Bayesian Hazard Change-Point Estimation with Incomplete Data

C. Deniz Yenigün * Ülkü Gürler †

Abstract
Hazard function is a fundamental tool in reliability and survival studies. Sometimes abrupt changes may be observed in the hazard function and there is a need for understanding the structure of this change. In this study we first focus on piecewise constant hazard functions with a single change-point when the observations are subject to truncation and censoring. From a Bayesian perspective, we discuss a method for estimating the time and size of the change. We then consider extending this work for piecewise linear hazard functions with a single change-point. The performance of the proposed estimation method is illustrated with a numerical study. Our results indicate that the Bayesian method performs well, and in some cases it may outperform the maximum likelihood method considered as a benchmark. An application to a real data set is also considered.

Key Words: Hazard function, change-point, left truncated right censored data, Bayesian statistics

1. Introduction

Hazard function is a fundamental tool in reliability and survival studies since it quantifies the instantaneous risk of failure of an item at a given time point. Sometimes abrupt changes may be observed in the hazard function, which may correspond to significant improvements in the health conditions of a patient due to a particular treatment, or an alarming deterioration in the physical conditions of an equipment due to fatigue. In such cases it is of interest to detect both the time and the size of the change.

One of the first works that consider changes in the hazard function is by Matthews and Farewell (1982) which studies a piecewise constant hazard model with a single change-point given by

$$\lambda(t) = \begin{cases} \beta & 0 \leq t < \tau \\ \beta + \theta & t \geq \tau \end{cases}, \tag{1}$$

where β and $\beta + \theta > 0$. Here β represents the initial constant hazard rate, θ represents the size of the change in the hazard rate, and τ is the location of the change-point, all of which are unknown. Matthews and Farewell (1982) used numerical techniques to obtain maximum likelihood estimates and simulated the sampling distribution of the likelihood ratio statistic for the null hypothesis that there is no change-point. Loader (1991) considered this model and discussed inference based on the likelihood ratio process. Several other authors considered hazard change-point models including Gijbels and Gürler (2003), Liu et al. (2008), Frobish and Ebrahimi (2009), and Dupuy (2009).

Most of the studies in the literature for hazard change-point estimation assume either complete observations or random censoring. In this study we consider hazard change-point models when the observations are subject to random censoring and truncation. The well-known form censoring naturally arises in survival data in various forms. Truncation is another form of incomplete data observed in survival studies, where the observed sample comes only from a subset of the population. Previous studies on truncation include Hyde (1977), Kalbfleisch and Lawless (1989), Uzunoğulları and Wang (1992), and Lim et al. (2002) among others. More insight on censoring and truncation will be given in the following section.

Our main reference in this paper is Gürler and Yenigün (2011), which considers the piecewise constant hazard change-point model (1) when the observations are subject to left truncation and right censoring, and proposes two estimation methods based on maximum likelihood ideas. Although the maximum likelihood approach is powerful, it has some drawbacks that will be explained below, which motivate us to focus on the same problem from a Bayesian perspective. In what follows, we summarize the maximum likelihood approach of Gürler and Yenigün (2011), then we propose an alternative Bayesian approach. Another contribution of our work is the generalization of piecewise constant hazard model to a more flexible one, piecewise linear hazard model, which we study using

*Istanbul Bilgi University, Department of Industrial Engineering, 34060, Istanbul, Turkey
†Bilkent University, Department of Industrial Engineering, 06800, Ankara, Turkey

both maximum likelihood and Bayesian approaches. Our numerical study indicates that in some instances the Bayesian approach outperform the maximum likelihood approach considered as a benchmark.

The rest of the paper is organized as follows: In Section 2 we present an insight for censoring and truncation, and introduce the notation for left truncated and right censored data, in Section 3 we discuss the estimation problem for piecewise constant hazard model and in Section 4 we discuss the piecewise linear hazard model. Section 5 contains numerical studies for evaluating the performances of the estimators, Section 6 contains an application to a real data set, and Section 7 concludes.

2. Left Truncated and Right Censored Data

In this paper we study survival models when the observarions are incomplete, i.e., they are subject to random left truncation and right censoring. Therefore in this section, after a short discussion on censoring and truncation, we introduce the notation for left truncated and right censored data.

Truncation and censoring are two common forms of incomplete data encountered in survival studies. When truncation is present, the observed sample comes only from a subset of the population. In survival studies a subject may not be included in the study if the time origin of event time precedes the chronological time that the study starts, hence a truncation occurs because the incidences that have occurred before the recordings have started are lost to observation. The effects of truncation becomes more significant when a newly discovered epidemic in survival studies or a new product launch in reliability are considered. Censoring is the most common form of incomplete data. Once a subject is included in the study, it may be subject to censoring due to drop out or other causes such as competing risks. Hence, left truncation and right censoring may naturally arise in cases where truncation excludes some subjects from the study.

Consider a random variable of interest X, representing the time until an event occurs, which may correspond to the survival time of a patient after a treatment or the time until failure of a component. Let Y and C be the truncation and censoring variables respectively, which prevent the complete observation of the variable X. We assume that X, Y, C are independent and nonnegative. Let $T = \min(X, C)$, and $\delta = I(T = X)$, where I is the indicator function. In the presence of left truncation and right censoring, instead of observing independent and identically distributed (i.i.d.) samples of X, we observe triplets (T, Y, δ) only if $Y \leq T$, otherwise nothing is observed. Thus the observations come from the conditional distribution of (T, Y, δ), given that $Y \leq T$. The observed data are given by a set of i.i.d. observations (t_i, y_i, δ_i) for $i = 1, ..., n$.

3. Piecewise Constant Hazard Change-Point Model

In this section we focus on the piecewise constant hazard change-point model (1), when the observations are subject to left truncation and right censoring. After constructing the likelihood function, we review the maximum likelihood estimation method of Gürler and Yenigün (2011), and then we discuss the Bayesian approach to the same problem.

3.1 Constructing the Likelihood

Suppose the random variable of interest X has the hazard function λ as given in (1). Then, the p.d.f. f, the c.d.f. F, and the survival function S of X are as given below, which are all piecewise functions.

$$f(x) = \begin{cases} \beta e^{-\beta x} \equiv f_1(x) & 0 \leq x < \tau \\ (\beta + \theta) e^{-\beta x - \theta(x-\tau)} \equiv f_2(x) & x \geq \tau \end{cases}, \tag{2}$$

$$F(x) = \begin{cases} 1 - e^{-\beta x} \equiv F_1(x) & 0 \leq x < \tau \\ 1 - e^{-\beta x - \theta(x-\tau)} \equiv F_2(x) & x \geq \tau \end{cases}, \tag{3}$$

$$S(x) = \begin{cases} e^{-\beta x} \equiv S_1(x) & 0 \leq x < \tau \\ e^{-\beta x - \theta(x-\tau)} \equiv S_2(x) & x \geq \tau \end{cases}. \tag{4}$$

Here the pieces of the functions before and after the change-point τ are denoted separately for the ease of derivations in the remaining of the paper.

The random variable X is subject to left truncation and right censoring, where the full observation of X is prevented by a right censoring variable C and a left truncation variable Y. We do not assume parametric families of distributions for the censoring and truncation variables since they are not of direct interest. Instead, we treat the data as a random sample given that it is subject to the observed values of censoring and truncation variables. Our main concern is again to estimate the location of the change-point, along with the hazard rate before the change-point and the size of the change. Let us denote the set of unknown parameters by $\Psi = \{\beta, \theta, \tau\}$.

Klein and Moeschberger (2003) summarized the likelihood construction techniques frequently used in survival analysis literature. According to this construction, various types of censoring and truncation schemes have different contributions to the likelihood function. For example, if X is a random variable of interest with p.d.f. f and survival function S, and if X is subject to right censoring, then the contribution of an observed exact lifetime x to the likelihood function is given by $f(x)$, and the contribution of an observed censoring time c to the likelihood function is given by $S(c)$. When we generalize this approach to the left truncation and right censoring model, we have the following.

Recall that in the left truncation and right censoring model, one observes the triplets (T, Y, δ) only if $Y \leq T$, otherwise nothing is observed. Consider an observed random sample (t_i, y_i, δ_i) for $i = 1, ..., n$. In this case, the contribution of an exact lifetime $(t_i = x_i)$ to the likelihood function is $f(x_i)/S(y_i)$, and the contribution of an observed censoring time $(t_i = c_i)$ to the likelihood function is $S(c_i)/S(y_i)$. Putting together all the components, one may write the conditional likelihood function as

$$L \propto \prod_{i\in D} \frac{f(x_i)}{S(y_i)} \prod_{i\in R} \frac{S(c_i)}{S(y_i)}, \tag{5}$$

where D is the set of observations where the real lifetimes are observed and R is the set observations where the censoring times are observed only.

Consider an observed random sample (t_i, y_i, δ_i) for $i = 1, ..., n$. Note that the p.d.f. and the survival function of X are piecewise functions as described in (2) and (4). Then for a given τ, there are six possible types of observations that have different contributions to the likelihood function, all of which are given in Table 1. For example, A denotes the set of observed triplets for which t is an actual lifetime x, and both x and the observed truncation variable y are less than τ. The contribution of such (x, y) pairs to the likelihood function is $f_1(x)/S_1(y)$, where f_1 and S_1 are as described in (2) and (4). We define the sets B, C, D, E, F, and their likelihood contributions similarly. Let n_A denote the number of observed triplets in set A, let $\prod_A$ denote the product over set A, let $n_{B,C}$ denote the total number of observed triplets in sets B and C, and let $\sum_{B,C}$ denote the sum over sets B and C. Define all the other related subscripts similarly. Then we can write the likelihood function as follows:

$$L(\beta, \theta, \tau|y, t, \delta) = \prod_A \frac{f_1(x_i)}{S_1(y_i)} \prod_B \frac{f_2(x_i)}{S_1(y_i)} \prod_C \frac{f_2(x_i)}{S_2(y_i)} \prod_D \frac{S_1(c_i)}{S_1(y_i)} \prod_E \frac{S_2(c_i)}{S_1(y_i)} \prod_F \frac{S_2(c_i)}{S_2(y_i)}. \tag{6}$$

After some algebra, we have

$$\begin{aligned} L(\beta, \theta, \tau|y, t, \delta) &= \beta^{n_A}(\beta+\theta)^{n_{B,C}} \\ &\quad \exp\left\{ n\beta(\bar{y} - \bar{t}) + \theta \left[\sum_{C,F} y_i - \sum_{B,C} x_i - \sum_{E,F} c_i + n_{B,E}\tau \right] \right\}. \end{aligned} \tag{7}$$

3.2 Maximum Likelihood Estimation

The likelihood function (7) is not differentiable with respect to τ, hence it is not possible to find the M.L.E.'s for Ψ using standard methods. Therefore, Gürler and Yenigün (2011) suggest that one may first fix the value of τ and find the M.L.E.'s for the remaining parameters as a function of τ. Then a search for the value of τ leads to an estimator, which maximizes the likelihood function over a number of grid points on a specific interval $[\tau_0, \tau_1]$.

Type of Observation	Contribution to Likelihood
$A = \{(t_i, y_i, \delta_i) : \delta_i = 1, y_i < x_i \leq \tau\}$	$f_1(x_i)/S_1(y_i)$
$B = \{(t_i, y_i, \delta_i) : \delta_i = 1, y_i \leq \tau < x_i\}$	$f_2(x_i)/S_1(y_i)$
$C = \{(t_i, y_i, \delta_i) : \delta_i = 1, \tau < y_i < x_i\}$	$f_2(x_i)/S_2(y_i)$
$D = \{(t_i, y_i, \delta_i) : \delta_i = 0, y_i < c_i \leq \tau\}$	$S_1(c_i)/S_1(y_i)$
$E = \{(t_i, y_i, \delta_i) : \delta_i = 0, y_i \leq \tau < c_i\}$	$S_2(c_i)/S_1(y_i)$
$F = \{(t_i, y_i, \delta_i) : \delta_i = 0, \tau < y_i < c_i\}$	$S_2(c_i)/S_2(y_i)$

Table 1: Likelihood contributions of different observation types.

For a fixed τ, let $\Psi_\tau = \{\beta, \theta\}_\tau$ be the parameter set to be estimated, and let L_τ denote the likelihood function (7) written for this fixed τ value. Then, the M.L.E. $\hat{\Psi}_\tau = (\hat{\beta}, \hat{\theta})_\tau$ for Ψ_τ is obtained by maximizing L_τ for the parameters β and θ. Please see Gürler and Yenigün (2011) for details. Let $\tau_i \in [\tau_0, \tau_1]$, $i = 1, ..., m$ denote the fixed grid points in the search interval and let L_{τ_i} denote the maximum of the likelihood function for $\tau = \tau_i$. That is

$$L_{\tau_i} = L(\{\hat{\beta}, \hat{\theta}\}_{\tau_i}, \tau_i).$$

Then the maximum likelihood estimators for the change-point τ and the rest of the parameters are given by

$$\hat{\tau} = \text{argmax}_{\tau_i} L_{\tau_i} \tag{8}$$

and

$$\hat{\Psi} = \{(\hat{\beta}, \hat{\theta})_{\hat{\tau}}, \hat{\tau}\}. \tag{9}$$

3.3 Bayesian Estimation

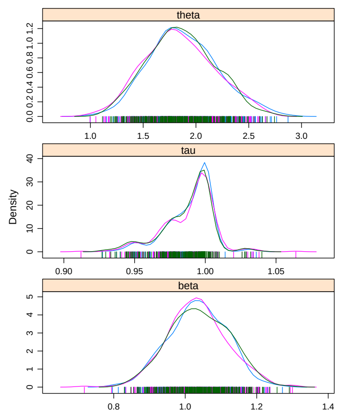
The major drawback of the maximum likelihood estimation method is the loss of dimension. Since the likelihood function is not differentiable with respect to the change-point τ, this method searches for the maximum of the likelihood function over fixed grid points for τ. Although this may give good estimation performance for τ itself, it does not guarantee the global maximum of the likelihood function, and the estimators for the rest of the parameters may suffer from this dimension loss. In order to overcome this disadvantage, we consider the Bayesian estimation of the model parameters, which is the main focus of this paper.

Consider the piecewise constant hazard change-point model (1), where the set of unknown parameters is $\Psi = \{\beta, \theta, \tau\}$. We now assume that these parameters are in fact random variables, and assign them some prior distributions denoted by f_β, f_θ, and f_τ. In the classical Bayesian approach, given the prior distributions, the likelihood function, and a set of observations, the posterior distribution of the model parameters are computed by the use of Bayes theorem. However, the likelihood function (7) does not allow the computation of closed form posterior distributions due to its structure depending on τ, the location of change-point. Therefore, we take the usual step and carry out the Bayesian estimation using the Markov Chain Monte Carlo (MCMC) methodology, where we generate random samples from the posterior distributions using WinBUGS, a statistical software for Bayesian analysis. We use the package R2WinBUGS for running WinBUGS from within the statistical software R. We illustrate the Bayesian estimation with the following numerical example.

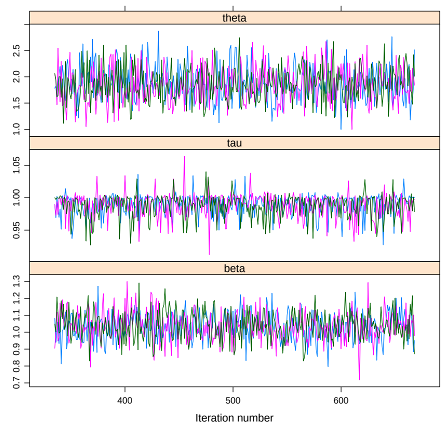
3.3.1 A Numerical Example for the Bayesian Estimation

We first generate a left truncated and right censored data set of size $n = 180$, where the random variable of interest X follows the piecewise constant hazard model (1) with $\beta = 1$, $\theta = 2$, $\tau = 1$. Left truncation and right censoring are introduced by two quantities, namely, $\alpha = P(Y < T)$ that corresponds to the proportion of untruncated observations, and $\alpha' = P(X < C|Y < T)$ that corresponds to the proportion of uncensored observations. Here Y and C are the truncation and censoring variables, respectively, which are assumed to follow exponential distributions. When we set $Y \sim \text{Exp}(4.1)$ and $C \sim \text{Exp}(0.31)$ we have $\alpha = \alpha' = 0.75$, a quantity that we refer to as *observation level*, which stands for the impact of censoring and truncation. Once we have generated the artificial data as triplets (t_i, y_i, δ_i), $i = 1, ..., n$ (see Section 2), we carry out the Bayesian estimation. We use uniform prior distributions such that $\beta \sim U(0, 2)$, $\theta \sim U(0, 3)$,

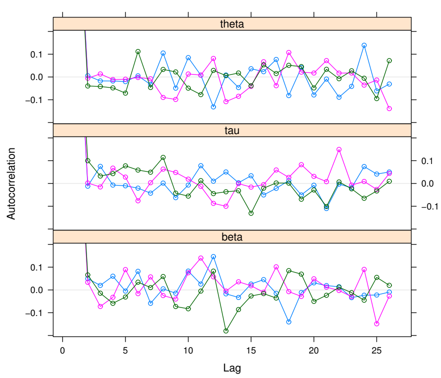
$\tau \sim U(0,2)$ along with the likelihood function (7) and the data, and have WinBUGS simulate from the posterior distributions of β, θ, and τ. For each parameter, three MCMC chains of size 10000 are simulated after a burn-in period of 2000 iterations. Density plots for the resulting samples are given in Figure 1-a, from which we may come up with 95% confidence intervals as $\tau \in (0.94, 1.01)$, $\beta \in (0.88, 1.20)$, $\theta \in (1.27, 2.53)$. Note that all three intervals cover the true parameter values. As for the Bayesian estimation diagnostics, we provide the MCMC iteration plots and autocorrelation plots in Figures 1-b and 1-c, respectively, where we observe that the MCMC sampling methodology seems to be satisfactory in terms of convergence and independence.



a) Density plots.



b) Iterations plots.



c) Autocorrelation plots.

Figure 1: Plots for the numerical example for the Bayesian estimation.

4. Piecewise Linear Hazard Change-Point Model

Before presenting the comparative performances of the maximum likelihood and Bayesian methods for the estimation of piecewise constant model (1) in the next section, we now consider a flexible alternative to this simple model. Although constant hazard models are of great importance in survival and reliability studies, sometimes there is a need for modeling variations in the hazard function. Therefore we introduce the piecewise linear hazard model with a single change-point given as follows:

$$\lambda(t) = \begin{cases} \beta + at & 0 \le t < \tau \\ \gamma + b(t-\tau) & t \ge \tau \end{cases}, \tag{10}$$

where β is the initial hazard rate, a is the initial slope, γ is the hazard rate right after the change-point, b is the slope after the change-point, τ is the location of the change-point.

Suppose that the random variable of interest X has the hazard function (10). Then, the p.d.f. f, the c.d.f. F, and the survival function S of X are as given below, which are all piecewise functions.

$$f(x) = \begin{cases} (\beta + ax)e^{-\beta x - \frac{a}{2}x^2} \equiv f_1(x) & 0 \le x < \tau \\ (\gamma - b\tau + bx)e^{-(\beta-\gamma)\tau - \frac{(a+b)}{2}\tau^2 - (\gamma - b\tau)x - \frac{b}{2}x^2} \equiv f_2(x) & x \ge \tau \end{cases}, \quad (11)$$

$$F(x) = \begin{cases} 1 - e^{-\beta x - \frac{a}{2}x^2} \equiv F_1(x) & 0 \le x < \tau \\ 1 - e^{-(\beta-\gamma)\tau - \frac{(a+b)}{2}\tau^2 - (\gamma - b\tau)x - \frac{b}{2}x^2} \equiv F_2(x) & x \ge \tau \end{cases}, \quad (12)$$

$$S(x) = \begin{cases} e^{-\beta x - \frac{a}{2}x^2} \equiv S_1(x) & 0 \le x < \tau \\ e^{-(\beta-\gamma)\tau - \frac{(a+b)}{2}\tau^2 - (\gamma - b\tau)x - \frac{b}{2}x^2} \equiv S_2(x) & x \ge \tau \end{cases}. \quad (13)$$

For the piecewise linear hazard model (10), the set of unknown parameters is $\Psi' = \{\beta, \gamma, \tau, a, b\}$. The likelihood may be constructed similarly to the piecewise constant hazard model, using the approach illustrated by Table 1 and the general structure in the likelihood function (6). Then the likelihood function is given by

$$\begin{aligned} L(\beta, \gamma, \tau, a, b | y, t, \delta) = & \prod_A (\beta + ax_i) \prod_{BC} (\gamma - b\tau + bx_i) \exp\Bigg\{ \beta \left(\sum_{ABDE} y_i - \sum_A x_i - \sum_D c_i \right) \\ & + (\gamma - b\tau) \left(\sum_{CF} y_i - \sum_{BC} x_i - \sum_{EF} c_i \right) + \frac{a}{2} \left(\sum_{ABDE} y_i^2 - \sum_A x_i^2 - \sum_D c_i^2 \right) \\ & + \frac{b}{2} \left(\sum_{CF} y_i^2 - \sum_{BC} x_i^2 - \sum_{EF} c_i^2 \right) + \tau n_{BE} \left(\beta - \gamma + \frac{a+b}{2}\tau \right) \Bigg\}. \quad (14) \end{aligned}$$

Note that here the sets A to F are defined exactly as in Table 1, and the notation such as $\prod_A$ is used the same as in Section 3.1. The structure of this likelihood function prevents us from carrying out the maximum likelihood or the Bayesian estimation in classical ways, therefore, we carry out the estimation procedures just as we did for the piecewise constant hazard model in Section 3. For the maximum likelihood estimation, we now have $\hat{\Psi}' = \{(\beta, \gamma, a, b)_\tau, \hat{\tau}\}$. As for the Bayesian methodology, we again use the MCMC approach where we now have two additional parameters a and b with prior distributions f_a and f_b. The mechanics of both estimation procedures are the same as the procedures in Section 3, therefore details are not given here.

5. Numerical Results

The maximum likelihood approach of Gürler and Yenigün (2011) is shown to be a powerful one, however, the dimension loss we have mentioned in Section 3 may result in poor estimation performance for model parameters other than the location of change-point τ. The Bayesian approach we take in this study does not seem to have this problem, however, since closed form expressions for the posterior distribution cannot be achieved, one needs to rely on MCMC samples. Therefore, it is of interest to compare the estimation performances of the two approaches, and in this section we present the results of a numerical study for this purpose. In order to be able to compare the two methods, we use the posterior sample means in the Bayesian approach as the point estimates of the model parameters. As in the numerical example above, all posterior samples are obtained by 10000 iterations after a burn-in period of 2000. We present our results for the piecewise constant hazard function first, followed by the results for the piecewise linear hazard model. In each case we simulate samples from the indicated scenarios, estimate the parameters with both methods, and report the empirical mean square error as our performance measure.

5.1 Numerical Results for the Piecewise Constant Hazard Model

We generate random samples from the piecewise constant hazard model (1) with $\beta = 1$, $\theta = 2$, $\tau = 1$. The impact of truncation and censoring is controlled by setting three *observation levels* (*OL*), 60%, 75%, and 90%. Here a 60% observation level corresponds to the case that both the proportion of untruncated observations, $\alpha = P(Y < T)$, and the proportion of the uncensored observations,

$\alpha' = P(X < C|Y < T)$, are both equal to 0.6. Simiarly for the other observation levels. In order to achieve these observation levels, we simulate the truncation and censoring variables as follows: $Y \sim$ Exp(2.2), $C \sim$ Exp(0.48) for 60% observation level, $Y \sim$ Exp(4.1), $C \sim$ Exp(0.31) for 75% observation level, and $Y \sim$ Exp(10), $C \sim$ Exp(0.1) for 90% observation level. In order to control the amount of information contained in the prior distributions, we use two prior sets. The less informative *Prior Set 1* assumes $\tau \sim U(0, 2)$, $\beta \sim U(0, 2)$, $\theta \sim U(0, 3)$, and the more informative *Prior Set 2* assumes $\tau \sim U(0.5, 1.5)$, $\beta \sim U(0.5, 1.5)$, $\theta \sim U(1.5, 2.5)$. For four sample sizes and three observation levels, the empirical mean squared errors based on $N = 100$ repetitions are given in Table 2. The numerical results indicate that the sensitivity to sample size, observation level, and the prior information seems to be as expecdet. An increase in sample size, observation level, or prior knowledge results in smaller empirical mean squared error. As for the comparison between the maximum likelihood and Bayesian methodology, we may say that two methods perform comparable, however, for estimating the size of the change, θ, Bayesian method performs much better than the maximum likelihood method.

OL		Bayes (Prior Set 1)			Bayes (Prior Set 2)			MLE		
	n	τ	β	θ	τ	β	θ	τ	β	θ
60%	60	0.013	0.045	0.165	0.010	0.024	0.008	0.008	0.075	2.237
	100	0.014	0.031	0.184	0.007	0.023	0.013	0.006	0.058	0.843
	140	0.005	0.019	0.125	0.005	0.015	0.015	0.004	0.042	0.457
	180	0.002	0.013	0.174	0.001	0.012	0.014	0.002	0.035	0.249
75%	60	0.018	0.051	0.165	0.011	0.020	0.009	0.007	0.044	1.080
	100	0.010	0.032	0.169	0.004	0.021	0.012	0.005	0.024	0.500
	140	0.003	0.017	0.143	0.003	0.014	0.016	0.003	0.017	0.279
	180	0.002	0.012	0.132	0.002	0.011	0.021	0.002	0.007	0.176
90%	60	0.015	0.033	0.122	0.010	0.024	0.008	0.006	0.030	0.993
	100	0.007	0.021	0.118	0.006	0.017	0.013	0.004	0.017	0.445
	140	0.003	0.014	0.132	0.003	0.012	0.019	0.003	0.013	0.274
	180	0.001	0.011	0.079	0.002	0.009	0.026	0.002	0.009	0.166

Table 2: Empirical mean square error comparisons for the Bayesian and maximum likelihood estimation methods for the piecewise constant hazard model (1).

5.2 Piecewise Linear Hazar Model

We generate random samples from the piecewise linear hazard model (10) with $\beta = 1$, $\gamma = 4$, $\tau = 1$, $a = 0.5$, $b = 0.31$. Truncation and censoring variables are generated from $Y \sim$ Exp(4.1) and $C \sim$ Exp(0.31). We use uniform priors for the Bayesian methodology, where $\tau \sim U(0, 2)$, $\beta \sim U(0, 2)$, $\gamma \sim U(3, 5)$, $a \sim U(0, 1)$, $b \sim U(2, 4)$. For four sample sizes the empirical mean squared errors based on $N = 100$ repetitions are given in Table 3. Here, due to increased number of parameters that need to be estimated, we observe a much better performance from the Bayesian method compared to the maximum likelihood method, especially for the parameters other than the change-point τ. This is what we were expecting as the maximum likelihood approach employed here is essentially a search over τ, however, Bayesian approach provides a holistic estimation method.

n	Method	τ	β	θ	a	b
60	MLE	0.006	0.100	2.079	0.334	4.735
	Bayesian	0.018	0.051	0.022	0.009	0.003
100	MLE	0.004	0.070	1.556	0.260	4.455
	Bayesian	0.008	0.028	0.037	0.012	0.005
140	MLE	0.003	0.043	0.986	0.148	4.193
	Bayesian	0.008	0.022	0.063	0.015	0.008
180	MLE	0.004	0.041	0.985	0.168	4.835
	Bayesian	0.004	0.020	0.068	0.017	0.008

Table 3: Empirical mean square error comparisons for the Bayesian and maximum likelihood methods for the piecewise linear hazard model (10).

6. An Application

In this section we apply the Bayesian estimation methodology studied above for the piecewise linear hazard model (10) to the Channing House data of Hyde (1977). The observations consists of survival data for male members of the Channing House retirement community in Palo Alto, California. The variable of interest X is the lifetime in months of the male retired people. The individuals are observed only if they survive long enough to participate in the retirement community. Once they come under observation, the participants are also subject to censoring. Therefore, the left truncation and right censoring model applies to this data set. Here the truncation variable Y is the age in months at entry into study, and $T = \min(X, C)$ is the variable referring to age in months when last seen in the study. The data consists of 97 observations, out of which 46 have died, 46 survived to the end of the study, and 5 withdrew from the retirement community.

Based on a previous analysis by Uzunoğulları and Wang (1992), we set the prior distributions such that $\tau \sim U(900, 1100)$, $\beta \sim U(0, 0.001)$, $\gamma \sim U(0, 0.01)$, $a \sim U(0, 0.001)$, $b \sim U(0, 0.01)$. Density plots of three chains of MCMC samples of size 10000, after a burning period of 2000, are given in Figure 2. 95 % confidence intervals are $a \in (2.68 \times 10^{-6}, 7.88 \times 10^{-6})$, $b \in (2.08 \times 10^{-5}, 1.09 \times 10^{-3})$, $\beta \in (2.94 \times 10^{-5}, 9.77 \times 10^{-4})$, $\gamma \in (5.5 \times 10^{-4}, 9.7 \times 10^{-3})$, $\tau \in (911.6, 1097)$. We find that there is a significant change in the slope, but the jump is found to be insignificant. The hazard rate increases rapidly in the earlier ages below 1000 months, i.e. 83 years old, then proceeds with a smaller rate of increase.

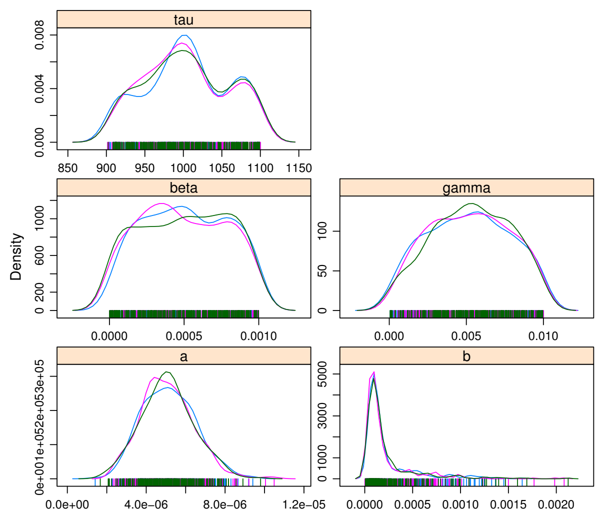


Figure 2: Channing house data, posterior density plots for the piecewise linear hazard model parameters.

7. Conclusions

In this paper we applied Bayesian estimation methodology to two hazard change-point models when the observations are subject to censoring and truncation. The first model is the well-known piecewise constant hazard model with a single change-point, and the second one is a flexible alternative we provide for this model, the piecewise linear hazard model with a single change-point. We study both models using the maximum likelihood and Bayesian estimation methods, and provide a comparative numerical study in order to illustrate the performances. Although both models estimate the location of the change-point successfully, when it comes to estimating the rest of the model parameters, such as the size of the change, the Bayesian method outperforms the maximum likelihood method which is considered as a benchmark.

REFERENCES

Dupuy, J-F. (2009), "Detecting change in a hazard regression model with right-censoring," *Journal of Statistical Planning and Inference*, 139, 1578-1586.

Frobish, D., Ebrahimi, N. (2009), "Parametric estimation of change-points for actual event data in recurrent events models," *Computational Statistics and Data Analysis*, 53, 671-682.

Gijbels, I., Gürler, Ü. (2003), "Estimation of a change-point in a hazard function based on censored data," *Lifetime Data Analysis*, 9, 395-411.

Gürler, Ü., Yenigün, C.D. (2003), "Full and conditional likelihood approaches for hazard change-point estimation with truncated and censored data," *Computational Statistics and Data Analysis*, 55, 2856-2870.

Hyde, J. (1977), "Testing survival under right censoring and left truncation," *Biometrika*, 64, 225-230.

Kalbfleisch, J.D., Lawless, J.F. (1989), "Inference based on retrospective ascertainment: An analysis of the data on transfusion-related AIDS," *Journal of the American Statistical Association*, 84, 360-372.

Klein, J.P., Moeschberger, M.L. (2003), *Survival Analysis*, New York:Springer.

Lim, H., Sun, J., Matthews, D.E. (2002), "Maximum likelihood estimation of a survival function with a change-point for truncated and interval-censored data," *Statistics in Medicine*, 21, 743-752.

Liu, M., Lu, W., Shao Y. (2008), "A Monte Carlo approach for change-point detection in the Cox proportional hazards model," *Statist. Med.*, 27, 3894-3909.

Loader, C.R. (1991), "Inference for hazard rate change-point," *Biometrika*, 78, 835-843.

Matthews, D.E., Farewell, V.T. (1982), "On testing for a constant hazard against a change-point alternative," *Biometrics*, 38, 463-468.

Uzunoğulları Ü., Wang, J-L. (1992), "A comparison of hazard rate estimators for left truncated and right censored data," *Biometrika*, 79, 297-310.