Hacettepe Journal of Mathematics and Statistics
Volume 32 (2003), 33 – 41

A BAYESIAN APPROACH TO THE CONSTANT HAZARD MODEL WITH A CHANGE POINT AND AN APPLICATION TO BREAST CANCER DATA

Durdu Sertkaya* and M. Tekin Sözer*

Received 18. 04. 2002 : Accepted 02. 04. 2003

Abstract

In this paper, a Bayesian approach to the problem of constant hazard with a change point is considered using noninformative priors. We apply the model to a data set gathered from a group of patients with breast cancer.

Keywords: Change point, Constant hazard, Hazard rate, Bayesian analysis.

1. Introduction

The distribution of survival times is usually described or characterized in terms of three functions:

(1) The survival function,
(2) The probability density function, and
(3) The hazard function.

These three functions are mathematically equivalent - if one of them is given, the other two can be derived [3,4,5].

Our aim in this study is to consider the constant hazard model with a change point and use Bayesian analysis to estimate the parameters in this model. We also consider an application with real data.

Often, researches in the medical area are interested in constant hazard models involving a single change point. Let T be a random variable representing the time to some event, the constant hazard model with change point is given by:

$$h(t) = \begin{cases} \lambda & t \leq \tau \\ \rho\lambda & t > \tau \end{cases}, \tag{1.1}$$

*Hacettepe University, Faculty of Science, Department of Statistics, 06532 Beytepe, Ankara, Turkey

where λ, ρ, $\tau > 0$. Here λ and $\rho\lambda$ are hazard rates and τ is the change point. The hazard model $h(t)$ is therefore assumed to be a constant λ until time τ and a constant $\rho\lambda$ after time τ.

Model (1.1) has been extensively studied in the literature. Matthews and Farewell [7] considered the model studied in this paper, using data obtained in the treatment of leukemia patients. They considered the problem of testing the hypothesis $\tau = 0$ based on the likelihood ratio test statistic and used simulations to find the distribution of the statistics of this model.

In another paper, Matthews, Farewell and Pyke [9] presented an asymptotic score statistic processes to test for constant hazard $\tau = 0$ against a change point alternative $\tau > 0$. Nguyen, Rogers and Walker [10] discussed the estimation of parameters in the constant hazard model with a change point. In their paper a consistent estimator of the change point was obtained by examining the properties of the density represented as a mixture.

Yao [13] considered the problem of estimation of parameters in hazard rate models with a change-point. Yao [13] proposed a maximum likelihood estimator of the change-point subject to a natural constraint. Worsley [11] used maximum likelihood methods to test for a change in a sequence of independent exponential family random variables, with particular emphasis on the exponential distribution. The exact null and alternative distributions of the test statistics were found, and the power was compared with a test based on a linear trend statistic. Exact and approximate confidence regions for the change-point were based on the values accepted by a level α likelihood ratio test.

Worsley [12] considered the test that the hazard rate of survival or failure-time data is constant against the alternative of a change in hazard after an unspecified time. The likelihood ratio was unbounded but the exact null distribution of a restricted likelihood-ratio test statistic was found.

Loader [6] discussed inference based on the likelihood ratio process for a hazard rate change point. Loader [6] derived approximate confidence regions for the change point and joint confidence regions for the change point and size of change. The effect of censorship was also discussed. The methods were illustrated using Stanford heart transplant data.

In the studies mentioned above the maximum likelihood estimate for the change point τ is considered. As an alternative to these classical solutions, Achcar and Bolfarine [1] presented a Bayesian analysis of the model (1.1) assuming τ as either known or unknown. Achcar and Loibel [2] presented Bayesian inferences for the model (1.1) using different prior densities.

2. The Probability Density, Survival and Likelihood Functions

The probability density and survival functions with change point for a random variable T are, respectively

$$f(t) = \begin{cases} \lambda \exp(-\lambda t) & t \le \tau \\ \rho\lambda \exp\{-\lambda\tau - \rho\lambda(t-\tau)\} & t > \tau \end{cases} \tag{2.1}$$

and

$$\mathrm{S}(t) = \begin{cases} \exp(-\lambda t) & t \le \tau \\ \exp\{-\lambda t - \rho\lambda(t-\tau)\} & t > \tau \end{cases} \tag{2.2}$$

[1,2,6,7,9,13].

Let $T_1, T_2, \ldots, T_n$ be a random sample from (1.1) and $T_1^0, T_2^0, \ldots, T_n^0$ the true survival times of n individuals considered as a random sample of size n, and let $C_1, \ldots, C_n$

be the fixed censoring times associated with each individual. The data are given by $T_i = \min(T_i^0, C_i)$, $i = 1, 2, \ldots, n$ and t_i is the observed survival time of the i th patient. An indicator variable δ_i is defined by $\delta_i = 1$ if $T_i = T_i^0$ (failure) and $\delta_i = 0$ if $T_i \leq T_i^0$ (censoring). Let $\epsilon_i = 1$ if $T_i \leq \tau$ and $\epsilon_i = 0$ if $T_i > \tau$, $i = 1, \ldots, n$. Then the likelihood function is

$$\begin{aligned} \mathrm{L}(\lambda, \rho, \tau) = \prod_{i=1}^{n} \left\{ (\lambda \mathrm{e}^{-\lambda t_i})^{\epsilon_i} \Big[\lambda\rho \exp(-\lambda\tau - \rho\lambda(t_i - \tau)) \Big]^{1-\epsilon_i} \right\}^{\delta_i} \times \\ \times \prod_{i=1}^{n} \left\{ (\mathrm{e}^{-\lambda t_i})^{\epsilon_i} \Big[\exp(-\lambda\tau - \rho\lambda(t_i - \tau)) \Big]^{1-\epsilon_i} \right\}^{1-\delta_i} \end{aligned} \tag{2.3}$$

that is,

$$\mathrm{L}(\lambda, \rho, \tau) = \lambda^{d_3} \rho^{d_3 - d_1(\tau)} \exp\{-\lambda W_1(\tau) - \rho\lambda W_2(\tau)\} \exp\{-\lambda\tau(1-\rho)(n - d_2(\tau)\} \tag{2.4}$$

where

$$d_1(\tau) = \sum_{i=1}^{n} \delta_i \epsilon_i \text{ is the number of uncensored observations which satisfy } T_i \leq \tau,$$

$$d_2(\tau) = \sum_{i=1}^{n} \epsilon_i \text{ is the number of observations which satisfy } T_1 \leq \tau \text{ and}$$

$$d_3 = \sum_{i=1}^{n} \delta_i \text{ is the number of uncensored observations,}$$

$$W_1(\tau) = \sum_{i=1}^{n} \delta_i \epsilon_i t_i + \sum_{i=1}^{n} (1 - \delta_i)\epsilon_i t_i \text{ is the sum of survival times of observations which satisfy } T_i \leq \tau.$$

Here,

$\sum_{i=1}^{n} \delta_i \epsilon_i t_i$ is the sum of failure times of uncensored observations satisfying $T_i \leq \tau$,

$\sum_{i=1}^{n} (1 - \delta_i)\epsilon_i t_i$ is the sum of censoring times of uncensored observations with $T_i \leq \tau$.

Also,

$$W_2(\tau) = \sum_{i=1}^{n} \delta_i (1 - \epsilon_i) t_i + \sum_{i=1}^{n} (1 - \delta_i)(1 - \epsilon_i) t_i \text{ is the sum of survival times of observations which satisfy } T_i \leq \tau$$

where

$\sum_{i=1}^{n} \delta_i (1 - \epsilon_i) t_i$ is the sum of failure times of uncensored observations satisfying $T_i \leq \tau$,

$\sum_{i=1}^{n} (1 - \delta_i)(1 - \epsilon_i) t_i$ is the sum of censoring times of censored observations which satisfy $T_i > \tau$.

[1, 2, 10, 13].

3. Maximum Likelihood Estimators for λ and ρ

Given τ, Achcar and Bolfarine [1], Loader [6], Nguyen, Rogers and Walker [10] and Yao [13] found $\widehat{\lambda}$ and $\widehat{\rho}$ that maximize $\mathrm{L}(\lambda, \rho, \tau)$ by deriving the loglikelihood with respect to λ and ρ. From $\partial L/\partial\lambda = 0$ and $\partial L/\partial\rho = 0$, they obtained the maximum likelihood estimators for λ and ρ as follows:

$$\widehat{\lambda} = \frac{d_1(\tau)}{W_1(\tau) + \tau(n - d_2(\tau))} \tag{3.1}$$

and

$$\widehat{\rho} = \frac{(d_3 - d_1(\tau))[W_1(\tau) + \tau(n - d_2(\tau))]}{d_1(\tau)[W_2(\tau) - \tau(n - d_2(\tau))]}, \tag{3.2}$$

[1,6,10,13].

4. Bayesian Analysis

4.1. A Bayesian Analysis assuming τ and λ known. With τ and λ known, and assuming that censoring is not present, the likelihood function for ρ is

$$\mathrm{L}(\rho) \propto \rho^{d_3 - d_1(\tau)} \exp\{-\lambda\rho Z(\tau)\}, \tag{4.1}$$

where $\rho > 0$, d_3, $\mathrm{d}_1(\tau)$ and $d_2(\tau)$ are as given in (2.4), and $Z(\tau) = W_2(\tau) - (n - d_2(\tau))$.

A noninformative prior density for ρ is ρ^{-1}. Thus, the posterior density for ρ is

$$\pi(\rho \mid D) \propto \rho^{d_3 - d_1(\tau) - 1} \exp\{-\rho\lambda Z(\tau)\}, \tag{4.2}$$

where $\rho > 0$, $Z(\tau)$ is as given in (4.1) and D denotes the data set. The value which maximizes the posterior density is known as the mode of posterior density, and this mode corresponds to the estimated value. Thus the mode of the posterior density given in (4.2) is

$$\widehat{\rho} = \frac{d_3 - d_1(\tau) - 1}{\lambda Z(\tau)}, \tag{4.3}$$

[1].

4.2. A Bayesian Analysis assuming τ known. With τ known, a noninformative prior density for λ and ρ is given by

$$\pi(\lambda, \rho) \propto 1/\lambda\rho,$$

where $\lambda, \rho > 0$ in the complete data case.

The posterior density for λ and ρ is then

$$\pi(\lambda, \rho \mid D) \propto \lambda^{d_3 - 1} \rho^{d_3 - d_1(\tau) - 1} \exp\{-\lambda W_1(\tau) - \rho\lambda W_2(\tau)\} \cdot \exp\{-\lambda\tau(1 - \rho)(n - d_2(\tau))\}, \tag{4.4}$$

from which it follows that the mode of posterior density for ρ is given by

$$\widehat{\rho} = \frac{(d_3 - d_1(\tau) - 1)[W_1(\tau) + \tau(n - d_2(\tau))]}{(d_1(\tau) + 1)[W_2(\tau) - \tau(n - d_2(\tau))]}. \tag{4.5}$$

From (4.4), it also follows that the mode of posterior density for λ is

$$\widehat{\lambda} = \frac{(d_1(\tau) - 1}{W_1(\tau) + \tau(n - d_2(\tau))}, \tag{4.6}$$

[1].

4.3. A Bayesian Analysis assuming τ unknown. Let us assume that τ is unknown and take discrete values $\tau_i = t_i$, with prior probabilities $\pi_0(\tau_i = t_i)$, $(i = 1, \ldots, n)$, where n is the sample size. Also assume that the data is uncensored. The prior density for λ, ρ and τ_i is

$$\pi(\lambda, \rho, \tau_i) = \pi(\lambda, \rho \mid \tau_i = t_i)\pi_0(\tau_i = t_i).$$

Hence the posterior density for λ, ρ and τ_i is

$$\pi(\lambda, \rho, \tau_i = t_i \mid D) \propto \lambda^{d_3-1}\rho^{d_3-d_1(\tau_i)-1}\pi_0(\tau_i = t_i)\exp\{-\lambda W_1(\tau_i) - \rho\lambda W_2(\tau_i)\}\cdot \exp\{-\lambda\tau_i(1-\rho)(n-d_2(\tau_i))\}, \tag{4.7}$$

where $\lambda, \rho > 0$.

The posterior density for τ_i is given by

$$\pi(\tau_i = t_i \mid D) \propto \frac{\Gamma(d_3 - d_1(\tau_i))\pi_0(\tau_i = t_i)}{[W_2(\tau_i) - \tau_i(n - d_2(\tau_i))]^{d_3-d_1(\tau_i)}} \cdot \frac{\Gamma(d_1(\tau_i))}{[W_1(\tau_i) - \tau_i(n - d_2(\tau_i))]^{d_1(\tau_i)}} \tag{4.8}$$

where $i = 1, 2, \ldots, n$. The value of t which maximizes this posterior density corresponds to the estimation of τ [1].

5. Test of Change Point

Consider the likelihood ratio statistic used to test the hypothesis of no change point, which Matthews and Farewell [7] denoted by $\tau = 0$. Let the loglikelihood statistic be $\mathrm{L}(\lambda_1, \rho, \tau)$. The loglikelihood test statistic for the null hypothesis $\tau = 0$ is, therefore,

$$\Delta_0 = \mathrm{L}(\widehat{\lambda}_1, \hat{\rho}, \hat{\tau}) - \mathrm{L}(\hat{\lambda}, \hat{\lambda}, 0), \tag{5.1}$$

where $\widehat{\lambda}_1$, $\hat{\rho}$ and $\hat{\tau}$ are the maximum likelihood estimators of λ_1, ρ and τ, respectively. Also $\hat{\lambda}$ is the maximum likelihood estimator of the hazard rate in a simple exponential model. Here, $2\Delta_0$ has a χ^2 distribution with two degrees of freedom [7].

6. Application: Breast Cancer Data

We apply the model to a data set for a group of patients with breast cancer from the Oncology Department in Hacettepe hospital, admitted between January 1980 and September 1991. There are 124 patients. Survival time was calculated from the date of diagnosis of breast cancer to date of death. Median follow-up is 43.5 (range 6-146) months. The total number of deaths is 80. The other 44 women being considered as censored in the analysis. We are interested in the age variable. The mean age is 50.59 $\pm$1.10 (range 25-78) years. The response variable is the time to death. Table 1 gives the survival times of the 124 breast cancer patients.

Table 1. Survival times of the breast cancer patients (in months).

Uncensored observations (80 patients)																			
7	12	12	13	14	14	17	18	18	19	20	21	22	22	24	24	27	28	30	30
30	36	36	36	36	36	36	37	37	37	41	43	43	47	48	48	48	53	57	57
60	63	63	67	69	72	75	80	81	84	84	84	84	84	84	91	95	96	96	104
106	108	108	108	108	108	108	111	114	120	120	120	120	120	122	122	123	124	126	132
Censored observations (44 patients)																			
6	7	12	12	13	15	16	17	18	18	20	22	23	24	24	24	24	24	24	24	27	28
29	30	32	36	37	37	42	44	45	60	60	60	60	65	72	72	72	84	87	96	143	146

The Kaplan-Meier method was used to obtain estimates of the survival functions. The estimated value of the survival function on a cumulative 5-years survival was found to be 59.35%.

The Kaplan-Meier method is a method for estimating the survival function developed by Kaplan and Meier [3, 4, 5]. The Kaplan-Meier estimate is also known as the product-limit estimate of the survival function [3].

In this study, for the cases of known and unknown τ, the estimations were first obtained as follows. We have $n = 124$ and $d_3 = 80$. Considering τ as unknown, the values of the loglikelihood statistic is given in Table 2.

Table 2. The values of the loglikelihood statistic for τ

τ	Loglikelihood statistic values	τ	Loglikelihood statistic values
12	-605.0025	63	-470.4756
13	-600.2445	67	-470.0089
14	-596.5686	69	-469.8421
17	-587.5395	72	-468.9179
18	-583.6025	75	-471.0959
19	-576.8237	80	-471.9082
20	-572.7831	81	-476.2543
21	-568.7370	84	-478.6781
22	-564.6804	91	-505.5099
24	-556.5950	95	-507.5332
27	-547.7425	96	-512.2455
28	-543.7958	104	-515.7777
30	-538.9656	106	-519.3068
36	-524.6235	108	-522.7629
37	-505.8578	111	-556.9126
41	-494.0894	114	-560.4926
43	-489.7784	120	-559.8295
47	-481.2809	122	-591.3748
48	-477.7125	123	-603.4668
53	-473.6280	124	-609.1618
57	-471.7205	126	-614.5192
60	-471.4931	132	-619.2683

In Table 2, the value of τ which maximizes the loglikelihood is 72. Then, assuming that τ is unknown, the mode of the loglikelihood for τ is $\hat{\tau} = 72$. Since $\hat{\tau} = 72$ corresponds to age 44, this shows that the risk of breast cancer above the age of 44 is high.

The data has been analyzed to see if a change point is present. For this purpose the test statistic given in equation (5.1), was used. We tested the null hypothesis of no change, $H_0 : \tau = 0$, against the alternative hypothesis of a change, $H_1 : \tau = 72$. We used $\mathrm{L}(\lambda, \lambda, 0) = \lambda^n \exp(-\lambda \sum t_i)$ to obtain the maximum likelihood estimator of the hazard

rate in a simple exponential model. This gave $\widehat{\lambda} = 0.0176915$ for the maximum likelihood estimator of λ. We thus obtained $\mathrm{L}(\widehat{\lambda}, \widehat{\lambda}, 0) = -624.298925$.

Using equations (3.1) and (3.2) we obtained $\hat{\lambda}_1 = 0.00787953$ and $\widehat{\rho} = 3.4221024$ for the maximum likelihood estimators of λ_1 and ρ, respectively. Then, $\mathrm{L}(\widehat{\lambda}_1, \hat{\rho}, \hat{\tau}) = -424.420026$ was obtained from equation (2.4), giving $\Delta_0 = 199.878899$. The null hypothesis H_0 was rejected because $2\Delta_0 > \chi^2_{(2)}$. That is, it can be said that the data has a change point with a 0.95 confidence level.

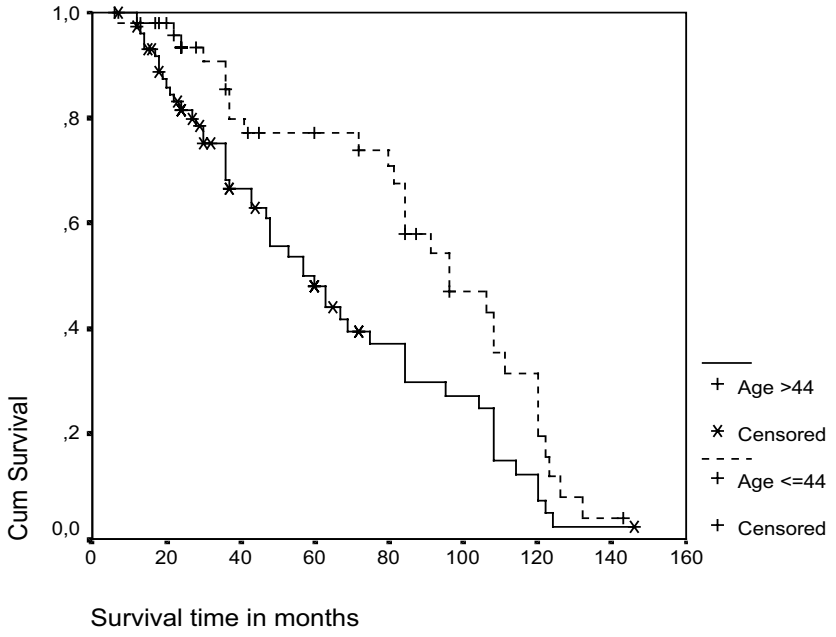
Assuming τ is known ($\tau = 72$), we have $d_1(\tau) = 45$, $d_2(\tau) = 81$, $W_1(\tau) = 2615$ and $W_2(\tau) = 4394$. Then, $\hat{\lambda} = 0.00787953$ and $\hat{\rho} = 3.4221024$ are found from equations (3.1) and (3.2). That is, the hazard rate is 0.00787953 until survival time 72 (age 44) and 0.0269646 after survival time 72 (age 44), with respect to the model (1.1).

Considering $\tau = 72$ and $\lambda = 0.00787953$ to be known and using equation (4.3), the estimation of the mode of posterior density for ρ is $\hat{\rho} = 0.991721$. That is, assuming the hazard rate is 0.00787953 until survival time 72, this rate is 0.0078143 after survival time 72 with respect to the model (1.1) using Bayesian analysis.

Assuming only that $\tau = 72$ is known, the mode of posterior density for ρ (from (4.5)) is $\hat{\rho} = 3.25206$. The mode of posterior density for λ (from (4.6)) is $\hat{\lambda} = 0.0077$. Then, with respect to the model (1.1), the hazard rate is 0.0077 until survival time 72 (age 44) and 0.025041 after survival time 72.

Finally, the ages of the patients were then separated into two groups according to the change point at 72 months. The Kaplan-Meier method was used to find estimates of the survival functions of each group. For the first group, the estimation value of the survival functions on a cumulative 5-year survival was found to be 77.09%. The other was 48.13%.

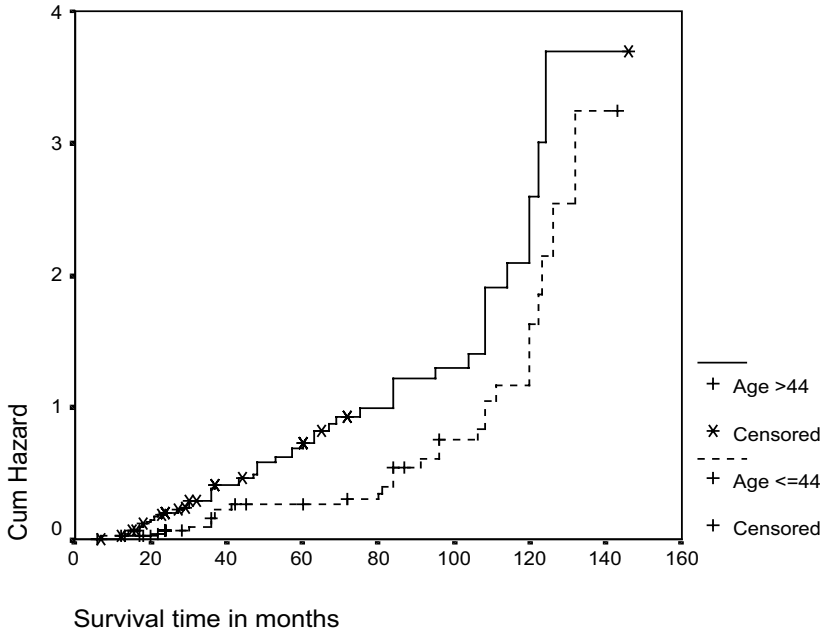
Figure 1. Kaplan-Meier estimates of the survival functions of the two groups defined by the estimated change point.



The significance of the observed difference between the groups was calculated by the log-rank test and its significance value found to be $p = 0.0082$.

The Kaplan-Meier estimates of the survival functions and hazard functions of the groups defined by the estimated change point are presented in Figure 1 and Figure 2, respectively.

Figure 2. Hazard functions of the two groups defined by the estimated change point.



7. Conclusion

In this paper, the constant hazard model with change point, the probability density function, survival function and likelihood function are considered. The use of the Bayesian approach and the likelihood ratio to test the hypothesis of a change point are considered. Also a Bayesian approach to the parameters and the use of the likelihood ratio test to examine the presence of a change point in the model are analyzed.

The likelihood ratio test, when applied to the data for 124 breast cancer patients in the Oncology Department of Hacettepe University Hospital, shows the presence of a change point in the data. Also, we obtained that the risk of breast cancer increases over the age of 44.

References

[1] Achcar, J. A. and Bolfarine, H. *Constant hazard against a change-point alternative: A Bayesian approach with censored data*, Commun. Statist.-Theory Meth. **18** (10), 3801–3819, 1989.

[2] Achcar, J. A. and Loibel, S. *Constant hazard function models with a change point: A Bayesian analysis using markov chain monte carlo methods*, Biometrical Journal **40** (5), 543–555, 1998.

[3] Collett, D. *Modelling survival data in medical research*, (Chapman and Hall, 1994).

[4] Kalbfleisch, J. D. and Prentice, R. L. *The statistical analysis of failure time data*, (John Wiley and Sons, 1980).
[5] Lee, E. T. *Statistical methods for survival data analysis*, (Lifetime Learning Publications, 1980).
[6] Loader, C. R. *Inference for a hazard rate change point*, Biometrika **78** (4), 749–757, 1991.
[7] Matthews, D. E. and Farewell, V. T. *On testing for a constant hazard against a change-point alternative*, Biometrics **38**, 463–468, 1982.
[8] Matthews, D. E. and Farewell, V. T. *On a singularity in the likelihood for a change-point hazard rate model*, Biometrika **72** (3), 703–704, 1985.
[9] Matthews, D. E. and Farewell, V. T. and Pyke, R. *Asymptotic score-statistic processes and tests for constant hazard against a change-point alternative*, The Annals of Statistics **13** (2), 583–591, 1985.
[10] Nguyen, H. T., Rogers, G. S. and Walker, E. A. *Estimation in change-point hazard rate models*, Biometrika **71** (2), 299–304, 1984.
[11] Worsley, K. J. *Confidence regions and tests for a change-point in a sequence of exponential family random variables*, Biometrika **73** (1), 91–104, 1986.
[12] Worsley, K. J. *Exact percentage points of the likelihood-ratio test for a change-point hazard-rate model*, Biometrics **44**, 259–263, 1988.
[13] Yao, Y. C. *Maximum likelihood estimation in hazard rate models with a change-point*, Commun. Statist.-Theory Meth. **15** (8), 2455–2466, 1986.