hat{\lambda}$ are the MLEs underH_1 and $\tilde{\alpha}$, $\tilde{\beta}$ and $\tilde{\lambda}$ are the estimates under H_0.

5. Application

Here, we present an application regarding the BWP model to the maintenance data with 46 observations reported on active repair times (hours) for an airborne communication transceiver discussed by Alven (1964), Chhikara and Folks (1977) and Dimitrakopoulou et al. (2007). We also fit a five-parameter beta Weibull geometric (BWG) distribution introduced by Cordeiro et al. (2011) to make a comparasion with the BWP model. The

BWG density function is given by

$$f(x;\boldsymbol{\theta}_1) = \frac{\alpha(1-p)^b\beta^\alpha x^{\alpha-1}\mathrm{e}^{-b(\beta x)^\alpha}(1-\mathrm{e}^{-(\beta x)^\alpha})^{a-1}(1-p\mathrm{e}^{-(\beta x)^\alpha})^{-(a+b)}}{B(a,b)},$$

where $\boldsymbol{\theta}_1 = (p,\alpha,\beta,a,b)$ and $x > 0$.

The data are: 0.2, 0.3, 0.5, 0.5, 0.5, 0.5, 0.6, 0.6, 0.7, 0.7, 0.7, 0.8, 0.8, 1.0, 1.0, 1.0, 1.0, 1.1, 1.3, 1.5, 1.5, 1.5, 1.5, 2.0, 2.0, 2.2, 2.5, 2.7, 3.0, 3.0, 3.3, 3.3, 4.0, 4.0, 4.5, 4.7, 5.0, 5.4, 5.4, 7.0, 7.5, 8.8, 9.0, 10.3, 22.0 e 24.5.

In Table 1, we list the MLEs of the model parameters and the bias-corrected Akaike information criterion (BAIC), Bayesian information criterion (BIC) and the Hannan-Quinn information criterion (HQIC). We observe that the value of the BAIC criterion is smaller for the BWP distribution as compared with those values of the other models. So, the new distribution seems to be a very competitive model to these data.

Table 1. MLEs of the parameters and BAIC, BIC and HQIC statistics of the BWP, BWG, WP and Weibull models for data of active repair times (hours) for an airborne communication transceiver.

Model	$\hat{a}$	$\hat{b}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\lambda}$	BAIC	BIC	HQIC
BWP	21.969	0.320	0.722	1.439	5.342	207.838	216.981	211.263
	(58.799)	(0.256)	(0.390)	(1.418)	(2.232)			
WP			1.101	0.092	3.522	210.927	216.413	212.982
			(0.120)	(0.052)	(1.917)			
Weibull			0.899	0.334		212.939	216.597	214.309
			(0.096)	(0.075)				
Model	$\hat{a}$	$\hat{b}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{p}$	BAIC	BIC	HQIC
BWG	3.269	0.587	1.417	0.212	0.988	208.205	217.348	211.630
	(4.599)	(0.323)	(0.642)	(0.076)	(0.017)			

The LR test statistic for testing $H_0 : a = b = 1$ against $H_1 : H_0$ *is not true* is $w = 7.08912$ (p-value $= 2.88 \times 10^{-2}$), which is statistically significant. Figure 4 displays the histogram of the data and the plots of the fitted BWP, WP, Weibull and BWG models.

6. Concluding remarks

The Weibull distribution is commonly used to model the lifetime of a system. However, it does not exhibit a bathtub-shaped failure rate function and thus it can not be used to model the complete lifetime of a system. We define a new lifetime model, called the beta Weibull Poisson (BWP) distribution, which extends the Weibull Poisson (WP) distribution proposed by Lu and Shi (2012), whose failure rate function can be increasing, decreasing and upside-down bathtub. The BWP distribution is quite flexible to analyse positive data instead of some other special models. Its density function can be expressed as a mixture of WP densities. We provide a mathematical treatment of the distribution including explicit expressions for the density function, generating function, ordinary and incomplete moments, Rényi entropy, order statistics and their moments. The estimation of the model parameters is approached by the method of maximum likelihood and the observed information matrix is determined. An application to real data reveals that the BWP distribution could provide a better fit than other well-known lifetime models.

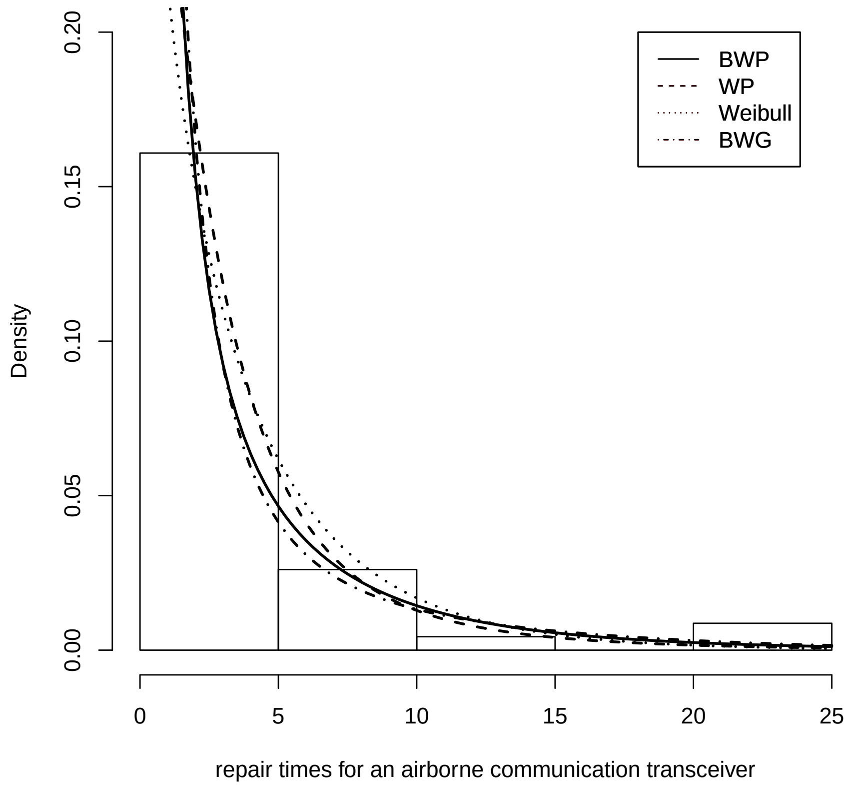

Figure 4. The density functions of the fitted BWP, WP, Weibull and BWG distributions.

Appendix A. The BWP density function

An expansion for the beta-G cumulative function is given by Cordeiro and Lemonte (2011) and follows from equation (1) as

$$F(x) = \frac{1}{B(a,b)} \sum_{r=0}^{\infty} t_r \, G(x)^r, \tag{A1}$$

where $t_r = \sum_{m=0}^{\infty} w_m \, s_r(a+m)$ for any real a, $w_m = (-1)^m (a+m)^{-1} \binom{b-1}{m}$ and $s_r(a+m) = \sum_{j=r}^{\infty} (-1)^{r+j} \binom{a+m}{j} \binom{j}{r}$. Differentiating equation (A1), we obtain an expansion for the BWP density function

$$f(x) = \sum_{r=0}^{\infty} v_r \, h_{r+1}(x), \tag{A2}$$

where $v_r = t_{r+1}/B(a,b)$. Note that $h_{r+1}(x) = (r+1)G(x)^r g(x)$ is the density function of the exponentiated G with power parameter $r+1$, say exp-G$(r+1)$, distribution. We can verify that $\sum_{r=0}^{\infty} v_r = 1$. In fact,

$$\sum_{r=0}^{\infty} v_r = \frac{1}{B(a,b)} \sum_{r=0}^{\infty}\sum_{m=0}^{\infty} w_m s_r(a) = 1$$

if and only if

$$\sum_{r=0}^{\infty}\sum_{m=0}^{\infty} w_m \; s_r(a) = B(a,b). \tag{A3}$$

But

$$B(a,b) = \int_0^1 t^{a-1}(1-t)^{b-1} dt = \sum_{j=0}^{\infty} \binom{b-1}{j} \frac{(-1)^j}{a+j},$$

and, consequently,

$$\sum_{r=0}^{\infty}\sum_{m=0}^{\infty} w_m \; s_r(a) = \sum_{m=0}^{\infty} \frac{(-1)^m}{(a+m)} \binom{b-1}{m} \sum_{r=0}^{\infty}\sum_{j=r}^{\infty} (-1)^{r+j} \binom{a}{j}\binom{j}{r} = B(a,b)$$

Consider the expressions of $g(x)$ and $G(x)$ from equations (3) and (4), respectively. Replacing them in (A2), we obtain an expansion for the BWP density function

$$f(x) = c\, x^{\alpha-1}\; u\, \mathrm{e}^{\lambda u} \sum_{r=0}^{\infty} v_r (r+1) \left(\frac{\mathrm{e}^{\lambda u} - \mathrm{e}^{\lambda}}{1-\mathrm{e}^{\lambda}}\right)^r. \tag{A4}$$

Hence, from this equation, the BWP density function can be expressed as a linear combination of WP density functions.

Appendix B. Special cases of the BWP distribution

Setting $b = 1$ in equation (5), we obtain the EWP density function

$$f(x) = c\, u\, x^{\alpha-1} \mathrm{e}^{\lambda u} \left(\frac{\mathrm{e}^{\lambda u} - \mathrm{e}^{\lambda}}{1-\mathrm{e}^{\lambda}}\right)^{a-1}, \quad c = \frac{\alpha\beta\lambda}{B(a,1)(1-\mathrm{e}^{-\lambda})}.$$

Using equation $G(x)^{\alpha} = \sum_{k=0}^{\infty} s_k(\alpha)\, G(x)^k$, we can write

$$\begin{aligned} f(x) &= c\, u\, x^{\alpha-1} \mathrm{e}^{\lambda u} \sum_{k=0}^{\infty} s_k(a-1) \left(\frac{\mathrm{e}^{\lambda u} - \mathrm{e}^{\lambda}}{1-\mathrm{e}^{\lambda}}\right)^k \\ &= c\, u\, x^{\alpha-1} \mathrm{e}^{\lambda u} \sum_{k=0}^{\infty} \frac{s_k(a-1)}{(1-\mathrm{e}^{\lambda})^k} \sum_{r=0}^{k} (-1)^r \binom{k}{r} \mathrm{e}^{\lambda u(k-r)} \mathrm{e}^{\lambda r}. \end{aligned} \tag{B1}$$

After some algebra, we obtain from (B1)

$$f(x) = \sum_{k=0}^{\infty}\sum_{r=0}^{k} v_{k,r}\, g(x;\alpha,\beta,\lambda_{k,r}), \tag{B2}$$

where $\lambda_{k,r} = \lambda(k-r+1)$ and

$$v_{k,r} = \frac{(-1)^r \binom{k}{r} B(a,b)\, s_k(a-1)\, \mathrm{e}^{\lambda r}(1-\mathrm{e}^{-\lambda_{k,r}})}{(k-r+1)\, B(a,1)\, (1-\mathrm{e}^{\lambda})^k (1-\mathrm{e}^{-\lambda})}.$$

Equation (B2) reveals that the density function $f(x)$ is a linear combination of the WP densities.

From equation (5) with $a = b = 1/2$, we obtain

$$f(x;\boldsymbol{\theta}) = \frac{c_1 x^{\alpha-1} u \mathrm{e}^{\lambda u}}{\pi}\left(\frac{\mathrm{e}^{\lambda u}-\mathrm{e}^{\lambda}}{1-\mathrm{e}^{\lambda}}\right)^{-1/2}\left(1-\frac{\mathrm{e}^{\lambda u}-\mathrm{e}^{\lambda}}{1-\mathrm{e}^{\lambda}}\right)^{-1/2},$$

where $c_1 = \dfrac{\alpha\beta\lambda\mathrm{e}^{-\lambda}(\mathrm{e}^{\lambda}-1)}{(1-\mathrm{e}^{-\lambda})}$ and $u = \mathrm{e}^{-\beta x^{\alpha}}$. Thus,

$$f(x;\boldsymbol{\theta}) = \frac{c_1 x^{\alpha-1} u \mathrm{e}^{\lambda u}}{\pi\sqrt{\left(\dfrac{\mathrm{e}^{\lambda u}-\mathrm{e}^{\lambda}}{1-\mathrm{e}^{\lambda}}\right)\left(\dfrac{1-\mathrm{e}^{\lambda u}}{1-\mathrm{e}^{\lambda}}\right)}}$$

If λ approaches to 0, then

$$\lim_{\lambda\to 0} f(x;\boldsymbol{\theta}) = \lim_{\lambda\to 0}\frac{c_1 x^{\alpha-1} u \mathrm{e}^{\lambda u}}{\pi\sqrt{\left(\dfrac{\mathrm{e}^{\lambda u}-\mathrm{e}^{\lambda}}{1-\mathrm{e}^{\lambda}}\right)\left(\dfrac{1-\mathrm{e}^{\lambda u}}{1-\mathrm{e}^{\lambda}}\right)}} = \frac{\alpha\beta x^{\alpha-1}u}{\pi\sqrt{u(1-u)}}$$

So, the BWP distribution reduces as a limiting case to a two-parameter arcsine Weibull-Poisson distribution.

Appendix C. Expansion for the Density Function of the order statistics

The density function $f_{i:n}(x)$ of the ith order statistic, say $X_{i:n}$, for $i = 1, 2, \ldots, n$, from data values $X_1, \ldots, X_n$ having the beta-G distribution can be obtained from (2) as

$$f_{i:n}(x) = \frac{g(x)\, G(x)^{a-1}\{1-G(x)\}^{b-1}}{B(a,b)\, B(i,n-i+1)}\sum_{j=0}^{n-i}(-1)^j\binom{n-i}{j} F(x)^{i+j-1}. \tag{C1}$$

By application of an equation in Section 0.314 of Gradshteyn and Ryzhik (2000) for a power series raised to any j positive integer

$$\left(\sum_{i=0}^{\infty} a_i\, u^i\right)^j = \sum_{i=0}^{\infty} c_{j,i}\, u^i, \tag{C2}$$

where the coefficients $c_{j,i}$ (for $i = 1, 2, \ldots$) can be obtained from the recurrence equation

$$c_{j,i} = (i a_0)^{-1} \sum_{m=1}^{i} [m(j+1) - i]\, a_m\, c_{j,i-m}, \tag{C3}$$

with $c_{j,0} = a_0^j$. The coefficient $c_{j,i}$ comes from $c_{j,0}, \ldots, c_{j,i-1}$ and then from $a_0, \ldots, a_i$. The coefficients $c_{j,i}$ can be given explicitly in terms of the quantities a_i's, although it is not necessary for programming numerically our expansions in any algebraic or numerical software.

For $a > 0$ real non-integer, we have

$$\begin{aligned} F(x)^{i+j-1} &= \left(\frac{1}{B(a,b)} \sum_{r=0}^{\infty} t_r(a,b) G(x)^r\right)^{i+j-1} \\ &= \left(\frac{1}{B(a,b)}\right)^{i+j-1} \left(\sum_{r=0}^{\infty} t_r G(x)^r\right)^{i+j-1}. \end{aligned}$$

We now use equations (C2)-(C3)

$$\begin{aligned} f_{i:n}(x) &= \sum_{j=0}^{n-i} (-1)^j \frac{g(x) G(x)^{a-1} [(1 - G(x)]^{b-1}}{B(a,b)^{i+j} B(i, n-i+1)} \binom{n-i}{j} \sum_{r=0}^{\infty} c_{i+j-1,r} G(x)^r \\ &= \sum_{j=0}^{n-i} \sum_{r=0}^{\infty} (-1)^j\, c_{i+j-1,r} \binom{n-i}{j} \frac{g(x)[(1 - G(x)]^{b-1} G(x)^{r+a-1}}{B(a,b)^{i+j} B(i, n-i+1)}, \end{aligned} \tag{C4}$$

where

$$c_{i+j-1,r} = (r t_0)^{-1} \sum_{m=1}^{r} ((i+j)m - r) t_m c_{i+j-1,r-m}. \tag{C5}$$

Equation (C4) can be written as

$$f_{i:n}(x) = \sum_{j=0}^{n-i} \sum_{r=0}^{\infty} (-1)^j c_{i+j-1,r} \binom{n-i}{j} \frac{g(x)[(1 - G(x)]^{b-1} [1 - (1 - G(x))]^{r+a-1}}{B(a,b)^{i+j} B(i, n-i+1)}.$$

For any $q > 0$ real, we have

$$G(x)^q = [1 - \{1 - G(x)\}]^q = \sum_{k=0}^{\infty} (-1)^k \binom{q}{k} [1 - G(x)]^k, \tag{C6}$$

and then

$$f_{i:n}(x) = \sum_{j=0}^{n-i} \sum_{r=0}^{\infty} g(x) \sum_{k=0}^{\infty} (-1)^k \binom{r+a-1}{k} [1-G(x)]^{k+b-1}.$$

In the same way, using equation (C6), it follows that

$$f_{i:n}(x) = \sum_{j=0}^{n-i} \sum_{r,k,l=0}^{\infty} \frac{(-1)^{j+k+l} \binom{n-i}{j} \binom{r+a-1}{k} \binom{k+b-1}{l} c_{i+j-1,r}}{B(a,b)^{i+j} \, B(i, n-i+1)} \, g(x) \, G(x)^l.$$

Replacing equations (3) and (4) in the above equation, $f_{i:n}(x)$ can be expressed as an infinite linear combination of WP density functions

$$\begin{aligned} f_{i:n}(x) &= \sum_{j=0}^{n-i} \sum_{r,k,l=0}^{\infty} \frac{(-1)^{j+k+l} \binom{n-i}{j} \binom{r+a-1}{k} \binom{k+b-1}{l} c_{i+j-1,r}}{B(a,b)^{i+j} B(i, n-i+1)} \left[\frac{\alpha\beta\lambda \mathrm{e}^{-\lambda}}{1-\mathrm{e}^{-\lambda}} \, u \, x^{\alpha-1} \, \mathrm{e}^{\lambda u} \right] \left[\frac{\mathrm{e}^{\lambda u} - \mathrm{e}^{\lambda}}{1-\mathrm{e}^{\lambda}} \right]^l \\ &= \sum_{j=0}^{n-i} \sum_{r,k,l=0}^{\infty} \frac{(-1)^{j+k+l} \binom{n-i}{j} \binom{r+a-1}{k} \binom{k+b-1}{l} c_{i+j-1,r}}{B(a,b)^{i+j} \, B(i, n-i+1)(1-\mathrm{e}^{\lambda})^l} \left[\frac{\alpha\beta\lambda \mathrm{e}^{-\lambda}}{1-\mathrm{e}^{-\lambda}} \, u \, x^{\alpha-1} \, \mathrm{e}^{\lambda u} \right] \\ &\times \sum_{s=0}^{l} (-1)^s \binom{l}{s} (\mathrm{e}^{\lambda u})^{l-s} \mathrm{e}^{s\lambda}. \end{aligned} \tag{C7}$$

Equation (C7) reduces to

$$f_{i:n}(x) = \sum_{l=0}^{\infty} \sum_{s=0}^{l} \gamma_{i:n}(l,s) \; g(x; \alpha, \beta, \lambda_{l,s}), \tag{C8}$$

where $\lambda_{l,s} = \lambda(l-s+1)$ and

$$\gamma_{i:n}(l,s) = \sum_{j=0}^{n-i} \sum_{r,k=0}^{\infty} \frac{(-1)^{l+s+j+k} \, \mathrm{e}^{s\lambda} \binom{l}{s} \binom{n-i}{j} \binom{r+a-1}{k} \binom{k+b-1}{l} (1-\mathrm{e}^{-\lambda_{l,s}}) \, c_{i+1-j,r}}{(l-s+1)(1-\mathrm{e}^{-\lambda})(1-\mathrm{e}^{\lambda})^l \, B(a,b)^{i+j} \, B(i, n-i+1)}.$$

Appendix D. Information Matrix

Let $u_i = \exp(-\beta x_i^{\alpha})$. The elements of the observed information matrix $J(\boldsymbol{\theta})$ for the parameters $(\alpha, \beta, \lambda, a, b)$ are

$$J_{\alpha\alpha} = -\frac{n}{\alpha^2} - \beta\sum_{i=1}^{n} x_i^{\alpha}\log^2(x_i) - \lambda\beta\sum_{i=1}^{n} u_i x_i^{\alpha}\log^2(x_i) + \lambda\beta^2\sum_{i=1}^{n} x_i^{2\alpha}u_i\log^2(x_i)$$
$$+(a-1)\sum_{i=1}^{n}\left[\frac{\lambda\beta u_i x_i^{\alpha}\mathrm{e}^{\lambda u_i}\log^2(x_i)}{\mathrm{e}^{\lambda u_i}-\mathrm{e}^{\lambda}}\right]\psi(x_i)$$
$$+(b-1)\sum_{i=1}^{n}\left[\frac{\lambda\beta u_i x_i^{\alpha}\mathrm{e}^{\lambda u_i}\log^2(x_i)}{1-\mathrm{e}^{\lambda u_i}}\right]\varphi(x_i),$$

$$J_{\alpha\beta} = J_{\beta\alpha} = -\sum_{i=1}^{n} x_i^{\alpha}\log(x_i) - \lambda\sum_{i=1}^{n} u_i x_i^{\alpha}\log(x_i) + \lambda\beta\sum_{i=1}^{n} u_i x_i^{2\alpha}\log(x_i)$$
$$+(a-1)\sum_{i=1}^{n}\left[\frac{\lambda u_i x_i^{\alpha}\mathrm{e}^{\lambda u_i}\log(x_i)}{\mathrm{e}^{\lambda u_i}-\mathrm{e}^{\lambda}}\right]\psi(x_i)$$
$$+(b-1)\sum_{i=1}^{n}\left[\frac{\lambda u_i x_i^{\alpha}\mathrm{e}^{\lambda u_i}\log(x_i)}{1-\mathrm{e}^{\lambda u_i}}\right]\varphi(x_i),$$

where

$$\psi(x_i) = \left(-1 + \beta x_i^{\alpha} + \lambda\beta u_i x_i^{\alpha} - \frac{\lambda\beta u_i x_i^{\alpha}\mathrm{e}^{\lambda u_i}}{\mathrm{e}^{\lambda u_i}-\mathrm{e}^{\lambda}}\right)$$

and

$$\varphi(x_i) = \left(1 - \beta x_i^{\alpha} - \lambda\beta u_i x_i^{\alpha} - \frac{\lambda\beta u_i x_i^{\alpha}\mathrm{e}^{\lambda u_i}}{1-\mathrm{e}^{\lambda u_i}}\right).$$

Further,

$$J_{\alpha\lambda} = J_{\lambda\alpha} = -\beta\sum_{i=1}^{n} u_i x_i^{\alpha}\log(x_i) + (a-1)\sum_{i=1}^{n}\rho(x_i)\Bigg[-1-\lambda u_i$$
$$+\frac{\lambda(u_i\mathrm{e}^{\lambda u_i}-\mathrm{e}^{\lambda})}{\mathrm{e}^{\lambda u_i}-\mathrm{e}^{\lambda}}\Bigg] + (b-1)\sum_{i=1}^{n}\phi(x_i)\left(1+\lambda u_i + \frac{\lambda u_i\mathrm{e}^{\lambda u_i}}{1-\mathrm{e}^{\lambda u_i}}\right),$$

$$J_{\alpha a} = J_{a\alpha} = -\sum_{i=1}^{n}\lambda\rho(x_i), \qquad J_{\alpha b} = J_{b\alpha} = \sum_{i=1}^{n}\lambda\phi(x_i),$$

where $\rho(x_i) = \dfrac{\beta u_i x_i^{\alpha}\mathrm{e}^{\lambda u_i}\log(x_i)}{\mathrm{e}^{\lambda u_i}-\mathrm{e}^{\lambda}}$ and $\phi(x_i) = \dfrac{\beta u_i x_i^{\alpha}\mathrm{e}^{\lambda u_i}\log(x_i)}{1-\mathrm{e}^{\lambda u_i}}$,

$$J_{\beta\beta} = -\frac{n}{\beta^2} + \lambda \sum_{i=1}^{n} u_i x_i^{2\alpha} + (a-1)\sum_{i=1}^{n}\left(\frac{\lambda u_i (x_i^{\alpha})^2 \mathrm{e}^{\lambda u_i}}{\mathrm{e}^{\lambda u_i} - \mathrm{e}^{\lambda}}\right)$$
$$\times\left(1 + \lambda u_i - \frac{\lambda u_i \mathrm{e}^{\lambda u_i}}{\mathrm{e}^{\lambda u_i} - \mathrm{e}^{\lambda}}\right) + (b-1)\sum_{i=1}^{n}\left(\frac{\lambda u_i x_i^{2\alpha}\mathrm{e}^{\lambda u_i}}{1 - \mathrm{e}^{\lambda u_i}}\right)$$
$$\times\left(-1 - \lambda u_i - \frac{\lambda u_i \mathrm{e}^{\lambda u_i}}{1 - \mathrm{e}^{\lambda u_i}}\right),$$

$$J_{\beta\lambda} = J_{\lambda\beta} = -\sum_{i=1}^{n} u_i x_i^{\alpha} + (a-1)\sum_{i=1}^{n}\gamma(x_i)\left[-1 - \lambda u_i + \frac{\lambda(u_i \mathrm{e}^{\lambda u_i} - \mathrm{e}^{\lambda})}{\mathrm{e}^{\lambda u_i} - \mathrm{e}^{\lambda}}\right]$$
$$+(b-1)\sum_{i=1}^{n}\delta(x_i)\left(1 + \lambda u_i + \frac{\lambda u_i \mathrm{e}^{\lambda u_i}}{1 - \mathrm{e}^{\lambda u_i}}\right),$$

$$J_{\beta a} = J_{a\beta} = -\sum_{i=1}^{n}\lambda\gamma(x_i), \qquad J_{\beta b} = J_{b\beta} = \sum_{i=1}^{n}\lambda\delta(x_i),$$

$\gamma(x_i) = \dfrac{u_i x_i^{\alpha}\mathrm{e}^{\lambda u_i}}{\mathrm{e}^{\lambda u_i} - \mathrm{e}^{\lambda}}$ and $\delta(x_i) = \dfrac{u_i x_i^{\alpha}\mathrm{e}^{\lambda u_i}}{1 - \mathrm{e}^{\lambda u_i}}$. Furthermore,

$$J_{\lambda\lambda} = -\frac{n}{\lambda^2} + \frac{n\mathrm{e}^{-\lambda}}{1 - \mathrm{e}^{-\lambda}} + \frac{n\mathrm{e}^{-2\lambda}}{(1 - \mathrm{e}^{-\lambda})^2} - \frac{n(a+b-2)\mathrm{e}^{\lambda}\log(1 - \mathrm{e}^{\lambda})^{a+b-2}}{1 - \mathrm{e}^{\lambda}\log(1 - \mathrm{e}^{\lambda})}$$
$$\times\left[1 - \frac{(a+b-2)\mathrm{e}^{\lambda}}{(1 - \mathrm{e}^{\lambda})\log(1 - \mathrm{e}^{\lambda})} + \frac{\mathrm{e}^{\lambda}}{1 - \mathrm{e}^{\lambda}} + \frac{\mathrm{e}^{\lambda}}{(1 - \mathrm{e}^{\lambda})\log(1 - \mathrm{e}^{\lambda})}\right]$$
$$+(a-1)\sum_{i=1}^{n}\left[\frac{u_i \mathrm{e}^{\lambda u_i} - \mathrm{e}^{\lambda}}{\mathrm{e}^{\lambda u_i} - \mathrm{e}^{\lambda}} - \frac{(u_i \mathrm{e}^{\lambda u_i} - \mathrm{e}^{\lambda})^2}{(\mathrm{e}^{\lambda u_i} - \mathrm{e}^{\lambda})^2}\right]$$
$$-(b-1)\sum_{i=1}^{n}\left[\frac{u_i^2 \mathrm{e}^{\lambda u_i}}{1 - \mathrm{e}^{\lambda u_i}} + \frac{u_i^2(\mathrm{e}^{\lambda u_i})^2}{(1 - \mathrm{e}^{\lambda u_i})^2}\right],$$

$$J_{\lambda a} = J_{a\lambda} = \left[\frac{n\mathrm{e}^{\lambda}\log(1 - \mathrm{e}^{\lambda})^{a+b-2}}{(1 - \mathrm{e}^{\lambda})\log(1 - \mathrm{e}^{\lambda})}\right]\Big\{1 + (a+b-2)\log[\log(1 - \mathrm{e}^{\lambda})]\Big\}$$
$$+\sum_{i=1}^{n}\left(\frac{u_i \mathrm{e}^{\lambda u_i} - \mathrm{e}^{\lambda}}{\mathrm{e}^{\lambda u_i} - \mathrm{e}^{\lambda}}\right),$$

$$J_{\lambda b} = J_{b\lambda} = \left[\frac{n\mathrm{e}^{\lambda}\log(1 - \mathrm{e}^{\lambda})^{a+b-2}}{(1 - \mathrm{e}^{\lambda})\log(1 - \mathrm{e}^{\lambda})}\right]\Big\{1 + (a+b-2)\log[\log(1 - \mathrm{e}^{\lambda})]\Big\}$$
$$+\sum_{i=1}^{n}\left(\frac{u_i \mathrm{e}^{\lambda u_i}}{1 - \mathrm{e}^{\lambda u_i}}\right),$$

$$J_{aa} = -\frac{n\ddot{B}_a(a,b)}{B(a,b)} + \frac{n\left[\dot{B}_a(a,b)\right]^2}{B(a,b)} - n\log(1-\mathrm{e}^{\lambda})^{a+b-2} \times \log^2[\log(1-\mathrm{e}^{\lambda})],$$

$$J_{ab} = J_{ba} = -\frac{n\ddot{B}(a,b)}{B(a,b)} + \frac{n\dot{B}_a(a,b)\dot{B}_b(a,b)}{[B(a,b)]^2} - n\log(1-\mathrm{e}^{\lambda})^{a+b-2} \times \log^2[\log(1-\mathrm{e}^{\lambda})],$$

$$J_{bb} = -\frac{n\ddot{B}_b(a,b)}{B(a,b)} + \frac{n\left[\dot{B}_b(a,b)\right]^2}{[B(a,b)]^2} - n\log(1-\mathrm{e}^{\lambda})^{a+b-2} \times \log^2[\log(1-\mathrm{e}^{\lambda})],$$

where $\dot{B}_a(a,b) = \frac{\partial}{\partial a}B(a,b)$, $\dot{B}_b(a,b) = \frac{\partial}{\partial b}B(a,b)$ and $\ddot{B}(a,b) = \frac{\partial^2}{\partial b \partial a}B(a,b)$.

References

Adamidis, K., Loukas, S., 1998. A lifetime distribution with decreasing failure rate, Statistics & Probability Letters, 39, 35-42.

Alven, W.H., 1964. Reliability Engineering by *ARINC*. Prentice-hall, New Jersey.

Barreto-Souza, W., de Morais, A.L., Cordeiro, G.M, 2010. Weibull-geometric distribution, Journal of Statistical Computation and Simulation, 81, 645-657.

Chhikara, R.S., Folks,J.L., 1977. The inverse Gaussian distribution as a lifetime model, Technometrics, 19, 461-468.

Cordeiro, G.M., Lemonte, A.J., 2011. The β-Birnbaum-Saunders distribution: An improved distribution for fatigue life modeling, Computational Statistics & Data Analysis, 55, 1445-1461.

Cordeiro, G.M., Silva, G.O., Ortega, E.M.M., 2011. The beta-Weibull geometric distribution, A Journal of Theoretical and Applied Statistics, 0, 1-18

Dimitrakopoulou, T., Adamidis, K., Loukas, S., 2007. A lifetime distribution with an upside-down bathtub-shaped hazard function, IEEE Transactions on Reliability, 56, 308-311.

Eugene, N., Lee, C., Famoye, F., 2002. Beta-normal distribution and its applications, Communication in Statistics - Theory and Methods, 31, 497-512.

Gradshteyn, I.S., Ryzhik, I.M., 2000. Table of Integrals, Series and Products. Academic Press, San Diego.

Gupta, R.D., Kundu, D., 1999. Generalized exponential distribution, Australian & New Zealand Journal of Statistics, 41, 173-188.

Gupta, R.D., Kundu, D., 2001. Generalized exponential distribution: Different method of estimations, Journal of Statistical Computation and Simulation, 69, 315-337.

Gupta, R.D., Kundu, D., 2001. Exponentiated exponential family: An alternative to gamma and Weibull distributions, Biometrical Journal, 43, 117-130.

Kus, C., 2007. A new lifetime distribution. Computational Statistics & Data Analysis, 51, 4497-4509.

Lee, C., Famoye F., Olumolade, O., 2007. Beta-Weibull Distribution: Some properties and applications to censored data. Journal of Modern Applied Statistical Methods, 6, 173-186.

Lu, W., Shi, D., 2012. A new compounding life distribution: the Weibull-Poisson distribution, Journal of Applied Statistics, 39, 21-38.

Nadarajah, S., Kotz, S., 2004. The beta Gumbel distribution, Mathematical Problems in Engineering, 10, 323-332.

Nadarajah, S., Gupta, A.K., 2004. The beta Fréchet distribution, Far East Journal of Theoretical Statistics, 14, 15-24.

Nadarajah, S., Kotz, S., 2006. The beta exponential distribution, Reliability Engineering and System Safety, 91, 689-697.