Confidence Interval for Change Point in Hazard Rate With Staggered Entry

Dong-yun Kim[1], Yanhong Wu[2]

[1] National Institutes of Health,Bethesda, Maryland 20892, USA

[2] Department of Mathematics, California State University Stanislaus, Turlock, CA, USA

Correspondence: Yanhong Wu, Department of Mathematics, California State University Stanislaus, Turlock, CA 95382, USA.

Abstract

We consider the construction of a con?dence region (interval) for a change point in hazard rate of the patients survival distribution when the patients enter the trial at random times. We show that the local- likelihood ratio process converges weakly to a certain process and obtain the maximum distribution of the process which does not depend on the change point, and thus can be used to construct the confidence region for the change point. We also compare the limiting density function to the empirical density and discuss the empirical coverage probability of the confidence interval by simulation. Stanford Heart Transplant data are used for illustration.

Keywords: change point, confidence region, hazard rate, likelihood ratio process, staggered entry

1. Introduction

Assume $Y_1, \cdots, Y_N$ are the survival times for the N patients that follows the density

$$
f_Y(y) = \begin{cases} \lambda_1 e^{-\lambda_1 y} & if \quad y < \nu \\ \lambda_2 e^{-\lambda_1 \nu - \lambda_2(y-\nu)} & if \quad y \geq \nu, \end{cases} \tag{1}
$$

where $0 < \lambda_1 < \lambda_2$, and ν is the change point. This change point problem in hazard rate was first proposed in Miller (1960). We refer to Müller and Wang(1994) for a review and Chen and Gupta (2011) for an introduction.

For the estimation problem under model (1), Nguyen et al. (1984), Yao (1986), Pham & Nguyen (1990), and Gosh & Joshi (1992) studied the MLE of ν and the limiting distribution of MLE. Chang,et al. (1994) considered the case with random censoring. Comparison between unconditonal and conditional estimations are considered by Gürler & Yenigün (2011). For the Bayes estimation, we referred to Achcar & Bolfarine(1989), Basu, et al.(1988), Gosh, et al.(1996), Groenwald & Shoeman (2001), Gosh & Ebrahimi (2008), and Karasoy & Kadilar (2007, 2009).

The corresponding testing problems have been considered by several approaches. For the likelihood ratio approach, we referred to Matthews & Farewell (1982), Worsley (1988), Henderson (1990), Loader (1991), and Luo, Turnbull et al. (1997). In particular, Loader (1991) also considered the random censoring case. A score-based test was considered by Mattews et al. (1985). A Bayes test was considered by Yao (1987).

In Kim et al. (2004), the testing problem is studied by showing the weak convergence of the log-likelihood ratio process under $C(-\infty, \infty)$. The same technique was extended to Weibull case in Williams & Kim (2011,2013a, 2013b).

For construction of confidence region or interval, one can use the MLE. Siegmund (1988) proposed to invert the log-likelihood ratio process to obtain confidence region and then construct the confidence interval by joining all the disconnected intervals and concluded that there is practically no difference between the confidence region and confidence interval. By using this approach, Loader (1991) obtained the confidence region by extending the method under model (1) by assuming the exponential random censoring mechanism.

In this article, we extend the model to the staggered entry case with type I censor. Assume that the patients arrive for treatment at times $0 < \tau_1 < \tau_2 < \cdots$, following a Poisson process with known rate γ. Let N be the total number of patients who arrived in the time interval $[0, T]$ and for $i = 1, ..., N$. The observations are

$$
\begin{cases} X_i = & \min(Y_i, T - \tau_i), \\ \delta_i = & I\{Y_i \leq T - \tau_i\}, \end{cases} \tag{2}
$$

where $I\{A\}$ denotes the indicator function of an event A. The log-likelihood function is

$$l(\lambda_1, \lambda_2, \nu) = K_1(\nu)log(\lambda_1) - \lambda_1 T_1(\nu) + K_2(\nu)log(\lambda_2) - \lambda_2 T_2(\nu),$$

where

$$K_1(\nu) = \sum_{i=1}^{N} \mathbf{1}\{X_i < \nu, \ \delta_i = 1\}, \quad K_2(\nu) = \sum_{i=1}^{N} \mathbf{1}\{X_i \geq \nu, \ \delta_i = 1\},$$

$$T_1(\nu) = \sum_{i=1}^{N} (X_i \wedge \nu), \quad T_2(\nu) = \sum_{i=1}^{N} (X_i - \nu)^+,$$

$x \wedge \nu = \min(x, \nu)$, and $x^+ = \max(x, 0)$. Given ν, the maximum likelihood estimator (MLE) of $\lambda_i(\nu)$ is

$$\hat{\lambda}_i = K_i(\nu)/T_i(\nu)$$

for $i = 1, 2$, respectively. So the profile log-likelihood is

$$
\begin{aligned}
\ell(\nu) &= \ell(\hat{\lambda}_1(\nu), \hat{\lambda}_2(\nu), \nu) \\
&= K_1(\nu)\left\{\log\left(\frac{K_1(\nu)}{T_1(\nu)}\right) - 1\right\} + K_2(\nu)\left\{\log\left(\frac{K_2(\nu)}{T_2(\nu)}\right) - 1\right\}.
\end{aligned}
$$

The MLE of ν is obtained by maximizing $\ell(\nu)$ with respect to ν. This must be done numerically.

The main contribution here is to give a systematic method of constructing confidence region/interval for ν by inverting the likelihood ratio process. For this, we need to prove the weak convergence of the local log-likelihood ratio process with a change point

$$Z_\gamma(u) = \ell\left(\nu + \frac{u}{\gamma T}\right) - \ell(\nu)$$

on $D(-\infty, \infty)$ and study the distribution of the maximum of log-likelihood ratio process

$$\ell(\hat{\nu}) - \ell(\nu) = \sup_{-\infty < u < \infty} Z_\gamma(u).$$

We obtain its limit distribution that is free of the change point ν. Thus, we can construct the confidence region of the form $\{\nu | \ell(\nu) \geq \ell(\hat{\nu}) - c\}$ by selecting c satisfying the designated coverage probability.

In Section 3, we use Monte Carlo simulation to compare the limiting density function to the empirical density function constructed from the simulated data. We also discuss how to construct the confidence interval and comment on the empirical coverage probability of the confidence interval with nominal confidence level. In Section 4, we use the Stanford Heart Transplant data to illustrate our method. We show that the data set fits the model assumptions quite well.

2. Main Results

In notation, let $\to^D$, $\to^p$, and $\Rightarrow$ denote the convergence in distribution, convergence in probability, and the weak convergence, respectively. $D[a, b]$ refers to the Skorohod space consisting of all functions on $[a, b]$ that are right-continuous on $[a, b)$ and possess left-hand limits at each point of $(a, b]$.

For a fixed ν, define a local likelihood ratio process Z_γ by

$$Z_\gamma(u) = I_1 + I_2,$$

where for $i = 1, 2$,

$$I_i = K_i\left(\nu + \frac{u}{\gamma T}\right)\left\{\log\left(\frac{K_i(\nu + \frac{u}{\gamma T})}{T_i(\nu + \frac{u}{\gamma T})}\right) - 1\right\} - K_i(\nu)\left\{\log\left(\frac{K_i(\nu)}{T_i(\nu)}\right) - 1\right\}.$$

We shall split the proof of weak convergence of $Z_\gamma(u)$ into several steps.

First, by denoting

$$\tilde{K}_i(u, \nu) = K_i\left(\nu + \frac{u}{\gamma T}\right) - K_i(\nu) \tag{3}$$

and

$$\tilde{T}_i(u, \nu) = T_i\left(\nu + \frac{u}{\gamma T}\right) - T_i(\nu). \tag{4}$$

and applying a two-term Taylor expansion, we can rewrite

$$I_i = \tilde{I}_i + R_i(\nu).$$

where

$$\tilde{I}_i = \log(\hat{\lambda}_i(\nu))\tilde{K}_i(u, \nu) - \hat{\lambda}_i(\nu)\tilde{T}_i(u, \nu).$$

Similar to the proof for Proposition 1 in Kim, Woodroofe, and Wu (2002), we can show

$$\sup_{|\nu| \le c} |R_i(\nu)| \to^p 0$$

as $\gamma \to \infty$ for any c.

Second, we show the weak convergence of $\tilde{K}_1(u, \nu)$ given in (3). Note that given $N = n$, the random vectors $(X_1, \delta_1), \ldots, (X_n, \delta_n)$ are a permutation of i.i.d. random vectors with a common distribution function $F_X(x) = P(X \le x)$ and sub-distribution function $\tilde{F}_X(x) = P(X \le x, \delta = 1)$ such that

$$F_X(x) = \begin{cases} 1 - e^{-\lambda_1 x}(1 - x/T) & if \quad 0 < x < \nu; \\ 1 - e^{-\lambda_1 \nu - \lambda_2(x-\nu)}(1 - x/T) & if \quad \nu \le x < T \end{cases}$$

and

$$\tilde{F}_X(x) = \begin{cases} \tilde{F}_X^1(x) & if \quad 0 < x < \nu; \\ \tilde{F}_X^2(x) & if \quad \nu \le x < T, \end{cases}$$

where

$$\tilde{F}_X^1(x) = \left(1 - \frac{1}{\lambda_1 T}\right) - e^{-\lambda_1 x}\left(1 - \frac{1}{\lambda_1 T} - \frac{x}{T}\right)$$

and

$$\tilde{F}_X^2(x) = \left(1 - \frac{1}{\lambda_1 T}\right) + e^{-\lambda_1 \nu}\left(\frac{1}{\lambda_1 T} - \frac{1}{\lambda_2 T}\right) - e^{-\lambda_1 \nu - \lambda_2(x-\nu)}\left(1 - \frac{1}{\lambda_2 T} - \frac{x}{T}\right).$$

By denoting $V^+(t)$, $V^-(t)$ as two independent copies of a random process such that

$$V^+(t) = \begin{cases} 0 & if \quad t < 0; \\ \text{Poisson process with rate } t & if \quad t > 0, \end{cases}$$

$$V^-(t) = \begin{cases} \text{Poisson process with rate } -t, & if \quad t > 0; \\ 0, & if \quad t < 0. \end{cases}$$

we have

Lemma 1. *For each fixed ν, $\tilde{K}_1(u, \nu)$ as in (3) converges weakly to a process of the form*

$$K_1^*(u, \nu) := \text{sgn}(u)\{V^+(\lambda_2 \beta u) + V^-(-\lambda_1 \beta u)\}$$

in $D[a, b]$ as $\gamma \to \infty$ for each $-\infty < a < b < \infty$, where $\text{sgn}(x)$ is the signum function, and

$$\beta = \left(1 - \frac{\nu}{T}\right)e^{-\lambda_1 \nu}. \tag{5}$$

Proof. For each $u > 0$, by using Taylor expansion and (3),

$$E(\tilde{K}_1(u, \nu)) \approx \lambda_2 \beta u$$

where β is as in (5). By the Poisson Convergence Theorem,

$$\tilde{K}_1(u, \nu) \to^D V^+(\lambda_2 \beta u).$$

For $u < 0$,

$$\tilde{K}_1(u, \nu) \to^D \text{sgn}(u)V^-(-\lambda_1 \beta u)$$

following essentially the same line as for $u > 0$ case. Therefore, $\tilde{K}_1(u, v) \to^D K_1^*(u, v)$ for each u. The convergence of all finite dimensional distributions may be similarly established. Finally, by Theorem 15.6 in Billingsley(1968), $\tilde{K}_1(u, v)$ is tight. $\square$

Third, we show the convergence of $\tilde{T}_i(u, v)$ as defined in (4). Let $F_N(x)$ denote the empirical distribution function of $F_X(x)$, i.e.,

$$F_N(x) = \frac{1}{N} \sum_{k=1}^{N} \mathbf{1}\{X_k \leq x\}.$$

Lemma 2. *With β as in (5), for each fixed v and any positive number c,*

$$\sup_{|u| \leq c} |\tilde{T}_i(u, v) - (-1)^{i+1} u\beta| \to^p 0$$

as $\gamma \to \infty$.

Proof. Since $\tilde{T}_2(u, v) = -\tilde{T}_1(u, v)$, we only need to show that $\tilde{T}_1(u, v) \to u\beta$ uniformly. Since

$$\sum_{i=1}^{n} (X_i \wedge v) = n \int_0^v (1 - F_n(x)) dx,$$

we can write

$$\begin{aligned}
\tilde{T}_1(u, v) &= n \int_v^{v + \frac{u}{\gamma T}} (1 - F_n(x)) \, dx \\
&= n \int_v^{v + \frac{u}{\gamma T}} (F_X(x) - F_n(x)) dx + n \int_v^{v + \frac{u}{\gamma T}} (F_X(v) - F_X(x)) dx + \frac{nu}{\gamma T}(1 - F_X(v)).
\end{aligned}$$

By uniform convergence of $F_n(x)$ and $E(N/(\gamma T)) = 1$, $\mathrm{Var}(N/(\gamma T)) \to 0$, and by continuity of F_X at v, $\tilde{T}_1(u, v)$ converges to $u\beta$ uniformly. $\square$

Combining Lemmas 1 and 2 and using Slutsky's theorem, we thus have

Theorem 1. *$Z_\gamma(u)$ and $Z_\gamma^*(u)$ weakly converge to a process*

$$Y(u) = \log\left(\frac{\lambda_1}{\lambda_2}\right) K_1^*(u, v) - (\lambda_1 - \lambda_2)\beta u \tag{6}$$

as $\gamma \to \infty$, in $D[a, b]$ for any $-\infty < a < b < \infty$ where $K_1^*(u, v)$ is given in (4) and β in (5).

The next theorem gives the limiting distribution of the supremum of $Y(u)$ process and its proof is given in Section 5.

Theorem 2. *As $x \to \infty$, the distribution function of the supremum of $Y(u)$ process can be approximated by*

$$P\left(\sup_{-\infty < u < \infty} Y(u) \leq x\right) = (1 - e^{-x})(1 - e^{-x} \kappa(\lambda_1/\lambda_2))(1 + o(1)) \tag{7}$$

where

$$\kappa(x) = \left(\frac{1 - x + x \log x}{x - 1 - \log x}\right)^{-\mathrm{sgn}(\log x)}. \tag{8}$$

Remark. The results are given for $\lambda_1 < \lambda_2$. For $\lambda_1 > \lambda_2$, similar results hold with only minor modifications in a few places. $f_Y(y)$ in (1), $F_X(x)$, and $\tilde{F}_X(x)$ are replaced by their left-continuous versions. $K_1(v)$ and $K_2(v)$ are replaced by

$$K_1(v) = \sum_{i=1}^{N} \mathbf{1}\{X_i \leq v, \ \delta_i = 1\}, \quad K_2(v) = \sum_{i=1}^{N} \mathbf{1}\{X_i > v, \ \delta_i = 1\}.$$

3. Simulation Study

As we have seen in Theorem 1, the local likelihood ratio process $Z_\gamma(u)$ converges weakly to a process $Y(u)$ whose supremum has distribution function that does not depend on the change point v; it depends only on the ratio of λ_1 and λ_2. In this section we compare the theoretical density and empirical density curves generated from $\sup Z_\gamma(u)$, which in turn are

based on data sets from (1) using three different v values. Then we describe how to construct confidence interval for v and show actual coverage probability of the confidence intervals for various combination of change points and nominal coverage level.

With T, λ_1, λ_2, and γ fixed, the data and density functions are generated in the following steps.

Step 1. Draw a random number N from Poisson (γT).

Step 2. Fix v and draw a random sample $Y_1, \ldots, Y_N$ from $f_Y(y)$ as in (1).

Step 3. Draw a random sample $U_1, \ldots, U_N$ from the Unif (0, 1).

Step 4. For $i = 1, \ldots, N$, obtain X_i and δ_i as in (2).

Step 5. Calculate $\ell(\hat{v}) - \ell(v)$ with the MLE $\hat{v}$.

Step 6. Repeat steps 1 through 5 many times.

For example, we set $T = 1, \lambda_1 = 1, \lambda_2 = 2$, and $\gamma = 5,000$. Limiting density in Figure 1 was directly calculated from the probability approximation formula (7).

For comparison, we generated 10,000 random samples for each of the change points $v = 0.3, 0.5$, and 0.7 and obtained the density curve of $\ell(\hat{v}) - \ell(v)$. They are shown together with the (approximated) limiting density function. The plot shows that the densities are quite similar in shape and all three are close to the theoretical density function. Besides, Kolmogorov-Smirnov goodness-of-fit tests indicate that the densities are not significantly different at 5%.

To construct a confidence interval for v from a given random sample, fix a suitable confidence level $1 - \alpha$ and find the $100(1 - \alpha)$th percentile from (7). Then follow Steps 1 through 4 described above for a fixed v to generate a random sample of (average) size 5,000 and do:

Step 5′. Cover a sufficiently large interval ($\hat{v} \pm 0.2$, for example) containing the MLE $\hat{v}$ with a fine set of grid points $v_1, \ldots, v_J$ and for each $j = 1, \ldots, J$, calculate $\ell_j = \ell(\hat{v}) - \ell(v_j)$.

Step 6′. Find $J_L = \inf_{1 \le j \le J}\{j : \ell_j \ge p_{1-\alpha} \text{ and } \ell_{j+1} < p_{1-\alpha}\}$ and $J_R = \sup_{1 \le j \le J}\{j : \ell_j < p_{1-\alpha} \text{ and } \ell_{j+1} \ge p_{1-\alpha}\}$. Then find C_L by linearly interpolating two points (v_{J_L}, ℓ_{J_L}) and $(v_{J_L+1}, \ell_{J_L+1})$ and finding the point crossing the horizontal line at $p_{1-\alpha}$. Do the same for C_R using (v_{J_R}, ℓ_{J_R}) and $(v_{J_R+1}, \ell_{J_R+1})$. Report (C_L, C_R) as the $100(1 - \alpha)\%$ confidence interval.

Table 1 shows the empirical coverage probabilities for 90%, 95%, and 99% confidence intervals based on 10,000 replications for each of change points ranging from 0.4 to 0.7. Standard errors for 90%:0.003; 95%:0.002; 99%:0.001. Apparently the coverage probabilities for 90% confidence intervals are all higher than the nominal level; in particular, the 90% interval covers the true change point 92.3% of the time when $v = 0.4$ and 91.9% when $v = 0.45$. The excessive coverage probability of 90% confidence interval is partially due to the approximation error caused by Theorem 2 in that (7) holds when x is large. Also, the jagged nature of the log-likelihood function seems to contribute to the excessive coverage.

Table 1. Coverage Probability for 90%, 95%, and 99% Confidence

nominal/v	0.40	0.45	0.50	0.55	0.60	0.65	0.70
0.90	0.923	0.919	0.917	0.913	0.916	0.912	0.909
0.95	0.954	0.955	0.956	0.956	0.956	0.956	0.956
0.99	0.989	0.989	0.992	0.992	0.991	0.993	0.992

4. Example: Stanford Heart Transplant Data

To illustrate the results that we developed in this paper, we consider the heart transplant data as reported in Crowley & Hu(1977). Out of 103 patients who participated in the Stanford Heart Transplantation Program, we consider the 69 patients who received heart transplantation. Of those, 45 patients died between the operation and the closing date on April 1, 1974, and 24 patients were alive by the closing date so their lifetimes are Type I censored. The published data contain the calendar dates for the transplant operation and dates of death as well as censoring indicator.

Recall that two underlying assumptions for our method are: (i) that patients enter the treatment following a Poisson process, and (ii) the patients' survival time after the treatment (subject to Type I censoring) are distributed as in (2). We regard the surgery dates as the treatment times, and the first assumption can be easily checked by comparing the distribution of inter-arrival times to the exponential distribution. The usual Kolmogorov-Smirnov test results in a p-value of 0.65 so the first assumption seems reasonable. Also, the heart transplant data seem to fit the second assumption equally
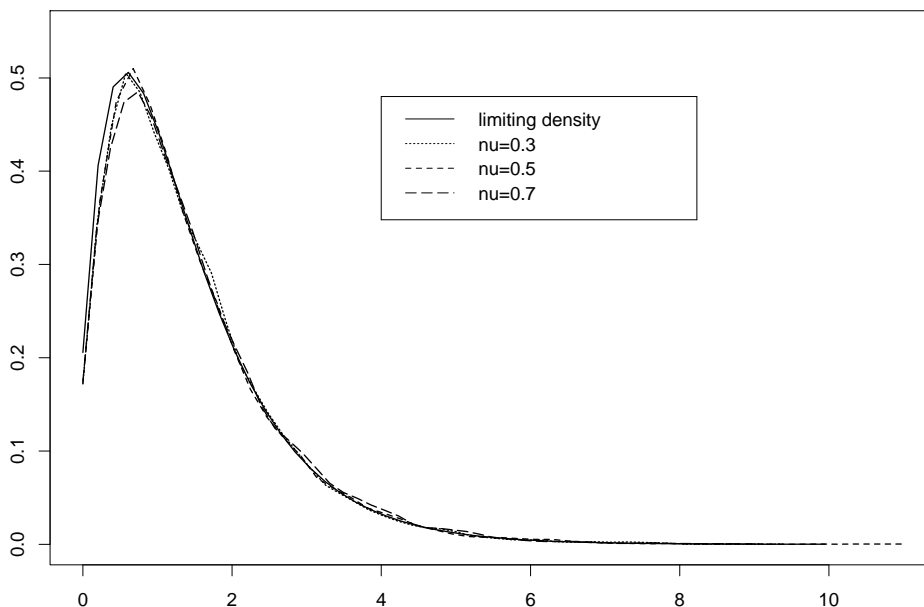
Figure 1. Simulated density function

well, with a p-value of 0.9 from the Kolmogorov-Smirnov test.

The empirical distribution function in Figure 2 strongly indicates the presence of a change point in the hazard at an early stage after the surgery. Not surprisingly the risk of death is quite high initially (steep slope at the beginning of the curve), then it quickly diminishes. From the data it is estimated that the change occurs approximately at 68 days, and the hazard rates before and after the change point are estimated to be $\hat{\lambda}_1 = 0.0075$, $\hat{\lambda}_2 = 0.00075$, which correspond to 132 days and 1320 days, respectively.

We can also formally test for a change point. Following the test procedure described in Kim et al. (2002), we first rescale the whole observed interval to [0,1] and the rescaled change point is 0.0298. The test statistic depends on the interval that is being tested, and for a particular interval [0.001, 0.03], the observed test statistic 6.622 is greater than the critical point 3.520, indicating that the test is significant at 1%. Notably, practically any subinterval containing the estimated change point is found to be highly significant.

To compute the confidence interval for the change point, we need the critical value of the distribution of the supremum of $Y(u)$. From Theorems 1 and 2, we see that

$$P\{\ell(\hat{\nu}) - \ell(\nu) \leq c\} \approx (1 - e^{-c})(1 - \kappa(\lambda_1/\lambda_2)e^{-c})$$

and for a given α, we can select the critical point c from the equation

$$(1 - e^{-c})(1 - \kappa(\lambda_1/\lambda_2)e^{-c}) = 1 - \alpha.$$

Since the upper 5% critical point is 3.37, by inverting $\ell(\hat{\nu}) - \ell(\nu)$ and reading off the two end points of the region that fall below the critical value (see Figure 3), we obtain a 95% confidence interval (65.5, 82.5) in days. Using the same method, 90% and 99% confidence intervals are computed as (65, 80), and (64, 90), respectively.
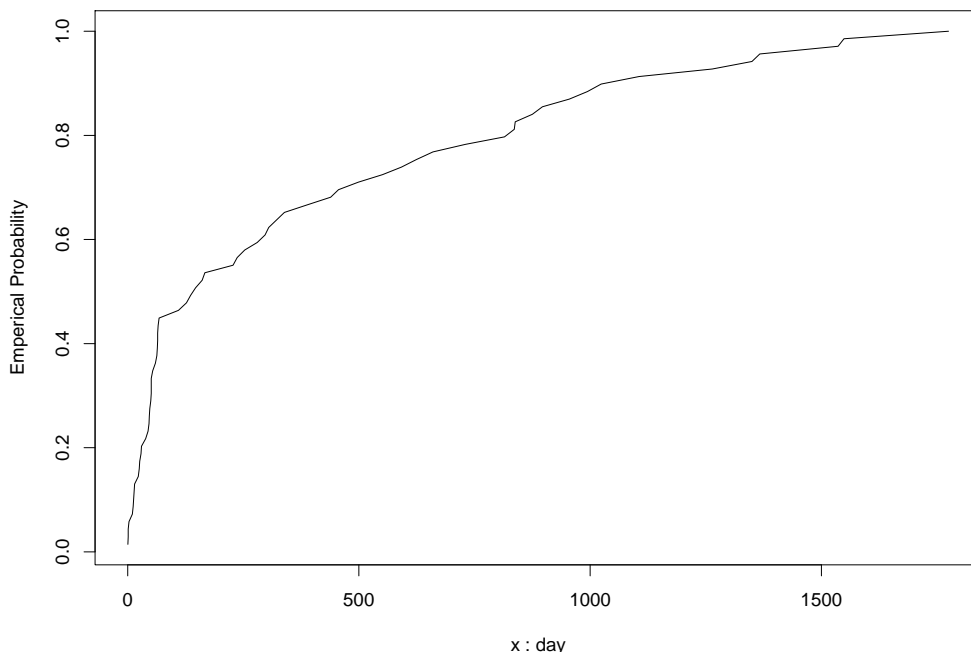
Figure 2. Empirical distribution for Stanford heart transplant data

5. Proof of Theorem 2

Suppose $\lambda_1 < \lambda_2$. Let $u^* = u(1 - v/T)$. Since $v < T$, $\mathrm{sgn}(u) = \mathrm{sgn}(u^*)$ and essential properties remain unchanged under the simple transformation, with abuse of notation we may write

$$Y(u) = \log\left(\frac{\lambda_1}{\lambda_2}\right)\left[\mathrm{sgn}(u)\left\{V^+\left(u\lambda_2 e^{-\lambda_1 v}\right) + V^-\left(-u\lambda_1 e^{-\lambda_1 v}\right)\right\}\right] + u\left(\lambda_2 - \lambda_1\right)e^{-\lambda_1 v}.$$

First consider the case $u \geq 0$. For each positive integer i, let W_i denote the "interarrival time", or the uncensored lifetime of patients associated with the Poisson process $K_1^*(u, v)$ as in (4). Then W_i is exponentially distributed with parameter $\lambda_2 e^{-\lambda_1 v}$ and

$$X_i = \lambda_2 e^{-\lambda_1 v} W_i \sim \mathrm{Exp}(1).$$

Define $A_1 = \{u : 0 \leq u < W_1\}$ and

$$A_k = \{u : \sum_{i=1}^{k-1} W_i \leq u < \sum_{i=1}^{k} W_i\}$$

for $k \geq 2$. Note that for $u \in A_k$,

$$Y(u) = u(\lambda_2 - \lambda_1)e^{-\lambda_1 v} + (k - 1)\log(\lambda_1/\lambda_2),$$

so

$$\sup_{u \in A_k} Y(u) = (\lambda_2 - \lambda_1)e^{-\lambda_1 v} \sum_{i=1}^{k} W_i + (k - 1)\log\left(\frac{\lambda_1}{\lambda_2}\right).$$

Therefore, by denoting $\delta = \lambda_2/\lambda_1$, we have

$$U_+ := \sup_{u \geq 0} Y(u) = \sup_{k \geq 1}\left\{\left(1 - \frac{1}{\delta}\right)\sum_{i=1}^{k} X_i + (k - 1)\left(\frac{1}{\delta}\right)\right\}.$$
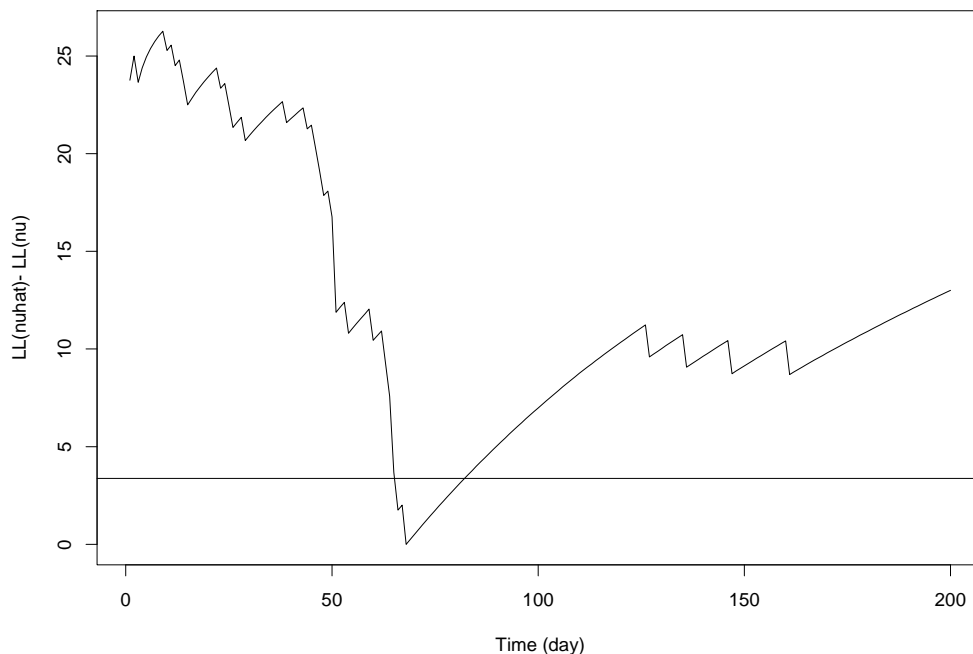
Figure 3. Confidence interval from local likelihood ratio process

Similarly, for $u < 0$,

$$U_- := \sup_{u<0} Y(u) = \sup_{k>1} (1 - \delta) \sum_{i=1}^{k-1} X'_i - (k - 1) \log\left(\frac{1}{\delta}\right),$$

where X'_i are independent copies of X_i. Thus,

$$\sup_u Y(u) = \max\{U_+, U_-\}.$$

To find the distribution function of U_-, Let

$$S_n = \sum_{i=1}^n \left\{(1 - \delta) X'_i + \log \delta\right\}.$$

Denote

$$\tau_+ = \inf\{n : n \geq 1, S_n > 0\}, \quad \tau_- = \inf\{n : n \geq 1, S_n \leq 0\},$$

and for $x > 0$,

$$\tau_x = \inf\{n : S_n > x\},$$

$R_x = S_{\tau_x} - x$ as the overshoot, and $R_\infty = \lim_{x \to \infty} R_x$. Also, let $P_\xi(.)$ denote the probability measure when the failure rate of X_i is ξ. By changing measure from $P_1(.0$ to $P_\delta(.)$, the renewal theorem implies that as $x \to \infty$,

$$
\begin{aligned}
P_1(U_- < x) &= P_1(\tau_x = \infty) \\
&= 1 - E_\delta e^{-S_{\tau_x}} \\
&\approx 1 - e^{-x} E_\delta e^{-R_\infty},
\end{aligned}
$$

where

$$P_\delta(R_\infty < y) = \frac{\int_0^y P_\delta(S_{\tau_+} > x)\, dx}{E_\delta S_{\tau_+}}.$$

Write

$$
\begin{aligned}
\kappa(\lambda_1/\lambda_2) &= E_\delta e^{-R_\infty} = \frac{1}{E_\delta S_{\tau_+}} \int_0^\infty e^{-y} P_\delta\left(S_{\tau_+} > y\right) dy \\
&= \frac{1 - E_\delta e^{-S_{\tau_+}}}{\left(\delta^{-1} - 1 + \log \delta\right) E_\delta \tau_+},
\end{aligned}
$$

by noting $E_\delta S_{\tau_+} = (\delta^{-1} - 1 + \log \delta) E_\delta \tau_+$. By changing measure and Corollary 8.39 in Siegmund(1985),

$$
E_\delta e^{-S_{\tau_+}} = P_1\left(\tau_+ < \infty\right),
$$

$$
E_\delta \tau_+ = 1/P_\delta\left(\tau_- = \infty\right).
$$

we have

$$
\kappa(\lambda_1/\lambda_2) = \frac{(1 - p^+)(1 - p^-)}{\left(\delta^{-1} - 1 + \log \delta\right)}
$$

where

$$
p^+ = P_1\left(\tau_+ < \infty\right), \quad p^- = P_\delta\left(\tau_- < \infty\right).
$$

By Theorem 24.5 in Billingsley(1986),

$$
p^+ = P\left(\sup_{n \geq 0} S_n > 0\right) = \frac{\log \delta}{\delta - 1}.
$$

On the other hand, by writing $\tau_+' = \inf\{n > 0 : -S_n > 0\}$, we have

$$
p^- = P_\delta\left(\tau_+' < \infty\right) = E_1 e^{-S_{\tau_+}} = E_1 e^{S_{\tau_-}} = \int_0^\infty e^{-(\delta-1)x} e^{-x} dx = 1/\delta.
$$

Finally, we get

$$
P_1(U_- < x) \approx 1 - \kappa(\lambda_1/\lambda_2) e^{-x}
$$

where $\kappa(x)$ is as in (8).

To find the distribution function of U_+, denote

$$
\tau_x^* = \inf\{n > 0 : S_n^+ > x\}
$$

where $S_n^+ = \sum_{i=1}^n \left\{\left(1 - \delta^{-1}\right) X_i - \log \delta\right\}$. Then

$$
\begin{aligned}
P_1\left(U_+ \leq x\right) &= P_1\left(\tau_{x-\log\delta}^* = \infty\right) \\
&= 1 - E_{\delta^{-1}} \exp\left(-S_{\tau_{x-\log\delta}}^+\right) \\
&= 1 - \exp\left(-x + \log\delta\right) E_{\delta^{-1}} \exp\left\{-\left(1 - \delta^{-1}\right) X_1\right\},
\end{aligned}
$$

by the memoryless property of the exponential distribution. Therefore,

$$
P_1\left(U_+ \leq x\right) = 1 - e^{-x}.
$$

By independence of U_+ and U_-,

$$
P\left(\sup_{-\infty < u < \infty} Y(u) \leq x\right) \approx \left(1 - e^{-x}\right)\left(1 - \kappa(\lambda_1/\lambda_2) e^{-x}\right).
$$

If on the other hand $\lambda_1 > \lambda_2$, U_+ is changed to

$$
U_+^* = \max\left\{\sup_{k \geq 1}\left(1 - \frac{\lambda_1}{\lambda_2}\right)\sum_{i=1}^{k-1} X_i + (k - 1)\log\left(\frac{\lambda_1}{\lambda_2}\right), 0\right\},
$$

which is similar to U_-. Applying the same method to U_+^*, we see

$$
P\left(U_+^* \leq x\right) = 1 - \kappa(\lambda_1/\lambda_2) e^{-x}
$$

where $\kappa(x)$ is in (8). On the other hand, U_- is changed to

$$U_-^* = \sup_{k \geq 1} \left(1 - \frac{\lambda_2}{\lambda_1} \right) \sum_{i=1}^{k} X_i' + (k - 1) \log \left(\frac{\lambda_1}{\lambda_2} \right)$$

and the approximated distribution of U_-^* is the same as the approximated distribution of U^+, i.e., $P(U_-^* \leq x) \approx 1 - e^{-x}$ so

$$P \left(\sup_{-\infty < u < \infty} Y(u) \leq x \right) \approx \left(1 - e^{-x} \right) \left(1 - \kappa \left(\lambda_1 / \lambda_2 \right) e^{-x} \right).$$

6. Conclusion

In this communication, we presented a method of constructing confidence region (interval) for a change point in hazard rate with staggered entry by using log-local likelihood ratio process. By using the weak convergence, we are able to calculate the limit distribution and use it to obtain the confidence region given confidence level. There are many possible extensions for staggered entry. We only mention some recent development. For estimations under regression models for hazard rate, we refer to Pons (2002), Na et al.(2005), Dupuy (2006), Goodman et al. (2011) and Li et al. (2013). For nonparametric estimations or under nonparametric models, we refer to Antoniadis et al. (2000), Wu et al.(2003) and Gijbels & Gürler (2003). Kosorok & Song (2007) considered a model where the change is caused by a covariate crosses a threshold. More general models under censoring mechanisms are considered in Zhao et al.(2009), Rabhi & Asgharian (2017), and Wang et al. (2019). Sequential tests for possible multiple change points are considered in He et al. (2012) and Zhang et al. (2014). Computing intensive techniques using Monte Carlo and Bootstrap technique are considered by Achcar & Loibel (1998), Pham & Nguyen (1990, 1993).

Acknowledgments

References

Achcar, J. A., & Bolfarine, H. (1989). Constant hazard against a change-point alternative: a Bayesian approach with censored data. *Communications in Statist: Theory and Methods*, *18*, 3801-3819. https://doi.org/10.1080/15598608.2008.10411892

Achcar, J. A., & Loibel, S. (1998). Constant hazard function models with a change point: A Bayesian analysis using markov chain Monte Carlo methods. *Biometrical Journal*, *40*, 543-555. http://dx.doi.org/10.1002/(SICI)1521-4036(199809)40:5

Antoniadis, A., Gijbels, I., & Macgibbon, B. (2000). Non-parametric estimation for the location of a change-point in an otherwise smooth hazard function under random censoring. *Scandinavian Journal of Statistics*, *27*(3), 501-519. http://dx.doi.org/10.1111/1467-9469.00203

Basu, A. P., Ghosh, J. K., & Joshi, S. N. (1988). On estimating change point in a failure rate. *Statistical Decision Theory and Related Topics* IV,2, Springer-Verlag, New York, 239-252. http://dx.doi.org/10.1007/978-1-4612-3818-8_19

Billingsley, P. (1968). *Convergence of Probability Measure.* Wiley, New York. http://dx.doi.org/10.1002/9780470316962

Billingsley, P. (1986). *Probability and Measure.* (2nd ed.). Wiley, New York.

Chang, I., Chen, C., & Hsiung, C. (1994). Estimation in change-point hazard rate models with random censorship. In: *Change-point Problems*. IMS Lecture Notes Monograph Series 23, Institute of Mathematical Statistics, Hayward, CA, 78-92. http://dx.doi.org/10.1214/lnms/1215463115

Chen, J., & Gupta, A. K. (2011). *Parametric Statistical Change Point Analysis*. Birkhauser, Boston. http://dx.doi.org/10.1007/978-1-4757-3131-6

Crowley, J., & Hu, M. (1977). Covariance analysis of heart transplant survival data. *Journal of American Statistical Association*, *72*, 27-36. http://dx.doi.org/10.1080/01621459.1977.10479903

Dupuy, J. F. (2006). Estimation in a change-point hazard regression model. *Statistics and Probability Letters*, *76*(2), 182-190. http://dx.doi.org/10.1016/j.spl.2005.07.013

Ghosh, S. K., & Ebrahimi, N. (2008). Bayesian Analysis of Change-Point Hazard Rate Problem. *Journal of Statistical Theory and Practice*, *2*(4), 523-533.

Ghosh, J. K., & Joshi, S. N. (1992). On the asymptotic distribution of an estimate of the change point in a failure rate. *Communications in Statistics: Theory and Methods*, *21*, 3571-3588. http://dx.doi.org/10.1080/03610929208830996

Ghosh, J. K., Joshi, S. N., & Mukhopadhyay, C. (1996). Asymptotics of a Bayesian approach to estimating change-point in a hazard rate. *Communications in Statistics: Theory and Methods*, *25*(12), 3147-3166. http://dx.doi.org/10.1080/03610929608831890

Gijbels, I., & Gürler, Ü. (2003). Estimation of a change point in a hazard function based on censored data. *Lifetime Data Analysis*, *9*, 395-411. http://dx.doi.org/10.1023/B:LIDA.0000012424.71723.9d

Goodman, M. S., Li, Y., & Tiwari, R. C. (2011). Detecting multiple change points in piecewise constant hazard functions. *Journal of Applied Statistics*, *38*(11), 2523-2532. http://dx.doi.org/10.1080/02664763.2011.559209

Gürler, Ü., & Yenigün, C. D. (2011). Full and conditional likelihood likelihood approaches for hazard change-point estimation with truncated and censored data. *Computational Statistics & Data Analysis*, *55*(10), 2856-2870.

Groenwald, P., & Schoeman, A. (2001). Bayesian estimation in change-point hazard rate models. *South African Statistics Journal*, *35*(1), 95-121.

He, P., Fang, L., & Su, Z. (2012). A sequential testing approach to detecting multiple change points in the proportional hazards model. *Statistics in Medicine*, *32*(7), 1239-1245. http://dx.doi.org/10.1002/sim.5605

Henderson, R. (1990). A problem with the likelihood ratio test for a change-point hazard rate model. *Biometrika*, *77*, 835-843. http://dx.doi.org/10.1093/biomet/77.4.835

Karasoy, D. S., & Kadilar, C. (2007). A new Bayes estimate of the change point in the hazard function. *Computational Statistics & Data Analysis*, *51*(6), 2993-3001. http://dx.doi.org/10.1016/j.csda.2006.07.007

Karasoy, D. S., & Kadilar, C. (2009). Modified estimators for the change point in hazard function. *Journal of Computational and Applied Mathematics*, *229*(1), 152-157. http://dx.doi.org/10.1016/j.cam.2008.10.019

Kosorok, M. R., & Song, R. (2007). Inference under right censoring for transformation models with a change-point based on a covariate threshold. *Annals of Statistics*, *35*, 957-989. http://dx.doi.org/10.1214/009053606000001244

Kim, D., Woodroofe, M., & Wu, Y. (2004). Testing for a change in the hazard rate with staggered entry. *Communications in Statistics: Theory and Methods*, *33*(9), 2041 - 2058. http://dx.doi.org/10.1081/STA-200026569

Li, Y., Qian, L., & Zhang, W. (2013). Estimation in a change-point hazard regression model with long-term survivors. *Statistics and Probability Letters*, *83*(7), 1683-1691. http://dx.doi.org/10.1016/j.spl.2013.03.026

Loader, C. R. (1991). Inference for a hazard rate change point. *Biometrika*, *78*, 749-757. http://dx.doi.org/10.1093/biomet/78.4.749

Luo, X., Turnbull, B. W., & Clark, L. C. (1997). Likelihood ratio tests for a changepoint with survival data. *Biometrika*, *84*(3), 555-565. http://dx.doi.org/10.1093/biomet/84.3.555

Matthews, D. E., & Farewell, V. T. (1982). On testing for a constant hazard against a change-point alternative. *Biometrics*, *38*, 463- 468. http://dx.doi.org/10.2307/2530460

Matthews, D. E., Farewell, V. T., & Pyke, R. (1985). Asymptotic score-statistic processes and tests for constant hazard against a change-point alternative. *Annals of Statistics*, *13*, 583 -591. http://dx.doi.org/10.1214/aos/1176349540

Miller, R. G. (1960). Early failure in life testing. *Journal of American Statistical Association*, *55*, 591-602. http://dx.doi.org/10.1080/01621459.1960.10482077

Müller, H. G., & Wang, J. L. (1994). Change-point models for hazard functions. In: *Change Point Problems*. IMS Lecture Estimation in change-point hazard rate models with random censorship. In: *Change-point Problems*. IMS Lecture Notes Monograph Series 23, Institute of Mathematical Statistics, Hayward, CA, 224-241. http://dx.doi.org/10.1214/lnms/1215463127

Na, M. H., Jeon, J., & Park, D. H. (2005). Testing whether failure rate changes its trend with unknown change points. *Journal of Statistical Planning and Inference*, *129*, 317-325.

Nguyen, H. T., Rogers, G. S., & Walker, E. A. (1984). Estimation in change-point hazard rate models. *Biometrika*, *71*, 299- 304. http://dx.doi.org/10.1093/biomet/71.2.299

Pham, D. T., & Nguyen, H. T. (1990). Strong consistency of the maximum likelihood estimator in the change-point hazard rate model. *Statistics*, *21*, 203-216.

Pham, D. T., & Nguyen, H. T. (1993). Bootstrapping the change-point of a hazard-rate. *Annals of Institute of Statistical Mathematics*, *45*, 331-340. http://dx.doi.org/10.1007/BF00775818

Pons, O. (2002). Estimation in a Cox regression model with a change-point at an unknown time. *Statistics*, *36*, 101-124. http://dx.doi.org/10.1080/02331880212043

Rabhi, Y., & Asgharian, M. (2017). Inference under biased sampling and right censoring for a change point in the hazard function. *Bernoulli*, *23*, 2720-2745. http://dx.doi.org/10.3150/16-BEJ825

Siegmund, D. (1985). *Sequential Analysis*. New York, Springer. http://dx.doi.org/10.1007/978-1-4757-1862-1

Wang, B., Wang, X., & Song, L. (2019). Estimation in the single change-point hazard function for interval-censored data with a cure fraction. *Journal of Applied Statistics*, *47*(2), 1-17. http://dx.doi.org/10.1080/02664763.2019.1635571

Williams, M. R., & Kim, D. (2011). Likelihood ratio tests for continuous monotone hazards with an unknown change point. *Statistics and Probability Letters*, *81*, 1599 - 1603. http://dx.doi.org/10.1016/j.spl.2011.06.013

Williams, M. R., & Kim, D. (2013a). A test for abrupt change in Weibull hazard functions with staggered entry and type I censoring. *Communication in Statistics: Theory and Methods*, *42*, 1922-1933. http://dx.doi.org/10.1080/03610926.2011.600505

Williams, M. R., & Kim, D. (2013b). A test for abrupt change in hazard regression models with Weibull baselines. *Journal of Statistical Planning and Inference*, *143*, 1566-1574. http://dx.doi.org/10.1016/j.jspi.2013.04.006

Worsley, K. J. (1988). Exact percentage points of the likelihood-ratio test for a change-point hazard-rate model. *Biometrics*, *44*, 259-263. http://dx.doi.org/10.2307/2531914

Wu, C. Q., Zhao, L. C., & Wu, Y. H. (2003). Estimation in change-point hazard function models. *Statistics and Probability Letters*, *63*(1), 41-48. http://dx.doi.org/10.1016/S0167-7152(03)00047-6

Yao, Y. (1986). Maximum likelihood estimation in hazard rate models with a change-point. *Communications in Statistics: Theory and Methods*, *15*(8), 2455 - 2466. http://dx.doi.org/10.1080/03610928608829261

Yao,Y. (1987). A note on testing for constant hazard against a change-point alternative. *Annals of Institute of Statistical Mathematics*, *39*, 377-383. http://dx.doi.org/10.1007/BF02491475

Zhang, W., Qian, L., & Li, Y. (2014). Semiparametric sequential testing for multiple change points in piecewise constant hazard functions with long-term survivors. *Communications in Statistics: Simulation and Computation*, *43*(7),1685-1699. http://dx.doi.org/10.1080/03610918.2012.742106

Zhao, X., Wu, X., & Zhou, X. (2009). A change-point model for survival data with long-term survivors. *Statistics Sinica*, *19*, 377-390.