Brazilian Journal of Probability and Statistics
2014, Vol. 28, No. 3, 333–352
DOI: 10.1214/12-BJPS210

A hidden competing risk model for censored observations

Chrys Caroni[a] **and Polychronis Economou**[b]

[a]*National Technical University of Athens*
[b]*University of Patras*

Abstract. We propose a model for lifetime data in which all units start out susceptible to the event of interest but may move into a non-susceptible group if another event intervenes. Practical examples include subjects dropping out of a study by leaving the study area. It is supposed that the investigator is unaware of each subject's status. The model results in the appearance of a cured fraction (long-term survivors). It is shown that certain parametric models for the lifetime and the intervening event are related to a cured fraction mixture model even though the non-susceptible group is fixed from the outset in the latter but not in the new model. A likelihood ratio test and a diagnostic plot are proposed and examples of applications are provided.

1 Introduction

A standard feature of lifetime data is the presence of right censored observations of the time until the event of interest. These could be described briefly as corresponding to experimental units that have not undergone the event by the time that data collection ceases, and this definition is adequate when the units are under close observation as in a laboratory study. More generally, however, right censoring corresponds to the fact that the event of interest is *not known to have occurred*. This includes two separate cases: either the event has not occurred, or it has occurred but not been recorded. Apart from possible failure of the recording mechanism, the latter case implies that because of some other intervening event, this unit was no longer in fact under observation. Possible intervening events include participants in a study leaving the area in which it is being carried out, or patients seeking treatment from another health care provider with the result that the original investigator no longer receives their data. The occurrence of this intervening event is unknown to the original investigator, otherwise these units could be simply right censored at the time of the intervening event. Another way of looking at this situation, is to suppose that the study sample consists of two subgroups: one subgroup is susceptible to the event of interest but the other is not, because the intervening event has occurred. As we will discuss extensively later on, this description is similar to "cured fraction" models which also contain susceptible and non-susceptible groups (Maller and Zhou, 1996). A basic difference is that, in the situation we have

Key words and phrases. Lifetime data, cured fraction, drop outs, censoring mechanism.
Received March 2012; accepted September 2012.

described, units do not become "non-susceptible" unless and until the intervening event occurs, whereas in a "cured fraction" model units are "cured" from the beginning. It is common to both models that the investigator does not know to which subgroup each unit belongs.

Actual examples of studies in which the proposed mechanism may be operating include the following. Lawless et al. (2001) analysed data on repeated surgical procedures in a population of children. They acknowledged the possibility that failure to return to the clinic after the last recorded visit did not necessarily indicate that the child continued trouble-free: it could mean movement out of the study area or to another treatment facility (Lawless et al., 2001, p. 452, p. 462). They handled this possibility by carrying out a sensitivity analysis in which the data were reanalysed after a chosen proportion of the right censored times based on the duration of the study had been selected at random and replaced by times based on their last observation time. Bring and Carling (2000) analysed unemployment durations and established, by conducting further sampling, that a large proportion (45%) of people who dropped out of unemployment registration (by failing to maintain contact with the agency) actually dropped out because they had found employment (the event of interest). Notice that in both of these examples, which are not in any way unusual, data were collected only when the subjected attended the clinic or the agency. Therefore, although the event of interest can be observed, the censoring event can not. We believe therefore that the situation we are considering in this paper arises rather commonly.

If units have, unknown to the investigator, moved into a non-susceptible group, the effect on the data is that their times until the event of interest are bound to appear as right censored, even though in some cases the event may have occurred much earlier. If the simple model in which all units are susceptible continues to be fitted to the data, it could be seriously affected by this inflation of event times. Caroni (2011) has considered some tests for excessive right censored times in certain models.

2 The model

We formulate the model as follows. Let the random variable T represent the time to the event E of interest. Another random variable R represents the time until the occurrence of an unobservable competing risk D such as dropping out of the study, for example, by moving out of the study area. Although we will refer to "dropping out," more generally the competing risk represents the transition from the susceptible group to the non-susceptible group. We emphasize that D is always unobservable; we will not consider the possibility that it is sometimes observed and sometimes not. If D occurs before E, then it will be impossible to observe E. However, the researcher does not know which group the individual belongs to, and therefore continues to treat the individual as being under observation right up until

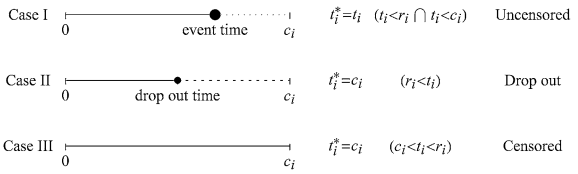

Figure 1 *Possible scenarios for the occurrence of the event of interest and dropping out.*

the end of the experiment or data collection period. Let $(0, t_0)$ be the duration of data collection. Then for an individual with (possibly latent) event times (t_i, r_i) for the pair (T, R), the data actually recorded are (t_i^*, δ_i) where t_i^* is defined as

$$t_i^* = \begin{cases} t_i, & \text{if } t_i < r_i \cap t_i < t_0, \\ t_0, & \text{otherwise} \end{cases} \tag{2.1}$$

and

$$\delta_i = I\left(t_i^* = t_i\right) \tag{2.2}$$

is the indicator taking the value $\delta_i = 1$ when the event of interest E has been observed and $\delta_i = 0$ when it has not.

The data collection period $(0, t_0)$ could be different for each individual. Introducing a censoring time variable C, t_i^* is now defined as

$$t_i^* = \begin{cases} t_i, & \text{if } t_i < r_i \cap t_i < c_i, \\ c_i, & \text{otherwise,} \end{cases} \tag{2.3}$$

where c_i is the censoring time for the ith individual. The time C is always observed if E has not occurred. Usually it will be the end of the study. The c_i need not be distinct and in fact they will all be identical if the subjects enter a study of fixed duration at the same time.

This structure describes three different possibilities (Figure 1). First, Case I is an uncensored observation. The other two cases cannot be distinguished from the data, because they both result in censoring at c_i for the event of interest. One (Case II) is a drop out: $r_i < t_i$ and the event of interest can no longer be recorded. This characterization is irrespective of whether $r_i < c_i$ or $r_i > c_i$. The other (Case III) is a censored observation with $c_i < t_i < r_i$.

3 Maximum likelihood estimation

Given a sample of n independent observations (t_i^*, δ_i), $i = 1, \ldots, n$, with t_i^* as in (2.3), the likelihood function under the proposed model is

$$L = \prod_{i=1}^{n} \left(f_T\left(t_i^*\right) S_R\left(t_i^*\right)\right)^{\delta_i} P\left(T^* = c_i\right)^{1-\delta_i} \tag{3.1}$$

assuming independence between R and T. As is usually done in most models for lifetime data, we assume that censoring is uninformative, in the sense of being "missing at random" in the terminology of the literature on missing or incomplete data. Consequently, we can avoid modelling the censoring process for inference (Lawless, 2003). If the censoring time is common for all individuals, that is, $c_i = t_0$, $\forall i$, then

$$L = \prod_{i=1}^{n} \left(f_T(t_i^*) S_R(t_i^*) \right)^{\delta_i} P(T^* = t_0)^{1-\delta_i}. \tag{3.2}$$

The last term in (3.1) is the probability that an observation appears to be right censored at c_i, and is given by

$$\begin{aligned} P(T^* = c_i) &= P(\text{drop out}) + P(\text{censored}) \\ &= 1 - P(T^* < c_i) \\ &= 1 - \int_0^{c_i} f_T(t) S_R(t)\, dt. \end{aligned} \tag{3.3}$$

This can be found analytically in certain cases. First suppose that T follows a Weibull distribution, which is a natural case to consider as this flexible distribution is one of the most widely used distributions in survival and reliability studies. If $T \sim \text{Weibull}(\alpha, \beta)$ with p.d.f. given by

$$f_T(t) = \alpha\beta(\alpha t)^{\beta-1} e^{-(\alpha t)^\beta} \tag{3.4}$$

and $R \sim \text{Weibull}(\gamma, \beta)$ (that is, with the same shape parameter as T but with a different scale), then

$$P(T^* = c_i) = 1 - \frac{\alpha^\beta}{\alpha^\beta + \gamma^\beta} \left(1 - e^{-(\alpha^\beta + \gamma^\beta) {c_i}^\beta}\right). \tag{3.5}$$

Another tractable choice is the combination of a Gamma distribution for T and an exponential for R. If $T \sim \text{Gamma}(\lambda, p)$ with p.d.f. given by

$$f_T(t) = \frac{\lambda^p t^{p-1} e^{-t\lambda}}{\Gamma(p)} \tag{3.6}$$

and $R \sim \text{Exp}(\gamma)$ with survival function

$$S_R(t) = e^{-\gamma t}, \tag{3.7}$$

then the probability of apparent censoring at c_i is given by

$$P(T^* = c_i) = 1 - \left(\frac{\lambda}{\gamma + \lambda}\right)^p \left(1 - \frac{\Gamma(p, (\gamma + \lambda)c_i)}{\Gamma(p)}\right). \tag{3.8}$$

These expressions can be substituted in the likelihood function (3.1) or, with t_0 replacing c_i, in (3.2), as appropriate.

4 Comparing the cured fraction mixture model and the proposed model

In cured fraction mixture models it is assumed that there is a proportion π_c of subjects who will never experience the event of interest (failure, death etc.) because they were not, at any time, actually at risk; see, for example, Maller and Zhou (1996). The survival function in this case is given by

$$S_c(t) = \pi_c + (1 - \pi_c) S_{T_c}(t), \tag{4.1}$$

where $S_{T_c}(t)$ is the survival function of the random variable T_c which represents the time to the event of interest conditional on belonging to the susceptible group. The proportion of non-susceptibles π_c is constant over time, expressing the assumption that these subjects are "cured" from the beginning. It should be noted that the same model can also be expressed as the improper survival function $S(t) = \exp(-\theta F(t))$, where F is a distribution function (Chen et al., 1999, Lambert, 2007). However, the motivation behind this alternative form is to consider an individual's event time to be generated as the minimum of an unknown number N of independent and identically distributed competing risks (Chen et al., 1999). The cured fraction, equal to $e^{-\theta}$, corresponds to individuals for whom $N = 0$ so that they face no risks. Because each individual's N is a characteristic that is fixed from the outset, "cure" is determined from the beginning in this formulation as well as in the simpler two-group mixture version.

In contrast, in our proposed model the size of the non-susceptible group (equivalent to the "cured fraction") grows continuously. The probability of belonging to this group at time t is

$$\pi(t) = P(R < T \cap R < t) = \int_0^t \int_r^\infty f_T(u) f_R(r)\, du\, dr \tag{4.2}$$

which can be expressed as

$$\pi(t) = \int_0^t f_R(r) S_T(r)\, dr. \tag{4.3}$$

This is an increasing function of t with $\lim_{t\to 0} \pi(t) = 0$ and $\lim_{t\to\infty} \pi(t) = P(\text{drop out})$.

In the model with $T \sim \text{Weibull}(\alpha, \beta)$ and $R \sim \text{Weibull}(\gamma, \beta)$, $\pi(t)$ is given by

$$\pi(t) = \frac{\gamma^\beta}{\alpha^\beta + \gamma^\beta}\left(1 - e^{-(\alpha^\beta + \gamma^\beta)t^\beta}\right) \tag{4.4}$$

while for the model with $T \sim \text{Gamma}(\lambda, p)$ and $R \sim \text{Exp}(\gamma)$ it is given by

$$\pi(t) = 1 - \left(\frac{\lambda}{\gamma + \lambda}\right)^p + \frac{(\lambda/(\gamma + \lambda))^p \Gamma(p, (\gamma + \lambda)t) - e^{-\gamma t}\Gamma(p, \lambda t)}{\Gamma(p)}. \tag{4.5}$$

Although the proposed model and the cured fraction model differ in how they conceive the nature of the group of subjects who will never experience the event of interest, they turn out to be closely related in some cases. More specifically based on the requirement that

$$\pi_c = P(\text{drop out})$$

—in other words, the same proportion of the population never experiences the event of interest under both models—the following theorem holds.

Theorem 1. *Suppose that the distribution $G(\alpha, \beta)$ of T under the proposed model belongs to a scale family of probability distributions with p.d.f. of the form*

$$f_T(t) = f_T(t; \alpha, \beta) = \frac{e^{-\psi_1(t;\alpha,\beta)-\psi_2(t;\beta)}}{\Phi(\alpha, \beta)}\xi(t), \qquad t > 0,$$

where $\psi_1(t; \alpha, \beta)$ is a monotonically increasing function with respect to t with $\psi_1(0; \alpha, \beta) = 0$ such that there is a $w > 0$

$$\psi_1(t; x, \beta) + \psi_1(t; y, \beta) = \psi_1(t; w, \beta),$$

for every $x > 0$ and $y > 0$, $\psi_2(t; \beta)$ is a function not depending on the scale parameter α, $\xi(t)$ is a non-negative function of t and $\Phi(\alpha, \beta)$ is a normalizing factor. Further suppose that the survival function of the random variable R can be expressed as

$$S_R(t) = e^{-\psi_1(t;\gamma,\beta)}.$$

Then the proposed model is equivalent to a cured fraction mixture model with distribution function of the time to event of interest belonging to the same distribution family, with the same shape parameter but with different scale parameter α_c.

Proof. In order to prove the theorem, we must prove that

$$f_T(t)S_R(t) = (1 - \pi_c) f_{T_c}(t),$$

where $f_{T_c}(t)$ is the p.d.f. of T_c under the cured fraction model. Since $\pi_c = P(\text{drop out})$ this relationship is equivalent to

$$f_T(t)S_R(t) = \big(1 - P(\text{drop out})\big) f_{T_c}(t)$$

or

$$\frac{f_T(t)}{1 - P(\text{drop out})} S_R(t) = f_{T_c}(t).$$

But $G(\alpha, \beta)$ belongs to a scale family of probability distributions, so we have that

$$F_T(t) \equiv F_T(t; \alpha, \beta) = F_T\big(\alpha^\beta t, 1, \beta\big)$$

from which

$$\begin{aligned}\frac{f_T(t)}{1-P(\text{drop out})}S_R(t) &= \frac{1}{1-P(\text{drop out})}\alpha^\beta f_T\big(\alpha^\beta t;1,\beta\big)S_R(t)\\ &= \frac{\alpha^\beta}{1-\pi_c}f_T\big(\alpha^\beta t;1,\beta\big)S_R(t)\\ &= \alpha_*^\beta f_T\big((1-\pi_c)\alpha_*^\beta t;1,\beta\big)S_R(t)\\ &= \frac{1}{1-\pi_c}f_T\big(t;\alpha_*\sqrt[\beta]{1-\pi_c},\beta\big)S_R(t),\end{aligned}$$

where

$$\alpha_*^\beta = \frac{\alpha^\beta}{1-P(\text{drop out})} = \frac{\alpha^\beta}{1-\pi_c}.$$

But

$$S_R(t) = e^{-\psi_1(t;\gamma,\beta)}$$

and so

$$\begin{aligned}f_T(t)S_R(t) &= f_T\big(t;\alpha_*\sqrt[\beta]{1-\pi_c},\beta\big)e^{-\psi_1(t;\gamma,\beta)}\\ &= \frac{e^{-\psi_1(t;\alpha_*\sqrt[\beta]{1-\pi_c},\beta)-\psi_2(t;\beta)}}{\Phi(\alpha_*\sqrt[\beta]{1-\pi_c},\beta)}\xi(t)e^{-\psi_1(t;\gamma,\beta)}\\ &= \frac{e^{-[\psi_1(t;\alpha_*\sqrt[\beta]{1-\pi_c},\beta)+\psi_1(t;\gamma,\beta)]-\psi_2(t;\beta)}}{\Phi(\alpha_*\sqrt[\beta]{1-\pi_c},\beta)}\xi(t)\\ &= \frac{e^{-\psi_1(t;\alpha_c,\beta)-\psi_2(t;\beta)}}{\Phi(\alpha_*\sqrt[\beta]{1-\pi_c},\beta)}\xi(t),\end{aligned}$$

where α_c is such that

$$\psi_1(t;\alpha_c,\beta) = \psi_1\big(t;\alpha_*\sqrt[\beta]{1-\pi_c},\beta\big) + \psi_1(t;\gamma,\beta).$$

But

$$\Phi\big(\alpha_*\sqrt[\beta]{1-\pi_c},\beta\big)E_{f_T}\big(e^{-\psi_1(t;\gamma,\beta)}\big) = \Phi(\alpha_c,\beta)$$

and

$$\begin{aligned}E_{f_T}\big(e^{-\psi_1(t;\gamma,\beta)}\big) &= E_{f_T}\big(S_R(t)\big)\\ &= \int_0^\infty f_T(t)S_R(t)\,dt\\ &= \int_0^\infty\int_t^\infty f_T(t)f_R(u)\,du\,dt\end{aligned}$$

$$
\begin{aligned}
&= P(R > T) \\
&= 1 - P(R < T) \\
&= 1 - P(\text{drop out}) \\
&= 1 - \pi_c .
\end{aligned}
$$

Therefore, finally

$$
\begin{aligned}
f_T(t) S_R(t) &= (1 - \pi_c) \frac{e^{-\psi_1(t;\alpha_c,\beta) - \psi_2(t;\beta)}}{\Phi(\alpha_c, \beta)} \xi(t) \\
&= (1 - \pi_c) f_{T_c}(t)
\end{aligned}
$$

as required. □

The two special cases that we considered earlier both satisfy the above theorem. When T and R both follow Weibull distributions with the same shape parameter then T_c also follows a Weibull distribution, still with the same shape but with scale given by

$$\alpha_c^\beta = \alpha^\beta + \gamma^\beta . \tag{4.6}$$

When T follows a Gamma distribution and R an Exponential, then T_c follows a Gamma distribution with the same shape as T but scale given by

$$\lambda_c = \lambda + \gamma . \tag{4.7}$$

Some other common lifetime distributions, on the other hand, do not meet the conditions of the theorem. For example, when we assume a Lognormal distribution for T, there is no distribution of R such that our model is equivalent to a cured fraction mixture model in which the time to the event of interest T_c is also Lognormal, with the same shape parameter but different scale parameter.

5 Diagnostic tools

We now consider ways of identifying the presence of unobserved drop-out. We will consider one formal test and one diagnostic plot. Both are based on $P(\text{drop out})$, the probability of never experiencing the event of interest, which is given by

$$P(\text{drop out}) = \int_0^\infty f_R(t) S_T(t)\, dt . \tag{5.1}$$

For the model with $T \sim \text{Weibull}(\alpha, \beta)$ and $R \sim \text{Weibull}(\gamma, \beta)$ model this is given by

$$P(\text{drop out}) = \frac{\gamma^\beta}{\alpha^\beta + \gamma^\beta} \tag{5.2}$$

letting $t \to \infty$ in (3.5). For the model with $T \sim \text{Gamma}(\lambda, p)$ and $R \sim \text{Exp}(\gamma)$ model, it is

$$P(\text{drop out}) = 1 - \left(\frac{\lambda}{\gamma + \lambda}\right)^p \tag{5.3}$$

from (3.8).

5.1 Likelihood ratio test

A likelihood ratio (LR) test can be based on the fact that under the proposed model $P(\text{drop out}) > 0$. Consequently, we may carry out a test of the null hypothesis

$$H_0 : P(\text{drop out}) = 0 \tag{5.4}$$

against

$$H_1 : P(\text{drop out}) > 0. \tag{5.5}$$

In both of the particular cases considered above (Weibull–Weibull and Gamma–Exponential), these hypotheses are equivalent to

$$H_0 : \gamma = 0, \tag{5.6}$$

$$H_1 : \gamma > 0, \tag{5.7}$$

so that under the null hypothesis the survival function of R satisfies $S_R(t) = 1, \forall t > 0$.

Because the null value of γ lies on the boundary of the parameter space, the asymptotic distribution of the LR statistic $-2(\hat{\ell}_0 - \hat{\ell}_1)$ under the null hypothesis is expected to be a 50 : 50 mixture of a X_1^2 distribution and a degenerate distribution at zero (Self and Liang, 1987). We will denote this mixture distribution by $X_{1:0}^2$.

5.2 Diagnostic plot

Suppose that data are collected over the period of time $(0, t_0)$. If the proposed model is correct, then $P(\text{drop out})$ does not depend on t_0. Therefore a different t_0 should not lead to a different estimate of $P(\text{drop out})$, apart from the effect of sampling error. A diagnostic plot for the appropriateness of the model can be based on this observation. For any chosen t_0, we modify the data to what they would have been with that stopping time. This means that known event times longer than t_0 are replaced by right censored values at t_0 and censored observations at times longer than t_0 are treated as censored at t_0. If units did not enter the experiment all at the same time and some therefore have stopping times less than t_0, these remain right censored at their actual stopping times. Given that the assumed model is correct, successive estimates of $P(\text{drop out})$ obtained by modifying the data in this way for different choices of t_0 should all be similar to each other.

The proposed diagnostic is a plot of estimates of $P(\text{drop out})$ against the t_0 on which they are based, and can be constructed by the following steps:

1. Select a set of k times $t_{01}, \ldots, t_{0k}$, where $t_k = t_0 = \max t_i^*$
2. Set $j = k$
3. Fit the model to obtain an initial estimate p_k of P(drop out) using the original data
4. Set $j = j - 1$
5. Define new data $(t_{ij}^*, \delta_{ij}), i = 1, \ldots, n$ as follows
   - if $t_i^* < t_{0j}$ then $(t_{ij}^*, \delta_{ij}) = (t_i^*, \delta_i)$
   - if $t_i^* \geq t_{0j}$ then $(t_{ij}^*, \delta_{ij}) = (t_{0j}, 0)$
6. Fit the model to the new data set and get the estimate of $p_j = P$(drop out)
7. Plot $(p_j - p_k)$ vs t_{0j}, $j = 1, \ldots, k$

If the plotted points are close to the horizontal axis, then the assumed model can not be rejected. The idea of looking at the stability of estimates obtained from modified sets of data for different time points was also used successfully for a diagnostic plot by Economou and Caroni (2005).

6 Simulations

A detailed simulation study was carried out in order to evaluate the tests presented in the previous section. More specifically, the performance of the proposed LR test was studied under various combinations of sample size ($n = 20, 50, 100, 200, 350$), shape parameter β of the null Weibull distribution ($\alpha = 0.03$, $\beta = 2/3, 1, 3/2$) and censoring rate (30%, 50%). For each combination, 2000 samples of n event times t_i were generated from the Weibull$(0.03, \beta)$ distribution. A censoring time c_i for each observation in the sample was generated from the uniform distribution on $(0, x_0)$ and if $c_i < t_i$ then the event time was right censored at c_i. The value of x_0 was chosen so that the resulting samples had approximately 30% or 50% censored observations. Both a simple Weibull distribution and the proposed Weibull–Weibull model were fitted to each sample, the LR test was computed each time and its value was compared to the upper $2\alpha\%$ point of the X_1^2 distribution in order to carry out an $\alpha\%$ level test under $X_{1:0}^2$. Table 1 presents the observed probability of exceeding this percentage point for the case $\alpha = 0.05$. The approximate binomial standard error of a proportion of about 5% is 0.5 for 2000 simulations. Apart from the combination of small sample size ($n = 20$) and high proportion of censoring (50%), the size of the test is acceptably close to the nominal level given by the asymptotic distribution. Results are systematically better for lower censoring rate and for larger sample size, both of which are of course expected, but do not appear to depend on the Weibull shape β except for small sample sizes.

Figure 2 shows the estimated exceedance probabilities for the LR test plotted against the nominal significance level for $n = 350$. This shows that the nominal

Table 1 *Simulated exceedance probabilities for the likelihood ratio test in the Weibull–Weibull model, at nominal 5% significance level*

		Sample size n				
	β	20	50	100	200	350
30% censoring	2/3	0.0630	0.0640	0.0600	0.0555	0.0445
	1.0	0.0615	0.0620	0.0495	0.0555	0.0495
	3/2	0.0435	0.0480	0.0430	0.0455	0.0545
50% censoring	2/3	0.1005	0.0665	0.0555	0.0545	0.0470
	1.0	0.0840	0.0650	0.0605	0.0595	0.0555
	3/2	0.0730	0.0655	0.0565	0.0600	0.0480

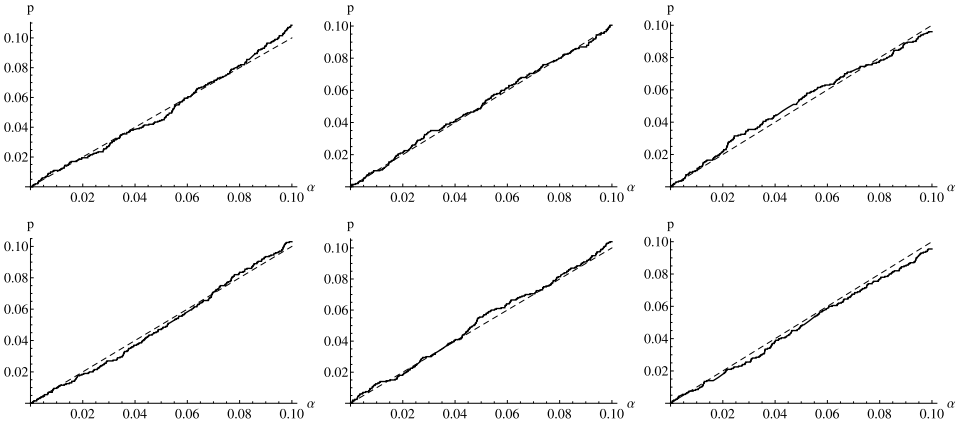

Figure 2 *Simulated exceedance probabilities* (*solid line*) *for the likelihood ratio test plotted against nominal significance level, with sample size* 350 *observations,* $\beta = 2/3$ (*plots on the left*), $\beta = 1$ (*center*), $\beta = 1$ (*right*) *and approximately* 30% (*upper plots*) *and* 50% (*lower plots*) *censoring.*

significance levels are accurate for all values of α of practical interest, and not only for $\alpha = 0.05$.

Further simulations were carried out to examine the power of the tests and the accuracy of estimation of the parameters of the non-null model, with Weibull distributions for both T and R. Values of n, parameters α and β of the distribution of T, and censoring rates were as above. The scale parameter γ of the distribution of R was 0.005 or 0.01, producing proportions of drop out observations in the samples varying from a minimum of 0.0637 to a maximum of 0.3246. The censoring times were generated by assuming a uniform distribution on $(0, x_0)$ where x_0 was determined so that the proportion of apparently censored observations (drop out and true censored observations) in the samples would be approximately equal to 30% or 50%. The only exception was made in the case of $\beta = 2/3$, $\gamma = 0.01$ with

censoring rate 30% in which no true censored observations were generated since the proportion of the drop out observations was slightly larger than 30%.

Tables 2–4 show the means and standard deviations of the estimates of the parameters α, β and γ, obtained from 2000 simulated samples in each case. The tendency is towards over-estimation of all parameters, but the size of this bias is large only when the sample size is very small ($n = 20$ for 30% censoring, $n = 50$ for 50%). For larger samples, the bias is small and in fact estimation is in effect unbiased for $n = 200$, or in many cases for $n = 100$. Table 5 shows the simulated power of the LR test in the same simulation runs. There is strong dependence on the underlying censoring rate.

In order to examine the behavior of the proposed diagnostic plot, samples were generated under the Weibull–Weibull model for various combinations of sample sizes and shape and scale parameters of the distributions. The parameter values were selected to give samples with a wide range of proportions of observations that were actually drop outs and not censored. Figure 3 presents a typical plot obtained under the $T \sim \text{Weibull}(0.03, 2/3)$–$R \sim \text{Weibull}(0.01, 2/3)$ model based on a sample with 100 observations and a $U(0, 250)$ censoring distribution. The behaviour of the proposed diagnostic plot was generally more clearly what was predicted whenever the proportion of the drop out observation was larger in the sample. Larger samples also resulted in better results, as expected.

7 Illustrative applications

In this section, the proposed model is applied to two data sets. The results are discussed and compared to the cured fraction mixture model.

7.1 Breast cancer data

For a simple illustration of the application of the model to an easily available set of data, we take data on breast cancer from Klein and Moeschberger (2003, page 7). Of these 45 observations, 24 (46.7%) have long survival. Fitting the Weibull–Weibull model by maximum likelihood gives

$$\hat{\ell}_1 = -140.625, \qquad \hat{\alpha} = 0.01246, \qquad \hat{\beta} = 2.428, \qquad \hat{\gamma} = 0.01178. \tag{7.1}$$

The null (no drop out) Weibull model fit is

$$\hat{\ell}_0 = -148.945, \qquad \hat{\alpha} = 0.00575, \qquad \hat{\beta} = 1.1441. \tag{7.2}$$

Hence the LR test statistic takes the value 16.64, which has p-value 0.000045 compared to X^2_1 and therefore half of this compared to $X^2_{1:0}$. The estimate of $P(\text{drop out})$ is 0.4661. The diagnostic plot is shown in Figure 4 and the fit of the proposed model (along with the Kaplan–Meier survival estimate is in Figure 5).

This is not intended as a serious analysis of these data but it is offered as an illustration of the method for the following reason. The description of the data says

Table 2 *Performance of the MLEs applied to samples simulated from the* Weibull($\alpha = 0.03, \beta = 2/3$)–Weibull($\gamma, \beta = 2/3$) *model*: 2000 *samples of size n with approximately* 30% *and* 50% *censored observations* (*drop out and true censored observations*). *The mean and the standard deviation* (*in parentheses*) *of the estimate of each parameter are shown*

		$\hat{\alpha}$		$\hat{\beta}$		$\hat{\gamma}$	
		Censoring rate		Censoring rate		Censoring rate	
	n	30%	50%	30%	50%	30%	50%
$\gamma = 0.005$	20	0.0352 (0.0180)	0.0453 (0.0291)	0.7409 (0.1959)	0.8328 (0.3245)	0.0081 (0.0091)	0.0238 (0.0410)
$P(\text{drop out}) = 0.2324$	50	0.0317 (0.0098)	0.0355 (0.0161)	0.6941 (0.1069)	0.7282 (0.1670)	0.0060 (0.0042)	0.0129 (0.0162)
	100	0.0311 (0.0068)	0.0330 (0.0114)	0.6773 (0.0733)	0.6947 (0.1062)	0.0055 (0.0028)	0.0096 (0.0107)
	200	0.0304 (0.0046)	0.0313 (0.0085)	0.6707 (0.0495)	0.6799 (0.0773)	0.0052 (0.0018)	0.0075 (0.0075)
	300	0.0304 (0.0039)	0.0308 (0.0069)	0.6706 (0.0413)	0.6752 (0.0646)	0.0052 (0.0015)	0.0068 (0.0059)
$\gamma = 0.01$	20	0.0361 (0.0182)	0.0424 (0.0302)	0.7566 (0.1823)	0.8085 (0.3030)	0.0139 (0.0108)	0.0258 (0.0376)
$P(\text{drop out}) = 0.3246$	50	0.0321 (0.0098)	0.0342 (0.0159)	0.6926 (0.0993)	0.7187 (0.1561)	0.0111 (0.0046)	0.0161 (0.0163)
	100	0.0312 (0.0068)	0.0323 (0.0115)	0.6777 (0.0654)	0.6899 (0.1074)	0.0105 (0.0032)	0.0132 (0.0110)
	200	0.0306 (0.0045)	0.0306 (0.0080)	0.6732 (0.0443)	0.6737 (0.0747)	0.0103 (0.0021)	0.0113 (0.0077)
	300	0.0304 (0.0037)	0.0305 (0.0065)	0.6724 (0.0380)	0.6711 (0.0608)	0.0102 (0.0017)	0.0107 (0.0062)

Table 3 *Performance of the MLEs applied to samples simulated from the* Weibull($\alpha = 0.03, \beta = 1$)–Weibull($\gamma, \beta = 1$): 2000 *samples of size n with approximately* 30% *and* 50% *censored observations* (*drop out and true censored observations*). *The mean and the standard deviation* (*in parentheses*) *of the estimate of each parameter are shown*

		$\hat{\alpha}$		$\hat{\beta}$		$\hat{\gamma}$	
		Censoring rate		Censoring rate		Censoring rate	
	n	30%	50%	30%	50%	30%	50%
$\gamma = 0.005$	20	0.0323 (0.0103)	0.0377 (0.0157)	1.1195 (0.2923)	1.2433 (0.4592)	0.0066 (0.0069)	0.0155 (0.0200)
$P(\text{drop out}) = 0.1429$	50	0.0305 (0.0062)	0.0332 (0.0093)	1.0404 (0.1645)	1.0961 (0.2332)	0.0056 (0.0041)	0.0099 (0.0113)
	100	0.0305 (0.0045)	0.0316 (0.0065)	1.0204 (0.1122)	1.0469 (0.1520)	0.0053 (0.0028)	0.0075 (0.0077)
	200	0.0301 (0.0031)	0.0308 (0.0049)	1.0088 (0.0778)	1.0246 (0.1094)	0.0052 (0.0020)	0.0064 (0.0058)
	300	0.0301 (0.0020)	0.0305 (0.0041)	1.0067 (0.0625)	1.0139 (0.0897)	0.0051 (0.0016)	0.0060 (0.0049)
$\gamma = 0.01$	20	0.0332 (0.0107)	0.0360 (0.0165)	1.1258 (0.2900)	1.2004 (0.4318)	0.0123 (0.0076)	0.0175 (0.0205)
$P(\text{drop out}) = 0.2500$	50	0.0309 (0.0058)	0.0321 (0.0096)	1.0389 (0.1478)	1.0646 (0.2184)	0.0105 (0.0039)	0.0128 (0.0115)
	100	0.0306 (0.0042)	0.0310 (0.0067)	1.0233 (0.0981)	1.0363 (0.1526)	0.0103 (0.0027)	0.0115 (0.0083)
	200	0.0302 (0.0030)	0.0302 (0.0050)	1.0131 (0.0684)	1.0126 (0.1098)	0.0102 (0.0019)	0.0102 (0.0059)
	300	0.0302 (0.0024)	0.0304 (0.0042)	1.0065 (0.0572)	1.0103 (0.0905)	0.0102 (0.0015)	0.0106 (0.0051)

Table 4 *Performance of the MLEs applied to samples simulated from the* Weibull($\alpha = 0.03, \beta = 3/2$)–Weibull($\gamma, \beta = 3/2$): 2000 *samples of size n with approximately* 30% *and* 50% *censored observations* (*drop out and true censored observations*). *The mean and the standard deviation* (*in parentheses*) *of the estimate of each parameter are shown*

		$\hat{\alpha}$		$\hat{\beta}$		$\hat{\gamma}$	
		Censoring rate		Censoring rate		Censoring rate	
	n	30%	50%	30%	50%	30%	50%
$\gamma = 0.005$	20	0.0313 (0.0060)	0.0342 (0.0082)	1.6773 (0.4332)	1.8698 (0.6580)	0.0052 (0.0064)	0.0107 (0.0132)
$P(\text{drop out}) = 0.0637$	50	0.0304 (0.0039)	0.0317 (0.0052)	1.5470 (0.2250)	1.6308 (0.3065)	0.0049 (0.0041)	0.0078 (0.0084)
	100	0.0301 (0.0026)	0.0310 (0.0037)	1.5302 (0.1592)	1.5711 (0.2155)	0.0049 (0.0029)	0.0064 (0.0066)
	200	0.0301 (0.0019)	0.0305 (0.0027)	1.5086 (0.1105)	1.5348 (0.1463)	0.0049 (0.0020)	0.0055 (0.0052)
	300	0.0300 (0.0015)	0.0304 (0.0023)	1.5074 (0.0883)	1.5274 (0.1201)	0.0049 (0.0016)	0.0053 (0.0046)
$\gamma = 0.01$	20	0.0310 (0.0063)	0.0330 (0.0091)	1.6772 (0.4205)	1.8073 (0.6276)	0.0106 (0.0068)	0.0132 (0.0138)
$P(\text{drop out}) = 0.1600$	50	0.0306 (0.0039)	0.0312 (0.0053)	1.5543 (0.2227)	1.6204 (0.3179)	0.0103 (0.0038)	0.0112 (0.0088)
	100	0.0302 (0.0027)	0.0304 (0.0040)	1.5260 (0.1492)	1.5398 (0.2083)	0.0101 (0.0026)	0.0101 (0.0067)
	200	0.0302 (0.0019)	0.0301 (0.0029)	1.5169 (0.1058)	1.5209 (0.1518)	0.0101 (0.0018)	0.0098 (0.0051)
	300	0.0301 (0.0016)	0.0300 (0.0023)	1.5107 (0.0856)	1.5157 (0.1223)	0.0100 (0.0015)	0.0099 (0.0041)

Table 5 *Simulated power of the LR test based on* 2000 *simulated samples from the* Weibull($\alpha = 0.03, \beta$)–Weibull(γ, β) *model of size n with approximately* 30% *and* 50% *censored observations* (*drop out and true censored observations*)

			Censoring rate				Censoring rate	
		n	30%	50%		n	30%	50%
$\beta = 2/3$	$\gamma = 0.005$	20	0.5375	0.1540	$\gamma = 0.01$	20	0.9985	0.2135
	$P(\text{drop out}) = 0.2324$	50	0.8410	0.1560	$P(\text{drop out}) = 0.3246$	50	1.0000	0.2895
		100	0.9860	0.2130		100	1.0000	0.4440
		200	1.0000	0.3005		200	1.0000	0.6425
		300	1.0000	0.3930		300	1.0000	0.7975
$\beta = 1$	$\gamma = 0.005$	20	0.2840	0.1310	$\gamma = 0.01$	20	0.9095	0.1705
	$P(\text{drop out}) = 0.1429$	50	0.5250	0.1370	$P(\text{drop out}) = 0.2500$	50	0.9995	0.2390
		100	0.7910	0.1375		100	1.0000	0.3590
		200	0.9600	0.1985		200	1.0000	0.5400
		300	0.9940	0.2570		300	1.0000	0.7055
$\beta = 3/2$	$\gamma = 0.005$	20	0.2000	0.1075	$\gamma = 0.01$	20	0.6510	0.1590
	$P(\text{drop out}) = 0.0637$	50	0.3635	0.0940	$P(\text{drop out}) = 0.1600$	50	0.9430	0.2140
		100	0.5595	0.1075		100	0.9985	0.2815
		200	0.8110	0.1345		200	1.0000	0.4445
		300	0.9235	0.1610		300	1.0000	0.5790

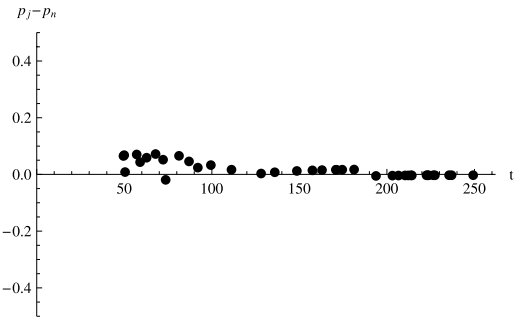

Figure 3 *Typical diagnostic plot under the proposed model.*

that patients required a minimum 10-year follow up for inclusion. Consequently, all the censored times are necessarily large and hence the data give a very strong impression of the existence of a cured fraction. Because of this, even though the sample is small, our methods worked very effectively.

Under the cured fraction mixture model the estimated distribution of the survival time T among the population group that will experience the event of interest in some point is a Weibull distribution with scale parameter equal to 0.01613 and shape parameter equal to 2.428 (see Theorem 1 and Eq. (5.2)). Meanwhile, under the drop out model it is a Weibull distribution with the same shape parameter but

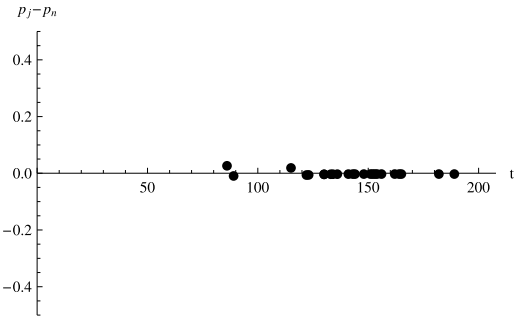

Figure 4 *Diagnostic plot for the breast cancer data.*

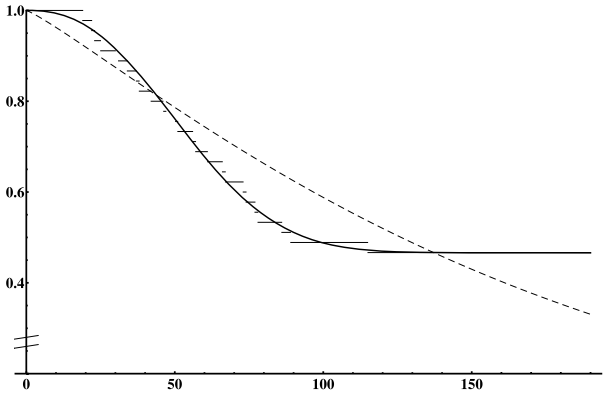

Figure 5 *Breast cancer data*: *Kaplan–Meier survival estimate* (*horizontal line segments*), *fitted survival under the drop out model* (*solid line*) *and fitted survival under a simple Weibull model* (*broken line*).

different scale equal to 0.01246. This difference is important and has a significant effect on all the characteristics of the survival time of the vulnerable group. For example, under the cured fraction model the expected value of the survival time of a member of the group that will experience the event of interest at some time is 55.0 months while under the drop out model it is 71.2 months, a substantial difference of more than 16 months greater than the expected lifetime given by the cured fraction model.

7.2 Recidivism data

The second illustration takes part of a set of data on the recidivism of offenders released from detention in North Carolina in 1978 and 1980. Various analyses of these data can be found in Chung et al. (1991) and earlier papers that are cited there. From the Kaplan–Meier estimate (Figure 6), it is evident that the survival function (meaning, in this case, the probability of staying out of prison) does not

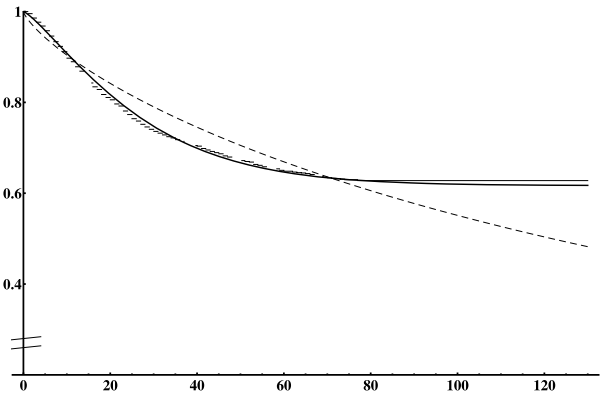

Figure 6 *Recidivism data*: *Kaplan–Meier survival estimate* (*horizontal line segments*), *fitted survival under the drop out model* (*solid line*) *and fitted survival under a simple Weibull model* (*broken line*).

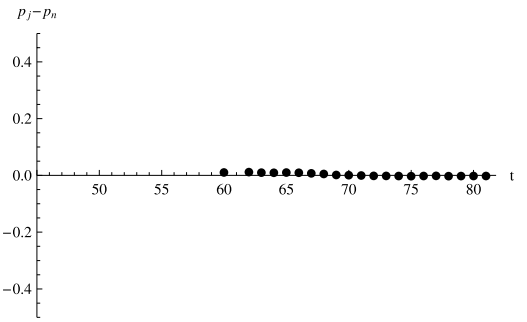

Figure 7 *Diagnostic plot for the recidivism data.*

tend to zero. For this reason, the models that have been tried include cured fraction mixture models, called split-population models by those authors. We fitted our Weibull–Weibull model to the subset of 1540 cases which had been randomly selected by the original authors to form the estimation sample from the 1978 data. The event time observations of 970 (63%) of these people were right censored, meaning that they had not been recorded as re-entering prison by the end of the study six years after its start. The data were obtained from the Inter-university Consortium for Political and Social Research (www.icpsr.umich.edu).

The value of the LR test statistic was 121.0 which is of course statistically highly significant, as expected given the appearance of the survival curve and the large sample. Parameter estimates for the Weibull–Weibull model were $\hat{\alpha} = 0.0164$, $\hat{\beta} = 1.2475$, $\hat{\gamma} = 0.0240$, with $P(\text{drop out})$ estimated as 0.6168. The fitted curve is shown in Figure 6. The parameter estimates for the equivalent cured fraction model are $\hat{\alpha}_c = 0.0155$, $\hat{\beta}_c = 1.2475$, $\hat{\pi} = 0.6168$. The diagnostic plot (Figure 7) supports the model.

Although, as we have seen, the two models are mathematically equivalent, they differ in the interpretation of the nature of the group that will never experience the event of interest (in this case returning to a prison in North Carolina). The cured fraction mixture model implies that, for some offenders, it is already fixed at the time of release that they will never return to prison. On the other hand, our model implies that throughout time there is always a chance of joining the group that will never re-offend, so long as you have avoided re-offending so far. The two models also differ in the distribution of the survival time of the group that will eventually, at some point, experience the event of interest. Because of the different scale estimates in the fitted Weibull distributions, the expected time to the event conditional on its occurrence is 60.3 months under the cured fraction model but only 56.9 months under the drop out model.

8 Comments

As stated in the Introduction, we believe that the model here corresponds to an important feature of the data that are collected in many studies. However, we are aware that this is only the first step in the analysis of the problem. In particular, most practical analyses of lifetime data include covariates. Consequently, the next stage should be to introduce covariates into our model too.

References

Bring, J. and Carling, K. (2000). Attrition and misclassification of drop-outs in the analysis of unemployment duration. *Journal of Official Statistics* **16**, 321–330.

Caroni, C. (2011). Testing for undetected withdrawal in a follow-up study of times to events. *Statistical Methodology* **8**, 213–220. MR2769281

Chen, M. H., Ibrahim, J. G. and Sinha, D. (1999). A new Bayesian model for survival data with a surviving fraction. *Journal of the American Statistical Association* **94**, 909–919. MR1723307

Chung, C. F., Schmidt, P. and Witte, A. D. (1991). Survival analysis: A survey. *Journal of Quantitative Criminology* **7**, 59–98.

Economou, P. and Caroni, C. (2005). Graphical tests for the assumption of Gamma and Inverse Gaussian frailty distributions. *Lifetime Data Analysis* **11**, 565–582. MR2213505

Klein, J. P. and Moeschberger, M. L. (2003). *Survival Analysis: Techniques for Censored and Truncated Data*. New York: Springer.

Lambert, P. C. (2007). Modeling of the cure fraction in survival studies. *Stata Journal* **7**, 351–375.

Lawless, J. F. (2003). *Statistical Models and Methods for Lifetime Data*, 2nd ed. Hoboken, NJ: Wiley. MR1940115

Lawless, J. F., Wigg, M. B., Tuli, S., Drake, J. and Lamberti-Pasculli, M. (2001). Analysis of repeated failures or durations, with application to shunt failures for patients with paediatric hydrocephalus. *Applied Statistics* **50**, 449–465. MR1871799

Maller, R. A. and Zhou, X. (1996). *Survival Analysis with Long-Term Survivors*. Chichester: Wiley. MR1453117

Self, S. G. and Liang, K. Y. (1987). Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions. *Journal of the American Statistical Association* **82**, 605–610. MR0898365

Department of Mathematics
School of Applied Mathematical and Physical Sciences
National Technical University of Athens
Athens
Greece
E-mail: ccar@math.ntua.gr

Department of Engineering Science
University of Patras
Rion-Patras
Greece
E-mail: peconom@upatras.gr