A new class of defective models based on the Marshall–Olkin family of distributions for cure rate modeling

DOI:
10.1016/j.csda.2016.10.001

Document Version
Accepted author manuscript

A new class of defective models based on the Marshall-Olkin family of distributions for cure rate modeling

Ricardo Rocha[a,*], Saralees Nadarajah[b], Vera Tomazella[a], Francisco Louzada[c]

[a] *Universidade Federal de São Carlos, Departamento de Estatística, São Carlos, SP, Brazil.*
[b] *University of Manchester, School of Mathematics, Manchester, United Kingdom.*
[c] *Universidade de São Paulo, Instituto de Ciências Matemáticas and de Computação, São Carlos, SP, Brazil.*

Abstract

Defective distributions model cure rates by changing the usual domain of its parameters in a way that their survival functions converge to a value $p \in (0, 1)$. A new way to generate defective distributions to model cure fractions is proposed. The new way relies on a property derived from the Marshall Olkin family of distributions. To exemplify this new result we use the extended Weibull distribution and introduce ten new defective distributions. A regression approach for these models is also proposed. Estimation by maximum likelihood is discussed and their asymptotes verified through simulations. Practical use is illustrated by applications to four real data sets.

Keywords: Defective distributions, Extended Weibull distribution, Long-term survivors, Regression modeling, Survival analysis.

1. Introduction

Modeling of a cure fraction, also known as long-term modeling, is a part of survival analysis. It studies cases where supposedly there are observations not susceptible to the event of interest. Such cases require special theoretical treatment, in a way that the modeling assumes the existence of such observations. In the standard theory of survival analysis the survival function $S(t)$ tends to zero as time increases. We need to use some strategy to make the survival function converge to a value $p \in (0, 1)$, representing the cure rate.

The method most commonly used is the standard mixture model, initially proposed by [8] and [7]. The model is described by $S(t) = p + (1 - p)S_0(t)$, where $S_0(t)$ is a proper survival function. Common choices for $S_0(t)$ are the Weibull, Gompertz and lognormal distributions [19]. [39] proposed a non-mixture model defined in terms of a cumulative hazard rate function. Its survival function has the form $S(t) = p^{F_0(t)}$, where $F_0(t)$ represents a proper distribution function. More about this method can be found in [28]. Many other methods are known for cure rate modeling, see, for example, [11], [34], [29] and the book [26].

Recently, there has been much interest with respect to cure rate models and many different approaches have been proposed to estimate quantities of interest. [10] proposed some Bayesian models to estimate cure fractions. [38]

*ricardorocha23@hotmail.com

discussed maximum likelihood techniques for cure models having a Cox proportional hazards structure. [35] used the Conway-Maxwell Poisson as the distribution of competing causes, as proposed in [34]. In [42], a unified approach is presented based on the Box-Cox transformation. In [30], an extension of the model presented in [42] is proposed and some model selection criteria are discussed. [1] proposed an expectation-maximization algorithm for estimation of the model proposed in [35], where the time-to-event was assumed to follow the exponential distribution. [3], [2] and [4] developed expectation-maximization algorithms with the time-to-event following the Weibull, lognormal and generalized gamma distributions.

Another way to model cure rates is to use defective distributions as explored in this paper. Defective distributions are characterized by having probability density functions which integrate to values less than 1 when the domain of some of their parameters is different from that usually defined. There is not so much literature about these distributions. There are at least two distributions in the literature that can be used for defective modeling: the Gompertz and inverse Gaussian distributions. The use of these defective distributions became more appealing after the works of [5] and [6], although some previous papers have used the same idea. In [41], the term "defective" was used to refer to the inverse Gaussian distribution that allows one of its parameters to be negative.

The Gompertz distribution becomes defective when its shape parameter is negative. It first appeared in [17], where it was used to model a breast cancer data set. [9] applied a modified version of this distribution to a pediatric cancer data set. [15] extended the distribution to include covariates. More recently, [33] performed Bayesian estimation of this distribution.

The inverse Gaussian distribution was first proposed in [37] for calculating the first time passage probability of a one-dimensional Brownian motion (Wiener process). More details were studied in [40] and [41]. Defective versions were investigated in [5] and [6], with classical and Bayesian approaches.

Having only two distributions is not enough to provide sufficient flexibility. In this paper, we derive a useful property of the Marshall Olkin family [27] of distributions which allows one to generate new defective distributions. The details are given in Section 2, including estimation by the method of maximum likelihood and an approach to include covariate information. Simulation studies are performed in Section 3 in order to check the usual asymptotic properties of maximum likelihood estimators and to assess the quality of maximum likelihood estimators. Four real data applications of the proposed methodology are illustrated in Section 4. Some concluding remarks are given in the last section.

In short, the contributions of this paper to the literature are: i) derive a new property of the Marshall Olkin family of distributions which allows for the construction of numerous defective distributions; ii) propose ten new defective distributions in order to exemplify the derived property; iii) illustrate the performance of such distributions through simulations and applications to real data sets.

2. Methodology

In this section, we present details about the Marshall Olkin family of distributions and derive a new property of this family that can be very useful for cure rate modeling. We also discuss the extended Weibull family of distributions, which together with the Marshall Olkin family can generate a whole set of new distributions for cure fraction estimation. Furthermore, we discuss details of maximum likelihood estimation and an approach to use the proposed distributions as regression models.

2.1. The Marshall Olkin family

Let $f(t)$, $S(t)$ and $\lambda(t)$ denote, respectively, the density, survival and hazard rate functions associated with a baseline distribution. The Marshall Olkin (MO) family, proposed in [27], extends the baseline distribution by adding an extra shape parameter, leading to a more flexible distribution often capable of providing better fits. The density, survival and hazard rate functions of the Marshall Olkin family are

$$f_{MO}(t;r) = \frac{rf(t)}{[1-(1-r)S(t)]^2}, \tag{1}$$

$$S_{MO}(t;r) = \frac{rS(t)}{1-(1-r)S(t)}, \tag{2}$$

$$\lambda_{MO}(t;r) = \frac{\lambda(t)}{1-(1-r)S(t)} \tag{3}$$

for $t > 0$ and $r > 0$.

There has been much work on the Marshall Olkin family of distributions. Many authors have derived details for particular Marshall Olkin distributions. For some examples, see [21] for the Marshall Olkin-uniform distribution, [12] for the Marshall Olkin-Pareto distribution, [14] for the Marshall Olkin-Weibull distribution, [13] for the Marshall Olkin-Lomax distribution, [32] for the Marshall Olkin-gamma distribution and [20] for the Marshall Olkin-beta distribution.

Theorem 2.1 derives a new property of the Marshall Olkin family that relates to the theory of defective distributions. This new property allows one to generate of new defective distributions.

Theorem 2.1. *Suppose $S(t)$ is an improper non-decreasing survival function satisfying $\lim_{t\to\infty} S(t) = \infty$, $S(t) \geq 1$, $\forall t \geq 0$, with the associated density function $f(t) \leq 0$, $\forall t \geq 0$. Then, the Marshall Olkin distribution given by (1) and (2), for $r < 0$, is a defective distribution.*

Proof: If $\lim_{t\to\infty} S(t) = \infty$ then

$$\begin{aligned}
\lim_{t\to\infty} S_{MO}(t;r) &= \lim_{t\to\infty} \frac{rS(t)}{1-(1-r)S(t)} \\
&\overset{L'H}{=} \frac{rS'(t)}{(r-1)S'(t)} \\
&= \frac{r}{r-1},
\end{aligned}$$

where $L'H$ indicates the use of the L'Hôpital rule. If $r < 0$ then $\frac{r}{r-1} \in (0,1)$. Also, if $f(t) \leq 0$, $\forall t \geq 0$ then $f_{MO}(t;r)$ is a positive density. The survival function

$$S_{MO}(t;r) = \frac{rS(t)}{1 - S(t) + rS(t)} = \frac{-rS(t)}{-rS(t) + [S(t) - 1]}$$

is positive, since r is negative and $S(t) \geq 1$. It is also non-increasing, as long as $S(t)$ is non-decreasing. The proof is complete. □

For example, the exponential distribution has survival function $S(t) = \exp(-at)$, $a > 0$. If $a < 0$, then $\lim_{t\to\infty} S(t) = \infty$, $S(t) \geq 1$, $\forall t \geq 0$, is non-decreasing and the respective density function is negative, for all t. Then the exponential distribution satisfies the conditions of Theorem 2.1. Therefore, the Marshall Olkin-exponential distribution is a defective distribution when $a < 0$ and $r < 0$.

Theorem 2.1 still holds for other values of $\lim_{t\to\infty} S(t)$. If $\lim_{t\to\infty} S(t) = M$, with $M \geq 0$, we have the cases: i) $M = 0$ then $S(t)$ is a proper survival function and therefore there is no cure rate; ii) $M \in (0,1)$, then the distribution is defective by definition; iii) $M = 1$, then the distribution is degenerate with the cure rate of 1 (no one would be susceptible to the event of interest); iv) $M > 1$, then the limiting cure rate will be $rM/(rM + 1 - M)$. If $\lim_{t\to\infty} S(t) = M$, with $M < 0$, or $\lim_{t\to\infty} S(t) = -\infty$, then Theorem 2.1 still holds if $S(t) \leq 1$, $\forall t > 0$ and $S(t)$ is non-increasing.

Note that $S_{MO}(t;r) = S(t)$ if and only if $r = 1$. If $S(t) > 1$ then $S_{MO}(t;r)$ increases as r becomes increasing negative. If $S(t) < 1$ then $S_{MO}(t;r)$ decreases as r becomes increasing negative. Also $\lim_{r\to-\infty} S_{MO}(t;r) = 1$.

Section 2.2 shows that a known family of extended Weibull distributions can give ideal choices for $S(t)$.

2.2. The extended Weibull distribution

The extended Weibull (EW) distribution, firstly proposed in [16], generalizes the Weibull distribution by means of a non-negative monotonically increasing function $H(t,\boldsymbol{\gamma})$, where $\boldsymbol{\gamma}$ is a vector of k parameters. Its density, survival and hazard rate functions are

$$f_{EW}(t;v,\boldsymbol{\gamma}) = v\, h(t,\boldsymbol{\gamma}) \exp\left[-vH(t,\boldsymbol{\gamma})\right], \tag{4}$$

$$S_{EW}(t;v,\boldsymbol{\gamma}) = \exp\left[-vH(t,\boldsymbol{\gamma})\right], \tag{5}$$

$$\lambda_{EW}(t;v,\boldsymbol{\gamma}) = v\, h(t,\boldsymbol{\gamma}) \tag{6}$$

for $t > 0$, $v > 0$ and $h(t,\boldsymbol{\gamma}) = dH(t,\boldsymbol{\gamma})/dt$.

Different choices for $H(t,\boldsymbol{\gamma})$ lead to different extended Weibull distributions. Table 1 lists ten extended Weibull distributions which will be used to illustrate Theorem 2.1. They were selected from [36].

Some more distributions for positive data can be obtained from the extended Weibull family: the Pareto distribution for $H(t,\boldsymbol{\gamma}) = \log(t/a)$, $t \geq a$; the log-logistic distribution for $H(t,\boldsymbol{\gamma}) = \log(1 + t^a)$; the Fréchet distribution for $H(t,\boldsymbol{\gamma}) = t^{-a}$; the exponential power distribution for $H(t,\boldsymbol{\gamma}) = \exp\left[(at)^b\right] - 1$; the Pham distribution for $H(t,\boldsymbol{\gamma}) = (a^t)^b - 1$, among others. For more details, see [36].

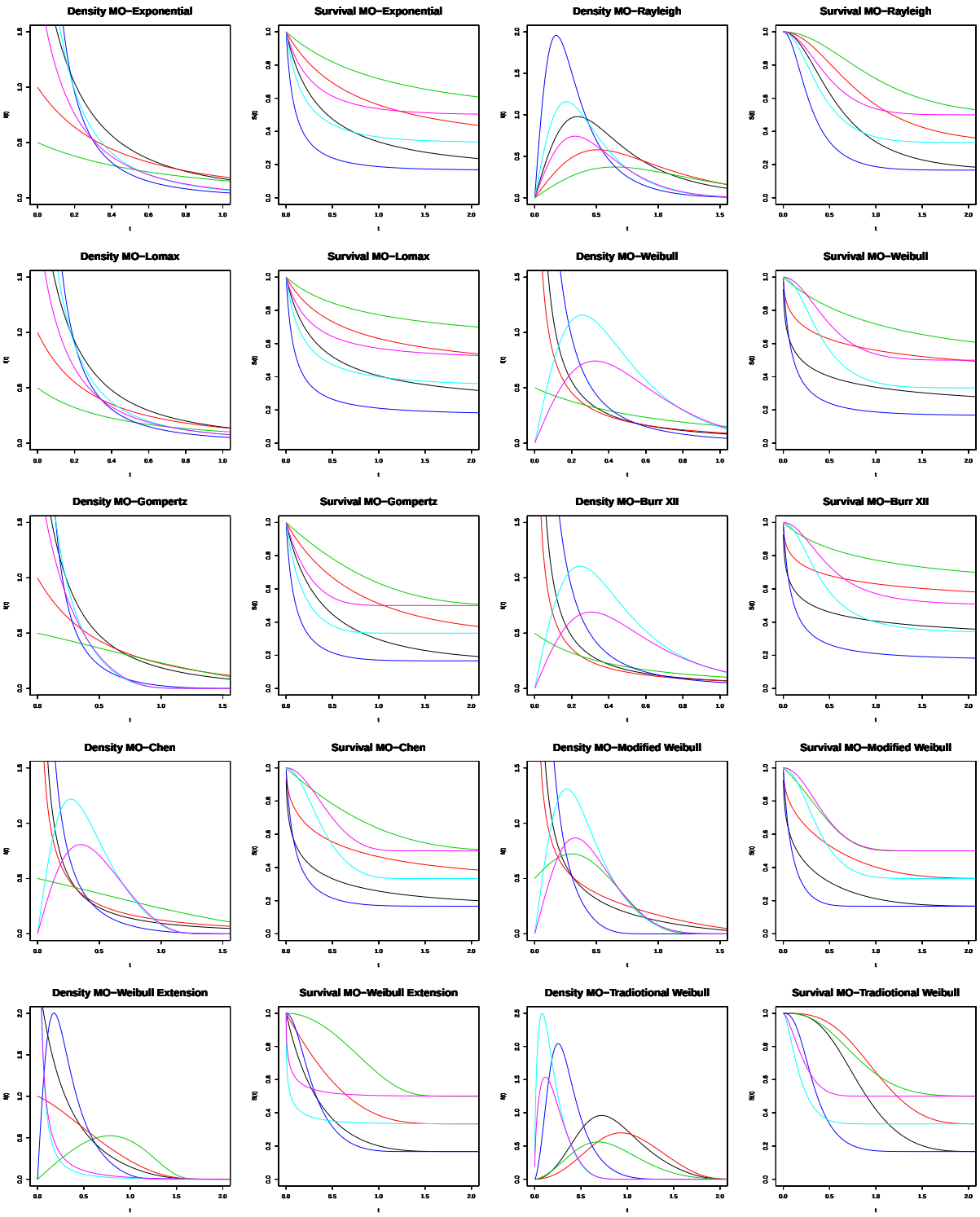

Figure 1: From the left to the right, from the top to the bottom, the density and survival functions of the proposed distributions, in the same order presented in Table 1. The parameter values used are $u = (-0.2, -0.5, -1, -0.2, -0.5, -1)$, $v = (-0.5, -0.5, -0.5, -2, -2, -2)$, $a = (0.5, 0.5, 1, 1, 2, 2)$, $b = (1, 1, 2, 2, 0.5, 0.5)$ and $c = (2, 2, 0.5, 0.5, 1, 1)$. The colors are (black, red, green, blue, light blue, pink).

Table 1: Some particular cases of the extended Weibull distribution.

Distribution	$H(t, \boldsymbol{\gamma})$	Parameters in $\boldsymbol{\gamma}$
Exponential	t	$\emptyset$
Rayleigh	t^2	$\emptyset$
Lomax	$\log(1+t)$	$\emptyset$
Weibull	t^a	$a > 0$
Gompertz	$[\exp(at) - 1]/a$	$a > 0$
Burr XII	$\log(1 + t^a)$	$a > 0$
Chen	$\exp(t^a) - 1$	$a > 0$
Modified Weibull	$t^a \exp(bt)$	$a \geq 0, b > 0$
Weibull extension	$a\left\{\exp\left[(t/a)^b\right] - 1\right\}$	$a > 0, b > 0$
Traditional Weibull	$t^b\left[\exp\left(at^c\right) - 1\right]$	$a \geq 0, b \geq 0, c > 0$

Note that some distributions in Table 1 are generalizations of others: the exponential and Rayleigh distributions are particular cases of the Weibull distribution for $a = 1$ and $a = 2$, respectively; the Lomax distribution is the particular case of the Burr XII distribution for $a = 1$; the Weibull distribution is the particular case of the modified Weibull distribution for $b = 0$; the Chen and Gompertz distributions are particular cases of the Weibull extension distribution for $a = 1$ and $a' = a^{-1}$, $b = 1$, respectively; the Weibull distribution is the particular case of the Chen distribution for $a = 1, b = 0$.

If $v < 0$ then $\lim_{t\to\infty} S_{EW}(t; v, \boldsymbol{\gamma}) = \infty$ provided that $H(t, \boldsymbol{\gamma})$ is non-negative and monotonically increasing. Also, it is easy to check that $S_{EW}(t; v, \boldsymbol{\gamma}) \geq 1$, is non-decreasing and $f_{EW}(t; v, \boldsymbol{\gamma}) \leq 0$, $\forall t \geq 0$, when $v < 0$. So, any member of the extended Weibull family that uses v as a parameter can be used to generate a defective distribution. This class of distributions is a good example on how one can build defective cure rate models. At first, there are no reasons why this class is more compelling than any other choice. But the class incorporates several common choices in the literature into one. Any function $H(t, \boldsymbol{\gamma})$ that satisfies the conditions of the class can be used to generate a cure rate model.

The Marshall Olkin-extended Weibull (MOeW) distributions are obtained by combining (4), (5), (6) and (1), (2), (3), i.e., by using the extended Weibull distribution as a baseline distribution for the Marshall Olkin family. It is important to note that, for $v < 0$, (4) and (5) are no longer density and survival functions. In this sense, the MOeW is not, precisely, a distribution belonging to the Marshall Olkin family. However, for convenience, we denote it as a distribution of the family. Moreover, properties of these special functional forms have not been studied in the literature, yet. Because of that, we cannot use properties of the extended Weibull family to derive any result for the MOeW distribution. The resulting density, survival and hazard rate functions are

$$f_{MOeW}(t; r, v, \boldsymbol{\gamma}) = \frac{r\, v\, h(t, \boldsymbol{\gamma}) \exp\left[-vH(t, \boldsymbol{\gamma})\right]}{\left\{1 - (1 - r)\exp\left[-vH(t, \boldsymbol{\gamma})\right]\right\}^2}, \tag{7}$$

$$S_{MOeW}(t; r, v, \boldsymbol{\gamma}) = \frac{r \exp[-vH(t, \boldsymbol{\gamma})]}{1 - (1 - r)\exp[-vH(t, \boldsymbol{\gamma})]}, \tag{8}$$
$$\lambda_{MOeW}(t; r, v, \boldsymbol{\gamma}) = \frac{v\, h(t, \boldsymbol{\gamma})}{1 - (1 - r)\exp[-vH(t, \boldsymbol{\gamma})]}.$$

The extra parameter of the Marshall Olkin family shifts the hazard rate of the new distribution above, or below, the hazard rate of the baseline distribution [**?**]. This means that $\lambda_{MO}(t) \leq \lambda(t)$, when $r \geq 1$ and $\lambda_{MO}(t) \geq \lambda(t)$, when $0 < r \leq 1$, for all t. The resulting distribution under the Marshall Olkin family has a more flexible hazard rate function than the baseline hazard rate function. On the other hand, one of the merits of the extended Weibull class is the flexibility of its hazard rate function, given by (6). Therefore, one can expect a duplicated effect when both classes are combined.

The ten different functions in Table 1 lead to ten different defective distributions. Figure 1 plots the density and survival functions of all distributions proposed in Table 1. This collection of distributions can be very flexible. The black and blue curves in the figure have the same cure rate of $-0.2/(-0.2 - 1) = 1/6$. The red and light blue curves have the curve rate of $-0.5/(-0.5 - 1) = 1/3$. The green and pink curves have the cure rate of $-1/(-1 - 1) = 1/2$.

We have used the extended Weibull family to generate defective distributions. The generated distributions give good fits to the data considered in this paper. But other distributions could have been used to generate defective versions via Theorem 2.1. As an example, consider the Maxwell-Boltzmann distribution specified by the density and survival functions

$$f_{MB}(t; a) = a^{-3} t^2 \exp\left(-\frac{t^2}{2a^2}\right) \sqrt{2\pi^{-1}},$$
$$S_{MB}(t; a) = 1 - \operatorname{erf}\left(\frac{t}{\sqrt{2}a}\right) + \frac{t \exp\left(-\frac{t^2}{2a^2}\right) \sqrt{2\pi^{-1}}}{a}$$

for $t > 0$, where $a > 0$ is a scale parameter and $\operatorname{erf}(t) = 2\pi^{-\frac{1}{2}} \int_0^t e^{-x^2} dx$ denotes the error function. The error function approaches 1 as $t \to \infty$ and approaches -1 as $t \to -\infty$. If $a < 0$ we have

$$\lim_{t \to \infty} S_{MB}(t; a) = 1 - (-1) + 0 = 2.$$

We also have that $f_{MB}(t; a) \leq 0$, $S_{MB}(t; a) \geq 1$ and $S_{MB}(t; a)$ is non-decreasing, $\forall t > 0$. So, this distribution under the Marshall Olkin family is defective when $a < 0$ and $r < 0$. Its cure rate is $2r/(2r + 1 - 2) = r/(r - 0.5)$.

2.3. Inference

Here, we present a procedure to obtain maximum likelihood estimates for the MOeW distribution. We consider data with right-censored information. Let$\mathbf{D} = (\boldsymbol{t}, \boldsymbol{\delta})$, where $\mathbf{t} = (t_1, \ldots, t_n)'$ are the observed failure times and $\boldsymbol{\delta} = (\delta_1, \ldots, \delta_n)'$ are the right-censored times. The δ_i is equal to 1 if a failure is observed and 0 otherwise. Suppose that the data are independently and identically distributed and come from a distribution with density and survival functions specified by $f(\cdot, \boldsymbol{\theta})$ and $S(\cdot, \boldsymbol{\theta})$, respectively, where $\boldsymbol{\theta} = (r, v, \boldsymbol{\gamma})'$ denotes a vector of $k + 2$ parameters. The

log-likelihood function of $\boldsymbol{\theta}$ can be written as

$$l(\boldsymbol{\theta}, \mathbf{D}) = \log L(\boldsymbol{\theta}, \mathbf{D}) = \text{const} + \sum_{i=1}^{n} \delta_i \log f(t_i, \boldsymbol{\theta}) + (1 - \delta_i) \log S(t_i, \boldsymbol{\theta}). \tag{9}$$

By (7) and (8), the log-likelihood function for the MOeW distribution is

$$l(\boldsymbol{\theta}, \mathbf{D}) = \text{const} + n \log(r) - v \sum_{i=1}^{n} H(t_i, \boldsymbol{\gamma}) - \sum_{i=1}^{n} (1 + \delta_i) \log\{1 - (1 - r) \exp[-vH(t_i, \boldsymbol{\gamma})]\} + \sum_{i=1}^{n} \delta_i \log[vh(t_i, \boldsymbol{\gamma})].$$

The maximum likelihood estimates are the simultaneous solutions of $\frac{\partial l(\boldsymbol{\theta}, \mathbf{D})}{\partial r} = 0$, $\frac{\partial l(\boldsymbol{\theta}, \mathbf{D})}{\partial v} = 0$ and $\frac{\partial l(\boldsymbol{\theta}, \mathbf{D})}{\partial \gamma_j} = 0$. The estimates are obtained using the BFGS algorithm of maximization, which is an option for the `optim` function in R [31].

If $\widehat{\boldsymbol{\theta}}$ denotes the maximum likelihood estimator of $\boldsymbol{\theta}$ then it is well known that the distribution of $\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}$ can be approximated by a $(k + 2)$-variate normal distribution with zero means and covariance matrix $\mathbf{I}^{-1}\left(\widehat{\boldsymbol{\theta}}\right)$, where $\mathbf{I}(\boldsymbol{\theta})$ denotes the observed information matrix defined by

$$\mathbf{I}(\boldsymbol{\theta}) = -\left(\frac{\partial^2 l(\boldsymbol{\theta}, \mathbf{D})}{\partial \theta_i \partial \theta_j}\right)$$

for i and j in $1, 2, \ldots, k + 2$. This approximation can be used to deduce confidence intervals and tests of hypotheses. For example, an approximate $100(1 - \alpha)$ percent confidence interval for θ_i is $\left(\widehat{\theta_i} - z_{\alpha/2}\sqrt{I^{ii}}, \widehat{\theta_i} + z_{\alpha/2}\sqrt{I^{ii}}\right)$, where I^{ii} denotes the ith diagonal element of the inverse of $\mathbf{I}$ and z_a denotes the $100(1 - a)$ percentile of a standard normal random variable.

Asymptotic normality of the maximum likelihood estimates holds only under certain regularity conditions. These conditions are not easy to check analytically for our models. Section 3 performs a simulation study to see if the usual asymptotes of the maximum likelihood estimates hold. Simulations have been used in many papers to check the asymptotic behavior of maximum likelihood estimates, especially when an analytical investigation is not trivial.

2.4. Defective Marshall Olkin-G regression model

The use of covariate information is essential when analysing survival data. Here, we discuss an approach on how to include covariate information to the proposed models. The approach has a simple interpretation as we shall see.

Suppose $\mathbf{x}' = \left(1, x_1, \ldots, x_p\right)$ is a vector of covariates from a data set and $\boldsymbol{\beta}' = \left(\beta_0, \beta_1, \ldots, \beta_p\right)$ a vector of regression coefficients. We are going to set $r(\boldsymbol{x}) = -\exp(\boldsymbol{\beta}'\boldsymbol{x})$ to link the cure rate to the covariates. This way, the Marshall Olkin-G regression model is given by

$$S(t|\boldsymbol{x}) = \frac{r(\boldsymbol{x})S(t)}{1 - [1 - r(\boldsymbol{x})]S(t)} = \frac{\exp(\boldsymbol{\beta}'\boldsymbol{x})S(t)}{[1 + \exp(\boldsymbol{\beta}'\boldsymbol{x})]S(t) - 1}$$

for $t > 0$. If $S(t)$ has a cure rate of p then that of $S(t|\boldsymbol{x})$ is

$$p = \lim_{t \to \infty} S(t|\boldsymbol{x}) = \frac{r(\boldsymbol{x})}{r(\boldsymbol{x}) - 1} = \frac{\exp(\boldsymbol{\beta}'\boldsymbol{x})}{1 + \exp(\boldsymbol{\beta}'\boldsymbol{x})}. \tag{10}$$

This way, the cure fraction is easily calculated through the logit function. This approach is attractive because of the way the cure rate depends on the regression coefficients, making it very easy to interpret. If $\boldsymbol{\beta}'\boldsymbol{x}$ increases, so does the cure rate (towards 1). If $\boldsymbol{\beta}'\boldsymbol{x}$ decreases, so does the cure rate (towards 0).

The MOeW regression model is given by

$$S(t|\boldsymbol{x}) = \frac{r(\boldsymbol{x})\exp\left[-vH(t,\boldsymbol{\gamma})\right]}{1-\left[1-r(\boldsymbol{x})\right]\exp\left[-vH(t,\boldsymbol{\gamma})\right]} = \frac{\exp(\boldsymbol{\beta}'\boldsymbol{x})\exp\left[-vH(t,\boldsymbol{\gamma})\right]}{\left[1+\exp(\boldsymbol{\beta}'\boldsymbol{x})\right]\exp\left[-vH(t,\boldsymbol{\gamma})\right]-1}.$$

An application is presented in Section 4.4. A much more detailed application of the regression models will be the focus of a future paper.

3. Simulation studies

Here, we assess the performance of the maximum likelihood estimates with respect to sample size to show, among other things, that the usual asymptotes of maximum likelihood estimators still hold for defective distributions. The assessment is based on simulations. The description of data generation and details of the distributions simulated from are described below. All computations were performed in R [31].

Suppose that the time of occurrence of an event of interest has cumulative distribution function $F(t)$. We want to simulate a random sample of size n containing real times, censored times and a cure fraction of p. An algorithm for this purpose is:

- Determine the desired parameter values, as well as the value of the cure fraction p;
- For each $i = 1, \ldots, n$, generate a random variable $M_i \sim \text{Bernoulli}(1-p)$;
- If $M_i = 0$ set $t_i' = \infty$. If $M_i = 1$ take t_i' as the root of $F(t) = u$, where $u \sim \text{uniform}(0, 1-p)$;
- Generate $u_i' \sim \text{uniform}\left(0, \max\left(t_i'\right)\right)$, for $i = 1, \ldots, n$, considering only the finite t_i';
- Calculate $t_i = \min\left(t_i', u_i'\right)$. If $t_i < u_i'$ set $\delta_i = 1$, otherwise set $\delta_i = 0$.

We took the sample size to vary from 50 to 1000 in steps of 50. Each sample was replicated 1000 times. The variance of the cure rate p was estimated using the delta method with first order Taylor's approximation. We chose only four of our proposed distributions: the Marshall Olkin-Lomax distribution with $(r, v) = (-1, -10)$, the simulation results for which are shown in Figure 2; the Marshall Olkin-Weibull distribution with $(r, v, a) = (-1, -2, 3)$, the simulation results for which are shown in Figure 3; the Marshall Olkin-Chen distribution with $(r, v, a) = (-1, -2, 2)$, the simulation results for which are shown in Figure 4; the Marshall Olkin-Burr XII distribution with $(r, v, a) = (-1, -2, 2)$, the simulation results for which are shown in Figure 5. For the purpose of comparison, we have fixed $r = -1$ for all simulations, which leads to a cure rate of 0.5.

We can observe the following from the figures: the biases for each parameter approach zero as sample size increases; the biases for each parameter appear small enough for all $n \geq 600$; the mean squared errors for each parameter decrease to zero as sample size increases; the mean squared errors for each parameter appear small enough

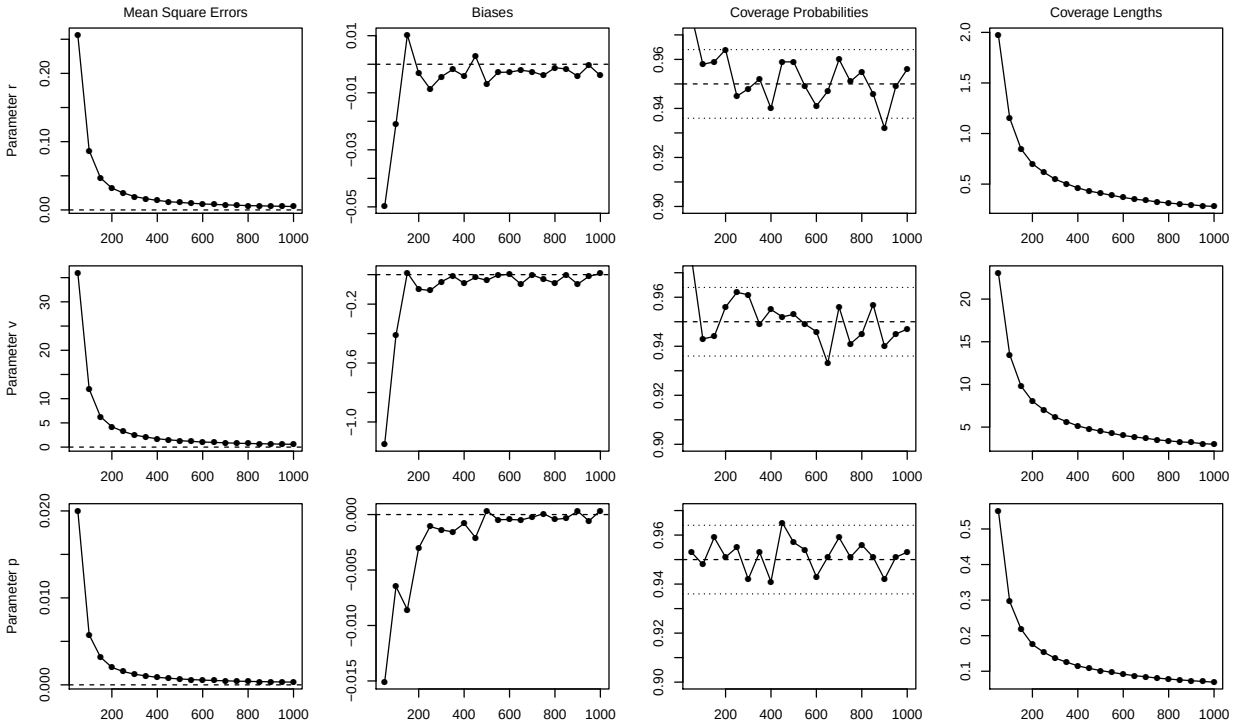

Figure 2: Mean squared errors, biases, coverage probabilities and coverage lengths of the estimators of r, v and p versus n for the Marshall Olkin-Lomax distribution with $(r, v) = (-1, -10)$.

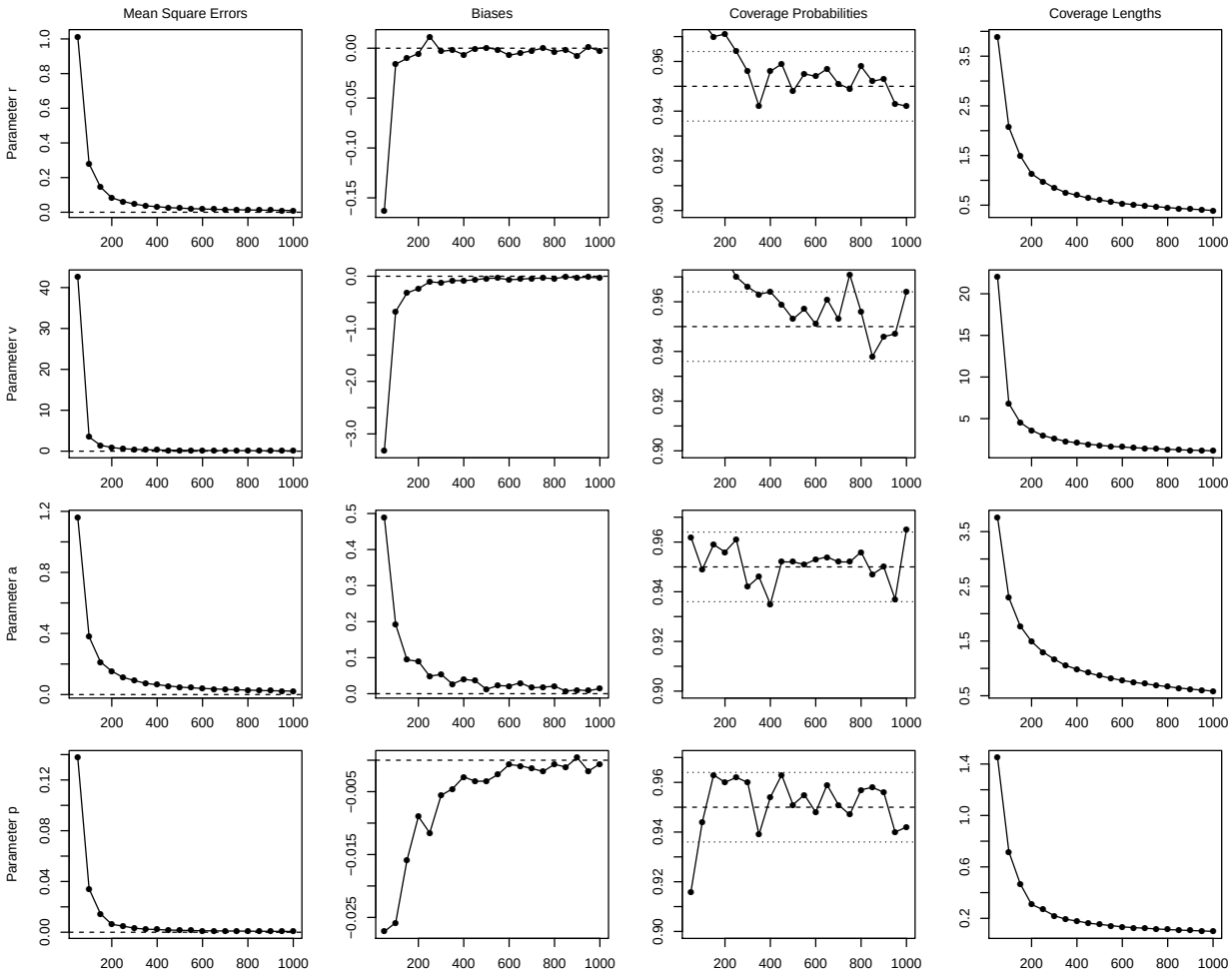

Figure 3: Mean squared errors, biases, coverage probabilities and coverage lengths of the estimators of r, v, a and p versus n for the Marshall Olkin-Weibull distribution with $(r, v, a) = (-1, -2, 3)$.

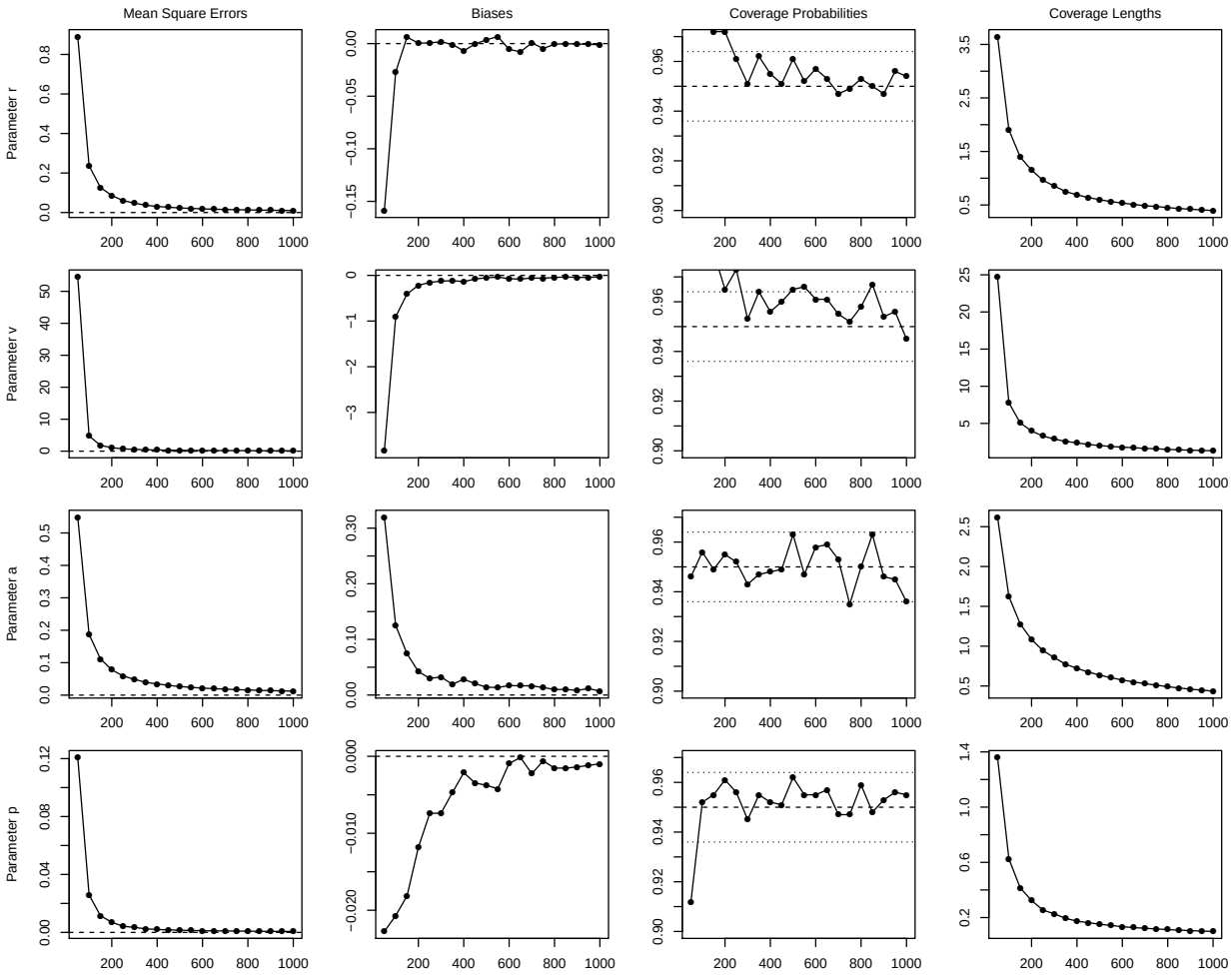

Figure 4: Mean squared errors, biases, coverage probabilities and coverage lengths of the estimators of r, v, a and p versus n for the Marshall Olkin-Chen distribution with $(r, v, a) = (-1, -2, 2)$.

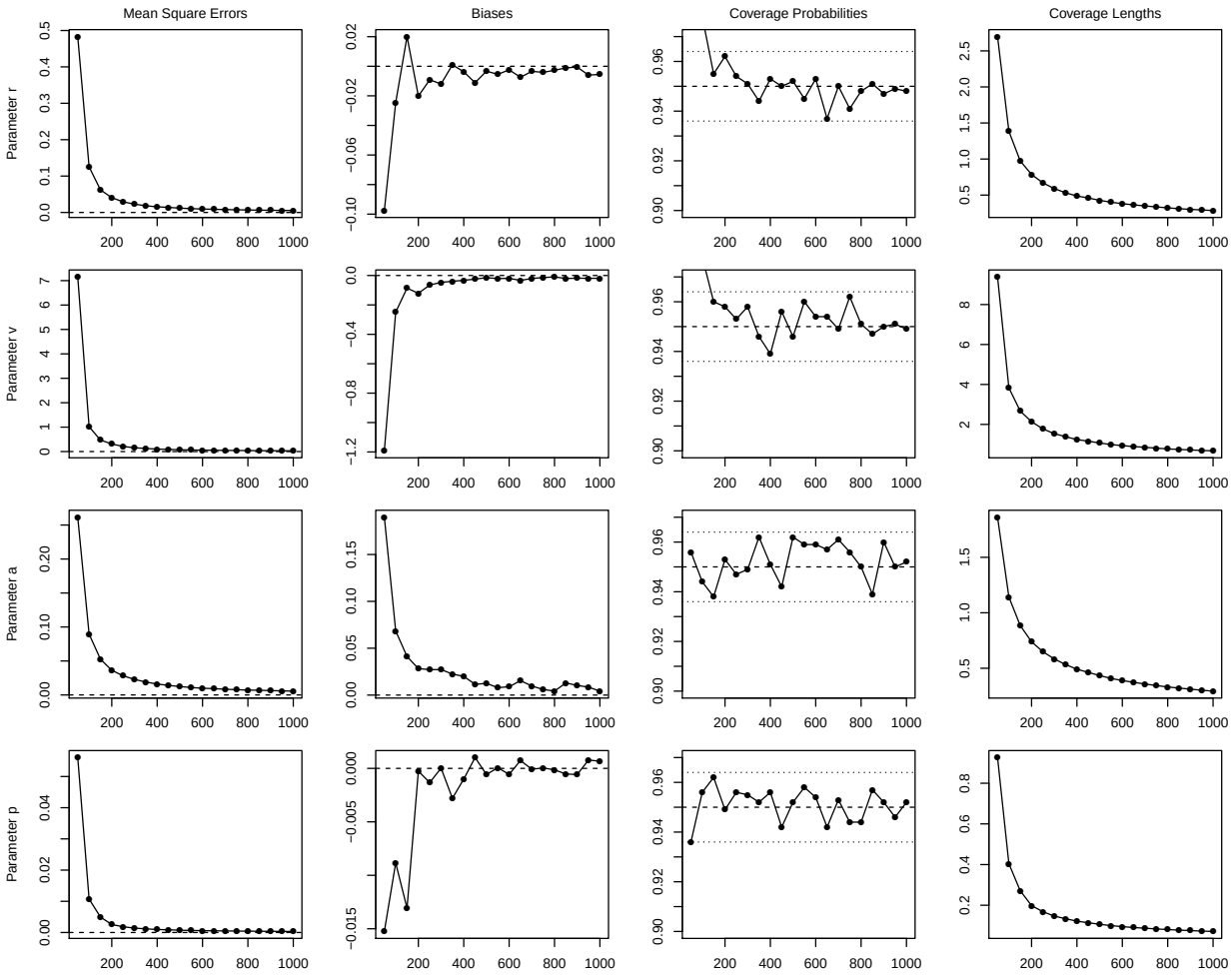

Figure 5: Mean squared errors, biases, coverage probabilities and coverage lengths of the estimators of r, v, a and p versus n for the Marshall Olkin-Burr XII distribution with $(r, v, a) = (-1, -2, 2)$.

for all $n \geq 600$; the coverage probabilities for each parameter stay mostly in the interval $(0.936, 0.964)$; the coverage lengths for each parameter decrease fast to zero as sample size increases; the coverage lengths for each parameter appear small enough for all $n \geq 600$.

Similar observations held when the simulations were repeated for other defective distributions and for a wide range of parameter values under the Marshall Olkin family. In particular, the biases always approached zero as sample size increased, the biases for each parameter always appeared small enough for all $n \geq 600$, the mean squared errors always approached zero as sample size increased, the mean squared errors for each parameter always appeared small enough for all $n \geq 600$, the coverage probabilities always stayed mostly in the interval $(0.936, 0.964)$, the coverage lengths always decreased fast to zero as sample size increased and the coverage lengths for each parameter always appeared small enough for all $n \geq 600$.

4. Real data applications

Here, we present applications to four real data sets. For the first three data sets, we are only considering the event times and censoring information, with no covariates. The fourth data set contains covariate information and is used to illustrate the model proposed in Section 2.4. The ten defective distributions discussed in Section 2.2 are fitted to each data set. The following are used to distinguish between the fitted distributions: the Akaike information criterion (AIC), the Bayesian information criterion (BIC), the consistent Akaike information criterion (CAIC) and visual comparison of the fitted survival curves and the Kaplan-Meier [22] curve. All the computations were performed using the R software [31]. `optim` was used to maximize the log-likelihood function. The algorithm "BFGS" was chosen for maximization. For computational stability, the observed times in each data set were divided by their maximum value. The parameters r and v were set free to take any value on the real line. Negative estimates of r and v correspond to a defective model. Positive estimates of r and v correspond to a proper survival model.

The four data sets were chosen to show a variety of survival curves and sample sizes. Each data set is supposed to contain observations not susceptible to the event of interest. In practice, it is unknown if the event of interest could be observed if enough time was given. An evidence of existence of cured individuals is when the Kaplan-Meier curve reaches a plateau between zero and one. In some cases that is more clear than others, as one can see in our examples. We can assume that some of the censored observations at the end of the study belong to the cured group. If everyone censored at the end are indeed cured, then the plateau reached by the Kaplan-Meier curve is a good estimate of the cure fraction. In general, a lower value of this plateau or a value close to it is an acceptable estimate.

4.1. Leukemia data

This data set relates to a study of recurrence of leukemia in patients who were submitted to a certain kind of transplantation. Leukemia is a type of cancer that affects the white blood cells produced by the bone marrow and can take several forms. There are 44 observed times, of which 9 were censored (20.45 percent). The overall survival is the

average observed time in a study, including the censored elements. For this data set, the overall survival is 0.99 years. For details of this data, see [23]. The Kaplan-Meier curve for this data stabilizes at 0.1988. It appears quite safe to say that this value is an asymptote of the curve.

Table 2: MLEs for the fitted distributions and some measures for the leukemia data set.

MO-Distribution	$\widehat{r}$	$\widehat{v}$	$\widehat{a}$	$\widehat{b}$	$\widehat{c}$	$\widehat{p}$	AIC	BIC	CAIC
Exponential	-0.2798	-3.5807	-	-	-	0.2186	-54.86	-51.29	-54.57
Rayleigh	-0.1932	-37.6941	-	-	-	0.1619	**-75.39**	**-71.82**	**-75.10**
Lomax	-0.2296	-2.9506	-	-	-	0.1867	-54.03	-50.46	-53.74
Weibull	-0.2064	-33.0993	1.9246	-	-	0.1711	-73.50	-68.15	-72.90
Gompertz	-0.2557	-2.5202	6.3571	-	-	0.2036	-58.31	-52.96	-57.71
Burr XII	-0.2103	-35.3517	1.9384	-	-	0.1738	-73.93	-68.58	-73.33
Chen	-0.2015	-30.8096	1.9111	-	-	0.1677	-73.00	-67.65	-72.40
Modified Weibull	-0.2064	-33.224	1.9259	0.0002	-	0.1711	-71.50	-64.37	-70.48
Weibull extension	-0.2053	-56.4223	1.935	1.9045	-	0.1703	-71.36	-64.22	-70.33
Traditional Weibull	-0.2057	-1.6413	3.1087	1.5997	0.1237	0.1706	-69.09	-60.17	-67.51

The fitted results are summarized in Table 2 and Figure 6. Every distribution is estimated as a defective distribution. The cure rate estimates are around 0.02 lower than the value suggest by the Kaplan-Meier curve. The Marshall Olkin-Rayleigh distribution gives the smallest values for AIC, BIC and CAIC, suggesting it fits better than the others. Its estimate of the cure fraction is furthest from the one suggested by the Kaplan-Meier curve, 0.1619, but still an acceptable estimate. The Marshall Olkin-Weibull, Marshall Olkin-Burr XII, Marshall Olkin-Chen and Marshall Olkin-Modified Weibull distributions also provide reasonable fits. All other distributions perform poorly. Visual comparison of the fitted survival curves and the Kaplan-Meier curve shows that the Marshall Olkin-Exponential, Marshall Olkin-Lomax, Marshall Olkin-Gompertz and Marshall Olkin-Weibull extension distributions provide the worst fits.

4.2. Colon data

This data set arises from one of the first successful trials of adjuvant chemotherapy for colon cancer. The event of interest here is the recurrence or death for the individual under the proposed treatment. There are 1858 observed times, of which 938 were censored (50.58 percent). For this data set, the overall survival is 4.21 years. For details of this data, see [24]. The Kaplan-Meier curve for this data stabilizes at 0.4651. However, we can see some subjects failing near the end of the study. So, values of the cure rate a little lower than what the Kaplan-Meier curve suggests are expected and acceptable.

The fitted results are summarized in Table 3 and Figure 7. All of the fitted distributions are estimated to being defective. The Marshall Olkin-Lomax distribution estimates the cure fraction as 0.1858, far lower than the Kaplan-Meier plateau. The Marshall Olkin-Exponential distribution gives the estimate 0.3699 and the Marshall Olkin-Weibull distribution gives the estimate 0.4198. All others give a value very close to the Kaplan-Meier estimate. The Marshall Olkin-Burr XII distribution has the smallest values for AIC, BIC and CAIC. This distribution gives a cure rate of

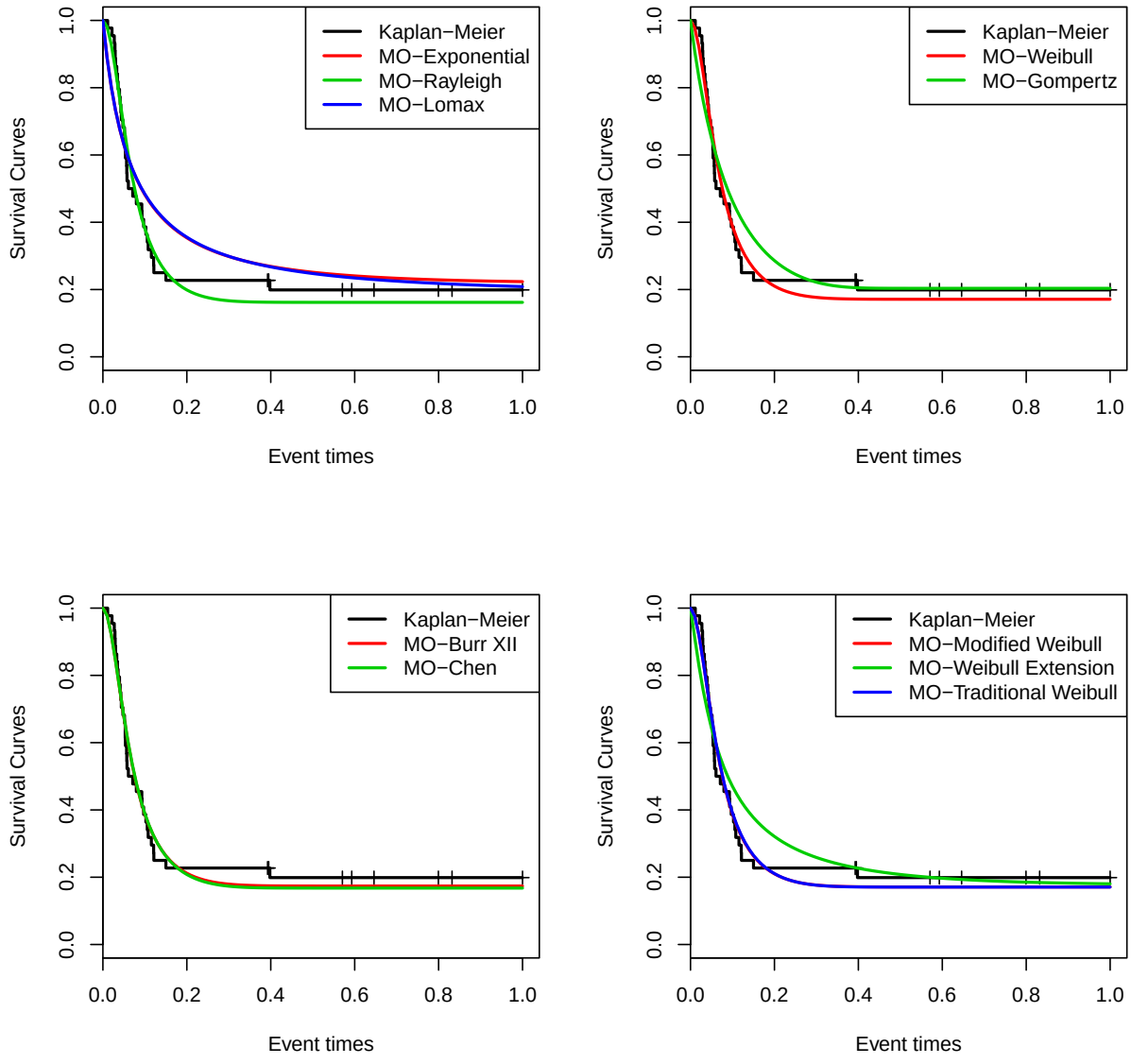

Figure 6: Fitted distributions for the leukemia data set.

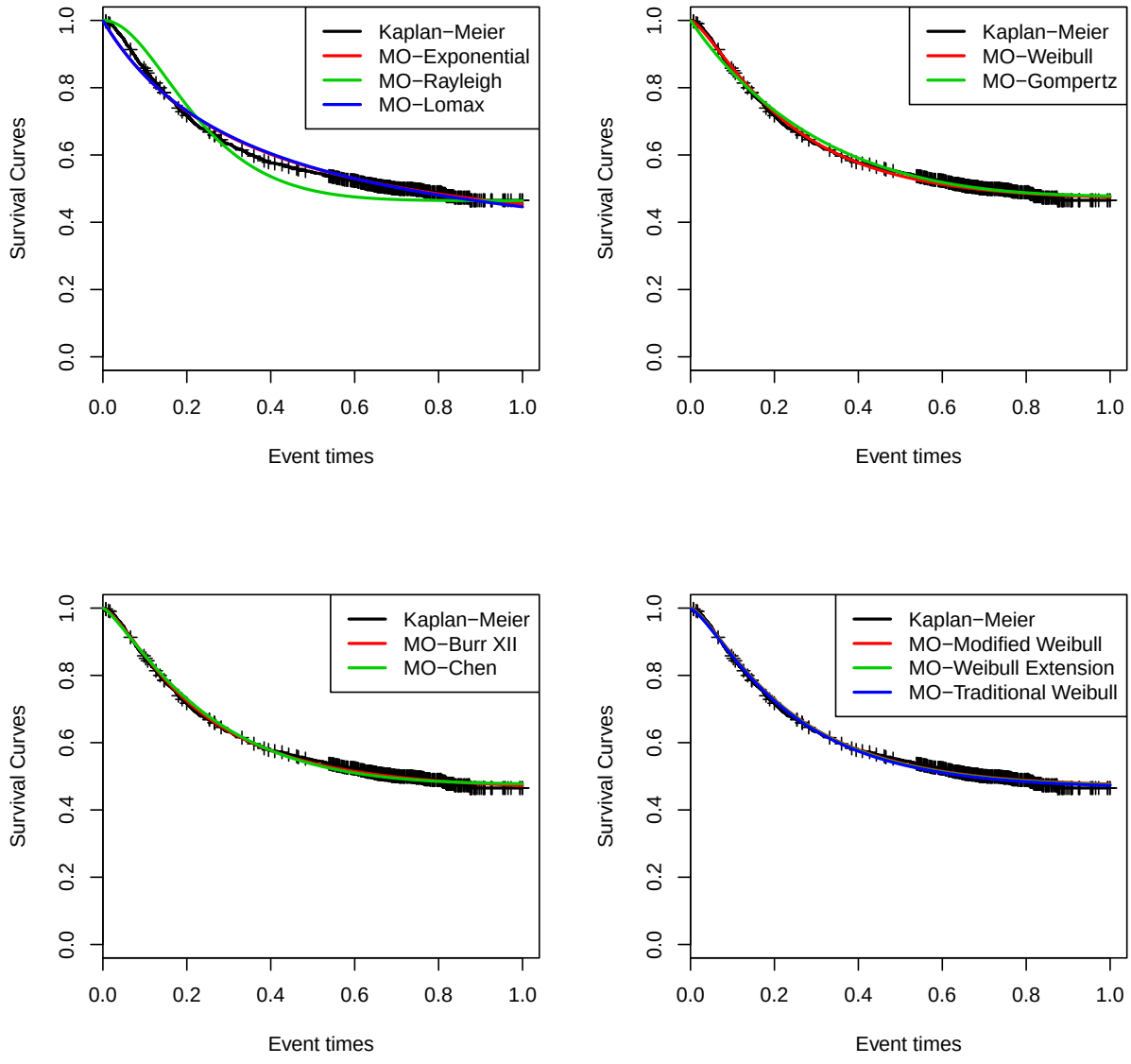

Figure 7: Fitted distributions for the colon data set.

Table 3: MLEs for the fitted distributions and some measures for the colon data set.

MO-Distribution	$\widehat{r}$	$\widehat{v}$	$\widehat{a}$	$\widehat{b}$	$\widehat{c}$	$\widehat{p}$	AIC	BIC	CAIC
Exponential	-0.5871	-1.2272	-	-	-	0.3699	1531.10	1542.15	1531.10
Rayleigh	-0.8655	-8.6495	-	-	-	0.464	1668.31	1679.36	1668.32
Lomax	-0.2282	-0.4812	-	-	-	0.1858	1537.00	1548.06	1537.01
Weibull	-0.8805	-3.6376	1.367	-	-	0.4682	1462.36	1478.94	1462.38
Gompertz	-0.9101	-1.6598	1.9054	-	-	0.4765	1516.68	1533.27	1516.70
Burr XII	-0.8381	-3.9545	1.4167	-	-	0.456	**1456.53**	**1473.11**	**1456.54**
Chen	-0.9114	-3.0808	1.2918	-	-	0.4768	1474.74	1491.32	1474.75
Modified Weibull	-0.8809	-3.6404	1.3672	0.0014	-	0.4683	1464.39	1486.50	1464.41
Weibull extension	-0.8805	-14.1338	40.9902	1.366	-	0.4682	1464.42	1486.52	1464.44
Traditional Weibull	-0.8805	-1.3754	1.2936	1.355	0.0068	0.4682	1466.37	1494.00	1466.40

0.456, slightly lower than the Kaplan-Meier estimate and is probably the best for this data. Note that the Marshall Olkin-Weibull, Marshall Olkin-Modified Weibull, Marshall Olkin-Weibull extension and Marshall Olkin-Traditional Weibull distributions give practically the same cure rate estimate of 0.4682, very close to the Kaplan-Meier estimate.

Visual comparison of the fitted survival curves and the Kaplan-Meier curve shows that the Marshall Olkin-Rayleigh distribution gives the worst fit (and the worst measures for AIC, BIC and CAIC too). The Marshall Olkin-Exponential and Marshall Olkin-Lomax distributions provide a better comparison, but their fits are worst than all others (also in agreement with the AIC, BIC and CAIC values). The remaining distributions seem to fit the Kaplan-Meier curve well. The cure rate asymptotes for the Marshall Olkin-Modified Weibull, Marshall Olkin-Exponential and Marshall Olkin-Lomax distributions are after the end of the study.

4.3. Divorce data

This data set collected in the USA describes married couples and the event of interest is the divorce. Of course, that event may never occur, there is a high censoring in this data set. The cure elements are those couples who will never divorce. There are 3371 observed times, of which 2339 were censored (69.38 percent). The maximum observed time was 73.07 years and the overall survival is 18.41 years. For details of this data, see [25]. The Kaplan-Meier curve for this data stabilizes at 0.5566. It appears quite safe to say that this value is an asymptote of the curve. Almost no failures were observed in the second half of the period of study. So, we can expect a real cure fraction quite close to the Kaplan-Meier estimate.

The fitted results are summarized in Table 4 and Figure 8. The Marshall Olkin-Chen distribution has the smallest values for AIC, BIC and CAIC. Its cure estimate is 0.5555, the closest to the Kaplan-Meier estimate. Its fit captures the Kaplan-Meier curve very well. Therefore, we can consider Marshall Olkin-Chen distribution as giving the most adequate fit. The Marshall Olkin-Modified Weibull, Marshall Olkin-Weibull extension and Marshall Olkin-Traditional Weibull distributions also give very close fits as the Marshall Olkin-Chen distribution. Their measures differ basically because of the difference in the number of parameters. The simplest Marshall Olkin-Exponential, Marshall Olkin-

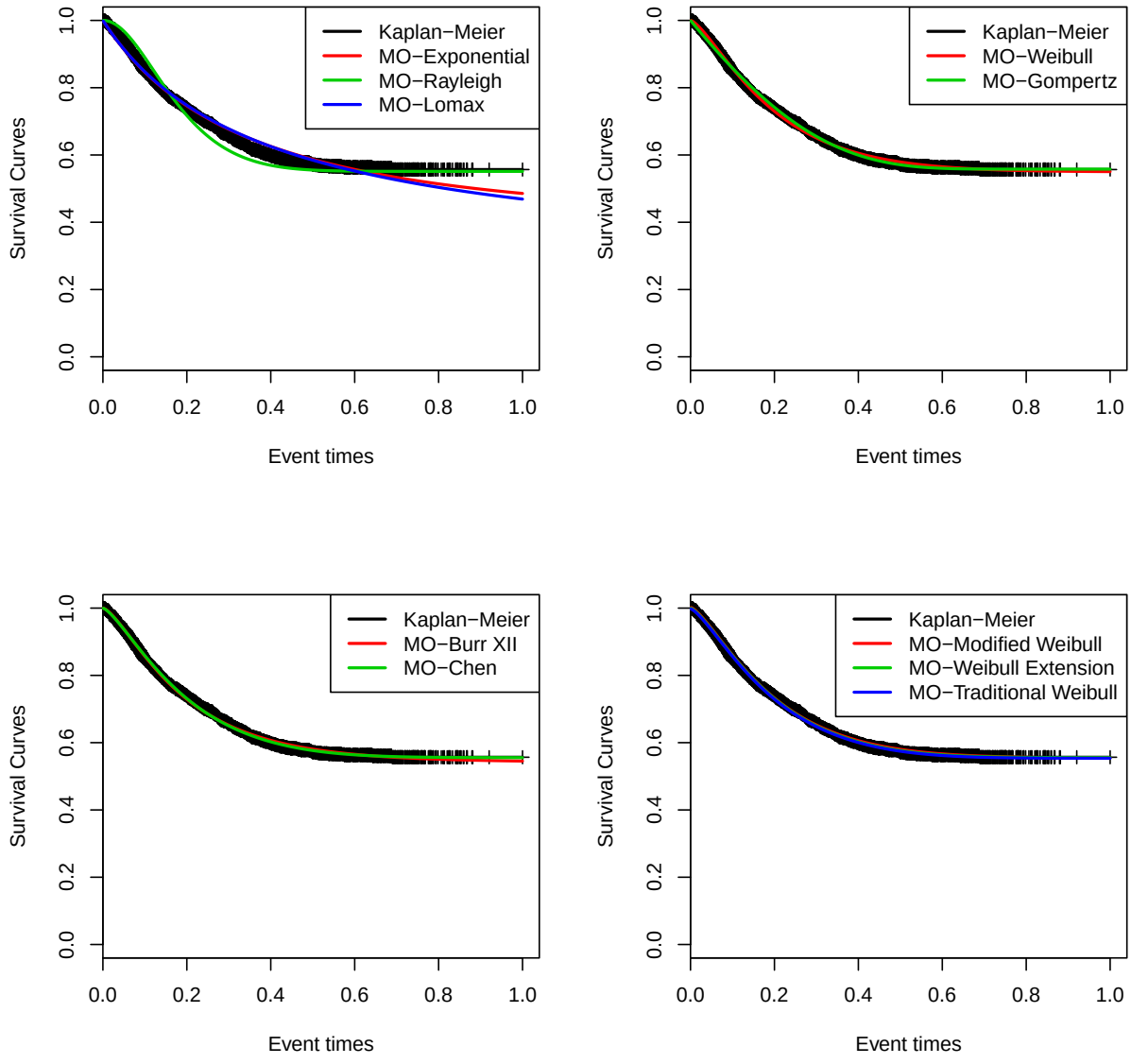

Figure 8: Fitted distributions for the divorce data set.

Table 4: MLEs for the fitted distributions and some measures for the divorce data set.

MO-Distribution	$\widehat{r}$	$\widehat{v}$	$\widehat{a}$	$\widehat{b}$	$\widehat{c}$	$\widehat{p}$	AIC	BIC	CAIC
Exponential	-0.7037	-1.3674	-	-	-	0.4130	1532.16	1544.41	1532.17
Rayleigh	-1.2283	-16.5389	-	-	-	0.5512	1633.88	1646.13	1633.89
Lomax	-0.2604	-0.5045	-	-	-	0.2066	1538.82	1551.07	1538.83
Weibull	-1.2215	-5.8018	1.4083	-	-	0.5499	1435.90	1454.27	1435.90
Gompertz	-1.2622	-1.9082	3.9220	-	-	0.5579	1471.46	1489.82	1471.46
Burr XII	-1.1819	-6.1051	1.4355	-	-	0.5417	1437.27	1455.64	1437.28
Chen	-1.2498	-5.2752	1.3694	-	-	0.5555	**1435.01**	**1453.38**	**1435.02**
Modified Weibull	-1.2350	-5.3745	1.3820	0.2300	-	0.5526	1437.73	1462.22	1437.74
Weibull extension	-1.2497	-5.2844	1.0033	1.3696	-	0.5555	1437.01	1461.51	1437.03
Traditional Weibull	-1.2499	-5.2052	1.0112	0.0036	1.3651	0.5555	1439.02	1469.64	1439.04

Rayleigh and Marshall Olkin-Lomax distributions all give poor fits. The remaining distributions provide reasonably good fits with respect to AIC, BIC and CAIC measures as well as visual comparison to the Kaplan-Meier curve. Their cure rate estimates are quite close to the value suggested by the Kaplan-Meier curve.

4.4. Melanoma data

This data set collected in the period 1991-1998 is related to a clinical study in which patients were observed for recurrence after a removal of a malignant melanoma. Melanoma is a type of cancer that develops in melanocytes, responsible for skin pigmentation. It is a potentially serious malignant tumor that may arise in the skin, mucous membranes, eyes and the central nervous system, with a great risk of producing metastases and high mortality rates in the later stages. There are 417 observed times, of which 232 were censored (55.63 percent). The overall survival is 3.18 years. This data set has covariate information, which is used to illustrate the regression model proposed in Section 2.4. The covariate taken represents the nodule category ($n_1 = 82$, $n_2 = 87$, $n_3 = 137$, $n_4 = 111$). The overall survival for the categories are 3.60, 3.27, 3.07, 2.55 years. For more details of this data, see [18].

Table 5: MLEs for the fitted regression models and the AIC measure for the melanoma data set.

MO-Distribution	$\widehat{v}$	$\widehat{a}$	$\widehat{b}$	$\widehat{c}$	$\widehat{\beta}_0$	$\widehat{\beta}_1$	$\widehat{p}_1$	$\widehat{p}_2$	$\widehat{p}_3$	$\widehat{p}_4$	AIC
Exponential	-0.03	-	-	-	-2.84	-0.43	0.0365	0.0240	0.0156	0.0102	354.12
Rayleigh	-5.99	-	-	-	1.22	-0.51	0.6697	0.5485	0.4213	0.3036	306.38
Lomax	-0.02	-	-	-	-3.34	-0.41	0.0230	0.0154	0.0102	0.0068	363.34
Weibull	-5.16	1.89	-	-	1.19	-0.50	0.6647	0.5455	0.4209	0.3056	307.57
Gompertz	-0.85	3.74	-	-	1.12	-0.47	0.6560	0.5434	0.4263	0.3169	338.02
Burr XII	-5.82	1.96	-	-	1.17	-0.50	0.6609	0.5413	0.4167	0.3019	305.87
Chen	-4.26	1.79	-	-	1.19	-0.50	0.6665	0.5474	0.4227	0.3071	310.96
Modified Weibull	-5.16	1.89	0.00	-	1.18	-0.50	0.6645	0.5454	0.4209	0.3058	309.58
Weibull extension	-31.25	7.66	1.89	-	1.19	-0.50	0.6645	0.5450	0.4201	0.3047	309.63
Traditional Weibull	-73.10	0.10	0.98	1.11	1.37	-0.51	0.7034	0.5875	0.4610	0.3393	335.90

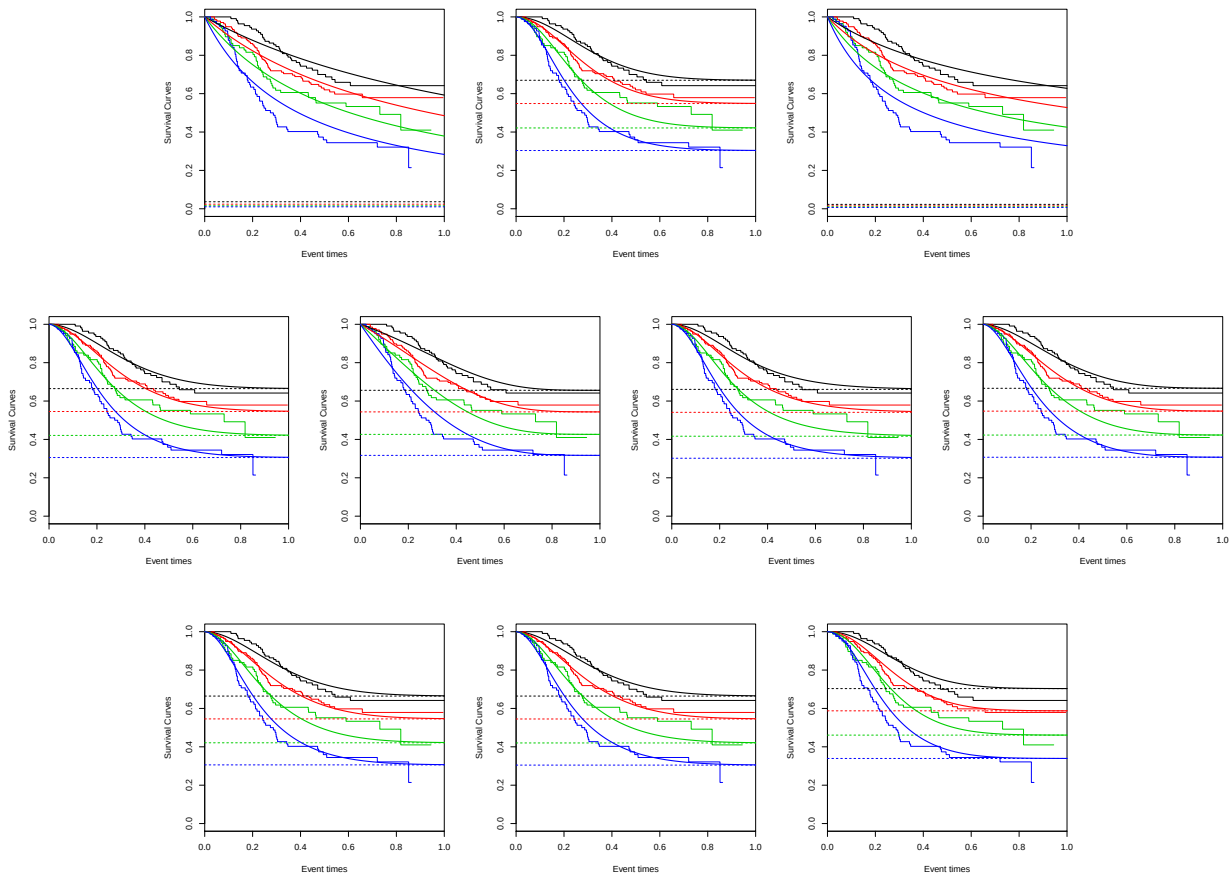

Figure 9: From the left to the right, top to bottom, the fitted regression models for the melanoma data set, in the same order as in Table 1. The colors black, red, green and blue represent the nodule categories 1, 2, 3 and 4, respectively.

The fitted results are summarized in Table 5 and Figure 9. The estimated cure rates $\widehat{p}_1$, $\widehat{p}_2$, $\widehat{p}_3$ and $\widehat{p}_4$ for groups 1, 2, 3 and 4, respectively, are calculated by (10). The Marshall Olkin-Lomax and Marshall Olkin-exponential distributions give cure rates very close to zero but they have the worst AIC. Better AIC values are given by the Marshall Olkin-Rayleigh, Marshall Olkin-Weibull, Marshall Olkin-Burr XII, Marshall Olkin-Chen, Marshall Olkin-modified Weibull and Marshall Olkin-modified Weibull extension distributions. The lowest AIC found was for the Marshall Olkin-Chen distribution with 305.87. The distributions giving the best AIC values capture the Kaplan-Meier curve relatively well, but not so well for nodule category 1 and nodule category 3 near the tails.

The estimates of β_0 and β_1 are in agreement in all models. For β_0, the value lies around 1.20 for most models (except for Marshall Olkin-Lomax and Marshall Olkin-exponential distributions) and for β_1, the value is around -0.50. That means that the cure rate decreases when the nodule category increases.

The estimated cure rate for nodule category 1 is around 0.66. For nodule category 2, it is around 0.54. For nodule category 3, it is 0.42. For nodule category 4, it is 0.30. The standard deviation of these cure rates can be estimated using the standard deviations of β_0 and β_1 by the delta method. For the Marshall Olkin-Chen distribution, the standard deviations of p_1, p_2, p_3 and p_4 are 0.0379, 0.0305, 0.0319 and 0.0395, respectively. The corresponding 95 percent (asymptotic) confidence intervals are (0.59, 0.74), (0.48,0.60), (0.36,0.48) and (0.23,0.38), respectively. These indicate a significant difference between nodules categories 1 and 3, 1 and 4 and 2 and 4. Similar results can be found for other models that performed well. These results agree with the results found in [35], [3] and [2].

4.5. Discussion

Here, we discuss some of the results in Sections 4.1, 4.2, 4.3, 4.4, a non-zero cure rate testing approach and compare the fitted distributions with their respective mixture model versions.

Table 6 compares the results in Tables 3, 4, 5 to the standard mixture model given by $S_{\text{mix}} = p + (1 - p)S(t)$, where $S(t)$ is the same baseline distribution as in the Marshall Olkin defective distributions. The distributions were compared in terms of the AIC and have the same number of parameters. The bold numbers represent the smaller AIC value. In all data sets, the defective approach performs better in seven out the ten cases. The baseline distributions performing better under a chosen approach are the same, regardless of the data analysed. The following distributions performed better under the defective approach for each of the three data sets: the Marshall Olkin-Rayleigh, Marshall Olkin-Weibull, Marshall Olkin-Burr XII, Marshall Olkin-Chen, Marshall Olkin-Modified Weibull, Marshall Olkin-Weibull extension and Marshall Olkin-Traditional distributions. The remaining performed better under the standard mixture approach. We can conclude that the defective distributions are good competitors for modelling cure rates. They provide better fits more often than the mixture model.

Table 7 gives 95 percent asymptotic confidence intervals for r based on the normal approximation. We check this table to see r is significantly lower than zero. Since the cure rate p only depends on r, the cure rate is significantly greater than zero, implying the existence of cure fraction, if r is significantly lower than zero. Almost all of the confidence intervals in Table 7 are in the negative side of the real line. The only exception is that for the Marshall

Olkin-Lomax distribution fitted to the leukemia and divorce data sets. Even this confidence interval is almost all negative. We can conclude therefore that the leukemia, colon and divorce data sets have non-zero cure rates.

Table 6: Comparison of the AIC value of the mixture and defective models.

Baseline	Leukemia data		Colon data		Divorce data	
distribution	Mixture	Defective	Mixture	Defective	Mixture	Defective
Exponential	**-64.41**	-54.86	**1509.62**	1531.10	**1503.66**	1532.16
Rayleigh	-55.71	**-75.39**	1879.82	**1668.31**	1770.51	**1633.88**
Lomax	**-63.77**	-54.03	**1518.33**	1537.00	**1518.61**	1538.82
Weibull	-68.54	**-73.50**	1481.29	**1462.36**	1439.79	**1435.90**
Gompertz	**-62.20**	-58.31	**1512.46**	1516.68	**1469.76**	1471.46
Burr XII	-69.58	**-73.93**	1470.14	**1456.53**	1439.09	**1437.27**
Chen	-67.18	**-73.00**	1503.11	**1474.74**	1442.66	**1435.01**
Modified Weibull	-63.19	**-71.50**	1464.69	**1464.39**	1441.13	**1437.73**
Weibull extension	-66.50	**-71.36**	1483.39	**1464.42**	1456.22	**1437.01**
Traditional Weibull	-64.54	**-69.09**	1485.29	**1466.37**	1443.79	**1439.02**

Table 7: Asymptotic 95 percent confidence intervals for r.

Baseline	Leukemia data		Colon data		Divorce data	
distribution	Lower CI	Upper CI	Lower CI	Upper CI	Lower CI	Upper CI
Exponential	-0.5486	-0.0109	-0.7249	-0.4493	-0.9083	-0.499
Rayleigh	-0.3302	-0.0562	-0.9467	-0.7844	-1.3366	-1.1199
Lomax	-0.5359	0.0767	-0.4271	-0.0294	-0.5376	0.0167
Weibull	-0.3702	-0.0426	-0.9776	-0.7834	-1.3542	-1.0889
Gompertz	-0.4619	-0.0495	-1.0096	-0.8107	-1.3825	-1.1419
Burr XII	-0.3767	-0.0440	-0.9445	-0.7318	-1.3252	-1.0385
Chen	-0.3620	-0.0409	-1.0032	-0.8197	-1.3752	-1.1244
Modified Weibull	-0.3702	-0.0426	-0.978	-0.7838	-1.3744	-1.0956
Weibull extension	-0.3679	-0.0397	-0.9777	-0.7833	-1.3797	-1.1197
Traditional Weibull	-0.3652	-0.0418	-0.9776	-0.7834	-1.3796	-1.1203

All of the examples provided here show that the newly introduced defective distributions can be used to provide adequate fits to several different kinds of data sets. The Marshall Olkin-Rayleigh distribution gives the best fit for the leukemia data set, but it does not perform so well for the colon data set. The Marshall Olkin-Burr XII distribution gives the best fit for the colon data set, while the Marshall Olkin-Chen distribution gives the most adequate fit for the divorce data set and the emelanoma data set, as a regression model.

This shows how competitive the newly proposed distributions can be, even when competing with the standard mixture models. More investigations are needed for these new distributions, but we hope we have provided strong evidence of the competitiveness of the proposed distributions.

5. Conclusions

The theory on defective distributions has been quite limited. In this paper, we have derived a new property of the Marshall Olkin family of distributions, allowing one to generate many new defective distributions as possible models for a wide variety of data sets. We have constructed ten new defective distributions based on the new property. The usual asymptotes of the maximum likelihood estimators for these distributions have been checked by simulation. An approach to include covariate information has been proposed and illustrated in one of the applications. In total, applications to four real data sets have been illustrated. We have presented sufficient evidence of the relevance and competitiveness of the proposed distributions, covering a range of different scenarios and showing that they can provide adequate fits. We have also shown that the proposed distributions can perform better than the standard mixture models. Future work is to explore in detail the properties of our proposed models, as well as to compare with other competitive cure rate models in the literature, in terms of interpretation and others measurements of model assessment and selection.

Acknowledgments

The authors would like to thank CNPq for financial support during the course of this project. The authors would also like to thank the two referees and the editors for their comments which greatly improved this paper.

References

[1] Balakrishnan, N., Pal, S., 2012. Em algorithm-based likelihood estimation for some cure rate models. Journal of Statistical Theory and Practice 6 (4), 698–724.

[2] Balakrishnan, N., Pal, S., 2013. Expectation maximization-based likelihood inference for flexible cure rate models with weibull lifetimes. Statistical Methods in Medical Research, 0962280213491641.

[3] Balakrishnan, N., Pal, S., 2013. Lognormal lifetimes and likelihood-based inference for flexible cure rate models based on com-poisson family. Computational Statistics and Data Analysis 67, 41–67.

[4] Balakrishnan, N., Pal, S., 2015. An em algorithm for the estimation of flexible cure rate model parameters with generalized gamma lifetime and model discrimination using likelihood- and information-based methods. Computational Statistics 30, 151–189.

[5] Balka, J., Desmond, A. F., McNicholas, P. D., 2009. Review and implementation of cure models based on first hitting times for wiener processes. Lifetime Data Analysis 15 (2), 147–176.

[6] Balka, J., Desmond, A. F., McNicholas, P. D., 2011. Bayesian and likelihood inference for cure rates based on defective inverse gaussian regression models. Journal of Applied Statistics 38 (1), 127–144.

[7] Berkson, J., Gage, R. P., 1952. Survival curve for cancer patients following treatment. Journal of the American Statistical Association 47 (259), 501–515.

[8] Boag, J. W., 1949. Maximum likelihood estimates of the proportion of patients cured by cancer therapy. Journal of the Royal Statistical Society. Series B (Methodological) 11 (1), 15–53.

[9] Cantor, A. B., Shuster, J. J., 1992. Parametric versus non-parametric methods for estimating cure rates based on censored survival data. Statistics in Medicine 11 (7), 931–937.

[10] Chen, M.-H., Ibrahim, J. G., Sinha, D., 1999. A new bayesian model for survival data with a surviving fraction. Journal of the American Statistical Association 94 (447), 909–919.

[11] Cooner, F., Banerjee, S., Carlin, B. P., Sinha, D., 2007. Flexible cure rate modeling under latent activation schemes. Journal of the American Statistical Association 102 (478).

[12] Ghitany, M., 2005. Marshall-olkin extended pareto distribution and its application. International Journal of Applied Mathematics 18 (1), 17.

[13] Ghitany, M., Al-Awadhi, F., Alkhalfan, L., 2007. Marshall-olkin extended lomax distribution and its application to censored data. Communications in Statistics—Theory and Methods 36 (10), 1855–1866.

[14] Ghitany, M., Al-Hussaini, E., Al-Jarallah, R., 2005. Marshall-olkin extended weibull distribution and its application to censored data. Journal of Applied Statistics 32 (10), 1025–1034.

[15] Gieser, P. W., Chang, M. N., Rao, P., Shuster, J. J., Pullen, J., 1998. Modelling cure rates using the gompertz model with covariate information. Statistics in Medicine 17 (8), 831–839.

[16] Gurvich, M., Dibenedetto, A., Ranade, S., 1997. A new statistical distribution for characterizing the random strength of brittle materials. Journal of Materials Science 32 (10), 2559–2564.

[17] Haybittle, J., 1959. The estimation of the proportion of patients cured after treatment for cancer of the breast. British Journal of Radiology 32 (383), 725–733.

[18] Ibrahim, J. G., Chen, M.-H., Sinha, D., 2001. Bayesian semiparametric models for survival data with a cure fraction. Biometrics 57 (2), 383–388.

[19] Ibrahim, J. G., Chen, M.-H., Sinha, D., 2005. Bayesian Survival Analysis. Wiley Online Library.

[20] Jose, K., Ancy, J., Ristić, M. M., 2009. A marshall-olkin beta distribution and its application. Journal of Probability and Statistical Science 7 (2), 173–186.

[21] Jose, K., Krishna, E., 2011. Marshall-olkin extended uniform distribution. Probability Statistics and Optimization 4, 78–88.

[22] Kaplan, E. L., Meier, P., 1958. Nonparametric estimation from incomplete observations. Journal of the American Statistical Association 53 (282), 457–481.

[23] Kersey, J., Weisdorf, D., Nesbit, M., LeBien, T., Woods, W., McGlave, P., Kim, T., Vallera, D., Goldman, A., Bostrom, B., 1987. Comparison of autologous and allogeneic bone marrow transplantation for treatment of high-risk refractory acute lymphoblastic leukemia. New England Journal of Medicine 317 (8), 461–467.

[24] Laurie, J. A., Moertel, C., Fleming, T., Wieand, H., Leigh, J., Rubin, J., McCormack, G., Gerstner, J., Krook, J., Malliard, J., 1989. Surgical adjuvant therapy of large-bowel carcinoma: An evaluation of levamisole and the combination of levamisole and fluorouracil. the north central cancer treatment group and the mayo clinic. Journal of Clinical Oncology 7 (10), 1447–1456.

[25] Lillard, L. A., Panis, C. W., 2000. aml multilevel multiprocess statistical software, release 1.0. Los Angeles: EconWare.

[26] Maller, R. A., Zhou, X., 1996. Survival analysis with long-term survivors. John Wiley and Sons, New York.

[27] Marshall, A. W., Olkin, I., 1997. A new method for adding a parameter to a family of distributions with application to the exponential and weibull families. Biometrika 84 (3), 641–652.

[28] Martinez, E. Z., Achcar, J. A., Jácome, A. A., 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. Computer Methods and Programs in Biomedicine 112 (3), 343–355.

[29] Nieto-Barajas, L. E., Yin, G., 2008. Bayesian semiparametric cure rate model with an unknown threshold. Scandinavian Journal of Statistics 35 (3), 540–556.

[30] Peng, Y., Xu, J., 2012. An extended cure model and model selection. Lifetime Data Analysis 18 (2), 215–233.

[31] R Core Team, 2013. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. URL `http://www.R-project.org/`

[32] Ristic, M. M., Jose, K., Ancy, J., 2007. A marshall-olkin gamma distribution and minification process. Stress Anxiety Research Society 11, 107–117.

[33] Rocha, R. F., Tomazella, V. L. D., Louzada, F., 2014. Inferência clássica e baysiana para o modelo de fração de cura gompertz defeituoso. Revista Brasileira de Biometria 32 (1), 104–114.

[34] Rodrigues, J., Cancho, V. G., de Castro, M., Louzada-Neto, F., 2009. On the unification of long-term survival models. Statistics and Proba-

bility Letters 79 (6), 753–759.

[35] Rodrigues, J., de Castro, M., Cancho, V. G., Balakrishnan, N., 2009. Com-poisson cure rate survival models and an application to a cutaneous melanoma data. Journal of Statistical Planning and Inference 139 (10), 3605–3611.

[36] Santos-Neto, M., Bourguignon, M., Zea, L. M., Nascimento, A. D., Cordeiro, G. M., 2014. The marshall-olkin extended weibull family of distributions. Journal of Statistical Distributions and Applications 1 (1), 9.

[37] Schrödinger, E., 1915. Zur theorie der fall-und steigversuche an teilchen mit brownscher bewegung. Physikalische Zeitschrift 16, 289–295.

[38] Sy, J. P., Taylor, J. M., 2000. Estimation in a cox proportional hazards cure model. Biometrics 56 (1), 227–236.

[39] Tsodikov, A., Ibrahim, J., Yakovlev, A., 2003. Estimating cure rates from survival data. Journal of the American Statistical Association 98 (464).

[40] Tweedie, M., 1945. Inverse statistical variates. Nature 155 (3937), 453–453.

[41] Whitmore, G., 1979. An inverse gaussian model for labour turnover. Journal of the Royal Statistical Society. Series A (General), 468–478.

[42] Yin, G., Ibrahim, J. G., 2005. Cure rate models: A unified approach. Canadian Journal of Statistics 33 (4), 559–570.