Ann. Inst. Statist. Math.
Vol. 54, No. 3, 517–530 (2002)

A GENERAL CLASS OF CHANGE POINT AND CHANGE CURVE MODELING FOR LIFE TIME DATA

KAUSHIK PATRA[1] AND DIPAK K. DEY[2]

[1]*Center for Biostatistics in AIDS Research, Department of Biostatistics, Harvard School of Public Health, 5th Floor, 651 Huntington Ave, Boston, MA 02115, U.S.A.*
[2]*Department of Statistics, University of Connecticut, U120, 215 Glen Brook Road, Storrs, CT 06269, U.S.A.*

(Received August 28, 2000; revised March 19, 2001)

Abstract. Change point hazard rate models arise in many life time data analysis, for example, in studying times until the undesirable side effects occur in clinical trials. In this paper we propose a general class of change point hazard model for survival data. This class includes and extends different types of change point models for survival data, e.g. cure rate model and lag model. Most classical approach develops estimates of model parameters, with particular interest in change point parameter and often the whole hazard function, but exclusively in terms of asymptotic properties. We propose a Bayesian approach, avoiding asymptotics and provide inference conditional upon the observed data. The proposed Bayesian models are fitted using Markov chain Monte Carlo method. We illustrate our proposed methodology with an application to modeling life times of the printed circuit board.

Key words and phrases: Change point, Gibbs sampling, hazard function, posterior inference, survival function.

1. Introduction

In life time data analysis, it is often reasonable to assume that early failures appear at one rate and later on after some threshold time they occur at a different rate. For example, in clinical trials, undesirable side effects may cause different failure rate after a threshold time. In cancer prevention trials, it is often reasonable to assume that any effect of a treatment is not immediate but affects the risk of failure only after a lag. There is a considerably good literature regarding the modeling of the hazard function under the change point scenario. See Nguyen *et al.* (1984), Basu *et al.* (1988), Ebrahimi (1991) and Loader (1991). In these papers an estimate of the change point and the initial hazard rate are proposed and also the study of their asymptotic properties are pursued. Bayesian approaches are considered by Raftery and Akman (1986), Achcar and Bolfarine (1989), Carlin *et al.* (1992), Gray (1994) and Stangl (1995). Most of them considered exponential likelihood with known or unknown change points. Ebrahimi *et al.* (1998) consider arbitrary likelihood and adopt a Bayesian approach. In this paper, we propose a general class of models which incorporates the change point model for hazard function and in general for curves which are functions of survival time. Such models incorporate change point in hazard rates, cure rates (Chen *et al.* (1999)) and time lag models (Zucker and Lakatos (1997)). We use Bayesian methodologies to infer about the

model parameters where the model for the hazard function is mainly governed by the empirical plot, for example, Kaplan- Meier plot. As a result the proposed model is very flexible regarding the choice of the likelihood and the empirical data are used to have an idea about the shape of the hazard function and the presence of the change point if there is any. Of course, the occurrence of a change point should be physically justified for all practical purpose. We consider a proper uniform prior for a change point over a bounded interval, see Raftery and Akman (1986) and Ghosh *et al.* (1993) in this regard.

Let T be a random variable denoting the life time with density function $f_T(t)$ and survival function $S_T(t)$. A suitable change point hazard function $h_T(t) = f_T(t)/S_T(t)$ is defined by

$$h_T(t) = h_1(t) I(0 \leq t \leq \eta) + h_2(t) I(t > \eta) \tag{1.1}$$

where $I(A) = 1$ if $x \in A$ and 0, otherwise, and $\eta \in R^+$ is the (possibly unknown) threshold parameter. From (1.1), the cumulative hazard function is given by $H_T(t) = \int_0^t h_T(u)\,du = H_1(\min(t,\eta)) + [H_2(t) - H_2(\eta)]^+$ where $H_i(t) = \int_0^t h_i(u)\,du$, $i = 1, 2$ and $[a]^+ = aI(a > 0)$. In most change point hazard models as described in (1.1), it is assumed that $h_1(t) \geq h_2(t)$ for all $t > 0$, for example, in clinical trials the event time may represent the time until desirable side effects occur, in which case we expect lower hazard rate after the treatment has been in place for some time.

Next we consider the lag model as proposed in Zucker and Lakatos (1997). Suppose $h_0(t)$ and $h_1(t)$ denote the hazard functions for the control and treatment group respectively, then a general class of lag models can be given by

$$h_1(t) = [\phi l(t) + \{1 - l(t)\}] h_0(t) \tag{1.2}$$

where $l(t)$ is a monotone function with $0 \leq l(t) \leq 1$. The value of $l(t)$ represents the proportion of the treatment effect at time t. So if $l(t) = I(t > \eta)$, the treatment has no detectable effect during the period $[0, \eta]$; afterwards, the treatment is fully effective. Clearly η is a threshold lag parameter and (1.2) is a change point curve model.

To define a change point cure rate model, first we describe a cure rate model. Following Chen *et al.* (1999), the survival function for the cure rate model can be expressed as $S_p(t) = \exp(-\theta F(t))$ where $\theta > 0$ is an unknown parameter. Note that, $\lim_{t\to\infty} S_p(t) = \exp(-\theta) > 0$, and hence $S_p(t)$ is not a proper survival function. The quantity $\exp(-\theta)$ is called "Cure Fraction" which tends to 0 or 1 as θ approaches ∞ or 0. The corresponding (pseudo) hazard function is $h_p(t) = \theta f(t)$. Here, $h_p(t)$ is not a hazard function corresponding to a proper probability distribution, since $S_p(t)$ is not a proper survival function. However, $h_p(t) \to 0$ as $t \to \infty$ and $\theta = \int_0^\infty h_p(t)\,dt < \infty$. Often in cure rate model, a second type of risk comes into play after a saturation level is reached. Such a situation can be handled by the following model $h(t) = \theta f(t) I(t < \eta) + [\theta f(t) + h_1(t)] I(t > \eta)$ where η is the threshold parameter, which is needed when the second type of risk, e.g. higher stress level is prevalent. Above examples demand the development of a general class of change point or change curve model.

The rest of the article is organized as follows. In Section 2, we propose the general change point model and study some of its properties. Modeling survival fraction is discussed in Section 3. Section 4 describes the model incorporating a baseline hazard. We provide simple examples relating to constant change point hazard rates. In Section 5, we discuss modeling lifetimes of printed circuit boards as described by Meeker and LuValle

(1995). A Bayesian approach is described in this section and the posterior inference of the model parameters are obtained. Section 6 provides a comparison of the proposed model with a standard model and we conclude the paper in Section 7.

2. General formulation

Let T be a random variable denoting the life distribution. We propose a general class of hazard function as follows:

$$h_T(t) = h_1(t)I(t < \eta) + h_2(t)I(t > \eta) \tag{2.1}$$

where $h_i(t)$, $i = 1, 2$ are non-negative locally integrable functions defined on positive real line. In presence of covariates $\boldsymbol{x}$ the formulae can be modified by replacing $h_T(t)$ by $h_T(t \mid \boldsymbol{x})$ and similarly for other components of the hazard function. We present the model by suppressing $\boldsymbol{x}$ throughout in the notation. The cumulative hazard function corresponding to (2.1) is given by

$$H_T(t) = H_1(t)I(t < \eta) + [H_1(\eta) + (H_2(t) - H_2(\eta))]I(t > \eta) \tag{2.2}$$

where $H_i(t) = \int_0^t h_i(u)du$ for $i = 1, 2$. Note that the functions $h_i(\cdot)$ can be any non-negative functions defined on the positive real line. The limits $\lim_{t\to\infty} H_i(t) = \int_0^\infty h_i(u)du$, $i = 1, 2$ are important as they can be finite or infinite depending on the value of the integral. For example, if $h_i(u)$ is a probability density function then the integral is finite. The corresponding density function of the life distribution is given by $g(t) = pf_1(t) + (1-p)f_2(t)$ where $p = 1 - \exp(-H_1(\eta))$. Note that the mixture proportion $p \equiv p(\eta)$ depends on the unknown parameter η and $g(t)$ can be thought of as a mixture of two density functions where $f_1(t) = \frac{h_1(t)\exp(-H_1(t))}{1-\exp(-H_1(\eta))} I(t < \eta)$ which corresponds to a right truncated distribution and $f_2(t) = h_2(t)\exp[-\{H_2(t) - H_2(\eta)\}]I(t > \eta)$ which corresponds to a left truncated distribution. Next we provide the explicit form of the survival function and hazard function corresponding to different components of the mixture life distribution. Note that $H_i(t_1) \leq H_i(t_2)$ for all $t_1 < t_2$, $i = 1, 2$. The survival function of the first component in (2.2) is given by $S_1^*(t) = \frac{\exp(-H_1(t))-\exp(-H_1(\eta))}{1-\exp(-H_1(\eta))} I(t < \eta)$ with $H_1(t) \leq H_1(\eta)$ and the corresponding hazard function is given by $h_1^*(t) = \frac{h_1(t)\exp(-H_1(t))}{\exp(-H_1(t))-\exp(-H_1(\eta))} I(t < \eta)$. Similarly, the survival function of the second component is given by $S_2^*(t) = \exp[-(H_2(t)-H_2(\eta))]I(t > \eta)$ with $H_2(t) \geq H_2(\eta)$ and the corresponding hazard function is given by $h_2^*(t) = h_2(t)I(t > \eta)$.

3. Survival fraction

Survival fraction is a recently developed concept in life time data analysis, especially in cure rate models or in accelerated life time experiments, where a certain fraction of the population never fails. A population of a life time distribution is said to have a survival fraction if the corresponding survival function does not tend to zero but to a strictly positive fraction as time approaches infinity. Naturally, this type of survival function is not proper and hence it will be, instead, referred to as an improper survival function. In this section we discuss how we can model such an improper life time distribution in this context. The survival function corresponding to (2.2) is given by $S_T(t) = \exp(-H_T(t))$ and hence $\lim_{t\to\infty} S_T(t) = \exp(-[\lim_{t\to\infty} H_T(t)])$. Let, $L_i = \lim_{t\to\infty} H_i(t) = \int_0^\infty h_i(u)du$

for $i = 1, 2$, then $\lim_{t\to\infty} H_T(t) = H_1(\eta) + (L_2 - H_2(\eta))$. So an improper or a proper survival function will be obtained depending on the finiteness of L_2. The value of the improper survival function at infinity will be called the survival fraction. So modeling survival fraction is solely based on the choice of $h_2(\cdot)$. Note that for a simple hazard function without a change point we simply plug in $\eta = 0$ and construct the survival fraction accordingly.

4. Incorporation of baseline hazard

Suppose we incorporate a baseline hazard in above models and replace $h_i(t)$ by $h_0(t)\, h_i(t)$, for $i = 1, 2$. Consequently the overall hazard function will be given as

$$h_T(t) = h_0(t)\, h_1(t) I(t < \eta) + h_0(t)\, h_2(t) I(t > \eta). \tag{4.1}$$

Let us define $M_{0i}(t) = \int_0^t h_0(u)\, h_i(u)\, du$, for $i = 1, 2$. Then the cumulative hazard function can be expressed as

$$H(t) = M_{01}(t) I(t < \eta) + [M_{01}(\eta) + (M_{02}(t) - M_{02}(\eta))] I(t > \eta). \tag{4.2}$$

Thus we can express the cumulative hazard function just by replacing $H_i(t)$ with $M_{0i}(t)$ in (2.2). Again the limit $\lim_{t\to\infty} M_{0i}(t) = \int_0^\infty h_0(u)\, h_i(u)\, du$ can be finite or infinite depending on the choice of both $h_0(u)$ and $h_i(u)$s. As it is evident from (4.2), we only need to concentrate on the limit corresponding to $M_{02}(t)$ and if the limit is finite then the survival fraction will be obtained. The density function is again a mixture given by $g_T(t) = p f_1(t) + (1-p) f_2(t)$ where $p = 1 - \exp(-M_{01}(\eta))$, $f_1(t) = \frac{h_0(t) h_1(t) \exp(-M_{01}(t))}{1-\exp(-M_{01}(\eta))} I(t < \eta)$ which is the pdf of a right truncated distribution and $f_2(t) = h_0(t)\, h_2(t) \exp[-\{M_{02}(t) - M_{02}(\eta)\}] I(t > \eta)$, a pdf of the left truncated distribution.

Example 4.1. Let us assume that each component of the mixture distribution follows an exponential distribution i.e., the associated hazard rates are constant. Suppose, $h_i(t) = \lambda_i$ for $i = 0, 1, 2$, then from (4.1), $h_T(t) = \lambda_0\lambda_1 I(t < \eta) + \lambda_0\lambda_2 I(t > \eta)$, and $M_{0i}(t) = \lambda_0\lambda_i t$ for $i = 1, 2$. Therefore, from (4.2), $H(t) = \lambda_0[\lambda_1 t I(t < \eta) + \{\lambda_1\eta + \lambda_2(t-\eta)\} I(t > \eta)]$ and $f_1(t) = \frac{\lambda_0\lambda_1 \exp(-\lambda_0\lambda_1 t)}{1-\exp(-\lambda_0\lambda_1\eta)} I(t < \eta)$ which is the pdf of a right truncated exponential distribution with parameters $\lambda_0\lambda_1$ and $f_2(t) = \lambda_0\lambda_2 \exp[-\lambda_0\lambda_2(t-\eta)]$ $I(t > \eta)$ which is the pdf of a left truncated exponential distribution.

Example 4.2. In this example we illustrate how to model the survival fraction. Let us consider $h_0(t) = \lambda_0$, $h_1(t) = \lambda_1$ and $h_2(t) = \lambda_2 \exp(-\lambda_2 t)$, note that $h_2(t)$ is the pdf of an exponential distribution with parameter λ_2. Then from (4.1), $h_T(t) = \lambda_0\lambda_1 I(t < \eta) + \lambda_0\lambda_2 \exp(-\lambda_2 t) I(t > \eta)$. Now $M_{01}(t) = \lambda_0\lambda_1 t$ and $M_{02}(t) = \lambda_0 (1 - \exp(-\lambda_2 t))$. Hence the survival fraction will be given by $\lim_{t\to\infty} S_T(t) = \exp(-\lambda_0[\lambda_1\eta + \exp(-\lambda_2\eta)])$.

Example 4.3. (Lag model) Two popular lag models, as reported by Zucker and Lakatos (1997) are linear lag and threshold lag model given by $l(t) = \frac{t}{\eta} I(t < \eta) + I(t > \eta)$ and $l(t) = I(t > \eta)$ respectively. The general idea behind suggesting such models is to assure the monotonicity (possibly increasing) of the lag function. The

rationale is that often the treatment starts to be effective at a lag rather than immediately after its administration. These models can be incorporated in the present set-up very easily, rather in a wide fashion. We take $h_0(t)$ as the hazard function for the control group and propose the hazard function for the treatment group with the help of our hazard modeling. For example, taking $h_1(t) = 1 - (1-\varphi)\frac{t}{\eta}$ and $h_2(t) = \varphi$, with $0 < \varphi < 1$, we get the "Linear Lag" model and taking $h_1(t) = 1$ and $h_2(t) = \varphi$, with $0 < \varphi < 1$ we have the "Threshold Lag" model. Note that in above two models, $h_1(t) I(t < \eta) > h_2(t) I(t > \eta)$ which is intuitively simpler because we can always propose a wide class of non-negative functions $h_1(t)$ and $h_2(t)$ satisfying above conditions beyond linear functions.

Example 4.4. (Marshall-Olkin model) Marshall and Olkin (1997) proposed a new class of survival model by introducing a parameter in a survival function as $S_0(t) = \alpha \bar{S}(t)$. It can be shown that the corresponding hazard function is given by $h_1(t) = \frac{1}{1-(1-\alpha)S_0(t)}$, $0 < \alpha < 1$ where $S_0(t) = \exp(-H_0(t))$. Now we describe how our model resembles with that by Marshall and Olkin (1997). Note that in this case, our model reduces to $g(t) = h_0(t) h_1(t) \exp(-M_{01}(t)) I(t > 0)$, where

$$\begin{aligned} M_{01}(t) &= \int_0^t h_0(u) h_1(u) \, dt \\ &= \int_0^t \frac{h_0(u)}{1-(1-\alpha) S_0(u)} du \\ &= \int_0^t \frac{h_0(u)}{1-(1-\alpha)\exp(-H_0(u))} du \\ &= \ln\left[\frac{1-(1-\alpha) S_0(t)}{\alpha S_0(t)}\right]. \end{aligned}$$

Hence $g(t) = \frac{\alpha S_0(t) h_0(t)}{[1-(1-\alpha)S_0(t)]^2} = \frac{\alpha f_0(t)}{[1-(1-\alpha)S_0(t)]^2}$ which is exactly the same density derived by Marshall and Olkin (1997).

5. A real data example

Meeker and LuValle (1995) discuss a data set on lifetimes of printed circuit boards tested using relative humidity (RH) as the accelerating stress. In particular, 72 circuit boards were tested at each of the four stress levels, 49.5% RH, 62.8% RH, 75.4% RH and 82.4% RH. Meeker and LuValle (1995) noted that due to problems with the test equipment there were several circuit boards that did not yield useful information. Therefore, the resulting data set consisted of 70 boards at the stress levels of 49.5% RH, 75.4% RH and 82.4% RH and 68 boards at the 62.8% RH level. The boards were monitored periodically, thus the data were interval censored. For each unit, only the interval of consecutive inspection times containing the failure time was recorded. This data had been previously analyzed by Meeker and LuValle (1995), Sinha *et al.* (2000) and Sinha *et al.* (1999) assuming that the failures had occurred at the midpoints of the observed intervals containing failures. In addition to interval censoring, there were several circuit boards that did not fail. Therefore, the data were also subject to right-censoring. In particular, there were 48 censored observations at 4,078 hours at the 49.5% RH and 11 at 3,067 hours at the 62.8% RH. All of the circuit boards did fail at 75.4% RH and

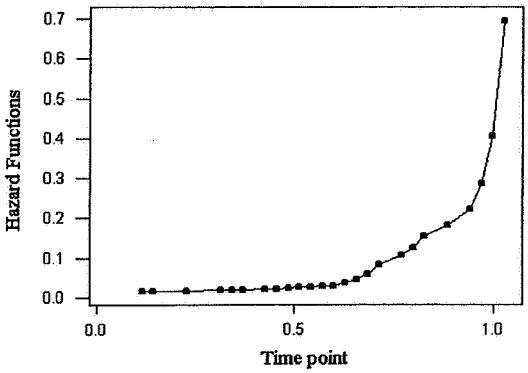

Fig. 1. Kaplan-Meier estimate of the hazard function.

82.4% RH. See the original data as described in Meeker and LuValle (1995). This data set was extensively investigated by Sinha *et al.* (1999) at all four humidity levels. They assumed that the failure to a unit is caused by several latent competing risks of manufacturing defects. In practice, this is an appropriate assumption for many manufacturing processes. For circuit board experiment, due to faulty manufacturing process, there are salt residues left on the surface of a circuit board. Under humidity, these salt residues ionize over time to create conductive filaments between two poles (cathode and anode) to cause short circuits. Now, humidity level affects this process in two ways—it may cause any salt particle to ionize sooner and it may even increase the number of latent risk factors (number of bridges which may potentially ionize to become conductive filaments) of a board. Even minor faults (small traces of salt bridges), which may not cause failure (ionize) under normal stress level, are at the risk of becoming fatal (at the risk of ionizing) under accelerated stress (humidity).

A circuit board may not have any salt bridge which are big enough to be at the risk of ionizing at the existing humidity level. For illustrative purpose we are only concentrating on the fourth humidity level where we do not have any survival fraction of the circuit board population. However, we can build up the model incorporating the survival fraction or so called LFP (Limited Failure Population) for the lower humidity levels by the technique discussed in previous sections. Figure 1 gives the Kaplan-Meier plot of hazard function for the fourth humidity level.The actual time points (hourly) were scaled by 140. The hazard plot is based on the minimum of several unknown life-times of the salt particle because a circuit board would fail if at least one salt particle gets ionized, also the number of such salt particles is random. We use the Lemma 1 given in the Appendix to construct the hazard function of the observed failure time for each of the circuit board. Figure 1 shows that there is a change in the hazard rate at around time approximately equal to 0.6.

Let x_{ij} represent the time when the j-th salt particle of the i-th board ionize, where $j = 1, \ldots, M_i$ and $i = 1, \ldots, n$. Clearly, the circuit board fails at $T_i = \min(x_{i_1}, \ldots, x_{i_{M_i}})$. Thus, T_i is the observable time to failure of the i-th board. In addition, the following assumptions will be made about the distributions for x_{ij} and M_i.

ASSUMPTIONS.

1. x_{ij} are iid with CDF $F^*(\cdot)$ for $j = 1, \ldots, M_i$ and $i = 1, \ldots, n$.
2. M_i are iid Poisson with parameter θ for $i = 1, \ldots, n$.

Figure 1 gives the estimated hazard for T and will be used to get a functional form of $h_T(t)$. The first component of the hazard (before the change at η occurred) seems to be linear in time and the second component is a power function in time. Observing these, we consider $h_0(t) = \theta$, $h_1(t) = \alpha + \beta t$ and $h_2(t) = \delta\gamma t^{\gamma-1}$ where θ, α, β, δ, γ and η are positive valued parameter. Then the hazard function given by (4.1) would be $h_T(t) = h_0(t) h_1(t) I(t < \eta) + h_0(t) h_2(t) I(t > \eta)$ and it follows that $M_{01}(t) = \theta\left(\alpha t + \beta t^2/2\right)$ and $M_{02}(t) = \theta\delta t^{\gamma}$. Consequently, the underlying probability distribution will be given by $g(t) = pf_1(t) + (1-p) f_2(t)$ with $p = 1 - \exp(-M_{01}(\eta))$, $f_1(t) = \frac{\theta(\alpha+\beta t)\exp\left[-\theta\left(\alpha t+\beta t^2/2\right)\right]}{1-\exp\left[-\theta\left(\alpha\eta+\beta\eta^2/2\right)\right]} I(t < \eta)$ and $f_2(t) = \theta\delta\gamma t^{\gamma-1}\exp\left[-\theta\delta\left(t^{\gamma} - \eta^{\gamma}\right)\right] I(t > \eta)$. We do not have any censoring for this data set, so the likelihood becomes

$$L(parameters \mid t) = \prod_{i=1}^{n} g(t_i).$$

To perform a Bayesian analysis we use priors as shown below:

$$\theta \sim Gamma(\theta_1, \theta_2), \quad \alpha \sim Gamma(\alpha_1, \alpha_2), \quad \beta \sim Gamma(\beta_1, \beta_2),$$
$$\delta \sim Gamma(\delta_1, \delta_2), \quad \gamma \sim Gamma(\gamma_1, \gamma_2) \quad \text{and} \quad \eta \sim Uniform\left(t_{(1)}, t_{(n)}\right),$$

where $Gamma(a, b)$ denotes the probability density of a gamma distribution given by $b^a \exp(-bx) x^{a-1} I(x > 0) / \Gamma(a)$. The Gamma prior for θ was chosen by keeping in mind the conjugate nature of the Gamma distribution for the Poisson parameter. For Weibull parameter δ and γ, the choice of Gamma priors are traditional. Even though α and β can be any real valued parameter, we choose Gamma prior to make them positive valued encouraged by the empirical observation. The hyperparameters are selected carefully to make the priors diffuse, for example, among several other choices, the results reported here correspond to shape and scale parameters equal to 0.001, even though the results are not sensitive to the choice of hyperparameters as along as the priors are sufficiently diffuse. Based on the empirical observation, we restrict the change point η in between the $t_{(1)}$ and $t_{(n)}$, the minimum and the maximum of the failure time. In order to simulate samples from the posterior distributions, we follow the approach proposed in Diebolt and Robert (1994). The set-up can be briefly described as follows. Suppose we have an unobserved random variable z_i associated with each observed lifetime t_i where z_i takes value 1, with probability $p = 1 - \exp\left[-\theta\left(\alpha\eta + \beta\eta^2/2\right)\right]$ if t_i is assumed to be distributed according to 1st component of the mixture and 0 if t_i follows 2nd component for $i = 1, \ldots, n$. Then the conditional likelihood function for given $\boldsymbol{z} = (z_1, \ldots, z_n)$ will be given by

$$L(\boldsymbol{t} \mid \boldsymbol{z}) = \left(\prod_{z_i=1} f_1(t_i)\right)\left(\prod_{z_i=0} f_2(t_i)\right).$$

The conditional distribution of z given t and the parameters is a Bernoulli distribution with probability $pf_1(t) / [pf_1(t) + (1-p) f_2(t)]$, but it is interesting to see that given η, the distribution of z is degenerate in the sense that the observations could be easily labeled as whether they are following the 1st or the 2nd component of the mixture. The full conditionals (listed in the Appendix) are non standard and Metropolis-Hastings algorithm is used to simlulate for α, β, θ and γ with suitably chosen Gamma proposal distribution (see Chib and Greenberg (1995), Gelfand and Smith (1990)). The full

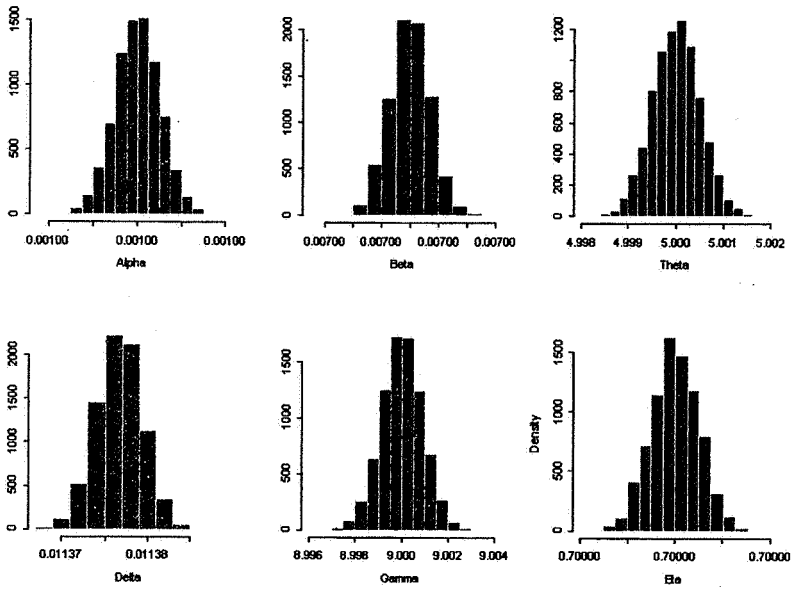

Fig. 2. Histogram of the model parameters.

Table 1. Estimates of the model parameters.

Parameter	Posterior Mean	Posterior Median	Std. Dev.	95% Credible Interval
α	0.000963	0.000999	$9.00e-5$	$(0.00098, 0.00102)$
β	0.009003	0.009002	$8.00-5$	$(0.00883, 0.0092)$
θ	5.0	5.0	0.00494	$(4.90716, 5.0989)$
δ	0.01181	0.01198	0.00026	$(0.01149, 0.0125)$
γ	6.0	5.99	0.06039	$(5.88382, 6.1207)$
η	0.6199	0.6199	0.00099	$(0.61796, 0.6217)$

conditional for δ is a Gamma distribution itself. The full conditional for the change point η is derived in the same fashion observing that if t_1 follows $f_1(t)$ and t_2 follows $f_2(t)$, then t_1 is stochastically smaller than t_2, hence $t_{(1)} < \eta < t_{(n)}$. From the continuity of the hazard function, it follows that $h_1(\eta) = h_2(\eta)$ i.e., $\alpha + \beta t_{(n_1)} < \alpha + \beta\eta = \delta\gamma\eta^{\gamma-1} < \delta\gamma t_{(n_1+1)}^{\gamma-1}$ which would be a good computational check while simulating η (see the Appendix for definition of n_1). A size of 10,000 simulation is obtained to get the posterior distribution of the parameters. Figure 2 shows the histograms of the posterior distributions of the model parameters and the posterior estimates of the parameters are given in Table 1.

The posterior distributions are roughly symmetric with small standard deviation. Since the time points are scaled to make them smaller in magnitude, the estimates of α and β are quite small. The shape parameter γ is greater than 1 which indicates that the hazard function after the change point is not constant over time. The mean number of unknown salt particle, θ agrees with the results obtained by Sinha *et al.* (1999). The estimate of the change-point falls well inside the data range, in fact, it is very close to

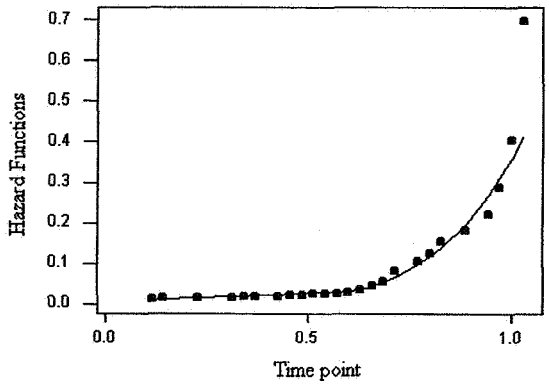

Fig. 3. Observed and predictive hazard functions.

what we have expected from the empirical hazard plot. Above all, none of the 95% credible interval includes 0 which indicates that the parameters included in the model are necessary. Figure 3 gives the empirical hazard plot given in Fig. 1 with posterior predictive hazard estimate superimposed on it. The posterior predictive hazard function is calculated as $\widehat{h(t)} = \frac{1}{N^*}\sum_{i=1}^{N^*} h(t \mid \mathbf{\Phi}_i)$ where $\mathbf{\Phi}_i = i$-th vector of model parameters from their posterior distributions and N^* is the size of the simulated sample.

6. Model Comparison

In this section we compare our change point model with a standard life time distribution known as Gumbel distribution which is typically used to model the life time distribution. Let us suppose that the life time distribution of x_{ij} is $Gumbel(\mu, \sigma)$ where μ is the location parameter and σ is the scale parameter and the hazard function of x_{ij} is given as $h(x) = \frac{1}{\sigma}\exp\left(\frac{x-\mu}{\sigma}\right)$ and consequently the hazard function of $T_i = \min(x_{i1}, \ldots, x_{iM_i})$, where $M_i \sim Poisson(\theta)$, will be $h_T(t) = \theta\exp((t-\mu)/\sigma)/\sigma = \phi\exp(t/\sigma)/\sigma$, with $\phi = \theta\exp(-\mu/\sigma) > 0$. The corresponding life time distribution is $g_T(t) = \phi\exp(t/\sigma)\exp[-\phi\exp(t/\sigma)]/\sigma$, $-\infty < t < \infty$. Suppose we transform $Y = \exp(T)$, then the probability distribution of Y would be

$$w_Y(y) = \phi y^{1/\sigma - 1}\exp[-\phi y^{1/\sigma}] I(y > 0)/\sigma,$$

which is a Weibull distribution with shape $1/\sigma$ and scale ϕ. Note that in this case θ and μ are non-identifiable, but the relationship, $\log\phi = \log\theta - \frac{\mu}{\sigma}$ indicates that $\log\phi$ and $1/\sigma$ would be linearly related. We again apply Bayesian approach to obtain the posterior of ϕ and σ. The prior distributions are taken as $\phi \sim Gamma(\phi_1, \phi_2)$ and $\sigma \sim Gamma(\sigma_1, \sigma_2)$, with $\phi_i, \sigma_i, i = 1, 2$ so chosen to make the priors diffuse. The likelihood of $y_i = \exp(t_i)$ can be written as

$$L(\boldsymbol{y}) = \left(\frac{\phi}{\sigma}\right)^n \left(\prod_{i=1}^n y_i\right)^{(1/\sigma)-1} \exp\left(-\phi\sum_{i=1}^n y_i^{1/\sigma}\right).$$

The full conditional distributions are listed in the Appendix. Figure 4 gives the histograms of the posterior distribution of the Gumbel parameters and Table 2 gives the estimates of the parameters.

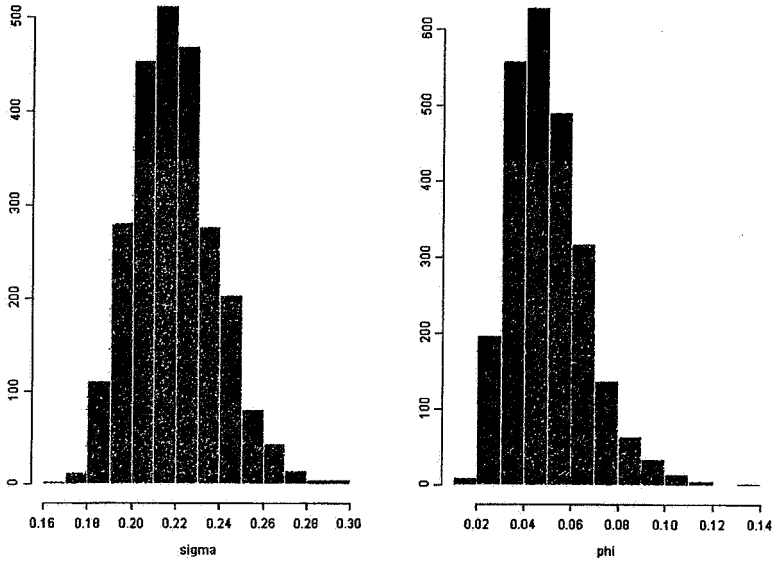

Fig. 4. Histogram of the Gumbel parameters.

Table 2. Posterior estimates of the Gumbel parameters.

Parameter	Posterior Mean	Posterior Median	Std. Dev.	95% Credible Interval
σ	0.2186	0.2173	0.01937	(0.18510, 0.26180)
ϕ	0.0497	0.0473	0.01669	(0.02497, 0.09030)

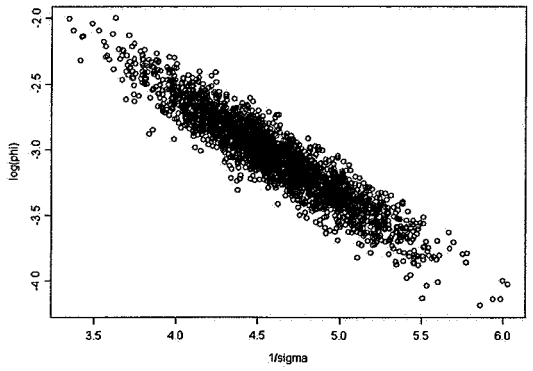

Fig. 5. Plot of $\log \phi$ vs $1/\sigma$.

Figure 5 shows that the posterior samples of $\log \phi$ and $1/\sigma$ follow a linear relationship as expected. But nonetheless the posterior predictive hazard function is not at all supportive to the original hazard function; in fact, the predictive hazard function is quite deviant from the empirical hazard function when compared to the predictive hazard function obtained from the change point model as seen in Fig. 6. The solid circles

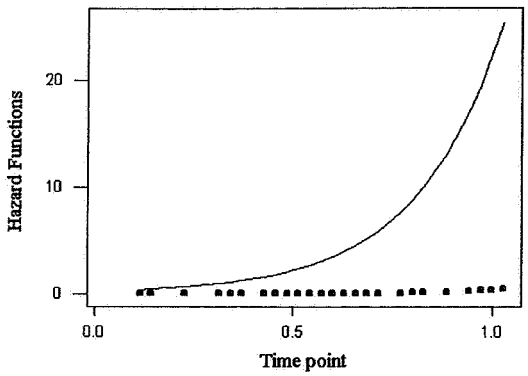

Fig. 6. Change point and Gumbel predictive hazard function superimposed on each other.

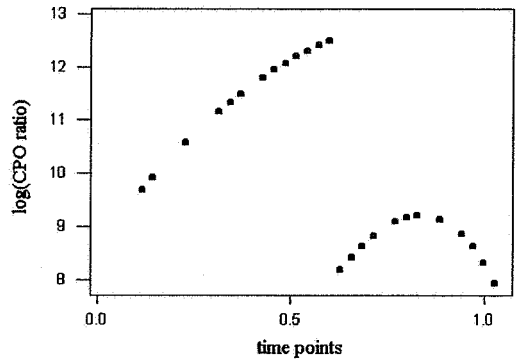

Fig. 7. Logarithm of the ratio of the CPOs.

are the Change point hazards and the solid line represents the Gumbel hazard. The lack of fitting is also evident from the logarithm of the conditional predictive ordinate (CPO) ratios given in Fig. 7. Let us denote the change point model by M_1 and the Gumbel model by M_2. The CPO (Gelfand *et al.* (1992)) of the r-th observation corresponding to model M_i is given by $f\left(y_{r,obs} \mid \boldsymbol{Y}_{(r),obs}, M_i\right)$ and the log ratio of the CPOs is given by

$$\log\left(C_r\right) = \log \frac{f\left(y_{r,obs} \mid \boldsymbol{Y}_{(r),obs}, M_1\right)}{f\left(y_{r,obs} \mid \boldsymbol{Y}_{(r),obs}, M_2\right)}, \qquad r = 1, \ldots, n$$

and $\boldsymbol{Y}_{(r),obs}$ is the observed data vector excluding only the r-th observation.

In Fig. 7 we observe that all the values of $\log\left(C_r\right)$ values are greater than 0 which gives a very strong evidence in favor of the change point model. We also observe how clearly the distribution before the change point differs from the distribution after the change point.

Thus our proposed Change point model performs much better than a traditional Gumbel model.

7. Concluding remarks

In this paper we have proposed a general class of change point hazard function model and established a well-defined relationship between change point model and mixture distribution. The class includes conventional change point model and extended version of it which includes cure rate and lag models. We also indicated how we can model survival fraction. A Bayesian model fitting using Markov chain Monte Carlo method is proposed. Model diagnostics and model selections were discussed through a real failure time data set.

Appendix

LEMMA A.1. *If* $T = \min(x_1, x_2, \ldots, x_N)$, *where* N *is a random variable, then the hazard function of* T *is given by* $h_T(t) = E(h_T(t) \mid N = n) = E(N) h_x(t)$. *Now further assume that* $N \sim Poisson(\theta)$ *then* $h_T(t) = \theta h_x(t)$.

Full conditionals of the model parameters

A. *Change point model parameters*

$$\pi(\alpha \mid rest) \propto \left(\prod_{z_i=1} f_1(t_i)\right) \pi(\alpha)$$
$$\propto \frac{\left(\prod_{z_i=1} (\alpha + \beta t_i)\right) \alpha^{\alpha_2 - 1} \exp\left[-\left(\theta \sum_{z_i=1} t_i + \alpha_1\right) \alpha\right]}{\left(1 - \exp\left[-\theta\left(\alpha\eta + \beta\eta^2/2\right)\right]\right)^{n_1}},$$

$$\pi(\theta \mid rest) \propto \left(\prod_{z_i=1} f_1(t_i)\right) \left(\prod_{z_i=0} f_2(t_i)\right) \pi(\theta)$$
$$\propto \frac{\theta^{n+\theta_1-1} \exp\left[-\theta\left(\sum_{z_i=1} \left(\alpha t_i + \beta t_i^2/2\right) + \delta \sum_{z_i=0} \left(t_i^\gamma - \eta^\gamma\right) + \theta_2\right)\right]}{\left(1 - \exp\left[-\theta\left(\alpha\eta + \beta\eta^2/2\right)\right]\right)^{n_1}},$$

$$\pi(\gamma \mid rest) \propto \left(\prod_{z_i=0} f_2(t_i)\right) \pi(\gamma)$$
$$\propto \prod_{z_i=0} t_i^{\gamma-1} \exp\left[-\left(\sum_{z_i=0} \theta\delta\left(t_i^\gamma - \eta^\gamma\right) + \gamma_2\gamma\right)\right] \gamma^{n_2+\gamma_1-1},$$

$$\pi(\delta \mid rest) \propto \left(\prod_{z_i=0} f_2(t_i)\right) \pi(\delta)$$
$$\equiv Gamma\left(n_2 + \delta_1 - 1, \sum_{z_i=0} \theta\left(t_i^\gamma - \eta^\gamma\right) + \delta_2\right)$$

$$\pi(\eta \mid rest) \propto \left(\prod_{z_i=1} f_1(t_i)\right) \left(\prod_{z_i=0} f_2(t_i)\right) \pi(\eta)$$
$$\propto \frac{\exp\left[-\theta\delta \sum_{z_i=0} \left(t_i^\gamma - \eta^\gamma\right)\right] I\left[\left(t_{(1)} < \eta < t_{(n)}\right)\right]}{\left(1 - \exp\left[-\theta\left(\alpha\eta + \beta\eta^2/2\right)\right]\right)^{n_1}}$$

where

$$n_1 = \sum_{i=1}^{n} z_i \quad \text{and} \quad n_2 = n - n_1.$$

B. *Gumbel model parameters*

$$\pi(\phi \mid \sigma) \equiv Gamma\left(n + \phi_2, \sum_{i=1}^{n} y_i^{1/\sigma} + \phi_1\right)$$

$$\pi(\sigma \mid \phi) \propto (1/\sigma)^{n-\sigma_2+1} \left(\prod_{i=1}^{n} y_i\right)^{1/\sigma-1} \exp\left(-\phi \sum_{i=1}^{n} y_i^{1/\sigma} - \sigma_1 \sigma\right).$$

REFERENCES

Achcar, J. A. and Bolfarine, H. (1989). Constant hazard against a change point alternative: A Bayesian approach with censored data, *Comm. Statist. Theory Methods*, **18**, 3801–3819.

Basu, A. P., Ghosh, J. K. and Joshi, S. N. (1988). On estimating change point in a failure rate, *Statist. Decisions*, **4**, 239–252.

Carlin, B. P., Gelfand, A. E. and Smith, A. F. M. (1992). Hierarchical Bayesian analysis of change point problems, *Applied Statistics*, **41**, 389–405.

Chen, M. H., Ibrahim, J. G. and Sinha, D. (1999). A Bayesian approach to survival data with a cure fraction, *J. Amer. Statist. Assoc.*, **94**, 909–919.

Chib, S. and Greenberg, E. (1995). Understanding the Metropolis-Hastings algorithm, *Amer. Statist.*, **49**, 327–335.

Diebolt, J. and P. Robert, C. P. (1994). Estimation of finite mixture distributions through Bayesian sampling, *J. Roy. Statist. Soc. Ser. B*, **56**, 363–375.

Ebrahimi, N. (1991). On estimating change point in a mean residual life function, *Sankhyā Ser. A*, **53**, 209–219.

Ebrahimi, N., Gelfand, A. E., Ghosh, M. and Ghosh, S. K. (1998). Bayesian analysis of change point hazard rate problem (unpublished manuscript).

Gelfand, A. E. and Smith A. F. M. (1990). Sampling based approaches to calculating marginal densities, *J. Amer. Statist. Assoc.*, **85**, 398–409.

Gelfand, A. E., Dey, D. K. and Chang, H. (1992). Model determination using predictive distributions with implementation via sampling-based methods, *Bayesian Statistics 4* (eds. J. M. Bernardo, J. O. Berger, A. P. Dawid and A. F. M. Smith), 147–167, Oxford University Press, New York.

Ghosh, J. K., Joshi, S. N. and Mukhopadhyay, C. (1993). A Bayesian approach to the estimation of change-point in a hazard rate, *Advances in Reliability* (ed. A. P. Basu), 141–170, North-Holland, Amsterdam.

Gray, R. (1994). A Bayesian analysis of institutional effects in a multicenter cancer clinical trial, *Biometrics*, **50**, 244–253.

Loader, C. R. (1991). Inference for a hazard rate change point, *Biometrika*, **78**, 749–757.

Marshall, A. W. and 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**, 641–652.

Meeker, W. Q. and LuValle, M. J. (1995). An accelerated life test model based on reliability kinetics, *Technometrics*, **37**, 133–146.

Nguyen, H. T., Rogers, G. S. and Walker, E. A. (1984). Estimation in the change point hazard rate models, *Biometrika*, **71**, 229–304.

Raftery, A. E. and Akman, V. E. (1986). Bayesian analysis of a Poisson process with a change point, *Biometrika*, **73**, 85–89.

Sinha, D., Patra, K. and Dey, D. K. (1999). New Bayesian approaches for accelerated life test data, Tech. Report 99–21, Department of Statistics, University of Connecticut.

Sinha, D., Capozzoli, M. H. and Owen, J. (2000). New Bayesian approach to modeling accelerated life test data, *Perspectives in Statistical Sciences* (eds. Basu *et al.*), 276–289, Oxford University Press.

Stangl, D. (1995). Prediction and decision making using Bayesian hierarchical models, *Statistics in Medicine*, **14**, 2173–2190.

Zucker, D. M. and Lakatos, E. (1997). Weighted log rank type statistics for comparing survival curves when there is a time lag in the effectiveness of a treatment, *Biometrika*, **77**, 853–864.