Computational Statistics and Data Analysis 53 (2009) 671–682

Contents lists available at ScienceDirect

Computational Statistics and Data Analysis

journal homepage: www.elsevier.com/locate/csda

Parametric estimation of change-points for actual event data in recurrent events models

Daniel Frobish [a,*], Nader Ebrahimi [b]

[a] Department of Statistics, Grand Valley State University, Allendale, MI 49401, United States
[b] Division of Statistics, Northern Illinois University, Dekalb, IL 60115, United States

ARTICLE INFO

Article history:
Available online 24 August 2008

ABSTRACT

Time to event data have long been important in many applied fields. Many models and analysis methods have been developed for this type of data in which each sample unit experiences at most a single end-of-life event.

In contrast, many applications involve repeated events, where a subject or sampling unit experiences more than one event. There is growing interest in the analysis of recurrent events data, also called repeated events and recurrence data. This type of data arises in many fields. For example, the repair history of manufactured items can be modeled as recurrent events. In medical studies, the times of recurrent disease episodes in patients can also be modeled as recurrent events. In this paper we focus on medical applications (e.g. seizures, heart attacks, cancerous tumors, etc.). However, our proposed methodologies can be applied to other areas as well.

For analyzing recurrence data, the first and perhaps most important step is to model the expected number of events, and sometimes this can be facilitated by modeling the cumulative intensity function or its derivative, the intensity rate function. One particular recurrent events scenario involves patients experiencing events according to a common intensity rate, and then a treatment may be applied. Assuming the treatment to be effective, the patients would be expected to follow a different intensity rate after receiving the treatment. Further, the treatment might be effective for a limited amount of time, so that a third rate would govern arrivals of the recurrent events after the effects of the treatment wore out. In this paper we model the intensity rate for such scenarios. In particular we allow models for the intensity rate, post-treatment, to be piecewise constant. Two estimators of the location of this change are proposed. Properties of the estimators are discussed. An example is studied for illustrative purposes.

1. Introduction

Repeated events processes, where the subject experiences the same type of event more than once, are common in various applied fields such as reliability, medicine, social sciences, business and criminology. Recurrence data consists of the times to any number of repeated events for each sample unit, for example, times to recurrent episodes of a disease in patients or times of repair of a manufactured product. The sample units are considered to be statistically independent, but the times between events within a sample unit are not necessarily independent nor identically distributed. The data are usually censored in the sense that sample units have different ends of histories. For simplicity of our presentation, in this paper we focus on medical applications (e.g. seizures, heart attacks, cancerous tumors). We refer to Pena and Stocker (2007), and Pena et al.

* Corresponding author. Tel.: +1 616 331 3028; fax: +1 616 331 2910.
E-mail address: frobishd@gvsu.edu (D. Frobish).

0167-9473/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.csda.2008.08.022

(2007) and the references cited there for more discussions of recurrent event data. Another good source for discussion is Cook and Lawless (2007).

The mathematics behind analysis of recurrence data involves the theory of counting processes. A counting process $\{N(t); t \geq 0\}$ is a nondecreasing stochastic process that has jumps of size one each time an individual experiences an event of interest. The process $N(t)$ is continuous (and constant) almost surely, except at event times. At these times, $N(t)$ is taken to be continuous from the right with a limit from the left (cadlag). Throughout this paper we assume that there are no simultaneous events for a given unit. Thus if t is an event time, then $N(t) - N(t-) = 1$. According to the Doob–Meyer decomposition theorem, see Andersen et al. (1993), $N(t)$ can be decomposed into a sum of two stochastic processes, $\Lambda(t)$ and $M(t)$, i.e. $N(t) = \Lambda(t) + M(t)$, where $\Lambda(t)$ is the compensator of $N(t)$ or the cumulative intensity process of $N(t)$. The theorem guarantees that $\Lambda(t)$ is a predictable, right-continuous process (but not necessarily deterministic), and $M(t)$ is a martingale. If it exists, define $\lambda(t) = \frac{d\Lambda}{dt}$ to be the intensity rate. Equivalently, one can define

$$\lambda(t) = \lim_{\delta \to 0} \frac{P(N[t, t+\delta) \geq 1 | \mathcal{F}_{t-})}{\delta} \tag{1.1}$$

where $\mathcal{F}_{t-}$ is the complete history of the process, sometimes called the filtration, just prior to time t. Thus given the history prior to time t, $\delta\lambda(t)$ is the approximate probability of observing at least one event in a small time interval of width δ.

Another related quantity is the expected number of events up to time t, referred to as the mean cumulative function, $E(N(t)|\mathcal{F}_{t-})$. The mean cumulative function is usually of interest in the analysis of recurrence data, just as the mean function in regression is modeled. Assuming that it exists, define $\mu(t) = dE(N(t)|\mathcal{F}_{t-})/dt$. Because we assume no two events can occur simultaneously, $\lambda(t) = \mu(t)$. That is, $\Lambda(t) = E(N(t)|\mathcal{F}_{t-})$. Thus when we model $\Lambda(t)$, we really are modeling the mean cumulative function.

In some practical situations, $\lambda(t)$ can change abruptly, so as to cause it to be discontinuous at one or more points. For example, application of a treatment, environmental change, or in general, any commonly experienced event can cause such a shift. The point in time of the shift is called a change-point. In this paper, methods are proposed to estimate such a change-point, when intensity rates are assumed to be piecewise constant. Many researchers have considered the estimating of a change-point for time to event data. We refer you to Matthews and Farewell (1982), Nguyen et al. (1984), Wu et al. (2003), Karasoy and Kadilar (2007) and the references cited there. Further, the use of piecewise constant functions is common when the true form is unknown (e.g. Lawless and Zhan (1998)). This seems natural in our context because our focus is on estimating the unknown change-point, to which significantly less time has been devoted than estimating the rate function itself.

More specifically, consider m patients experiencing recurrent events according to a common intensity rate

$$\lambda(t) = \lambda_0 I(0 \leq t < \tau_1) + \lambda_1 I(\tau_1 \leq t < \tau_2) + \lambda_2 I(\tau_2 \leq t < \infty) \tag{1.2}$$

and $\tau_1 \leq \tau_2 \leq \tau_u$, where τ_u is some known upper bound on τ_2, and $I(x)$ is the usual indicator function. For ease of notation, we define $I_0(t) = I(0 \leq t < \tau_1)$, $I_1(t) = I(\tau_1 \leq t < \tau_2)$ and $I_2(t) = I(\tau_2 \leq t < \infty)$. Without loss of generality, we have assumed that each patient begins the study at time 0. Thus integrating we get the cumulative intensity

$$\Lambda(t) = \lambda_0 t I_0(t) + (\lambda_0\tau_1 + \lambda_1(t - \tau_1))I_1(t) + (\lambda_0\tau_1 + \lambda_1(\tau_2 - \tau_1) + \lambda_2(t - \tau_2))I_2(t). \tag{1.3}$$

Here we assume that τ_1 is known while τ_2 is considered unknown in keeping with the idea that patients undergo a treatment (at time τ_1), and then the effectiveness of the treatment may cease to exist (at time τ_2). In this particular context, $\lambda_0 \geq \lambda_1$ and $\lambda_1 \leq \lambda_2$, though we do not need to know this information for the proposed likelihood procedure to work. For example, if it is believed that a treatment is not immediately effective, but later reduces the frequency of the recurrent event, then we would expect $\lambda_0 \approx \lambda_1$ and $\lambda_1 \geq \lambda_2$. The known parameter τ_1 can further be used as a lower bound on τ_2, or set equal to the starting time of the study, if λ_0 is not of interest.

Because τ_1 is known, it serves as a natural lower bound for τ_2. We impose a known upper bound, τ_u, which is required for consistency of the proposed estimator. Thus, the practitioner will have to have some intuition about range of reasonable values for τ_2. In most cases, this restriction should not be overly burdensome, because the practitioner will probably have some experience with the treatment prior to implementing a study. For example, τ_u could be the second to last event or even the minimum censoring time. It is important to choose τ_u before the last event time to avoid a potentially unbounded likelihood function.

Further, notice we are taking the intensity rate to be piecewise constant. This implies that we are assuming a nonhomogeneous Poisson process to be appropriate. Thus, increments of time behave independently of each other, a fact we use when writing the likelihood function.

Recently, Aschar et al. (2007) proposed a likelihood approach to estimate the change-point τ_2 for only one individual process. However, their paper differs from ours in that they only propose a likelihood approach (while we propose an alternative), and they promote a Bayesian method for acquiring confidence intervals for τ_2 (while we use the bootstrap).

In similar fashion, West and Odgen (1997) studied estimating τ_2 for only one individual process in the context of exact event times. However, like Aschar et al. (2007), they did not offer an alternative to the likelihood estimator, and also used Bayesian methods to get confidence intervals for τ_2.

In Section 2, we propose two methodologies for estimating the unknown change-point, τ_2. The first is a likelihood-based method, and the second is based on the Nelson–Aalen estimator (see Andersen et al. (1993)). We should emphasize that our likelihood estimator is offered only as a comparison to the more robust, Nelson–Aalen-based estimator. Properties of

our estimators are discussed in Section 2 as well. In Section 3, these two estimators are compared through simulations. In Section 4, we illustrate our methodologies using an example.

2. Estimation of τ_2

A sample will consist of arrival times y_{ij}, $i = 1, \ldots n_j$ for the jth patient, $j = 1, \ldots m$, as well as the end of study times, C_j. The C_j play the same role as censoring times in traditional Survival Analysis.

2.1. Likelihood procedure

In order to estimate the parameters, we write the likelihood (Thompson, 1988) for the jth patient,

$$L_j = \exp(-\Lambda(C_j)) \prod_{i=1}^{n_j} \lambda(y_{ij}). \tag{2.1}$$

Thus the log-likelihood, denoted LL, for all patients combined, using Eqs. (1.2) and (1.3), is

$$\begin{aligned} LL = &- \sum_{j=1}^{m} \left[\lambda_0 C_j I_0(C_j) + (\lambda_0\tau_1 + \lambda_1(C_j - \tau_1))I_1(C_j) + (\lambda_0\tau_1 + \lambda_1(\tau_2 - \tau_1) + \lambda_2(C_j - \tau_2)I_2(C_j))\right] \\ &+ \sum_{j=1}^{m}\sum_{i=1}^{n_j} \{I_0(y_{ij}) \log(\lambda_0) + I_1(y_{ij}) \log(\lambda_1) + I_2(y_{ij}) \log(\lambda_2)\}. \end{aligned}$$

We require that $C_j > \tau_2$ for all j, so that each patient contributes to the estimation of λ_2. Thus, $I_0(C_j) = I_1(C_j) = 0$ and $I_2(C_j) = 1$. This implies

$$LL = -\sum_{j=1}^{m}(\lambda_0\tau_1 + \lambda_1(\tau_2 - \tau_1) + \lambda_2(C_j - \tau_2)) + \sum_{j=1}^{m}\left(\sum_{y_{ij}<\tau_1} \log(\lambda_0) + \sum_{\tau_1 \le y_{ij} < \tau_2} \log(\lambda_1) + \sum_{\tau_2 \le y_{ij} < \infty} \log(\lambda_2)\right).$$

Now let $r_{0j} = \#\{y_{ij} | y_{ij} < \tau_1\} = \sum_{i=1}^{n_j} I_0(y_{ij})$, $r_{1j} = \#\{y_{ij} | \tau_1 \le y_{ij} < \tau_2\} = \sum_{i=1}^{n_j} I_1(y_{ij})$, and $r_{2j} = \#\{y_{ij} | \tau_2 \le y_{ij} < \infty\} = \sum_{i=1}^{n_j} I_2(y_{ij})$ for the jth patient. These count the number of events that occur in the three respective pieces of the study. Then

$$LL = -\sum_{j=1}^{m}(\lambda_0\tau_1 + \lambda_1(\tau_2 - \tau_1) + \lambda_2(C_j - \tau_2)) + \log(\lambda_0)\sum_{j=1}^{m} r_{0j} + \log(\lambda_1)\sum_{j=1}^{m} r_{1j} + \log(\lambda_2)\sum_{j=1}^{m} r_{2j}.$$

Next, differentiating with respect to λ_0, λ_1 and λ_2 yields

$$\frac{\partial LL}{\partial \lambda_0} = \frac{1}{\lambda_0}\sum_{j=1}^{m} r_{0j} - m\tau_1 = 0,$$

$$\frac{\partial LL}{\partial \lambda_1} = \frac{1}{\lambda_1}\sum_{j=1}^{m} r_{1j} - m(\tau_2 - \tau_1) = 0$$

and

$$\frac{\partial LL}{\partial \lambda_2} = \frac{1}{\lambda_2}\sum_{j=1}^{m} r_{2j} - \sum_{j=1}^{m}(C_j - \tau_2) = 0.$$

Solving each of these for λ_0, λ_1 and λ_2 respectively gives

$$\hat{\lambda}_0 = \frac{\sum_{j=1}^{m} r_{0j}}{m\tau_1}, \qquad \hat{\lambda}_1 = \frac{\sum_{j=1}^{m} r_{1j}}{m(\tau_2 - \tau_1)}, \qquad \hat{\lambda}_2 = \frac{\sum_{j=1}^{m} r_{2j}}{\sum_{j=1}^{m}(C_j - \tau_2)}. \tag{2.2}$$

Of course, these are just the average number of events per unit time per person, as we would expect. However, we will need an estimate of τ_2 in order to get numerical values for $\hat{\lambda}_1$ and $\hat{\lambda}_2$, since they are functions of τ_2, both explicitly as well as implicitly through r_{1j} and r_{2j} respectively. This leads to the following theorem.

Theorem 1. *The value of $\hat{\tau}_2$ that maximizes $LL(\hat{\lambda}_0, \hat{\lambda}_1, \hat{\lambda}_2)$ occurs at one of the event times.*

Proof. To establish this, we plug the estimates of λ_0, λ_1 and λ_2 into LL. Doing so (with some simplification) yields the profile log-likelihood

$$LL(\hat{\lambda}_0, \hat{\lambda}_1, \hat{\lambda}_2) = \log\left(\frac{\sum_{j=1}^{m} r_{0j}}{m\tau_1}\right)\sum_{j=1}^{m} r_{0j} + \log\left(\frac{\sum_{j=1}^{m} r_{1j}}{m(\tau_2-\tau_1)}\right)\sum_{j=1}^{m} r_{1j} + \log\left(\frac{\sum_{j=1}^{m} r_{2j}}{\sum_{j=1}^{m}(C_j-\tau_2)}\right)\sum_{j=1}^{m} r_{2j} - N, \tag{2.3}$$

where $N = \sum_{j=1}^{m} n_j = \sum_{j=1}^{m} r_{0j} + \sum_{j=1}^{m} r_{1j} + \sum_{j=1}^{m} r_{2j}$.

Unfortunately, $LL(\hat{\lambda}_0, \hat{\lambda}_1, \hat{\lambda}_2)$ is not continuous as a function of τ_2. There are jump discontinuities at every event time because of the indicator functions. Thus, we order $\{y_{ij} | 1 \le j \le m, 1 \le i \le n_j\}$, labeling $y_{(1)}, y_{(2)}, \ldots, y_{(N)}$. For any fixed c, and for $y_{(c)} < \tau_2 < y_{(c+1)}$, LL is convex. To see this, note that r_{0j}, r_{1j} and r_{2j} are fixed between any two event times, and so $LL(\hat{\lambda}_0, \hat{\lambda}_1, \hat{\lambda}_2)$ is continuous and differentiable between any two event times. Thus

$$\frac{\partial LL(\hat{\lambda}_0, \hat{\lambda}_1, \hat{\lambda}_2)}{\partial \tau_2} = \frac{-\sum_{j=1}^{m} r_{1j}}{\tau_2 - \tau_1} + \frac{\sum_{j=1}^{m} r_{2j}}{\sum_{j=1}^{m}(C_j - \tau_2)}$$

and

$$\frac{\partial^2 LL(\hat{\lambda}_0, \hat{\lambda}_1, \hat{\lambda}_2)}{\partial \tau_2^2} = \frac{\sum_{j=1}^{m} r_{1j}}{(\tau_2 - \tau_1)^2} + \frac{m\sum_{j=1}^{m} r_{2j}}{(\sum_{j=1}^{m}(C_j - \tau_2))^2},$$

which is clearly positive for all values of τ_2. This establishes the fact that $LL(\hat{\lambda}_0, \hat{\lambda}_1, \hat{\lambda}_2)$ attains a maximum on each interval $[y_{(c)}, y_{(c+1)}]$ at one of the endpoints, which are by definition event times. This completes the proof.

Applying Theorem 1, one can find $\hat{\tau}_2$, by plugging every event time between τ_1 and τ_u into $LL(\hat{\lambda}_0, \hat{\lambda}_1, \hat{\lambda}_2)$ and choosing the event time that yields the maximum. Then this value can be used to get numerical values for $\hat{\lambda}_1$ and $\hat{\lambda}_2$ in Eq. (2.2).

Next, we argue that $\hat{\tau}_2$ is consistent. Consistency of $\hat{\tau}_2$ is extremely important because $\hat{\lambda}_1$ and $\hat{\lambda}_2$ are functions of $\hat{\tau}_2$. It should be noted that the usual regularity conditions for MLEs to be consistent (e.g. Serfling (1980, pp. 144–5)) are violated, because $LL(\hat{\lambda}_0, \hat{\lambda}_1, \hat{\lambda}_2)$ is not even continuous as a function of τ_2. There are jump discontinuities in r_{1j} and r_{2j} at each observed event time. □

Theorem 2. *The MLE of τ_2 on $[\tau_1, \tau_u]$ is a consistent estimator of τ_2.*

Proof. Since we are maximizing $LL(\hat{\lambda}_0, \hat{\lambda}_1, \hat{\lambda}_2)$ only on $[\tau_1, \tau_u]$, and it is finite on this interval, the conditions listed in Van Der Vaart (1998, p. 48, Theorem 5.14) apply, and guarantee consistency. □

2.2. Robust estimation of τ_2 based on Nelson–Aalen estimator

As before consider m patients experiencing recurrent events according to a common intensity rate as in (1.2) or the cumulative intensity in (1.3). In order to acquire a consistent estimator of the unknown τ_2, we make the following assumptions. For the purposes of consistency, $\tau_1 \le \tau_2 \le \tau_u \le T$, where τ_u is a fixed, known upper bound on the unknown change-point τ_2, and T is any fixed time at or after τ_u. If all of the end of observation times $\{C_j\}$ occur after τ_u, which we assume here, then we can take $T = \min\{C_j | j = 1, \ldots, m\}$. Another choice is $T = \tau_u$. Also, in keeping with the notion of treatment application and wearout, we assume $\lambda_2 \ge \lambda_1$. It is worth noting that the estimation procedure proposed here will still work if we stipulate $\lambda_2 \le \lambda_1$, with a minor adjustment in the procedure.

Lawless and Nadeau (1995) note that the Nelson–Aalen estimator of $\Lambda(t)$ is appropriate (unbiased and consistent) in the context of recurrent events as argued by Nelson (1988). If we let $g(x) = x^p$, $0 \le p \le 1$, then define the following function of $\Lambda(t)$ (as a slight modification of what Chang et al. (1994) proposed in the context of hazard rates).

$$Y(t) = \left(\frac{\Lambda(T) - \Lambda(t)}{T - t} - \frac{\Lambda(t) - \Lambda(\tau_1)}{t - \tau_1}\right) g((t - \tau_1)(T - t)), \quad \tau_1 < t < T. \tag{2.4}$$

Now we estimate $Y(t)$ by

$$Y_n(t) = \left(\frac{\hat{\Lambda}(T) - \hat{\Lambda}(t)}{T - t} - \frac{\hat{\Lambda}(t) - \hat{\Lambda}(\tau_1)}{t - \tau_1}\right) g((t - \tau_1)(T - t)), \quad \tau_1 < t < T \tag{2.5}$$

where $\hat{\Lambda}(t)$ is the Nelson–Aalen estimator which is given by

$$\hat{\Lambda}(t) = \sum_{y_{(i)} \le t} \frac{d_i}{R_i}. \tag{2.6}$$

Here, $y_{(i)}$ is the ith ordered event time, d_i is the number events at $y_{(i)}$ and R_i is the number of events that have yet to happen (at risk set for recurrent events).

We will use $Y_n(t)$ as an estimating function for τ_2. The rational is found in viewing (2.5) as a difference in slopes of the Nelson–Aalen estimator. We would expect the difference to be largest near the true value of τ_2. The purpose of $g(x)$ is to guarantee consistency of $\hat{\tau}_2$. Though consistency of the estimator we will propose will be maintained for any value of p between 0 and 1, Chang et al. (1994) note that values of $p < 0.5$ yield relatively large root mean square errors, and should not be used in practice. We shall follow their recommendation.

Next, we show that $Y(t)$ is maximized at τ_2, a fact crucial to the consistency of the estimator we will propose. In fact, the estimator will be consistent for any intensity rate that has this property, which is why it will be robust to departures from the piecewise constant intensity rate.

We are assuming that $\tau_2 \in (\tau_1, \tau_u)$, so we will first consider $Y(t)$ on (τ_1, τ_2). Using (1.3) and (2.4), we can write

$$Y(t) = (t - \tau_1)^p (T - t)^{p-1} (T - \tau_2)(\lambda_2 - \lambda_1).$$

Then differentiating,

$$Y'(t) = (T - \tau_2)(\lambda_2 - \lambda_1)(t - \tau_1)^{p-1} (T - t)^{p-2} [p(T - t) + (1 - p)(t - \tau_1)]$$

which is positive for all values of $t \in (\tau_1, \tau_2)$. Note that it would be negative if we assumed $\lambda_2 \le \lambda_1$. Similarly, for $t \in (\tau_2, \tau_u)$,

$$Y(t) = (t - \tau_1)^{p-1} (T - t)^p (\lambda_2 - \lambda_1)(\tau_2 - \tau_1)$$

and

$$Y'(t) = (\lambda_2 - \lambda_1)(\tau_2 - \tau_1)(t - \tau_1)^{p-2} (T - t)^{p-1} [(p - 1)(T - t) - p(t - \tau_1)]$$

which is negative for the prescribed values of t.

Thus if $\lambda_2 \ge \lambda_1$, $Y(t)$ is maximized at τ_2 and, as in Chang et al. (1994), we define

$$\hat{\tau}_2 = \inf\{t \in [\tau_1, \tau_u] : Y_n(t\pm) = \sup_{[\tau_1, \tau_u]} Y_n(t)\} \tag{2.7}$$

where $Y_n(t\pm)$ represents the right- or left-hand limit of $Y_n(t)$ at t. If we believe that $\lambda_2 \le \lambda_1$, $Y(t)$ will have its minimum at τ_2 and we use

$$\hat{\tau}_2 = \inf\{t \in [\tau_1, \tau_u] : Y_n(t\pm) = \inf_{[\tau_1, \tau_u]} Y_n(t)\}. \tag{2.8}$$

How to acquire $\hat{\tau}_2$ in practice is given in the following theorem.

Theorem 3. *$\hat{\tau}_2$ must occur at one of the observed event times.*

Proof. We demonstrate this by showing $Y_n(t)$ is increasing between any two event times. After ordering the event times to get $\{y_{(c)} : 1 \le c \le N\}$, between any two of them $[y_{(c)}, y_{(c+1)})$ consider

$$Y_n(t) = \left[\frac{\hat{\Lambda}(T) - \hat{\Lambda}(y_{(c)})}{T - t} - \frac{\hat{\Lambda}(y_{(c)}) - \hat{\Lambda}(\tau_1)}{t - \tau_1} \right] (t - \tau_1)^p (T - t)^p$$

because $\hat{\Lambda}(t)$ is a step function. On each of these intervals,

$$\begin{aligned} Y_n'(t) = {} & (T - t)^{p-2} (t - \tau_1)^{p-2} \{(1 - p)(t - \tau_1)[c_1(t - \tau_1) - c_2(T - t)] \\ & + (T - t)[(p - 1)(c_1(t - \tau_1) - c_2(T - t)) + (t - \tau_1)(c_1 + c_2)]\} \end{aligned}$$

where

$$c_1 = \hat{\Lambda}(T) - \hat{\Lambda}(y_{(c)})$$

and

$$c_2 = \hat{\Lambda}(y_{(c)}) - \hat{\Lambda}(\tau_1).$$

Discarding $(T - t)^{p-2} (t - \tau_1)^{p-2}$ because it is positive for all values of t, and collecting terms we get

$$Y_n'(t) \propto c_1\{(1 - p)(t - \tau_1)^2 + p(T - t)(t - \tau_1)\} + c_2\{(1 - p)(T - t)^2 + p(T - t)(t - \tau_1)\}.$$

Thus the derivative is positive for $0 \le p \le 1$, and $Y_n(t)$ is increasing between observed event times. In practice, we can compute $Y_n(t)$ at each of these times and choose the time that yields the maximum if we assume $\lambda_2 \ge \lambda_1$. Conversely, we can pick the time corresponding to the minimum if $\lambda_2 \le \lambda_1$. This completes the proof. □

Table 1
Piecewise constant model 1

$\lambda_1 = 0.2, \lambda_2 = 0.3, \tau_2 = 5$	m	20	50	100
λ_1	Bias	0.01	0.01	0.01
	RMSE	0.14	0.08	0.04
λ_2	Bias	0.01	0.01	0
	RMSE	0.04	0.02	0.01
τ_2^{like}	Bias	0.89	0.39	0.21
	RMSE	3.74	2.78	1.75
$\tau_2^{p=0.5}$	Bias	4.18	3.59	3.06
	RMSE	5.35	4.64	3.95
$\tau_2^{p=0.75}$	Bias	3.16	2.74	2.52
	RMSE	4.05	3.43	3.1
$\tau_2^{p=1.0}$	Bias	2.81	2.53	2.41
	RMSE	3.49	3.05	2.83

Next, we establish the consistency of the estimator of τ_2.

Theorem 4. *$\hat{\tau}_2$ is consistent for τ_2.*

Proof. Because $Y(t)$ is maximized at τ_2, we can find a sufficiently small $\epsilon > 0$ such that there exists a constant c_1 where $Y(\tau_2) - Y(\hat{\tau}_2 \pm \epsilon) \geq c_1$. Then for sufficiently large n,

$$\begin{aligned} P(|\hat{\tau}_2 - \tau_2| > \epsilon) &\leq P(|Y(\hat{\tau}_2) - Y(\tau_2)| > c_1) \\ &\leq P(|Y(\hat{\tau}_2) - Y_n(\hat{\tau}_2)| + |Y_n(\hat{\tau}_2) - Y(\tau_2)| > c_1). \end{aligned}$$

First note that $|Y(\hat{\tau}_2) - Y_n(\hat{\tau}_2)| \to 0$ because $\hat{\Lambda}(t) \to \Lambda(t)$ (Lawless and Nadeau, 1995). Thus we can bound this term by $c_1/2$. Also, the second term is bounded by $\sup_{t\in[\tau_1,\tau_u]} |Y_n(t) - Y(t)|$ because $Y_n(t)$ and $Y(t)$ are maximized at $\hat{\tau}_2$ and τ_2 respectively.

Now define $U_m(t) = \hat{\Lambda}(t) - \Lambda(t)$. In the following, all suprema are taken over $[\tau_1, \tau_u]$. Then

$$\begin{aligned} P(|\hat{\tau}_2 - \tau_2| > \epsilon) &\leq P\left(\sup |Y_n(t) - Y(t)| > \frac{c_1}{2}\right) \\ &= P\left(\sup |t^{p-1}(T-t)^{p-1}[t\hat{\Lambda}(T) - T\hat{\Lambda}(t)] - t^{p-1}(T-t)^{p-1}[t\Lambda(T) - T\Lambda(t)]| > \frac{c_1}{2}\right) \\ &= P\left(\sup |t^p(T-t)^{p-1}U_m(T) - Tt^{p-1}(T-t)^{p-1}U_m(t)| > \frac{c_1}{2}\right) \\ &\leq P\left(\sup\{t^p(T-t)^{p-1}|U_m(T)|\} + \sup\{Tt^{p-1}(T-t)^{p-1}|U_m(t)|\} > \frac{c_1}{2}\right) \\ &\leq P\left(\tau_u^p(T-\tau_u)^{p-1}|U_m(T)| + T\tau_1^{p-1}(T-\tau_u)^{p-1}\sup|U_m(t)| > \frac{c_1}{2}\right) \\ &\leq P(|U_m(T)| > c_3) + P(\sup|U_m(t)| > c_2), \end{aligned}$$

where c_3 and c_2 are constants that depend on τ_1, τ_u, p, T and c_1. Since $\hat{\Lambda}(t)$ is known to be uniformly consistent (Andersen et al., 1993), both $P(|U_m(T)| > c_3)$ and $P(\sup|U_m(t)| > c_2)$ go to 0, which completes the proof. □

Again, note that if we use this estimate of τ_2 in $\hat{\lambda}_1$ and $\hat{\lambda}_2$ given in (2.2), we still have consistent estimates for these parameters. It is worth noting that the choice of estimation method (likelihood or robust) for τ_2 has little effect on the values of $\hat{\lambda}_1$ and $\hat{\lambda}_2$.

Remark 1. Asymptotic distributions of our estimators are very complicated and they are difficult to obtain. Throughout this paper we have chosen to use bootstrap methods instead to construct confidence intervals for the change-point.

3. Simulation results

Tables 1 through 8 are created using a piecewise constant model. The later Tables 9–14, come from a perturbation of that model. The study starts at time 0, and we also use $\tau_1 = 0$. Because the focus is on the behavior of $\hat{\tau}_2$, we take τ_2 to be 5 or 10, and $\tau_u = T = 15$, which mimics a practitioner having prior information about possible values for τ_2. Censoring times are randomly generated from between 20 and 25.

Without loss of generality, we always choose $\lambda_1 < \lambda_2$. We use $\lambda_1 = 0.2$ paired with $\lambda_2 = 0.3$, 0.6 and 1.0, and also $\lambda_1 = 0.6$ with $\lambda_2 = 1.0$ to maintain $\lambda_1 < \lambda_2$. As previously stated, for the purposes of simplicity, we avoid simulating times prior to τ_1 so that $\lambda_0 = 0$.

Table 2
Piecewise constant model 2

$\lambda_1 = 0.2, \lambda_2 = 0.6, \tau_2 = 5$	m	20	50	100
λ_1	Bias	0.01	0	0
	RMSE	0.05	0.03	0.02
λ_2	Bias	0.01	0	0
	RMSE	0.04	0.03	0.02
τ_2^{like}	Bias	0.14	0.05	0.02
	RMSE	0.76	0.27	0.13
$\tau_2^{p=0.5}$	Bias	2.36	1.33	0.73
	RMSE	3.53	2.26	1.37
$\tau_2^{p=0.75}$	Bias	1.86	1.35	1.03
	RMSE	2.55	1.92	1.48
$\tau_2^{p=1.0}$	Bias	1.84	1.55	1.36
	RMSE	2.36	1.97	1.69

Table 3
Piecewise constant model 3

$\lambda_1 = 0.2, \lambda_2 = 1.0, \tau_2 = 5$	m	20	50	100
λ_1	Bias	0.01	0	0
	RMSE	0.05	0.03	0.02
λ_2	Bias	0	0	0
	RMSE	0.05	0.04	0.02
τ_2^{like}	Bias	0.06	0.02	0.01
	RMSE	0.24	0.09	0.04
$\tau_2^{p=0.5}$	Bias	1.27	0.49	0.21
	RMSE	2.25	1.03	0.49
$\tau_2^{p=0.75}$	Bias	1.27	0.77	0.52
	RMSE	1.87	1.22	0.85
$\tau_2^{p=1.0}$	Bias	1.43	1.12	0.97
	RMSE	1.87	1.47	1.25

Table 4
Piecewise constant model 4

$\lambda_1 = 0.6, \lambda_2 = 1.0, \tau_2 = 5$	m	20	50	100
λ_1	Bias	0.03	0.01	−0.01
	RMSE	0.14	0.06	0.04
λ_2	Bias	0.01	0	0
	RMSE	0.06	0.03	0.02
τ_2^{like}	Bias	0.25	0.06	0.02
	RMSE	1.72	0.61	0.28
$\tau_2^{p=0.5}$	Bias	3.63	2.84	2.48
	RMSE	4.53	3.57	3.12
$\tau_2^{p=0.75}$	Bias	2.82	2.5	2.39
	RMSE	3.41	2.94	2.74
$\tau_2^{p=1.0}$	Bias	2.58	2.46	2.41
	RMSE	3	2.75	2.63

5000 replications were made for each combination of parameters for sample sizes of 20, 50 and 100. The bias and root mean squared error given for $\hat{\lambda}_1$ and $\hat{\lambda}_2$ are based on the likelihood estimate of τ_2, but they do not change significantly if the Nelson–Aalen-based estimate of τ_2 is used. This is because they are simply the average number of events per person per time.

The tables suggest that if τ_2 occurs near the front of the study, then $p = 1$ should be used unless λ_1 is believed to be very different from λ_2, in which case, p should be 0.5. However if τ_2 takes place later in the study, then $p = 1$ should still be used except if the rates are close. If it is believed that the rates are close, then p should be chosen to be 0.75.

We expect the MLE to outperform the robust estimator because the data are generated from the correct model. This is certainly true when τ_2 occurs early in the study ($\tau_2 = 5$). However, when τ_2 is later ($\tau_2 = 10$), the robust estimator does essentially as well as the MLE.

Table 5
Piecewise constant model 5

$\lambda_1 = 0.2, \lambda_2 = 0.3, \tau_2 = 10$	m	20	50	100
λ_1	Bias	0.01	0.01	0.01
	RMSE	0.13	0.07	0.02
λ_2	Bias	0.01	0.01	0
	RMSE	0.04	0.03	0.02
τ_2^{like}	Bias	1.35	0.37	0
	RMSE	4.11	2.66	1.52
$\tau_2^{p=0.5}$	Bias	0.8	0.72	0.51
	RMSE	2.54	1.73	1.17
$\tau_2^{p=0.75}$	Bias	0.07	0.03	0.03
	RMSE	1.88	1.18	0.74
$\tau_2^{p=1.0}$	Bias	0.48	0.32	0.22
	RMSE	1.68	1.12	0.74

Table 6
Piecewise constant model 6

$\lambda_1 = 0.2, \lambda_2 = 0.6, \tau_2 = 10$	m	20	50	100
λ_1	Bias	0.01	0	0
	RMSE	0.03	0.02	0.01
λ_2	Bias	0.01	0	0
	RMSE	0.05	0.03	0.02
τ_2^{like}	Bias	0.11	0.04	0.02
	RMSE	0.69	0.25	0.12
$\tau_2^{p=0.5}$	Bias	0.65	0.26	0.13
	RMSE	1.20	0.53	0.28
$\tau_2^{p=0.75}$	Bias	0.24	0.11	0.05
	RMSE	0.60	0.27	0.14
$\tau_2^{p=1.0}$	Bias	0.07	0.03	0.01
	RMSE	0.44	0.21	0.11

Table 7
Piecewise constant model 7

$\lambda_1 = 0.2, \lambda_2 = 1.0, \tau_2 = 10$	m	20	50	100
λ_1	Bias	0	0	0
	RMSE	0.03	0.02	0.01
λ_2	Bias	0.01	0	0
	RMSE	0.06	0.04	0.03
τ_2^{like}	Bias	0.05	0.02	0.01
	RMSE	0.23	0.08	0.04
$\tau_2^{p=0.5}$	Bias	0.31	0.12	0.06
	RMSE	0.6	0.24	0.12
$\tau_2^{p=0.75}$	Bias	0.14	0.06	0.03
	RMSE	0.30	0.12	0.06
$\tau_2^{p=1.0}$	Bias	0.07	0.03	0.01
	RMSE	0.2	0.09	0.05

In Tables 9 through 14, we study the behavior of the estimators of τ_2 if the piecewise constant model is violated. Specifically, we conservatively take λ_1 to be 0.1 (constant) while generating the rate linearly after τ_2 with slopes 0.05, 0.01 and 0.005 and at the same time, ensuring continuity of $\lambda(t)$. This is done for $\tau_2 = 5$ and 10, and for sample sizes 20, 50 and 100.

When the unknown change-point is early, then $p = 1$ is the universal choice. When τ_2 occurs later, the robust Nelson–Aalen-based estimator does much better.

It is important to note that the consistency of the robust estimator depends on the piecewise constant form (or at least a form where $\lambda(t)$ is optimized at τ_2). This explains why the bias does not decrease monotonically for Tables 9 through 14. However the robust estimator still outperforms the likelihood estimator in practically every scenario considered.

In conclusion, if it is believed the data follow a piecewise constant model and the change-point occurs early in the study, the practitioner should use the MLE. However, in all other cases, we recommend the use of the robust estimator.

Table 8
Piecewise constant model 8

$\lambda_1 = 0.6, \lambda_2 = 1.0, \tau_2 = 10$	m	20	50	100
λ_1	Bias	0.02	0.01	0
	RMSE	0.09	0.04	0.03
λ_2	Bias	0.01	0.01	0
	RMSE	0.07	0.04	0.03
τ_2^{like}	Bias	0.01	0.05	0.03
	RMSE	1.47	0.58	0.27
$\tau_2^{p=0.5}$	Bias	0.70	0.37	0.21
	RMSE	1.35	0.74	0.44
$\tau_2^{p=0.75}$	Bias	0.16	0.09	0.06
	RMSE	0.76	0.4	0.23
$\tau_2^{p=1.0}$	Bias	0.11	0.07	0.04
	RMSE	0.68	0.38	0.23

Table 9
Perturbation of piecewise constant model 1

$\tau_2 = 5, \lambda_1 = 0.1, \lambda_2(\text{slope}) = 0.05$	m	20	50	100
τ_2^{like}	Bias	9.6	9.83	9.92
	RMSE	9.61	9.83	9.92
$\tau_2^{p=0.5}$	Bias	5.17	4.87	4.73
	RMSE	5.52	5.09	4.87
$\tau_2^{p=0.75}$	Bias	4.27	4.13	4.06
	RMSE	4.53	4.28	4.16
$\tau_2^{p=1.0}$	Bias	4.53	4.28	4.16
	RMSE	4.03	3.89	3.79

Table 10
Perturbation of piecewise constant model 2

$\tau_2 = 5, \lambda_1 = 0.1, \lambda_2(\text{slope}) = 0.01$	m	20	50	100
τ_2^{like}	Bias	8.68	9.34	9.67
	RMSE	8.81	9.39	9.68
$\tau_2^{p=0.5}$	Bias	4.83	5.17	5.11
	RMSE	5.93	5.85	5.55
$\tau_2^{p=0.75}$	Bias	4.06	4.21	4.2
	RMSE	4.94	4.72	4.54
$\tau_2^{p=1.0}$	Bias	3.66	3.77	3.76
	RMSE	4.36	4.18	4.04

Table 11
Perturbation of piecewise constant model 3

$\tau_2 = 5, \lambda_1 = 0.1, \lambda_2(\text{slope}) = 0.005$	m	20	50	100
τ_2^{like}	Bias	7.59	8.71	9.31
	RMSE	8.08	8.89	9.36
$\tau_2^{p=0.5}$	Bias	4.18	5.03	5.17
	RMSE	5.83	6.01	5.85
$\tau_2^{p=0.75}$	Bias	3.62	4.09	4.12
	RMSE	4.89	4.85	4.64
$\tau_2^{p=1.0}$	Bias	3.33	3.66	3.69
	RMSE	4.38	4.29	4.09

4. Analysis of example

A randomized double-blinded placebo-controlled study was performed by Dibley et al. (1996) to assess the effectiveness of Vitamin A supplements to reduce the incidence of acute respiratory illness (ARI) in children. The authors also studied the impact on acute lower respiratory illness as well as diarrhea, but we limit our analysis to ARI.

An occurrence of ARI was defined to be two or more adjoining days for which a child was reported to have a cough. We take the start of each event to be our recurrent event. Though this means a child is not at risk for another event until the

Table 12
Perturbation of piecewise constant model 4

$\tau_2 = 10, \lambda_1 = 0.1, \lambda_2(\text{slope}) = 0.05$	m	20	50	100
τ_2^{like}	Bias	4.47	4.74	4.86
	RMSE	4.51	4.75	4.86
$\tau_2^{p=0.5}$	Bias	2.15	2.24	2.15
	RMSE	2.86	2.55	2.34
$\tau_2^{p=0.75}$	Bias	1.24	1.42	1.4
	RMSE	2.18	1.77	1.6
$\tau_2^{p=1.0}$	Bias	0.63	0.84	0.9
	RMSE	1.8	1.33	1.15

Table 13
Perturbation of piecewise constant model 5

$\tau_2 = 10, \lambda_1 = 0.1, \lambda_2(\text{slope}) = 0.01$	m	20	50	100
τ_2^{like}	Bias	3.42	4.22	4.52
	RMSE	3.93	4.33	4.57
$\tau_2^{p=0.5}$	Bias	0.57	0.71	1.32
	RMSE	4.51	3.71	3.13
$\tau_2^{p=0.75}$	Bias	1.01	0.19	0.24
	RMSE	3.9	3.06	2.44
$\tau_2^{p=1.0}$	Bias	1.37	0.74	0.4
	RMSE	3.57	2.77	2.19

Table 14
Perturbation of piecewise constant model 6

$\tau_2 = 10, \lambda_1 = 0.1, \lambda_2(\text{slope}) = 0.005$	m	20	50	100
τ_2^{like}	Bias	1.31	3.24	4.02
	RMSE	4.23	4.1	4.24
$\tau_2^{p=0.5}$	Bias	1.31	0.16	0.59
	RMSE	4.84	4.15	3.51
$\tau_2^{p=0.75}$	Bias	1.6	0.85	0.45
	RMSE	4.25	3.46	2.88
$\tau_2^{p=1.0}$	Bias	1.81	1.23	0.94
	RMSE	3.91	3.13	2.59

current bout of cough subsides, most episodes ended in a matter of days. In the context of a two year study, these durations are taken to be negligible.

Thirty-four villages in Indonesia were surveyed, and 1036 children aged six to forty-seven months were identified for the study, with more than 90% completing treatment. Some of these did not enter into the study until the second year, but the time scale used in the analysis accounts for this. Blocking was used to ensure balance of groups by village and distance from health services. Seasonality should not explain any drop in events as individuals entered the study at different times, but the start time in for each individual in the data is set to zero.

Children were randomly assigned to receive a vitamin A supplement or placebo every 4 months for 2 years. Though this design may suggest the possibility of multiple unknown change-points, we use a model with only one, occurring at some time after the initial treatment. The subsequent treatments could be thought to only maintain the level of vitamin A in patients.

A total of 10,735 episodes of ARI were observed, with 88% of children experiencing two or more. Interestingly, the incidence of ARI was 8% higher in the treatment group. The investigators theorize that the immune systems of children with preexisting sufficient levels of vitamin A are suppressed when levels of vitamin A are too high, leading to more frequent infections. Thus we focus on the children in the poor nutrition group. Also, we require individuals to be under observation for at least 1.5 of the 2 years, so that each included child contributes to the estimation of the unknown change-point. Individuals that do not stay on study past τ_2 do not contribute to the estimation of τ_2, and so are excluded. These restrictions yield a sample of size 102.

The main question we answer is, when does the effect of the Vitamin A supplement wear out? That is, when do we expect the supplement to stop suppressing the incidence of ARI, if not permanently? Is there a time at which the supplement starts to lose its effectiveness?

Examination of a histogram (Fig. 1) of these times suggests that the piecewise constant model is not applicable. It seems that, after the initial dose period of 4 months, there is a drop in the number of events. Later, around one year (or 365 days),

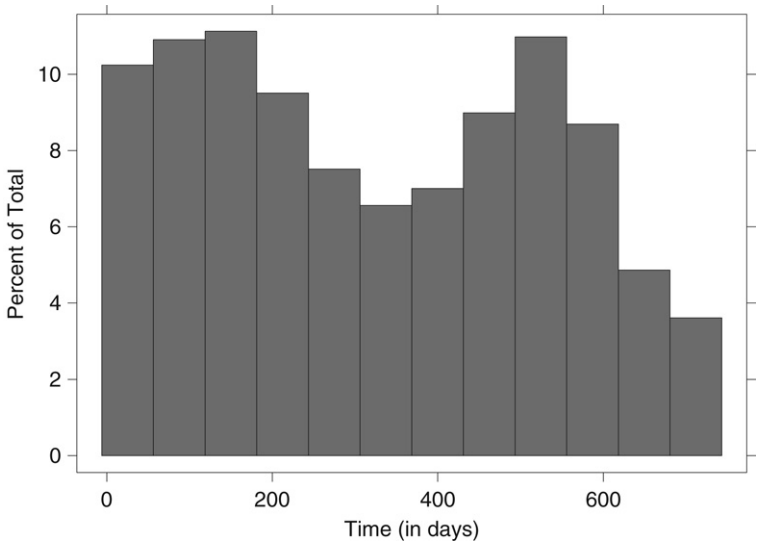

Fig. 1. Time to incidence of acute respiratory illness.

Table 15
Analysis of acute respiratory illness data with piecewise constant model

Method	Lower bound	$\hat{\tau}_2$	$SE(\hat{\tau}_2)$	$\hat{\lambda}_1$	$\hat{\lambda}_2$
Likelihood	0	520	171.424	0.019	0.025
Likelihood	120	490	72.963	0.018	0.025
$NA^{p=1}$	0	422	15.316	0.019	0.023
$NA^{p=0.75}$	0	422	11.505	0.019	0.023
$NA^{p=0.5}$	0	489	32.809	0.019	0.025
$NA^{p=1}$	120	422	10.120	0.018	0.023
$NA^{p=0.75}$	120	428	18.556	0.018	0.023
$NA^{p=0.5}$	120	489	33.942	0.018	0.025

the frequency of events appears to go back up. We hope to detect the location of this minimum frequency, as it may be the time at which the Vitamin A treatment ceases to be effective. At this point, we should emphasize again that, though the likelihood estimator of τ_2 is not expected to perform well, the robust estimator may still give a reasonable estimate. This is the reason we have chosen an example that seems to violate the piecewise constant model.

It seems from the graph that we should choose a time greater than zero for τ_1, since there is an initial increase in the number of events, which levels out at some point before 200 days. In the analysis, we use 120 days as a lower bound for the unknown τ_2 because patients receive a dose of Vitamin A every 4 months. 540 days is chosen as an upper bound, as that is when we start to see patients leaving the study (due to different ends of time on study). Both likelihood and the more robust Nelson–Aalen estimators are calculated, and given in Table 15. We expect the likelihood estimator to perform poorly as the piecewise constant model is violated. We include the robust estimator for $p = 0.5, 0.75$ and 1.0. Also, using a lower bound of 0 is tried for comparison, and does not seem to impact the value of $\hat{\tau}_2$ for the robust procedure. For the choice of T, we use 540. The standard errors for $\hat{\tau}_2$ are found by generating 250 bootstrap resamples.

Note that estimates of the rates are not sensitive to the choice of method of estimation for τ_2. This is because they average the number of events per person per time. Thus, if you use a different location for τ_2, the change in number of events will be offset by the change of time in the denominator. We are primarily interested in the behavior of $\hat{\tau}_2$ anyway.

As we mentioned in Remark 1, asymptotic distributions of our proposed estimators are very complicated. If we want a confidence interval for τ_2, then we can use the hybrid bootstrap method (Shao and Tu, 1996). We also computed intervals based on the normal approximation and the bootstrap percentiles. Because there was disagreement among the intervals, we choose the hybrid method which has better coverage probability (Shao and Tu, 1996). They let

$$H_{boot}(x) = P\{\sqrt{n}(\hat{\tau}_2^* - \hat{\tau}_2) \leq x\} \tag{4.1}$$

where $\hat{\tau}_2^*$ represents the bootstrap analog of $\hat{\tau}_2$. Because the graph clearly shows the piecewise constant form is violated, we will use the Nelson–Aalen version of $\hat{\tau}_2$ with $p = 1$, which is 422. The simulations strongly suggest that the MLE will not perform well for this data set, so there is no reason to construct a confidence interval for it.

Then an approximate lower confidence bound for τ_2 is given by

$$\underline{\tau_2}_{HB} = \hat{\tau}_2 - \frac{H_{boot}^{-1}(1-\alpha)}{\sqrt{n}} \tag{4.2}$$

with $\alpha = 0.025$ for a 95% confidence interval. The upper bound can be found by choosing $\alpha = 0.975$. For this example, $H^{-1}_{boot}(1-0.025) = 171.6916$ and $H^{-1}_{boot}(1-0.975) = -293.8956$ found with 2500 resamples, each with the original sample size of 102. Thus the hybrid bootstrap 95% confidence interval for τ_2 is (405.0, 451.1).

5. Concluding remarks

Two estimators of an unknown change-point (τ_2) are proposed in the context of recurrent events. Though the given formulation includes a known change-point τ_1 (e.g. for time of intervention or treatment), such a feature is not central to the use of the estimators. However, we do require that a lower (τ_1) and upper bound (τ_u) are known for τ_2.

The post-treatment intensity rate is separated into two (constant) pieces, λ_1 and λ_2. Though this may not be entirely accurate for every practical example, approximating a function with a piecewise constant function has long been used in mathematics and statistics. We allow either $\lambda_1 > \lambda_2$ or $\lambda_1 < \lambda_2$, and only require knowledge of which restriction is expected for the robust estimator in Section 2.2.

The likelihood estimator will obviously be sensitive to model misspecification, and so we use it for a comparison for our robust estimator. The robust estimator is based on maximizing or minimizing the difference in slopes of the Nelson–Aalen estimator over two adjacent pieces of the target interval (τ_1 to t and t to τ_u). We search for the value of t to optimize this slope difference and use it as our estimator of the unknown change-point, and show it to be consistent.

A natural extension of the model (1.2) is to include covariates in the model, and this is currently under investigation. Traditionally this is accomplished by making the intensity rate dependent on covariates. We will allow the change-point to depend directly on covariates in future work. Further, we hope to develop methods that will relax the dependence on the piecewise constant model.

Acknowledgement

We wish to acknowledge the MORVITA team (Dibley et al., 1996) that allowed us to use their data to demonstrate the proposed methods.

References

Andersen, P., Borgan, O., Gill, R., Keiding, N., 1993. Statistical Models Based on Counting Processes. Springer-Verlag, New York.
Aschar, J.A., Loibel, S., Andrade, M., 2007. Interfailure data with constant hazard function in the presence of change-points. REVSTAT 5, 209–226.
Chang, I., Chen, C., Hsiung, C., 1994. Estimation in Change-Point Hazard Rate Models with Random Censorship. In: IMS Lecture Notes-Monograph Series, vol. 23. pp. 78–92.
Cook, R.J., Lawless, J.F., 2007. The Statistical Analysis of Recurrent Events. Springer, New York.
Dibley, M.J., Sadjimin, T., Kjolhede, C.L., Moulton, L.H., 1996. Vitamin A supplementation fails to reduce incidence of acute respiratory illness and diarrhea in preschool-age indonesian children. Journal of Nutrition 126, 434–442.
Karasoy, D.S., Kadilar, C., 2007. A new Bayes estimate of change-point in the hazard function. Computational Statistics and Data Analysis 51, 2993–3001.
Lawless, J., Nadeau, C., 1995. Some simple robust methods for the analysis of recurrent events. Technometrics 37, 158–168.
Lawless, J., Zhan, M., 1998. Analysis of interval-grouped recurrent event data using piecewise constant rate functions. Canadian Journal of Statistics 26, 549–565.
Matthews, D., Farewell, V., 1982. On testing for a constant hazard against a change-point alternative. Biometrics 38, 463–468.
Nelson, W., 1988. Graphical analysis of system repair data. Journal of Quality Technology 20, 24–35.
Nguyen, H., Rogers, G., Walker, E., 1984. Estimation in change-point hazard rate models. Biometrika 71, 299–304.
Pena, E., Slate, E., Gonzalez, J., 2007. Semiparametric inference for a general class of models for recurrent events. Journal of Statistical Planning and Inference 137, 1727–1747.
Pena, E., Stocker, R., 2007. A general class of parametric models for recurrent event data. Technometrics 49, 210–221.
Serfling, R., 1980. Approximation Theorems of Mathematical Statistics. Wiley, New York.
Shao, J., Tu, D., 1996. The Jackknife and Bootstrap. Springer-Verlag, New York.
Thompson, W., 1988. Point Process Models with Applications to Safety and Reliability. Chapman and Hall.
Van Der Vaart, A., 1998. Asymptotic Statistics. Cambridge University Press, UK.
West, R., Odgen, R., 1997. Continuous-time estimation of a change-point in a poisson process. Journal of Statistical Computational Simulation 56, 293–302.
Wu, C.Q., Zhao, L.C., Wu, Y.H., 2003. Estimation in change-point hazard function models. Statistics and Probability Letters 63, 41–48.