

Chapter 24

Mixture and Non-mixture Cure Rate Model Considering the Burr XII Distribution

Emílio Augusto Coelho-Barros, Jorge Alberto Achcar, and Josmar Mazucheli

Abstract This paper presents estimates for the parameters included in long-term mixture and non-mixture lifetime models, applied to analyze survival data when some individuals may never experience the event of interest. We consider the case where the lifetime data have a three-parameter Burr XII distribution, which includes the popular Weibull mixture model as a special case.

24.1 Introduction

A long-term survivor mixture model, also known as standard cure rate model, assumes that the studied population is a mixture of susceptible individuals, who experience the event of interest and non susceptible individuals that will never experience it. These individuals are not at risk with respect to the event of interest and are considered immune, non susceptible or cured [9]. Following Maller and Zhou [9], the standard cure rate model assumes that a certain fraction p in the population is cured or never fail with respect to the specific cause of death or failure, while the remaining $(1 - p)$ fraction of the individuals is not cured, leading to the survival function for the entire population written as:

$$S(t) = p + (1 - p)S_0(t), \quad (24.1)$$

where $p \in (0, 1)$ is the mixing parameter and $S_0(t)$ denotes a proper survival function for the non cured group in the population. Considering a random sample of

E.A. Coelho-Barros

COMAT, Universidade Tecnológica Federal do Paraná, Cornélio Procópio, PR, Brazil

e-mail: ebarros@utfpr.edu.br

J.A. Achcar (✉)

DMS/FMRP, Universidade de São Paulo, Ribeirão Preto, SP, Brazil

e-mail: achcar@fmrp.usp.br

J. Mazucheli

DES, Universidade Estadual de Maringá, Maringá, PR, Brazil

e-mail: jmazucheli@uem.br

lifetimes $(t_i, \delta_i, i = 1, \dots, n)$, under the assumption of right censored lifetime, the contribution of the i th individual for the likelihood function is:

$$L_i = [f(t_i)]^{\delta_i} [S(t_i)]^{1-\delta_i}, \tag{24.2}$$

where δ_i is a censoring indicator variable, that is, $\delta_i = 1$ for an observed lifetime and $\delta_i = 0$ for a censored lifetime.

From the mixture survival function, (24.1), the probability density function is obtained from $f(t_i) = -\frac{d}{dt}S(t_i)$ and given by:

$$f(t_i) = (1 - p)f_0(t_i), \tag{24.3}$$

where $f_0(t_i)$ is the probability density function for the susceptible individuals.

An alternative to a long-term survivor mixture model is the long-term survivor non-mixture model suggested by [7, 12, 13] which defines an asymptote for the cumulative hazard and hence for the cure fraction. The survival function for a non-mixture cure rate model is defined as:

$$S(t) = p^{1-S_0(t)}, \tag{24.4}$$

where, like in (24.1), $p \in (0, 1)$ is the mixing parameter and $S_0(t)$ denotes a proper survival function for the non cured group. Observe that if the probability of cure is large, then the intrinsic survival function $S(t)$ is large – $S_0(t)$ will be large which implies in $F_0(t) = 1 - S_0(t)$ small. Larger values of $F_0(t)$ at a fixed time t imply lower values of $S(t)$. This model was derived under the threshold model for tumor resistance (cancer research) where, $F_0(t)$ refers to the distribution of division time for each cell in a homogeneous clone of cells. The non-mixture model (24.4) or the promotion time cure fraction has been used by Lambert et al. [7, 8] to estimate the probability of cure fraction in cancer lifetime data.

From (24.4), the survival and hazard function for the non-mixture cure rate model can be written, respectively, as:

$$S(t_i) = \exp[\log(p)F_0(t_i)] \tag{24.5}$$

and

$$h(t_i) = -\log(p)f_0(t_i). \tag{24.6}$$

Since $f(t) = h(t)S(t)$, the contribution of the i th individual for the likelihood function is given by:

$$L_i = h(t_i)^{\delta_i} S(t_i) \tag{24.7}$$

that is:

$$L_i = [-\log(p)f_0(t_i)]^{\delta_i} \exp[\log(p)F_0(t_i)]. \tag{24.8}$$

A Bayesian formulation of the non-mixture cure rate model is given in Chen et al. [2]. A model which includes a standard mixture model for cure rate was considered in Yin and Ibrahim [14]. Rodrigues et al. [10] extended the long-term survival model proposed by Chen et al. [2].

In this paper, considering the Burr XII distribution, we compare the performance of the mixture and non-mixture cure fraction formulation when the scale and shape parameters are dependent of covariates. The Burr XII distribution provides more flexibility than the Weibull distribution which could be a special case of the Burr XII distribution if its parameters are extended to a limiting case. It is also important to point out that the Burr XII distribution is mathematically tractable with a closed form for its cumulative distribution function.

24.2 The Burr XII Distribution Cure Model

Burr [1] suggested a number of cumulative distributions, where the most popular one is the so-called Burr XII distribution, whose three-parameter probability density function is given by:

$$f_0(t \mid \mu, \alpha, \lambda) = \frac{\alpha}{\mu^\alpha} t^{\alpha-1} \left[1 + \lambda \left(\frac{t}{\mu} \right)^\alpha \right]^{-(1+\frac{1}{\lambda})}, \tag{24.9}$$

where $\mu > 0$ is the scale parameter; $\alpha > 0$ and $\lambda > 0$ are shape parameters. For $\lambda \rightarrow +0$ we have the Weibull distribution as a particular case. The hazard function of a Burr XII distribution is decreasing if $\alpha \leq 1$ and is unimodal with the mode at $t = \frac{(\alpha-1)^{1/\alpha}}{\mu^{-1}\lambda^{1/\alpha}}$ when $\alpha > 1$. The three-parameter Burr XII distribution is much more flexible than the standard two-parameter Weibull distribution.

From (24.9), the survival function is written by:

$$S_0(t \mid \mu, \alpha, \lambda) = \left[1 + \lambda \left(\frac{t}{\mu} \right)^\alpha \right]^{-\frac{1}{\lambda}}. \tag{24.10}$$

From (24.10), the Burr XII model in the presence of long-term survivors or immunes has a probability density function and a survival function given, respectively, as follows:

$$f(t \mid \theta) = (1 - p) \frac{\alpha}{\mu^\alpha} t^{\alpha-1} \left[1 + \lambda \left(\frac{t}{\mu} \right)^\alpha \right]^{-(1+\frac{1}{\lambda})}, \tag{24.11}$$

$$S(t \mid \theta) = p + (1 - p) \left[1 + \lambda \left(\frac{t}{\mu} \right)^\alpha \right]^{-\frac{1}{\lambda}}, \tag{24.12}$$

where $\theta = (\mu, \alpha, \lambda, p)$, μ is the scale parameter, α and λ are shape parameters and p is the proportion of immunes or non susceptible.

Under the non-mixture formulation and using (24.10), the probability density function and the survival function are given respectively by:

$$f(t \mid \theta) = -\log(p) \frac{\alpha}{\mu^\alpha} t^{\alpha-1} \left[1 + \lambda \left(\frac{t}{\mu} \right)^\alpha \right]^{-(1+\frac{1}{\lambda})} p^{\{1-[1+\lambda(\frac{t}{\mu})^\alpha]^{-\frac{1}{\lambda}}\}} \tag{24.13}$$

$$S(t \mid \theta) = p^{\{1-[1+\lambda(\frac{t}{\mu})^\alpha]^{-\frac{1}{\lambda}}\}}. \tag{24.14}$$

In the presence of one covariate x_i , $i = 1, \dots, n$, we can assume a link function for μ , α , λ and p , that is, $\log(\mu_i) = \beta_0 + \beta_1 x_i$, $\log(\alpha_i) = \alpha_0 + \alpha_1 x_i$, $\log(\lambda_i) = \gamma_0 + \gamma_1 x_i$ and $\log(\frac{p_i}{1-p_i}) = \eta_0 + \eta_1 x_i$, where x_i , for example, taking the value 0 if individual i is in the treatment group 1 or the value 1 if individual i is in the treatment group 2. In this way, we can have interest in test the following hypothesis: $H_0: \beta_1 = 0$ (no treatment effect in the susceptible patients), $H_0: \alpha_1 = 0$ (no treatment effect in the shape of the lifetime distribution), $H_0: \gamma_1 = 0$ (no treatment effect in the shape of the lifetime distribution) or $H_0: \eta_1 = 0$ (no treatment effect in the proportion of cured individuals).

24.3 A Bayesian Analysis

For a Bayesian analysis of the mixture and non-mixture models introduced in Sect. 24.1, we assume an prior uniform distribution defined in the interval $(0, 1)$, $U(0, 1)$, for the probability of cure p and $Gamma(0.001, 0.001)$ prior distributions for the scale parameter μ and shape parameters α and λ , where $Gamma(a, b)$ denotes a gamma distribution with mean a/b and variance a/b^2 . We further assume prior independence among p , μ , α and λ . Observe that we are using approximately non-informative priors for the parameters of the models.

Assuming the mixture and non-mixture models introduced in Sect. 24.1, let us consider a gamma prior distribution $Gamma(0.001, 0.001)$ for the regression parameters β_0 and α_0 and a normal prior distribution $N(0, 100)$ for the regression parameters β_l and α_l , $l = 1, \dots, k$, where $N(\mu, \sigma^2)$ denotes a normal distribution with mean μ and variance σ^2 . We also assume prior independence among the parameters.

Posterior summaries of interest are obtained from simulated samples for the joint posterior distribution using standard Markov Chain Monte Carlo (MCMC) methods as the Gibbs sampling algorithm [4] or the Metropolis–Hastings algorithm [3].

24.4 An Application

In this section we analyze a leukaemia data set consisting of 90 observations introduced by Kersey et al. [6] and reproduced by Maller and Zhou [9]. In this data 46 patients were treated by allogeneic transplant (Group I) and the other 44 by autologous transplant (Group II). The survival time refers to the number of days to recurrence of leukaemia for patients after one of the two treatments. The medical problems of interest include: the existence of “cured” patients (who will never suffer a recurrence of leukaemia) and the estimation of their proportion; the failure distributions of susceptible patients; and comparison between the effects of the two treatments.

In Tables 24.1 and 24.2, we have the inference results considering the Bayesian approach for mixture and non-mixture models, respectively. We also have the Monte

Table 24.1 Posterior means (standard deviation) for μ , α , λ and p in each group—mixture model

Group	$\hat{\mu}$	$\hat{\alpha}$	$\hat{\lambda}$	\hat{p}	DIC
I	170.2	1.3224	1.5235	0.2046	495.3
	(15.5727)	(0.3386)	(1.2280)	(0.0984)	
II	114.4	3.2585	1.8328	0.2073	457.3
	(22.5142)	(1.1278)	(1.3489)	(0.0622)	

Table 24.2 Posterior means (standard deviation) for μ , α , λ and p in each group — non-mixture model

Group	$\hat{\mu}$	$\hat{\alpha}$	$\hat{\lambda}$	\hat{p}	DIC
I	302.0	1.3200	1.1538	0.2497	494.0
	(60.1777)	(0.2091)	(0.5350)	(0.0673)	
II	158.4	2.7506	1.3057	0.2141	455.8
	(25.4148)	(0.5098)	(0.4480)	(0.0603)	

Fig. 24.1 Fitted models for the data

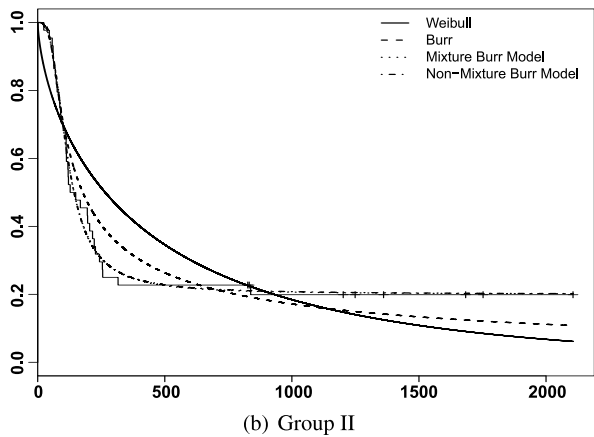
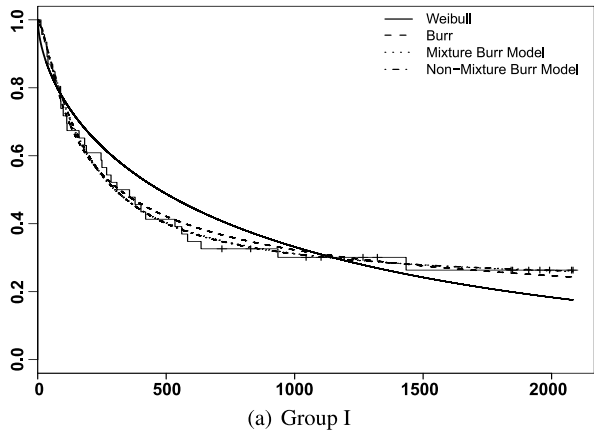


Table 24.3 Posterior Means (PM) and Standard Deviation (SD) for regression models — mixture model

Model	Parameter	PM	SD	Credible Interval
Model 1	$\widehat{\mu}$	122.6	28.7453	(77.2546; 194.9)
	$\widehat{\alpha}$	2.1934	0.5451	(1.2602; 3.4459)
	$\widehat{\lambda}$	2.2280	1.3879	(0.3819; 5.9824)
	\widehat{p}	0.2009	0.0632	(0.0646; 0.3176)
Model 2	$\widehat{\alpha}$	1.8272	0.4815	(1.0778; 2.9824)
	$\widehat{\lambda}$	1.3434	1.1575	(0.0332; 4.5494)
	\widehat{p}	0.2173	0.0593	(0.0851; 0.3261)
	$\widehat{\beta}_0$	174.9	52.0182	(81.5222; 277.9)
	$\widehat{\beta}_1$	-0.3018	0.2916	(-0.8292; 0.3199)
Model 3	$\widehat{\lambda}$	1.7100	1.0094	(0.3734; 4.1672)
	\widehat{p}	0.2005	0.0528	(0.0997; 0.3093)
	$\widehat{\beta}_0$	175.8	52.2661	(88.4162; 298.2)
	$\widehat{\beta}_1$	-0.4070	0.2806	(-0.9384; 0.1674)
	$\widehat{\alpha}_0$	1.4214	0.3461	(0.8879; 2.2550)
	$\widehat{\alpha}_1$	0.8004	0.2449	(0.3295; 1.2986)

Carlo estimates of DIC (Deviance Information Criterion) used as a discrimination criterion for different models. Smaller values of DIC indicates better models.

To obtain the Bayesian estimates we have used MCMC (Markov Chain Monte Carlo) methods available in SAS software 9.2, SAS/MCMC [11]. A single chain has been used in the simulation of samples for each parameter of both models considering a “burn-in-sample” of size 15,000 to eliminate the possible effect of the initial values. After this “burn-in” period, we simulated other 200,000 Gibbs samples taking every 100th sample, to get approximated uncorrelated values which result in a final chain of size 2,000. Usual existing convergence diagnostics available in the literature for a single chain using the SAS/MCMC procedure indicated convergence for all parameters.

In Fig. 24.1, we have the plots of the estimated survival functions considering mixture and non-mixture models in presence of cure fraction and the plot of the non-parametric Kaplan–Meier estimate for the survival function [5]. We also have in Fig. 24.1, the plot of the estimated survival function based on the Weibull and Burr XII distributions not considering the cure fraction modeling.

From the fitted survival models (see Fig. 24.1), we conclude that the survival times are very well fitted by the mixture and non mixture cure fraction models. From the results of Tables 24.1 and 24.2, the obtained DIC discrimination values from both models also give similar results.

We can also consider a binary variable related to the different groups where $x_{1i} = 1$ for Group II and 0 for the Group I. Then we consider three cases: model without covariates (Model 1), regression model for μ (Model 2) and regression model for μ and α (Model 3).

Table 24.4 Posterior Means (PM) and Standard Deviation (SD) for regression models — non-mixture model

Model	Parameter	PM	SD	Credible Interval
Model 1	$\widehat{\mu}$	191.3	47.1324	(118.4; 309.2)
	$\widehat{\alpha}$	2.0741	0.4720	(1.2757; 3.0983)
	$\widehat{\lambda}$	3.4496	3.0633	(0.2589; 12.2087)
	\widehat{p}	0.1879	0.0668	(0.0432; 0.3062)
Model 2	$\widehat{\alpha}$	1.6350	0.2545	(1.1474; 2.1417)
	$\widehat{\lambda}$	1.3797	1.8359	(0.00255; 7.8565)
	\widehat{p}	0.2170	0.0674	(0.0482; 0.3320)
	$\widehat{\beta}_0$	295.4	9.5409	(276.6; 313.4)
	$\widehat{\beta}_1$	-0.3886	0.1929	(-0.7404; 0.0111)
Model 3	$\widehat{\lambda}$	1.6695	1.1204	(0.1919; 4.6356)
	\widehat{p}	0.2044	0.0499	(0.1049; 0.3055)
	$\widehat{\beta}_0$	338.9	15.2548	(308.1; 356.9)
	$\widehat{\beta}_1$	-0.7833	0.1830	(-1.1088; -0.4185)
	$\widehat{\alpha}_0$	1.2762	0.2040	(0.9006; 1.7011)
	$\widehat{\alpha}_1$	0.7752	0.2432	(0.2887; 1.2575)

Table 24.5 Deviance Information Criterion (DIC)

Model	Mixture model	Non-mixture model
Model 1	959.5	958.2
Model 2	959.8	958.7
Model 3	949.5	948.7

In Tables 24.3 and 24.4, we have the inference results considering the Bayesian approach for regression models considering mixture and non-mixture models, respectively.

In Bayesian context using MCMC methods, we have used the DIC given automatically by the SAS software (see, Table 24.5).

From the results of Table 24.5, we conclude that Model 3 (regression model for μ and α) is better fitted by the data. Since DIC is a little bit smaller considering the non-mixture Model 3 when compared to the other models, we use this model to get our final inferences of interest. From Table 24.4 and using the non-mixture Model 3, we conclude that the parameters β_1 and α_1 have significant treatment effect in the ratio of susceptible patients.

References

1. Burr IW (1942) Cumulative frequency functions. Ann Math Stat 13:215–232
2. Chen MH, Ibrahim JG, Sinha D (1999) A new Bayesian model for survival data with a surviving fraction. J Am Stat Assoc 94(447):909–919

3. Chib S, Greenberg E (1995) Understanding the Metropolis-Hastings algorithm. *Am Stat* 49(4):327–335
4. Gelfand AE, Smith AFM (1990) Sampling-based approaches to calculating marginal densities. *J Am Stat Assoc* 85(410):398–409
5. Kaplan EL, Meier P (1958) Nonparametric estimation from incomplete observations. *J Am Stat Assoc* 53:457–481
6. Kersey JH, Weisdorf D, Nesbit ME, LeBien TW, Woods WG, McGlave PB, Kim T, Vallera DA, Goldman AI, Bostrom B (1987) Comparison of autologous and allogeneic bone marrow transplantation for treatment of high-risk refractory acute lymphoblastic leukemia. *N Engl J Med* 317(8):461–467
7. Lambert PC, Dickman PW, Weston CL, Thompson JR (2010) Estimating the cure fraction in population-based cancer studies by using finite mixture models. *J R Stat Soc, Ser C, Appl Stat* 59(1):35–55
8. Lambert PC, Thompson JR, Weston CL, Dickman PW (2007) Estimating and modeling the cure fraction in population-based cancer survival analysis. *Biostatistics* 8(3):576–594
9. Maller RA, Zhou X (1996) Survival analysis with long-term survivors. *Wiley series in probability and statistics: applied probability and statistics*. Wiley, Chichester
10. Rodrigues J, Cancho VG, de Castro M, Louzada-Neto F (2009) On the unification of long-term survival models. *Stat Probab Lett* 79(6):753–759
11. SAS (2010) The MCMC procedure, SAS/STAT[®] user's guide, version 9.22. SAS Institute Inc., Cary, NC
12. Tsodikov AD, Ibrahim JG, Yakovlev AY (2003) Estimating cure rates from survival data: an alternative to two-component mixture models. *J Am Stat Assoc* 98(464):1063–1078
13. Yakovlev AY, Tsodikov AD, Asselain B (1996) Stochastic models of tumor latency and their biostatistical applications. World Scientific, Singapore
14. Yin G, Ibrahim JG (2005) Cure rate models: a unified approach. *Can J Stat* 33(4):559–570