A new compounded four-parameter lifetime model: Properties, cure rate model and applications

Arslan Nasir, Farrukh Jamal, Christophe Chesneau, Akbar Ali Shah

A new compounded four-parameter lifetime model: Properties, cure rate model and applications

M. Arslan Nasir[1], Farrukh Jamal[2], Christophe Chesneau[3] and Akbar Ali Shah[4]

[1,2,4]Department of Statistics, The Islamia University Bahawalpur, Pakistan.
[3]LMNO, University of Caen-Normandie, France.
[1]arslannasir147@gmail.com [2]drfarrukh1982@gmail.com
[3]christophe.chesneau@gmail.com
[4]akbar.ali@iub.edu.pk

March 7, 2019

Abstract

We propose a new four-parameter lifetime distribution obtained by compounding two useful distributions: the Weibull and Burr XII distributions. Among interesting features, it shows a great flexibility with respect to its crucial functions shapes; the probability density function can exhibit unimodal (symmetrical and right-skewed), bimodal and decreasing shapes, and the hazard rate function can accommodate increasing, decreasing, bathtub, upside-down bathtub and decreasing-increasing-decreasing shapes. Some mathematical properties of the new distribution are obtained such as the quantiles, moments, generating function, stress-strength reliability parameter and stochastic ordering. The maximum likelihood estimation is employed to estimate the model parameters. A Monte Carlo simulation study is carried out to assess the performance of the maximum likelihood estimates. We also propose a flexible cure rate survival model by assuming that the number of competing causes of the event of interest has the Poisson distribution and the time for the event follows the proposed distribution. Four empirical illustrations of the new distribution are presented to real-life data sets. The results of the proposed model are better in comparison to those obtained with the exponential-Weibull, odd Weibull-Burr and Weibull-Lindley models.

Keywords— Weibull distribution; Burr distribution; compounding; maximum likelihood method; Poisson distribution; cure rate model.

2000 Mathematics Subject Classification:— 60E05; 62N05; 62F10

1 Introduction

In probability theory, the act of compounding often refers to the mixing or joining of two distributions with possible different nature, that is, (i) discrete with discrete, (ii) discrete with continuous, and (iii) continuous and continuous. From a statistical point of view, this mixing may leads to an increase flexibility in the subjacent model, which can accommodate a variety of data sets generated from simple to complex phenomenons. A complete survey addressing trends in compounding can be found in [22]. The objective of this paper is to introduce a new promising distribution defined by mixing two useful continuous univariate distributions: the Weibull and Burr XII distributions. We thus follow the spirit of [9] with the exponential-Weibull (EW) distribution and [3] with the Weibull-Lindley (WL) distribution, both showing remarkable properties in the modelling real life data sets of various kinds.

Let us now present the distribution of interest, with motivations. Let Y be a random variable following the Burr distribution with parameters $c > 0$ and $k > 0$, i.e. having the cumulative distribution function (cdf) given by $G_B(y) = 1 - (1 + y^c)^{-k}$, $y > 0$ (recall that the associated survival function (sf) is given by $\bar{G}_B(y) = 1 - G_B(y) = (1 + y^c)^{-k}$ and the hazard rate function (hrf) is given by $h_B(y) = cky^{c-1}(1 + y^c)^{-1}$). Let Z be a random variable following the Weibull distribution with parameters $a > 0$ and $b > 0$, i.e. having the cdf given by $G_W(z) = 1 - \exp(-az^b)$, $z > 0$ (recall that the associated sf is given by $\bar{G}_W(z) = 1 - G_W(z) = \exp(-az^b)$ and the associated hrf is given by $h_W(z) = abz^{b-1}$). Consider that Y and Z are independent random variables. We define the minimum Weibull-Burr (minWB) distribution by the distribution of the random variable X defined by the minimum of Y and Z, i.e. $X = \min\{Y, Z\}$. Hence the cdf of the minWB distribution is given by

$$\begin{aligned} F(x) &= 1 - \bar{G}_B(x)\bar{G}_W(x) \\ &= 1 - (1 + x^c)^{-k} \exp(-ax^b), \quad x > 0, \end{aligned} \tag{1.1}$$

where $a > 0$ is the scale parameter, and $b > 0$, $c > 0$ and $k > 0$ are shape parameters. The associated sf, probability density function (pdf) and hrf of the minWB distribution are respectively given by $S(x) = (1 + x^c)^{-k} \exp(-ax^b)$,

$$f(x) = (1 + x^c)^{-k} \exp(-ax^b) \left[abx^{b-1} + ckx^{c-1} (1 + x^c)^{-1}\right] \tag{1.2}$$

and

$$h(x) = abx^{b-1} + ckx^{c-1} (1 + x^c)^{-1}. \tag{1.3}$$

Henceforth, a random variable X with pdf (1.2) is denoted by $X \sim minWB(a, b, c, k)$. Some special distributions of the minWB distribution are: (i) Weibull when $k = 0$, (ii) Burr when $a = 0$, (iii) Weibull-log-logistic (WLL) when $k = 1$, (iv) Rayleigh-Burr (RB) when $b = 2$, (v) exponential-Burr (EB) when $b = 1$, and (vi) Weibull-Lomax (WLx) when $c = 1$.

The major motivations for the new distribution are fivefold: (i) the cdf of the minWB distribution is quite simple, giving simple expressions for the pdf, sf and hrf; (ii) the new distribution is very flexible with respect to the pdf and hrf shapes. In particular, the possible pdf shapes are decreasing, unimodal (right-skewed or symmetrical) and bimodal. This means that the minWB model can show suitable fit to those data sets, whose histograms are similar to the minWB pdf shapes. Furthermore, the minWB distribution exhibits monotone [increasing (IFR) and decreasing (DFR)], non-monotone [bathtub (BT) and upside-down bathtub (UBT)] and decreasing-increasing-decreasing (DID) failure rate shapes to cope with all types of lifetime data sets; (iii) the WLL, WLx, RB, EB, Weibull and Burr distributions are special cases of the proposed distribution; (iv) the minWB pdf shows bimodal feature as well; (v) suppose a system has two sub-systems functioning in series independently at a given time, so that the system will fail when the first sub-system fails. Consider that the failure times of the sub-systems follow the Weibull and Burr distributions. Then, the time-to-failure of the system has cdf (1.1).

The paper is unfolded as follows. In Section 2, we derive some mathematical properties of minWB distribution including shapes of pdf and hrf, quantiles, moments, mean deviations, generating function, stress-strength reliability parameter and stochastic ordering. In Section 3, the model parameters are estimated by maximum likelihood and a simulation study is performed. In Section 4, we formulate the Poisson-Weibull-Burr regression model with cure fraction by defining the pdf, cdf and hrf. In Section 5, the usefulness of the new distribution is illustrated by means of four real data sets, where we prove empirically that our proposed model outperforms some well-known lifetime distributions. Finally, Section 6 offers some concluding remarks.

2 Properties of the minWB distribution

Hereafter, X denotes a random variable such that $X \sim minWB(a, b, c, k)$, $F(x)$ is the cdf of X given by (1.1), $f(x)$ is the pdf of X given by (1.2) and $h(x)$ is the hrf of X given by (1.3).

2.1 Quantile

For $p \in (0, 1)$, the pth quantile of X, say x_p, is defined by $F(x_p) = p$. It is then the root of

$$x_p = \left\{ \left[(1-p) \exp(a x_p^b) \right]^{-1/k} - 1 \right\}^{1/c}. \tag{2.1}$$

For a given p, it can be computed numerically. This is the case for the median of X defined by $M = x_{1/2}$.

2.2 Asymptotics

The asymptotics of the cdf, pdf and hrf of X when $x \to 0$ are given below.

- $F(x) \sim a x^b$.
- For $f(x)$, we must distinguish the cases $b > c$, $b = c$ and $c > b$. If $b > c$ then $f(x) \sim ckx^{c-1}$ (note that it tends to $+\infty$ if $c \in (0, 1)$, k if $c = 1$ and 0 if $c > 1$). Moreover, if $c > b$ then $f(x) \sim abx^{b-1}$, if $b = c$ then $f(x) \sim (ab + ck)x^{c-1}$.
- For $h(x)$, we have $h(x) \sim f(x)$, and the results aboves are still valid.

The asymptotics of the cdf, pdf and hrf of X when $x \to +\infty$ are given below.

- $1 - F(x) \sim x^{-kc} \exp(-ax^b)$.
- $f(x) \sim abx^{-kc+b-1} \exp(-ax^b)$ (note that it tends to 0 in all case).
- $h(x) \sim abx^{b-1}$ (note that it tends to 0 if $b \in (0, 1)$, a if $b = 1$ and $+\infty$ if $b > 1$).

These results show the effects of the parameters on the tails of the minWB distribution.

2.3 Shapes of the pdf

The critical points of the minWB pdf are the roots of the equation $df(x)/dx = 0$, i.e., after calculus,

$$\begin{aligned} &a^2 b^2 x^{2b} (x^c + 1)^2 - abx^b (x^c + 1) \left[x^c (b - 2ck - 1) + b - 1 \right] \\ &+ ckx^c \left[c(kx^c - 1) + x^c + 1 \right] = 0. \end{aligned} \tag{2.2}$$

By using any numerical software, we can examine Equation (2.2) to determine the local maximum and minimum and inflexion points.

Figures 1 and 2 display some plots of the minWB pdf for selected values of a, b, c and k (some of them are fixed, the others differ). The plots in Figures 1 and 2 reveal that the shapes of the minWB pdf are decreasing, unimodal (right-skewed or symmetric) and bimodal.

2.4 Shapes of the hrf

The critical points of the minWB hrf are obtained from the equation $dh(x)/dx = 0$, i.e., after calculus,

$$a(b-1)bx^{b-2}(x^c+1)^2 - ck\left[(x^c+1)x^{c-2} + c(x^{2c-2} - (x^c+1)x^{c-2})\right] = 0. \quad (2.3)$$

By using any numerical software, we can examine Equation (2.3) to determine the local maximum and minimum and inflexion points.

Figures 3 and 4 display some plots of the minWB hrf for different values of a, b, c and k (some of them are fixed, the others differ). These plots indicate that the hazard rate shapes of the minWB model are IFR, DFR, BT, UBT and DID.

(a) (b) (c)

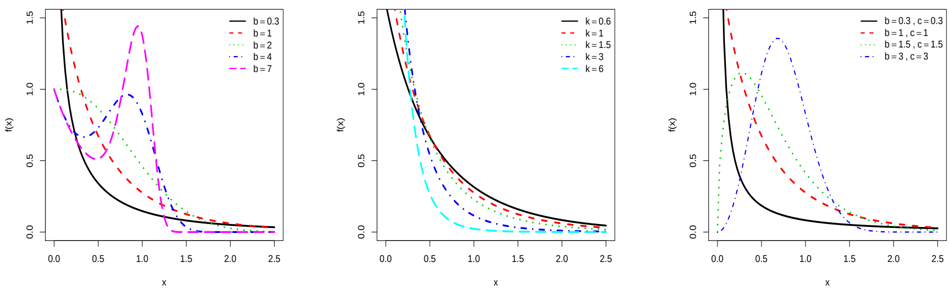

Figure 1: Plots of the minWB pdf for some parameter values.

(a) (b) (c)

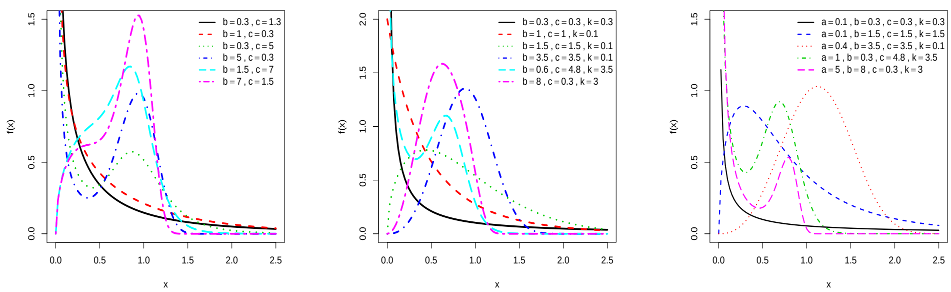

Figure 2: Plots of the minWB pdf for some parameter values.

2.5 Moments and generating function

First, we obtain general expressions for the following two integrals, which are used to determine some structural properties of the minWB distribution. There are no closed-form expressions for the integrals and then they can be computed numerically.

Integral 1: We define the integral $\mathbb{J}_1(p, c, k, a, b)$ by

$$\mathbb{J}_1(p, c, k, a, b) = \int_0^{+\infty} x^p \,(1 + x^c)^{-k} \exp(-ax^b) dx.$$

(a) (b) (c)

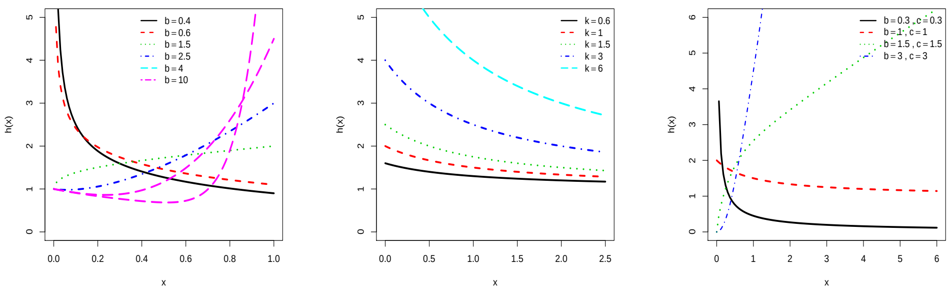

Figure 3: Plots of the minWB hrf for some parameter values.

(a) (b) (c)

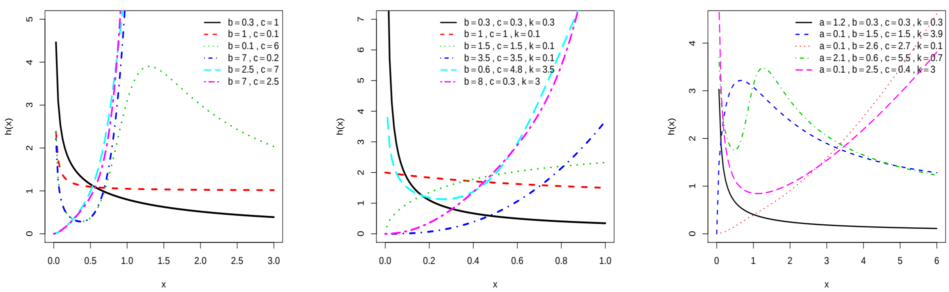

Figure 4: Plots of the minWB hrf for some parameter values.

Let us now determine a series expansions for this integral. Using the generalized binomial expansion, we have

$$(1+x^c)^{-k} = \sum_{i=0}^{+\infty} \binom{-k}{i} x^{ic}\mathbf{1}_{(0,1)}(x) + 2^{-k}\mathbf{1}_{\{1\}}(x) + \sum_{i=0}^{+\infty} \binom{-k}{i} x^{-c(i+k)}\mathbf{1}_{(1+\infty)}(x),$$

where $\mathbf{1}_A(x)$ denotes the indicator function over a given set of real numbers A, i.e. $\mathbf{1}_A(x) = 1$ if $x \in A$ and $\mathbf{1}_A(x) = 0$ elsewhere, and $\binom{-k}{i}$ is the (generalized) binomial coefficient defined by $\binom{-k}{i} = (-k)(-k-1)\ldots(-k-i+1)/i!$ and $\binom{-k}{0} = 1$.

By the change of variable $y = ax^b$, we can write

$$\mathbb{J}_1(p,c,k,a,b) = \sum_{i=0}^{+\infty} \alpha_i \gamma\Big(\frac{p+ic+1}{b}, a\Big) + \sum_{i=0}^{+\infty} \alpha_i^* \Gamma\Big(\frac{p-c(i+k)+1}{b}, a\Big), \tag{2.4}$$

where $\gamma(m,x) = \int_0^x t^{m-1}e^{-t}dt$, $\Gamma(m,x) = \int_x^{+\infty} t^{m-1}e^{-t}dt$, $m > 0$, $\alpha_i = b^{-1}\binom{-k}{i}a^{-(p+ic+1)/b}$ and $\alpha_i^* = b^{-1}\binom{-k}{i}a^{-(p-c(i+k)+1)/b}$.

Integral 2: Let $y > 0$. We define the integral function $\mathbb{J}_2(y; p, c, k, a, b)$ (function according to y) by

$$\mathbb{J}_2(y;p,c,k,a,b) = \int_0^y x^p\,(1+x^c)^{-k}\exp(-ax^b)dx.$$

By using algebraic developments similar to those used for the previous integral, we can write

$$\mathbb{J}_2(y;p,c,k,a,b) = \begin{cases} \displaystyle\sum_{i=0}^{+\infty} \alpha_i \gamma\Big(\frac{p+ic+1}{b}, ay^b\Big) \mathbf{1}_{(0,1)}(y) \\ \displaystyle\sum_{i=0}^{+\infty} \alpha_i \gamma\Big(\frac{p+ic+1}{b}, a\Big) + \sum_{i=0}^{+\infty} \alpha_i^* \Gamma\Big(\frac{p-c(i+k)+1}{b}, ay^b\Big) \mathbf{1}_{(1,+\infty)}(y). \end{cases} \tag{2.5}$$

Equations (2.4) and (2.5) are the main tools to obtain some mathematical properties of the minWB distribution such as the ordinary and incomplete moments, mean deviations and moment generating function.

The nth ordinary moment of X can be determined from Equations (1.2) and (2.4) as

$$\mu_n' = \int_0^{+\infty} x^n f(x)dx = ab\mathbb{J}_1\Big(n+b-1, c, k, a, b\Big) + ck\mathbb{J}_1\Big(n+c-1, c, k+1, a, b\Big). \tag{2.6}$$

Hence, the mean of X is given by μ_1' and the variance of X is given by $V = \mu_2' - (\mu_1')^2$. The central moments of X can follow from Equation (2.6). Indeed, we have

$$\mu_s = \sum_{k=0}^{p} \binom{s}{k} (-1)^k \mu_1'^s \mu_{s-k}'.$$

Similarly, the cumulants of X are given by the recursive equation: $\kappa_s = \mu_s' - \sum_{k=1}^{s-1} \binom{s-1}{k-1} \kappa_k \mu_{s-k}'$, with as initial value: $\kappa_1 = \mu_1'$. The skewness and kurtosis of X can be calculated from the third and fourth standardized cumulants. Indeed, they are respectively given by

$$\gamma_1 = \frac{\kappa_3}{\kappa_2^{3/2}}, \qquad \gamma_2 = \frac{\kappa_4}{\kappa_2^2}.$$

The rth incomplete moment of X can be determined from Equations (1.2) and (2.5) as

$$m_r(y) = \int_0^{y} x^r f(x)dx = ab\mathbb{J}_2\Big(y; r+b-1, c, k, a, b\Big) + ck\mathbb{J}_2\Big(y; r+c-1, c, k+1, a, b\Big). \tag{2.7}$$

The Bonferroni and Lorenz curves, useful in several fields, involve the first incomplete moment. For a given $\pi \in (0,1)$, they are given by $B(\pi) = m_1(q)/(\pi\mu_1')$ and $L(\pi) = m_1(q)/\mu_1'$, respectively, where $m_1(q)$ comes from Equation (2.7) with $r = 1$ and $q = Q(\pi)$ follows from Equation (2.1).

The amount of scatter in a population is measured to some extent by the totality of deviations from the mean and median defined by $\delta_1 = \int_0^{+\infty} |x-\mu_1'| f(x)dx$ and $\delta_2(x) = \int_0^{+\infty} |x-M| f(x)dx$, respectively. These measures can be expressed as $\delta_1 = 2\mu_1' F(\mu_1') - 2m_1(\mu_1')$ and $\delta_2 = \mu_1' - 2m_1(M)$, where $F(\mu_1')$ is given by Equation (1.1).

The moment generating function of X can be expressed as

$$M(t) = \int_0^{+\infty} \exp(tx)\exp(-ax^b)(1+x^c)^{-k}\left[abx^{b-1} + ckx^{c-1}(1+x^c)^{-1}\right] dx, \qquad t < 0.$$

By expanding $\exp(-ax^b)$ in power series and using Equation (2.5), we obtain

$$M(t) = \sum_{i=0}^{+\infty} \frac{(-a)^i}{i!} \Big[ab\mathbb{J}_1\Big(b(i+1)-1, c, k, -t, 1\Big) + ck\mathbb{J}_1\Big(bi+c-1, c, k+1, -t, 1\Big)\Big].$$

Table 1 provides the values of the mean, median, variance, skewness and kurtosis of X for selected values of a, b, c and k. One can observe that the mean and variance of the minWB model are decreasing functions of a and b and increasing functions of c and k. Also, these values indicate that the minWB distribution can be left-skewed or right-skewed.

Table 1: Mean, median, variance, skewness and kurtosis of X for some combinations. I: a, b=5.0, c=1.3, k=0.6, II: a=0.5, b, c=0.3, k=1.6, III: a=1.5, b=1.5, c, k=2.6, IV: a=2.5, b=0.5, c=2.5, k.

	Varying Parameter ↓	Mean μ_1'	Median M	Variance V	Skewness γ_1	Kurtosis γ_2
I	a					
	0.5	0.8505	0.9104	0.1361	-0.4228	2.3799
	1.5	0.7123	0.7580	0.0806	-0.5181	2.6236
	5.0	0.5819	0.6143	0.0451	-0.5965	2.8988
	25.0	0.4390	0.4586	0.0208	-0.6539	3.2271
II	b					
	0.5	1.5241	0.0595	38.3800	13.6801	39.1397
	1.5	0.5237	0.1158	0.6336	2.0954	8.0115
	5.0	0.4274	0.1299	0.2392	0.6878	1.9001
	25.0	0.4137	0.1387	0.2041	0.4431	1.3432
III	c					
	0.5	0.3201	0.3008	0.0999	2.3395	9.9112
	1.5	0.3805	0.3165	0.0792	1.4144	5.9816
	5.0	0.5306	0.5301	0.0702	0.0341	2.4914
	25.0	0.5903	0.5977	0.0892	-0.1702	1.7217
IV	k					
	0.5	0.2528	0.0766	0.2033	4.6483	48.6909
	1.5	0.2059	0.0763	0.0910	2.7029	15.1240
	5.0	0.1593	0.0751	0.0384	1.6819	6.0251
	25.0	0.1086	0.0700	0.0126	1.1186	3.3692

2.6 Stress-strength reliability

The reliability parameter R is defined as $R = \mathbb{P}(X_1 > X_2)$, where X_1 and X_2 are independent random variables. An amount of applications of this parameter have been investigated in the literature (such as the area of classical stress-strength model, the breakdown of a system having two components...). Let us now study it in the context of the minWD distribution. Let $X_1 \sim minWB(a_1, b, c, k_1)$ and $X_2 \sim minWB(a_2, b, c, k_2)$ with cdfs denoted by $F_1(x)$ and $F_2(x)$ and pdfs $f_1(x)$ and $f_2(x)$, respectively, the reliability R is given by

$$R = \mathbb{P}(X_1 > X_2) = \int_0^{+\infty} f_1(x)F_2(x)dx. \tag{2.8}$$

Theorem 2.1 *Suppose that X_1 and X_2 are two independent random variables as defined above with fixed parameters b and c. Then we can express R as*

$$\begin{aligned} R &= 1 - a_1 b \mathbb{J}_1\Big(b-1, c, k_1+k_2, a_1+a_2, b\Big) \\ &\quad -ck_1\mathbb{J}_1\Big(c-1, c, k_1+k_2+1, a_1+a_2, b\Big). \end{aligned} \tag{2.9}$$

Proof of Theorem 2.1. Putting Equations (1.1) and (1.2) in Equation (2.8), we have

$$\int_0^{+\infty} f_1(x)F_2(x)dx = 1 - a_1 b \int_0^{+\infty} x^{b-1}(1+x^c)^{-(k_1+k_2)} \exp[-(a_1+a_2)x^b]dx$$
$$- ck_1 \int_0^{+\infty} x^{c-1}(1+x^c)^{-1}(1+x^c)^{-(k_1+k_2)} \exp[-(a_1+a_2)x^b]dx.$$

Equation (2.9) follows immediately after using Equation (2.4). □

2.7 Stochastic ordering

Stochastic ordering has been recognized as an important tool in reliability theory and other fields to assess comparative behavior. Here we present a stochastic ordering result related to the minWD distribution. Let X_1 and X_2 be two random variables having cdfs, sfs and pdfs $F_1(x)$ and $F_2(x)$, $\bar{F}_1(x) = 1 - F_1(x)$ and $\bar{F}_2(x) = 1 - F_2(x)$, and $f_1(x)$ and $f_2(x)$, respectively. The random variable X_1 is said to be smaller than X_2 in the following ordering as:

1. stochastic order (denoted by $X_1 \leq_{st} X_2$) if $\bar{F}_1(x) \leq \bar{F}_2(x)$ for all x;
2. likelihood ratio order (denoted by $X_1 \leq_{lr} X_1$) if $f_1(x)/f_2(x)$ is decreasing in $x \geq 0$;
3. hazard rate order (denoted by $X_1 \leq_{hr} X_2$) if $\bar{F}_1(x)/\bar{F}_2(x)$ is decreasing in $x \geq 0$;
4. reversed hazard rate order (denoted by $X_1 \leq_{rhr} X_2$) if $F_1(x)/F_2(x)$ is decreasing in $x \geq 0$.

All these four stochastic orders defined in (1)–(4) are related to each other due to [20] and the following implications hold:

$$(X_1 \leq_{rhr} X_2) \Leftarrow (X_1 \leq_{lr} X_2) \Rightarrow (X_1 \leq_{hr} X_2) \Rightarrow (X_1 \leq_{st} X_2).$$

The following theorem reveals that the minWB distributions are ordered with respect to strongest likelihood ratio ordering when appropriate assumptions hold.

Theorem 2.2 *Let $X_1 \sim minWB(a_1, b, c, k_1)$ and $X_2 \sim minWB(a_2, b, c, k_2)$. If $b > c$ and $k_1/k_2 > a_1/a_2$ then $X_1 \leq_{lr} X_2$.*

Proof of Theorem 2.2. First, we have

$$\frac{f_1(x)}{f_2(x)} = \left(\frac{\exp(-a_1x^b)(1+x^c)^{-k_1}}{\exp(-a_2x^b)(1+x^c)^{-k_2}}\right)\left[\frac{a_1bx^{b-1} + ck_1x^{c-1}(1+x^c)^{-1}}{a_2bx^{b-1} + ck_2x^{c-1}(1+x^c)^{-1}}\right].$$

After simplification, we get

$$\frac{f_1(x)}{f_2(x)} = \exp[-(a_1+a_2)x^b](1+x^c)^{-(k_1+k_2)}\left[\frac{a_1bx^{b-1} + ck_1x^{c-1}(1+x^c)^{-1}}{a_2bx^{b-1} + ck_2x^{c-1}(1+x^c)^{-1}}\right].$$

Next,

$$\begin{aligned}\log\left[\frac{f_1(x)}{f_2(x)}\right] &= -(a_1+a_2)x^b - (k_1+k_2)\log(1+x^c)\\ &\quad + \log\left[a_1bx^{b-1} + ck_1x^{c-1}(1+x^c)^{-1}\right]\\ &\quad - \log\left[a_2bx^{b-1} + ck_2x^{c-1}(1+x^c)^{-1}\right].\end{aligned}$$

Therefore

$$\begin{aligned}\frac{d}{dx}\log\left[\frac{f_1(x)}{f_2(x)}\right] &= -b(a_1+a_2)x^{b-1} - c(k_1+k_2)x^{c-1}(1+x^c)\\ &\quad -\frac{bcx^{b+c-1}(bx^c+b-c)(k_1a_2-k_2a_1)}{[a_1bx^b(x^c+1)+ck_1x^c]\,[ba_2x^b(x^c+1)+ck_2x^c]}.\end{aligned}$$

The two first terms are always negative. The denominator of the ratio in the third term is always positive and $-\left[bcx^{b+c-1}(bx^c+b-c)(k_1a_2-k_2a_1)\right] < 0$ if $b > c$ and $k_1/k_2 > a_1/a_2$. Thus $d[f_1(x)/f_2(x)]/dx < 0$ and $f_1(x)/f_2(x)$ is decreasing in x if $b > c$ and $k_1/k_2 > a_1/a_2$, implying that $X_1 \leq_{lr} X_2$. $\square$

The rest of the study is devoted to the study of the minWD distribution as statistical model, with applications.

3 Estimation of parameters

3.1 Maximum likelihood estimation

Inference can be carried out in three different ways: point estimation, interval estimation and hypothesis tests. Several approaches for parameter point estimation were proposed in the literature but the maximum likelihood method is the most commonly employed. The maximum likelihood estimates (MLEs) enjoy desirable properties that can be used when constructing confidence intervals for the model parameters. Large sample theory for these estimates delivers simple approximations that work well in finite samples. The normal approximation for the MLEs in distribution theory is easily handled either analytically or numerically.

We consider the estimation of the unknown parameters of the new distribution by the maximum likelihood method. Let $x_1, \cdots, x_n$ be n observations from the minWB distribution given by (1.2) with parameter vector $\boldsymbol{\theta} = (a, b, c, k)^\top$. The log-likelihood $\ell = \ell(\boldsymbol{\theta})$ for $\boldsymbol{\theta}$ is given by

$$\ell = -a\sum_{i=1}^{n} x_i^b - k\sum_{i=1}^{n}\log(1+x_i^c) + \sum_{i=1}^{n}\log\left[abx_i^{b-1} + ckx_i^{c-1}(1+x_i^c)^{-1}\right]. \tag{3.1}$$

Equation (3.1) can be maximized either directly by using the R (`optim function`), SAS (`NLMixed` procedure) or Ox (`MaxBFGS` function), or then by solving the nonlinear likelihood equations by differentiating it. The components of the score vector $U(\boldsymbol{\theta})$ are

$$\begin{aligned}
U_a &= -\sum_{i=1}^{n} x_i^b + \sum_{i=1}^{n} \frac{A_i}{a}\,(A_i+B_i)^{-1},\\
U_b &= -\frac{1}{b}\sum_{i=1}^{n}(A_i x_i \log x_i) + \sum_{i=1}^{n}\left(\frac{A_i}{b} + A_i\log x_i\right)(A_i+B_i)^{-1},\\
U_c &= -k\sum_{i=1}^{n}\frac{x_i^c\log x_i}{1+x_i^c} + k\sum_{i=1}^{n}\frac{x_i^{c-1}(1+x_i^c+c\log x_i)}{(1+x_i^c)^2}\,(A_i+B_i)^{-1},\\
U_k &= -\sum_{i=1}^{n}\log(1+x_i^c) + \frac{1}{k}\sum_{i=1}^{n} B_i\,(A_i+B_i)^{-1},
\end{aligned}$$

where $A_i = abx_i^{b-1}$ and $B_i = ckx_i^{c-1}\,(1+x_i^c)^{-1}$.

Setting these equations to zero and solving them simultaneously yields the MLEs $\hat{\boldsymbol{\theta}}$ of the model parameters.

Under standard regularity conditions, the multivariate normal $N_4(0, J(\hat{\boldsymbol{\theta}})^{-1})$ distribution, where $J(\hat{\boldsymbol{\theta}})^{-1}$ is the observed information evaluated at $\hat{\boldsymbol{\theta}}$, can be used to construct approximate confidence intervals for the model parameters. Further, we can compare the minWB model with any of its special models using likelihood ratio (LR) statistics.

3.2 Monte Carlo simulation study

We evaluate the performance of the MLEs of the model parameters of the minWB distribution using Monte Carlo simulations for selected parameter values varying the sample size. The simulation is repeated 3, 000 times each for sample size n= 50, 100, 200, 300, 500. The parametric values are I: $a = 0.5$, $b = 3$, $c = 1.5$, $k = 1.8$ and II: $a = 1.0$, $b = 3$, $c = 1.5$, $k = 2$. The MLEs are evaluated by maximizing Equation (3.1) using the optim routine in the R software. Table 2 gives the MLEs, average biases (Biases), mean square errors (MSEs), coverage probabilities (CPs), average lower bounds (LBs), average upper bounds (UBs) for the estimates of the parameters a, b, c and k for different sample sizes. The figures in this table indicate that the biases and MSEs decrease when the sample size increases and the MLEs tend to be close to the true parameter values. The CPs of the confidence intervals are quite close to the nominal level of 95% thus indicating that the asymptotic results for the MLEs can be used for estimating and constructing confidence intervals.

4 The Poisson Weibull Burr distribution with cure fraction

4.1 Motivations

We define the Poisson Weibull Burr (PWB) distribution by assuming that the latent number of failure causes has a Poisson distribution and that the time for these causes to be activated follows the minWB model. Also, we propose the inclusion of covariates in the model formulation in order to study their effects on the hrf. Inferential aspects based on the maximum likelihood method is discussed. Models for survival data with a cure fraction (also known as cure rate models or long-term survival models) play an important role in reliability and survival analysis. Cure rate models cover situations where there are sampling units not susceptible to the occurrence of the event of interest. The proportion of such units is called the cured fraction. These models have become very popular due to significant progress in treatment therapies leading to enhanced cure rates. The proportion of these "cured units" is termed the cure fraction.

The literature on the subject is by now rich and growing rapidly. The books by [13] and [12], as well as the review paper by [6], [23] and [7], could be mentioned as key references. Alternatively, other works dealt with cure rate models. For example, [11] proposed the long-term survival model with interval-censored data, [17] introduced the power series beta-Weibull regression model for predicting breast carcinoma, [24] investigated the Weibull-negative-binomial regression model with cure rate under latent failure causes, [16] studied the regression models generated by gamma random variables with long-term survivors and [21] defined the general long-term aging model with different underlying activation mechanisms.

4.2 The PWB cure rate model

The PWB cure rate model is derived as follows. For an individual in the population, let N denote the unobservable number of causes of the event of interest for this individual. We assume that N has a Poisson distribution with mean τ. The time for the jth cause to produce the event of interest is denoted by Z_j, $j = 1, \ldots, N$. Further, we consider that, conditional on N, the Z_j's are independent and identically random variables having cumulative function (1.1) and that $Z_1, Z_2, \ldots$ are independent of N. The observable time to the event of interest is defined by $X = \min\{Z_1, \ldots, Z_N\}$, and $T = +\infty$ if $N = 0$ with $\mathbb{P}(X = +\infty|N = 0) = 1$.

Under this setup, the survival function for the population is given by

$$S_{\text{pop}}(x) = \mathbb{P}(N = 0) + \mathbb{P}(Z_1 > x, \ldots, Z_N > x|N \geq 1)\mathbb{P}(N \geq 1).$$

Among others, [23] and [18] demonstrated that $S_{\text{pop}}(t) = A\big[S(t)\big]$, where $A(\cdot)$ is the probability generating function (pgf) of the number of competing causes (N). Then, the sf for the population reduces to

$$S_{\text{pop}}(x) = \exp\left\{-\tau\left[1 - e^{-ax^b}\,(1 + x^c)^{-k}\right]\right\} \tag{4.1}$$

and the cured fraction is given by $S_{\rm pop}(\infty) = \pi_0 = e^{-\tau}$ (not a proper survival function). The corresponding pdf reduces to

$$f_{\rm pop}(x) = \frac{\tau e^{-ax^b}}{(1+x^c)^k}\left[abx^{b-1} + \frac{ckx^{c-1}}{(1+x^c)}\right]\exp\left\{-\tau\left[1 - e^{-ax^b}(1+x^c)^{-k}\right]\right\}. \tag{4.2}$$

The hrf for the population is given by

$$h_{\rm pop}(x) = \frac{\tau e^{-ax^b}}{(1+x^c)^k}\left[abx^{b-1} + \frac{ckx^{c-1}}{(1+x^c)}\right]. \tag{4.3}$$

Equations (4.1), (4.2) and (4.3) are referred to as the PWB model with cure fraction in competitive-risk structure.

Then, the sf for the non-cured population, so-called the PWB survival function, is given by

$$S(x) = P(X > x|N \geq 1) = \frac{\exp\left\{-\tau\left[1 - e^{-ax^b}(1+x^c)^{-k}\right]\right\} - e^{-\tau}}{1 - e^{-\tau}}. \tag{4.4}$$

We note that $S(0) = 1$ and $S(+\infty) = 0$, so that it is a proper survival function. Henceforth, the model (4.4) will be referred to as the PWB survival function. The new pdf for the non-cured population reduces to

$$f(x) = \frac{\tau e^{-ax^b}}{(1+x^c)^k}\left[abx^{b-1} + \frac{ckx^{c-1}}{(1+x^c)}\right]\frac{\exp\left\{-\tau\left[1 - e^{-ax^b}(1+x^c)^{-k}\right]\right\}}{1 - e^{-\tau}}. \tag{4.5}$$

In Equation (4.5), the parameter $a \geq 0$ controls the scale of the distribution while the parameters $b > 0$, $c > 0$, $k > 0$ and $\tau > 0$ control its shape.

Some news special cases of Equation (4.5) are: (i) The Poisson Weibull (PW) model when $k = 0$, (ii) The Poisson Burr (PB) model when $a = 0$, (iii) The Poisson Weibull-log-logistic (PWLL) model when $k = 1$, (iv) The Poisson Rayleigh Burr (PRB) model when $b = 2$, (v) The Poisson exponential-Burr (PEB) model when $b = 1$, and (vi) The Poisson Weibull-Lomax (PWLx) model when $c = 1$.

4.3 Inference

Consider the situation where the time to the event is not completely observed and is subjected to right censoring. Let D_i denote the censoring time. We then observe $x_i = \min\{X_i, D_i\}$ and $\delta_i = \mathrm{I}(X_i \leq D_i)$, where $\delta_i = 1$ if X_i is the observed time to the event defined before and $\delta_i = 0$ if it is right censored (for $i = 1, \ldots, n$).

In many medical problems, the lifetimes are affected by explanatory variables such as the cholesterol level, blood pressure, weight and many others. Parametric models to estimate univariate survival functions for censored data regression problems are widely used. The parameter τ in (4.1) is now linked to a vector $\mathbf{v}_i$ of explanatory variables by $\tau_i = \exp(\mathbf{v}_i^T\boldsymbol{\beta})$, for $i = 1, \ldots, n$, where $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_p)^T$ denotes the vector of regression coefficients. The vector of model parameters is denoted by $\boldsymbol{\theta} = (a, b, c, k, \boldsymbol{\beta}^T)^T$.

We have the following special PWM regression models obtained from Equation (4.1): (i) The PW regression model when $k = 0$, (ii) The PB regression model when $a = 0$, (iii) The PWLL regression model when $k = 1$, (iv) The PRB regression model when $b = 2$, (v) The PEB regression model when $b = 1$, and (vi) The PWLx regression model when $c = 1$.

Let $\boldsymbol{\theta}$ denote the parameter vector of the distribution function $F(x)$ of the time-to-event. From n triples of times and censoring indicators $(x_1, \delta_1, \mathbf{v}_i), \ldots, (x_n, \delta_n, \mathbf{v}_n)$, the observed full log-likelihood function under non-informative censoring is given by

$$
\begin{aligned}
l(\theta) &= \sum_{i=1}^{n} \delta_i(\mathbf{v}_i^T\boldsymbol{\beta} - a x_i^b) + k\sum_{i=1}^{n}\delta_i \log(1+x_i^c) + \sum_{i=1}^{n}\delta_i \log\left(abx_i^{b-1} + \frac{ckx_i^{c-1}}{1+x_i^c}\right) \\
&\quad - \sum_{i=1}^{n}\exp(\mathbf{v}_i^T\boldsymbol{\beta})\left[1-\exp(-ax_i^b)(1+x_i^c)^{-k}\right].
\end{aligned}
$$

The MLE $\widehat{\boldsymbol{\theta}}$ of $\boldsymbol{\theta}$ is obtained by solving the nonlinear equations $U_a(\boldsymbol{\theta}) = 0$, $U_b(\boldsymbol{\theta}) = 0$, $U_c(\boldsymbol{\theta}) = 0$, $U_k(\boldsymbol{\theta}) = 0$ and $U_{\beta_j}(\boldsymbol{\theta}) = 0$, $\quad j = 1, \ldots, p$. These equations cannot be solved analytically and statistical software can be used to solve them numerically. We can use iterative techniques such as Newton-Raphson type algorithms to calculate the estimate $\widehat{\boldsymbol{\theta}}$. We use the software SAS (NLMixed procedure) to evaluate the MLE $\widehat{\boldsymbol{\theta}}$.

The inference procedures for $\boldsymbol{\theta} = (a, b, c, k, \boldsymbol{\beta}^T)^T$ can be based on the multivariate normal approximation

$$(\hat{a}, \hat{b}, \hat{c}, \hat{k}, \widehat{\boldsymbol{\beta}}^T)^T \sim \mathrm{N}_{p+4}\Big\{(a, b, c, k, \boldsymbol{\beta}^T)^T, -\ddot{\mathbf{L}}^{-1}(\widehat{\boldsymbol{\theta}})\Big\},$$

where $-\ddot{\mathbf{L}}(\boldsymbol{\theta}) = \left\{\frac{\partial^2 l(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}\boldsymbol{\theta}^T}\right\}$, the $(p+4)\times(p+4)$ observed information matrix, can be calculated numerically.

Besides estimation of the model parameters, hypothesis tests can be taken into account. Let $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$ be proper disjoint subsets of $\boldsymbol{\theta}$. Consider the test of $H_0 : \boldsymbol{\theta}_1 = \boldsymbol{\theta}_{01}$ against $H_1 : \boldsymbol{\theta}_1 \neq \boldsymbol{\theta}_{01}$, where $\boldsymbol{\theta}_2$ is an unspecified vector. Let $\widehat{\boldsymbol{\theta}}_0$ maximize the the log-likelihood $l(\boldsymbol{\theta})$ constrained to H_0 and define the likelihood ratio (LR) statistic by $w = 2[l(\widehat{\boldsymbol{\theta}}) - l(\widehat{\boldsymbol{\theta}}_0)]$. Under H_0 and some regularity conditions, the statistic w converges in distribution to a chi-square distribution with $\dim(\boldsymbol{\theta}_1)$ degrees of freedom.

5 Applications of the minWB distribution and PWB models

In this section, we provide some applications of the minWB and PWB models.

5.1 Application 1: WB distribution

Here, we compare the fits of the minWB, WL, EW and OWB (see [2]) distributions by means of four real data sets to illustrate the potentiality of the minWB model. The densities of the competitive models are, respectively, given by

$$
\begin{aligned}
f_{EW}(x) &= (\lambda + abx^{b-1})\exp[-(\lambda x + ax^b)], \quad x, a, b, \lambda > 0, \\
f_{WL}(x) &= \frac{1}{1+\lambda}\exp[-(\lambda + ax^b)]\left[(1+\lambda+\lambda x)(\lambda + abx^{b-1}) - \lambda\right], \quad x, a, b, \lambda > 0, \\
f_{OWB}(x) &= abckx^{c-1}(1+x^c)^{bk-1}\left[1-(1+x^c)^{-k}\right]^{b-1}\exp\left\{-a[(1+x^c)^k - 1]^b\right\}, \\
&\quad x, a, b, c, k > 0.
\end{aligned}
$$

We estimate the unknown parameters of the distributions by maximum likelihood. We compute the log-likelihood function evaluated at the MLEs ($\hat{\ell}$) using a limited-memory quasi-Newton code for bound-constrained optimization (L-BFGS-B). For model comparison, we consider five well-known statistics: the maximized log-likelihood ($\hat{\ell}$), Akaike information criterion (AIC), Anderson-Darling (A^*), Cramér–von Mises (W^*) and Kolmogorov-Smirnov (K-S) measures, where lower values of these statistics and higher p-values of K-S indicate good fits. The required computations are carried out using the R script `AdequacyModel` which is freely available from http://cran.r-project.org/web/packages/AdequacyModel/AdequacyModel.pdf.

The following data sets are considered for analysis:

Data set 1: Drilling Machine data. The first data set refers to the 50 observations with hole and sheet thickness of 12 mm and 3.15 mm (see [10]): 0.04, 0.02, 0.06, 0.12, 0.14, 0.08, 0.22, 0.12, 0.08, 0.26, 0.24, 0.04, 0.14, 0.16, 0.08, 0.26, 0.32, 0.28, 0.14, 0.16, 0.24, 0.22, 0.12, 0.18, 0.24, 0.32, 0.16, 0.14, 0.08, 0.16, 0.24, 0.16, 0.32, 0.18, 0.24, 0.22, 0.16, 0.12, 0.24, 0.06, 0.02, 0.18, 0.22, 0.14, 0.06, 0.04, 0.14, 0.26, 0.18, 0.16.

Data set 2: Guinea Pigs data. The second data set represents the survival times (in days) of 72 guinea pigs infected with virulent tubercle bacilli (see [5]). Guinea pigs are known to have high susceptibility of human tuberculosis, which is one of the reasons for choosing this species. The survival times of the Guinea pigs in days are: 0.1, 0.33, 0.44, 0.56, 0.59, 0.72, 0.74, 0.77, 0.92, 0.93, 0.96, 1, 1, 1.02, 1.05, 1.07, 07, .08, 1.08, 1.08, 1.09, 1.12, 1.13, 1.15, 1.16, 1.2, 1.21, 1.22, 1.22, 1.24, 1.3, 1.34, 1.36, 1.39, 1.44, 1.46, 1.53, 1.59, 1.6, 1.63, 1.63, 1.68, 1.71, 1.72, 1.76, 1.83, 1.95, 1.96, 1.97, 2.02, 2.13, 2.15, 2.16, 2.22, 2.3, 2.31, 2.4, 2.45, 2.51, 2.53, 2.54, 2.54, 2.78, 2.93, 3.27, 3.42, 3.47, 3.61, 4.02, 4.32, 4.58, 5.55.

Data set 3: Stress Level data. The third data set (see [8]) represents the failure times of Kevlar 49/epoxy strands when the pressure is at 90% stress level: 0.01, 0.01, 0.02, 0.02, 0.02, 0.03, 0.03, 0.04, 0.05, 0.06, 0.07, 0.07, 0.08, 0.09, 0.09, 0.10, 0.10, 0.11, 0.11, 0.12, 0.13, 0.18, 0.19, 0.20, 0.23, 0.24, 0.24, 0.29, 0.34, 0.35, 0.36, 0.38, 0.40, 0.42, 0.43, 0.52, 0.54, 0.56, 0.60, 0.60, 0.63, 0.65, 0.67, 0.68, 0.72, 0.72, 0.72, 0.73, 0.79, 0.79, 0.80, 0.80, 0.83, 0.85, 0.90, 0.92, 0.95, 0.99, 1.00, 1.01, 1.02, 1.03, 1.05, 1.10, 1.10, 1.11, 1.15, 1.18, 1.20, 1.29, 1.31, 1.33, 1.34, 1.40, 1.43, 1.45, 1.50, 1.51, 1.52, 1.53, 1.54, 1.54, 1.55, 1.58, 1.60, 1.63, 1.64, 1.80, 1.80, 1.81, 2.02, 2.05, 2.14, 2.17, 2.33, 3.03, 3.03, 3.34, 4.20, 4.69, 7.89.

Data set 4: Failure Times data. The fourth data set (see [15]) represents the failures times of 50 items: 0.032, 0.035, 0.104, 0.169, 0.196, 0.260, 0.326, 0.445, 0.449, 0.496, 0.543, 0.544, 0.577, 0.648, 0.666, 0.742, 0.757, 0.808, 0.857, 0.858, 0.882, 1.138, 1.163, 1.256, 1.283, 1.484, 1.897, 1.944, 2.201, 2.365, 2.531, 2.994, 3.118, 3.424, 4.097, 4.100, 4.744, 5.346, 5.479, 5.716, 5.825, 5.847, 6.084, 6.127, 7.241, 7.560, 8.901, 9.000, 10.482, 11.133.

The numerical measures of some statistics for the data sets 1–4 and for the fitted minWB model to these data are given in Tables 3 and 4, respectively.

We also analyzed the hazard rates of these four data sets. In order to identify the shapes of data, we consider the graphical method based on total time on test (TTT) transformed pioneered by [4]. The empirical illustration of the TTT-transform is given by [1]. The TTT plot is obtained by plotting $G(r/n) = \big[\sum_{i=1}^{r} T_{i:n} + (n-r)T_{r:n}\big] / \big[\sum_{i=1}^{n} T_{i:n}\big]$ versus $r/n (r = 1, 2, \ldots, n)$, where the observed variables $T_{i:n}$ (for $i = 1, 2, \ldots, n$) are the order statistics of the sample.

The TTT plots for four data sets are given in Figures 5 and 6. The TTT-plots for the data sets 1 and 2 in Figure 5(a) and 5(b) reveal that the hrf is concave giving an indication of an increasing hazard rate. The TTT-plot for the data set 3 in Figure 6(a) shows that the hrf is first convex, then concave and lastly convex giving an indication of decreasing-increasing-decreasing (DID) shape. The TTT-plot in Figure 6(b) for the data set 4 shows that the hrf is first concave then convex giving an indication of increasing-decreasing (UBT) shape. Hence, the minWB model could be in principle an appropriate model for fitting these data sets.

In Figures 7 and 8, we consider kernel density estimation (a non-parametric approach) with

Gaussian Filter. The kernel density estimator of $f(x)$ is given by

$$\hat{f}_h(x) = \frac{1}{nh} \sum_{i=1}^{n} K\Big(\frac{x - x_i}{h}\Big),$$

where $K(x)$ is the kernel function usually symmetric, $\int_{-\infty}^{+\infty} K(x)dx = 1$ and $h > 0$ is a smoothing parameter. Here we use a Gaussian kernel and the so called rule-of-thumb for the choice of h (see [19]).

(a) (b)

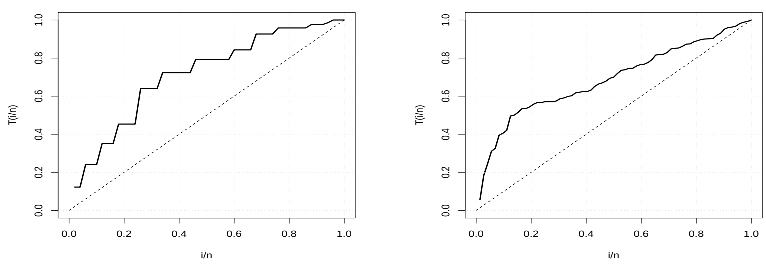

Figure 5: TTT plots for (a) Data set 1 (b) Data set 2.

(a) (b)

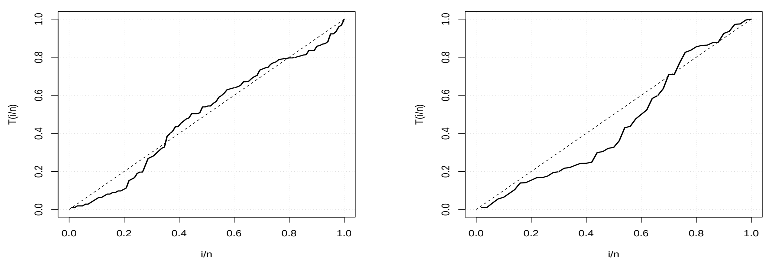

Figure 6: TTT plots for (a) Data set 3 (b) Data set 4.

(a) (b)

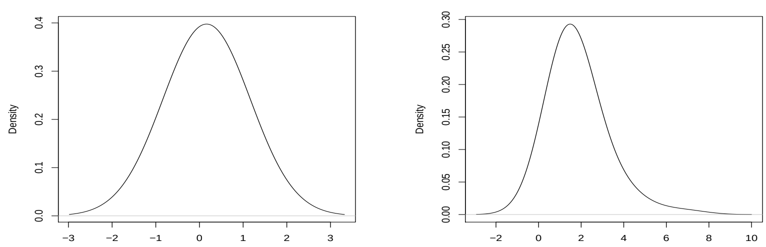

Figure 7: Gaussian kernel density estimation for data sets. (a) Data set 1 (b) Data set 2.

Table 6 lists the MLEs and their corresponding standard errors (in parentheses) of the model parameters for the fitted models to data sets 1–4. The results in Table 5 indicate that the minWB model provides the best fit as compared to the other models.

(a) (b)

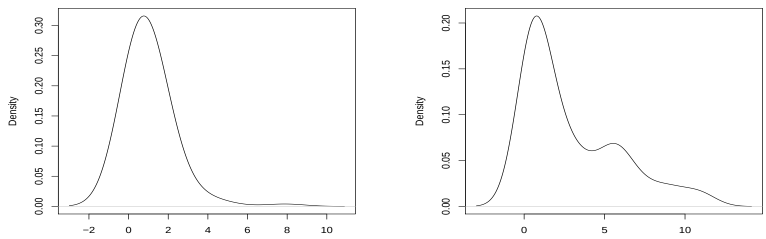

Figure 8: Gaussian kernel density estimation for data sets (a) Data set 3 (b) Data set 4.

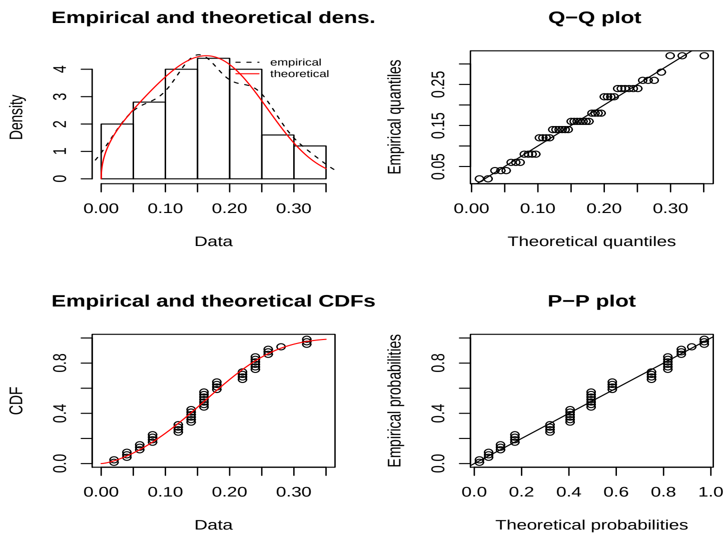

Figure 9: PP, QQ, epdf and ecdf plots of the minWB distribution for data set 1.

5.2 Application 2: PWB model with cure fraction: Gastric cancer data

The data set refers to $n = 201$ patients observed with gastric adenocarcinoma. Gastric (stomach) cancer is a disease in which malignant (cancer) cells form in the lining of the stomach. Almost all gastric cancers are adenocarcinomas (cancers that begin in cells that make and release mucus and other fluids). Other types of gastric cancer are gastrointestinal carcinoid tumors, gastrointestinal stromal tumors and lymphomas. These data sets have been analyzed by [14] and [16]. The response variable is the time x_i in months after surgery until death. The patients who die from other causes and the patients that are still alive at the end of the study are censored observations (53%). The only covariate is the type of therapy: v_{i1} (0=adjuvant chemoradiotherapy, $n = 125$; 1=surgery alone, $n = 76$). We are interested in the effect of the explanatory variable on the cure fraction.

For the PWB regression model with cure fraction, we consider ($i = 1, \ldots, 201$):

$$\tau_i = \exp(\beta_0 + \beta_1 v_{i1}).$$

Recently, [16] analyze these data using the family called the *Poisson-gamma-G* (PG-G) model with cure fraction in competitive-risk structure. The authors estimate of the parameters of the the following models: Poisson-gamma Weibull (PGW), Poisson-gamma log-logistic (PGLL),

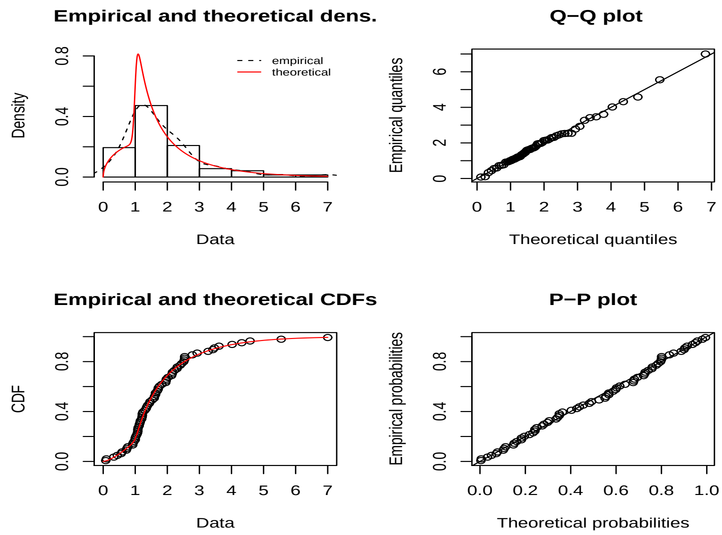

Figure 10: PP, QQ, epdf and ecdf plots of the minWB distribution for data set 2.

Poisson-gamma Birnbaum-Saunders (PGBS) and Poisson-gamma generalized half-normal (PG-GHN) regression model with cure fraction. In this application, we compare all these regression models with the PWB regression model with cure fraction.

In Table 7, we list the values of the AIC, Consistent Akaike Information Criterion (CAIC) and Bayesian Information Criterion (BIC) for all models discussed in Section 4. So, we will have more evidence to be able to discriminate and choose the most suitable model. The lowest values of these information criteria correspond to the PWLx regression model with cure fraction, which provides the best fit to the current data among these models.

Table 8 gives the MLEs for the fitted PWLx regression model with cure fraction. At a 5% significance level, the regression coefficient is significant for the type of therapy (v_1).

Goodness-of-fit. We adopt a regression structure for the cure probability in long-term survivor models (see Section 4). We now estimate the cure rate (π_0). Note that

$$\hat{\tau} = \frac{1}{201}\sum_{i=1}^{201}\hat{\tau}_i = 0.7242,$$

where

$$\hat{\tau}_i = \exp(-0.6282 + 0.4539 v_{i1}),$$

and then

$$\hat{\pi}_0 = \mathrm{e}^{-\hat{\tau}} = 0.4847.$$

In order to assess if the model is appropriate, Figure 13a displays the empirical survival function and the estimated marginal survival functions given by Equation (4.1) from the fitted PWLx model with long-term survivors.

The estimates of the cure rate for patients stratified by type of therapy (v_1) are:

- For Chemoradiotherapy ($v_1 = 0$)

  $\hat{\tau}_0 = \exp(-0.6282)$ and the cured fraction is $\hat{\pi}_{00} = \mathrm{e}^{-\hat{\tau}_0} = 0.5865$.

- For Surgery alone ($v_1 = 1$)

  $\hat{\tau}_1 = \exp(-0.6282 + 0.4539)$ and the cured fraction is $\hat{\pi}_{01} = \mathrm{e}^{-\hat{\tau}_1} = 0.4317$.

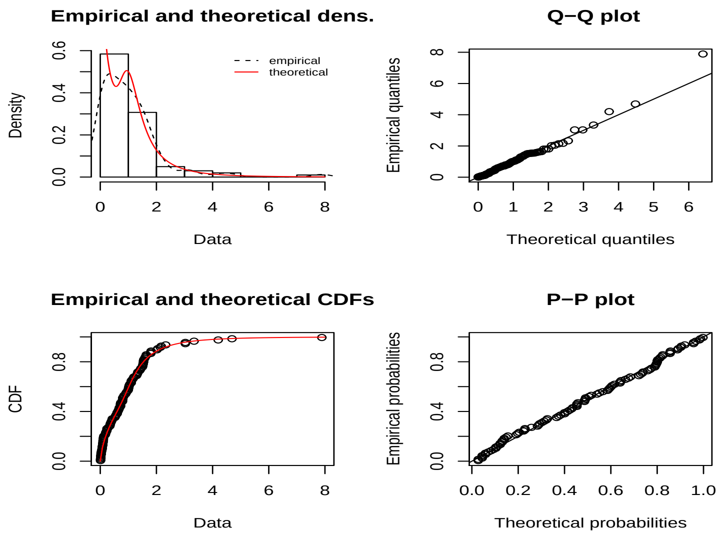

Figure 11: PP, QQ, epdf and ecdf plots of the minWB distribution for data set 3.

Also, the estimated survival function and cure fraction stratified by v_1 are displayed in Figure 13b, from which a significant fraction of survivors can be observed. Note that the proportion of cured is greater for patients receiving chemoradiotherapy.

6 Concluding remarks

We propose and study the minimun Weibull-Burr (minWB) model and obtain some mathematical properties such as quantile function, ordinary and incomplete moments, mean deviations, generating function, stress-strength reliability and stochastic ordering. The model parameters are estimated by the method of maximum likelihood. Some simulations are performed to check the asymptotic properties of the estimates. We define the Poisson-Weibull-Burr regression model with cure fraction as a competitor to other existing regression models. Some applications to real data set are presented to illustrate the potentiality of the proposed models. We expect the utility of the proposed models in different fields especially in lifetime and reliability.

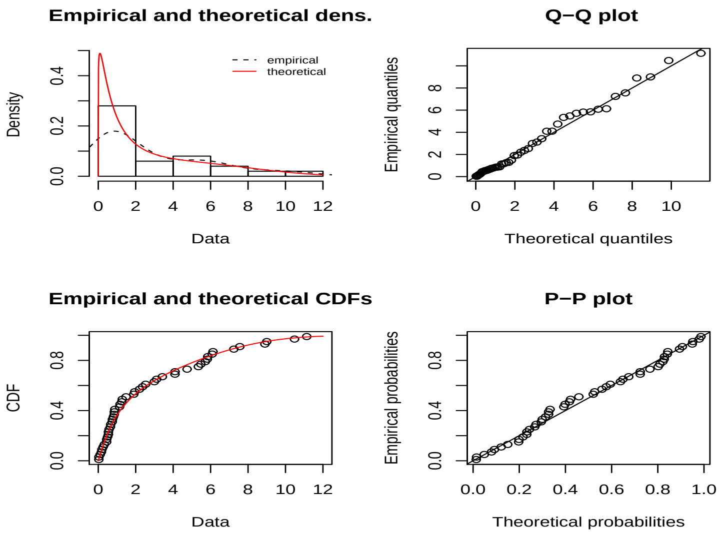

Figure 12: PP, QQ, epdf and ecdf plots of the minWB distribution for data set 4.

(a) (b)

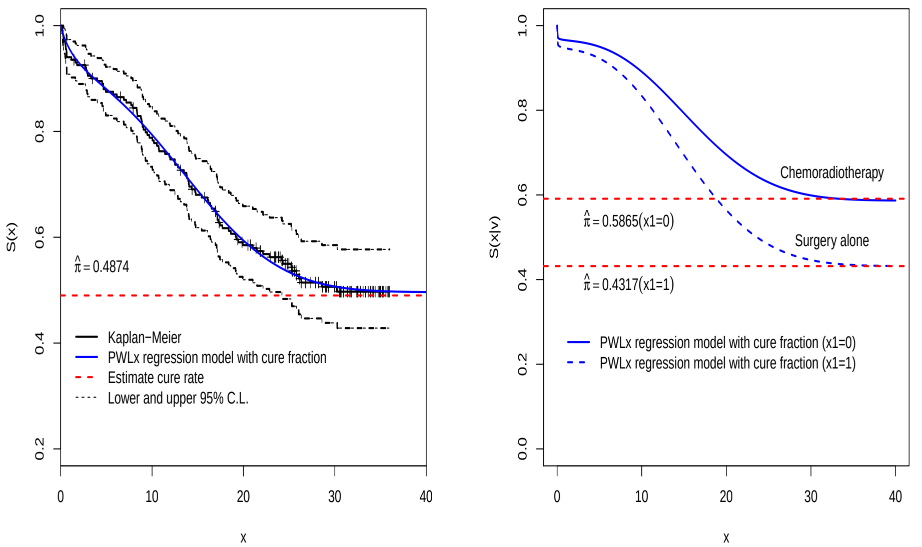

Figure 13: (a) Kaplan-Meier curves (solid lines), the estimated PWLx survival function and the estimated cure fraction for the gastric cancer data. (b) Estimates of the survival function and cure fraction of model stratified by type of therapy for the gastric cancer data.

Table 2: Monte Carlo simulation results: MLEs, Biases, MSEs, CPs, LBs and UBs.

n	Parameter	MLE	Bias	MSE	CP	LB	UB
			I				
50	a	0.4482	0.0598	0.2501	0.9403	0.3097	0.5868
	b	4.8071	1.8071	8.9206	0.8358	4.1446	5.4696
	c	1.5154	0.0154	0.1432	0.9552	1.4100	1.6207
	k	1.9229	0.1229	0.5944	0.8806	1.7109	2.1350
100	a	0.4943	0.0557	0.2848	0.9437	0.3892	0.5994
	b	4.1836	1.1836	4.1348	0.8451	3.8579	4.5093
	c	1.5172	0.0172	0.0994	0.9718	1.4552	1.5792
	k	1.8253	0.0253	0.5617	0.8998	1.6777	1.9729
200	a	0.5562	0.0562	0.2382	0.9787	0.4887	0.6237
	b	3.7267	0.7267	2.6368	0.8553	3.5244	3.9290
	c	1.4830	0.0170	0.0589	0.9574	1.4493	1.5167
	k	1.7619	0.0381	0.4219	0.7447	1.6716	1.8523
300	a	0.5397	0.0397	0.1978	0.9895	0.4893	0.5901
	b	3.5230	0.5230	1.6379	0.8610	3.3901	3.6558
	c	1.4924	0.0076	0.0480	0.9579	1.4675	1.5173
	k	1.7245	0.0255	0.3598	0.8737	1.6568	1.7922
500	a	0.5059	0.0159	0.1323	0.9501	0.5445	0.6072
	b	3.2088	0.2088	0.8549	0.9498	3.1295	3.2882
	c	1.4695	-0.0305	0.0296	0.9500	1.4546	1.4844
	k	1.7997	0.0103	0.2433	0.9400	1.6572	1.7423
			II				
50	a	0.8610	0.1390	0.7500	0.9125	0.6229	1.0991
	b	6.3291	3.3291	6.6283	0.8000	5.2307	7.4274
	c	1.5619	0.0619	0.8992	0.9167	1.4388	1.6851
	k	2.2878	0.4878	1.6848	0.7500	1.9527	2.6229
100	a	0.9934	0.0266	0.6577	0.8675	0.8337	1.1531
	b	4.4732	1.4732	5.7956	0.8747	3.9661	4.9802
	c	1.5075	0.0075	0.1388	0.9518	1.4342	1.5809
	k	2.0309	0.0309	1.1146	0.8554	1.8230	2.2388
200	a	1.0195	0.0195	0.4561	0.9565	0.9255	1.1136
	b	4.0912	1.0912	4.0493	0.9065	3.7841	4.3982
	c	1.4922	0.0058	0.0697	0.9348	1.4555	1.5290
	k	1.9767	0.0233	0.7875	0.8478	1.8532	2.1003
300	a	0.9893	0.0107	0.4083	0.8969	0.9166	1.0619
	b	3.7425	0.7425	3.4347	0.9113	3.5494	3.9356
	c	1.4831	0.0039	0.0474	0.9054	1.4495	1.5166
	k	1.9639	0.0161	0.7634	0.8823	1.8646	2.0632
500	a	1.0094	0.0094	0.1971	0.9489	0.9703	1.0485
	b	3.1607	0.3607	0.8446	0.9502	3.2863	3.4352
	c	1.4942	-0.0018	0.0366	0.9501	1.4575	1.4909
	k	2.0007	0.0007	0.3833	0.9568	1.9461	2.0552

Table 3: Summary statistics for the data sets 1–4.

Data	n	Mean	Median	Var.	Skewness	Kurtosis	Min	Max
Set 1:	50	0.163	0.160	0.007	0.072	2.216	0.02	0.32
Set 2:	72	1.837	1.560	1.478	1.755	7.152	0.08	7.0
Set 3:	101	1.025	0.800	1.253	3.002	16.70	0.01	7.89
Set 4:	50	2.897	1.383	9.047	1.118	3.240	0.032	11.133

Table 4: Summary statistics for the minWB distribution fitted to data sets 1–4.

Data	Mean	Median	Var.	S. D.	Skewness	Kurtosis
Set 1:	0.1633	0.1605	0.0065	0.0804	0.0413	2.5005
Set 2:	1.8380	1.4866	1.5402	1.2410	1.9570	8.8400
Set 3:	1.0216	0.8186	1.1759	1.0844	3.6427	34.6871
Set 4:	2.8910	1.3886	8.5375	2.9219	1.2355	3.8592

Table 5: The statistics $\hat{\ell}$, AIC,A^*, W^*, K-S and P-value for the data sets 1–4.

Distribution	$\hat{\ell}$	AIC	A^*	W^*	K-S	P-value
Data set 1						
minWB	-57.2249	-107.8497	0.4436	0.0730	0.0889	0.8239
WL	-56.7418	-107.4836	0.4763	0.0788	0.0976	0.7275
EW	-56.7323	-107.4646	0.4821	0.0802	0.1002	0.6971
OWB	-55.8928	-103.7855	0.6430	0.1051	0.1091	0.5916
Data set 2						
minWB	97.2426	202.4853	0.1536	0.0232	0.0514	0.9913
WL	108.6217	223.2434	0.8644	0.1380	0.2168	0.0023
EW	104.0168	214.0336	0.9758	0.1603	0.1135	0.3121
OWB	104.0160	216.0320	0.9757	0.1602	0.1128	0.3190
Data set 3						
minWB	98.2315	204.4629	0.2816	0.0347	0.0523	0.945
WL	103.6844	213.3688	0.8413	0.1373	0.1069	0.198
EW	102.9160	211.8320	1.0448	0.1841	0.0886	0.406
OWB	102.9772	213.9544	1.1118	0.1988	0.0902	0.384
Data set 4						
minWB	99.6313	207.2626	0.2519	0.0437	0.0866	0.816
WL	105.2206	216.4412	0.7036	0.1320	0.2248	0.011
EW	102.5311	211.0622	0.5493	0.0971	0.1109	0.533
OWB	102.5315	213.0631	0.5477	0.0967	0.1116	0.526

Table 6: MLEs and their standard errors (in parentheses) for the data sets 1–4.

Distribution	a	b	c	k	λ
Data set 1					
minWB	3.4425	115.7659	7.0668	1.4788	-
	(0.9872)	(114.3137)	(9.8228)	(0.5072)	-
WL	3.2015	4.4621	-	-	2.6462
	(0.9125)	(0.4891)	-	-	(1.2208)
EW	109.0637	3.1182	-	-	1.8680
	(114.3953)	(0.8582)	-	-	(1.13877)
OWB	0.0260	1.0340	0.8446	80.5528	-
	(0.0784)	(0.1803)	(7.2285)	(248.5867)	-
Data set 2					
minWB	27.1932	0.0355	0.17694	1.4943	-
	(19.9536)	(0.0300)	(0.0526)	(0.2236)	-
WL	0.1427	80.3697	-	-	0.8345
	(0.0018)	(85.0955)	-	-	(0.0738)
EW	0.3117	1.6173	-	-	0.0000
	(0.1470)	(0.2256)	-	-	(0.1436)
OWB	0.0189	0.9721	9.0398	86.2455	-
	(0.0211)	(0.0234)	(30.2416)	(96.2013)	-
Data set 3					
minWB	5.4645	0.2652	0.7096	0.7029	-
	(1.0051)	(0.1505)	(0.1152)	(0.1001)	-
WL	61.4381	0.1263	-	-	1.3775
	(72.3178)	(0.0025)	-	-	(0.1066)
EW	0.2788	0.7413	-	-	0.7237
	(0.5996)	(0.4483)	-	-	(0.5755)
OWB	0.0109	0.9807	10.0264	85.2866	-
	(0.0079)	(0.0239)	(38.5478)	(62.1627)	-
Data set 4					
minWB	1.1152	0.6451	0.0027	2.8371	-
	(0.1854)	(0.1250)	(0.0019)	(0.3356)	-
WL	56.0004	0.0900	-	-	0.5572
	(133.3242)	(0.0026)	-	-	(0.0589)
EW	0.3763	0.8690	-	-	0.0366
	(1.2059)	(0.4681)	-	-	(1.1929)
OWB	0.0105	0.9731	10.0811	84.8302	-
	(0.0153)	(0.0288)	(41.7357)	(124.2766)	-

Table 7: Some statistics from the fitted regression models with cure fraction to the gastric cancer data.

	Statistics		
Model proposed	*AIC*	*CAIC*	*BIC*
PWB	886.4	886.8	906.2
PW	898.2	898.4	911.4
PB	944.3	944.5	957.5
PWLL	1011.3	1011.6	1027.8
PRB	887.6	887.9	904.1
PEB	899.2	899.5	915.7
PWLx	**884.4**	**884.7**	**900.9**
Model proposed by	Statistics		
[16]	*AIC*	*CAIC*	*BIC*
PGW	900.3	900.6	916.8
PGLL	900.1	900.4	916.7
PGBS	893.9	894.2	910.4
PGGHN	892.9	893.2	909.4

Table 8: MLEs for the full PWLx regression model with cure rate fraction fitted to the gastric cancer data.

Parameter	Estimate	Standard Error	95% C.L.	p-value
a	0.0003	0.00004	(0.0002, 0.0004)	–
b	2.6880	0.3940	(1.9111, 3.4649)	–
k	0.0957	0.02442	(0.0475, 0.1438)	–
β_0	-0.6282	0.1805	(-0.9842, -0.2722)	0.0006
β_1	0.4539	0.2179	(0.0241, 0.8837)	0.0385

References

[1] Aarset, M.V. (1987). How to identify bathtub hazard rate. IEEE Trans Reliab, 36, 106–108.

[2] Afify, A.Z., Cordeiro, G.M., Ortega, E.M.M., Yousof, H.M. and Butt, N.S. (2016). The four-parameter Burr XII distribution: Properties, regression model and applications. Commun Stat Theory Methods, DOI: 10.1080/03610926.2016.1231821.

[3] Asgharzadeh, A., Nadarajah, S. and Sharafi, F. (2017). The Weibull-Lindley distribution. REVSTAT, To appear.

[4] Barlow, R.E. and Campo, R.A. (1975). Total time on test processes and applications to failure data analysis. In: Barlow RE, Fussel JB, Singpurwalla ND (eds.), Reliability and Fault Tree Analysis, SIAM, Philadelphia, 451-481.

[5] Bjerkedal, T. (1960). Acquisition of resistance in guinea pigs infected with different doses of virulent tubercle bacilli. Amer J Hygiene, 72, 130–148.

[6] Chen, M.H., Ibrahim, J.G. and Sinha, D. (1999). A new Bayesian model for survival data with a surviving fraction. J Amer Stat Assoc, 94, 909–919.

[7] Cooner, F., Banerjee, S., Carlin, B.P. and Sinha, D. (2007). Flexible cure rate modeling under latent activation schemes. J Amer Stat Assoc, 102, 560–572.

[8] Cooray, K. and Ananda, M.M.A. (2008). A generalization of the half-normal distribution with applications to lifetime data. Commun Stat Theory Method, 37, 1323–1337.

[9] Cordeiro, G.M. and Lemonte, A.J. (2014). The exponential-Weibull lifetime model. J Stat Comput Simul, 84, 2592–2606.

[10] Dasgupta, R. (2011) On the distribution of Burr with applications. Sankhyã, 73, 1–19

[11] Hashimoto, E.M., Ortega, E.M.M., Cancho, V.G. and Cordeiro, G.M. (2015). A new long-term survival model with interval-censored data. Sankhyã B, 77, 207–239.

[12] Ibrahim, J.G. and Chen, M.H. and Sinha, D. (2001). Bayesian Survival Analysis. Springer, New York.

[13] Maller, R. and Zhou, X. (1996). Survival analysis with long-term survivors. Wiley, New York.

[14] Martinez, E.Z., Achcar, J.A., Jacome, A.A.A. and Santos, J.S. (2013). Mixture and non-mixture cure fraction models based on the generalized modified Weibull distribution with an application to gastric cancer data, Comput Methods Prog Biomed, 112, 343–355.

[15] Murthy, D.N.P, Xie, M. and Jiang, R. (2004). Weibull Models. Wiley, New York.

[16] Ortega, E.M.M., Cordeiro, G.M., Hashimoto, E.M. and Suzuki, A.K. (2017). Regression models generated by gamma random variables with long-term survivors. Commun Stat Applic Methods, 24, 43–65.

[17] Ortega, E.M.M., Cordeiro, G.M., Campelo, A.K., Kattan, M.W. and Cancho, V.G. (2015). A power series beta Weibull regression model for predicting breast carcinoma. Statist Med, 34, 1366–1388.

[18] Rodrigues, J., Cancho, V.G., de-Castro, M. and Louzada-Neto, F. (2009). On the unification of long-term survival models. Statist Probab Lett, 79, 753–759.

[19] Silverman, B. (1986). Density Estimation for Statistics and Data Analysis, Chapman and Hall, London.

[20] Shaked, M. and Shanthikumar, J.G. (2017). Stochatic Orders. Wiley, New York.

[21] Suzuki, A.K., Barriga, G.D.C., Louzada-Neto, F. and Cancho, V.G. (2017). A general long-term aging model with different underlying activation mechanisms: Modeling, Bayesian estimation, and case influence diagnostics. Commun Stat Theory Methods, 46, 3080–3098.

[22] Tahir, M.H. and Cordeiro, G.M. (2016). Compounding of distributions: a survey and new generalized classes. J Stat Dist Applic, 3, 13–16.

[23] Tsodikov, A.D., Ibrahim, J.G. and Yakovlev, A.Y. (2003). Estimating cure rates from survival data: an alternative to two-component mixture models. J Amer Stat Assoc, 98, 1063–1078.

[24] Yiqi, B., Cancho, V.G. and Louzada-Neto, F. (2017). On the Bayesian estimation and influence diagnostics for the Weibull-Negative-Binomial regression model with cure rate under latent failure causes. Commun Stat Theory Methods, 46, 1462–1489.