Local Likelihood for Interval-Censored Recurrent Event Data

David Tolusso and Richard J. Cook

Department of Statistics and Actuarial Science, University of Waterloo,
Waterloo, Ontario, Canada N2L 3G1
†Corresponding Author: rjcook@uwaterloo.ca

May 30, 2008
SUMMARY. Interval-censored recurrent event data arise when event occurrence is only determined at periodic assessment times. Nonparametric estimates of rate and mean functions are readily obtained in this setting, but lack the smoothness usually anticipated for the underlying event process. We describe local likelihood methods which generate smooth estimates of rate and mean functions, and associated covariate effects in the context of multiplicative intensity-based models. A modified EM algorithm is described which incorporates kernel smoothing of the rate function under the assumption it can be modeled locally using polynomials. Both Poisson models and mixed Poisson models are considered. Simulations suggest the proposed methods work well, and applications to data from a study of feedwater flow losses in nuclear plants and a study of joint damage in patients with psoriatic arthritis illustrate their use.

KEY WORDS: local likelihood; mean function; multiplicative model; rate function; recurrent events
1. Introduction

Recurrent event data arise when individuals are at risk of an event that can occur repeatedly over time. In studies of patients with chronic bronchitis, for example, interest lies in the occurrence of acute exacerbations (Calverley et al., 2007), in studies of patients with herpes simplex virus infection outbreaks of symptoms can occur repeatedly over time (Romanowski et al., 2003), and neurological studies interest lies in modeling the occurrence of seizures in patients with epilepsy (Fuchs et al., 1994). In many settings involving recurrent events, the objectives of analyses are to characterize the event process in terms of gap-time distributions, marginal rate and mean functions, or more general intensity-based analyses. Identification and characterization of important covariate effects are also typically of importance. All of these analysis strategies are relatively straightforward when individuals are observed continuously and event times are known exactly during the observation period (Cook and Lawless, 2007).

In many settings, however, individuals are only seen at periodic assessment times, and because of the nature of the events it is only possible to detect their occurrence at assessment times. In clinical trials of patients with osteoporosis, for example, interest often lies in the occurrence of vertebral fractures which may only be documented upon radiographic examination of the spine at study assessments (Riggs et al., 1990). In rheumatology, interest often lies in the development of joint damage detected at periodic radiographic assessments (Gladman et al., 1995). Other examples include the assessment of skeletal metastases in patients with breast cancer (Hortobagyi et al., 1996; Chen et al., 2005), the detection of polyps among patients at risk of colon cancer through periodic gastroenterological examination (Avidan et al., 2002), and the recording of lesions among patients with multiple sclerosis through periodic CT scans (Johnson et al., 1995). In this situation, when one or more events have been determined to occur since the last visit, the event times are unknown, resulting in interval-censored recurrent event data; this type of data are also often referred to as panel count data (Sun and Kalbfleisch, 1995) or interval grouped count data (Lawless and Zhan, 1998). An extreme form of interval censoring occurs when the cumulative number of events is observed at a single inspection time, and the resulting data are called current status data.

Thall and Lachin (1988) propose simple nonparametric estimators of the rate function with interval-censored recurrent event data and Wellner and Zhang (2000) derive the non-parametric maximum likelihood estimate for the general case of interval-censored recurrent
event data under a Poisson assumption. Sun and Kalbfleisch (1993) discuss a tumorgenecity study in which the number of tumors developing in laboratory animals was assessed at the time of death yielding current status data, and Sun and Kalbfleisch (1995) discuss nonparametric estimation in another problem involving losses in feedwater flow in nuclear power plants; this nonparametric estimate of the mean function is only defined for time points at which an event was recorded. A semiparametric approach was adopted by Staniswalis et al. (1997) who used the profile likelihood arguments of Severini and Wong (1992) to obtain estimators for the regression coefficients in a multiplicative Poisson model and nondecreasing mean function estimates via smoothing at a second stage. Lawless and Zhan (1998) considered a “weakly” parametric piecewise constant approach for estimation of the baseline intensity function. These authors also developed a robust approach using generalized estimating equations defined in terms of first and second moments, but this method only yields consistent estimates under completely independent censoring. Sun (2006) and Cook and Lawless (2007) give recent reviews of methods for the analysis of interval-censored recurrent event data.

The purpose of this paper is to develop smooth estimators of the mean and rate functions by the use of local likelihood methods. While two-stage estimation procedures are possible in which nonparametric estimates are smoothed (e.g. see Staniswalis et al. (1997) for recurrent events, or Li et al. (1997) in the context of survival data) we prefer to proceed directly to estimation of a smooth function. Loader (1996) and Hjort and Jones (1996) describe the use of local likelihood methods for obtaining smooth estimates of density functions, and local likelihood methods have been adapted for density estimation with interval-censored data by Braun et al. (2005). Local likelihood methods for estimation of the hazard function with right-censored or interval-censored data have been developed by Betensky et al. (1999) and Bebchuk and Betensky (2001). Here we develop local likelihood methods to estimate parameters under multiplicative intensity functions under Poisson and mixed-Poisson settings. The primary objective is to provide a method of analysis that gives smooth estimates of the baseline rate and mean functions with interval-censored.

The remained of the paper is organized as follows. In Section 2 we describe local likelihood methods for the analysis of right-censored Poisson and mixed-Poisson data for estimation of rate and mean functions, and covariate effects. Adaptations to handle overdispersed Poisson data are given based on a simple variance inflation factor and under the framework
of a mixed-Poisson model. In Section 3 an EM-type algorithm is adopted for mixed Poisson interval-censored recurrent event data in which a kernel smoothed estimate of the conditional expectation is used at the E-step in the spirit of Braun et al. (2005). Profile likelihood is advocated for interval estimation of regression coefficients. Simulation studies are reported on in Section 4 and applications to the nuclear safety data of Sun and Kalbfleisch (1995) and data from a psoriatic arthritis clinic are given in Section 5. General remarks are made in Section 6.

2. Local Likelihood for Rate Function Analysis with Right-Censored Data

2.1 Notation and Model Specification

Consider a sample of \( m \) individuals in which individual \( i \) is observed over \( (0, \tau_i] \), \( i = 1, \ldots, m \). Let \( \{N_i(s), 0 < s \} \) denote the counting process for individual \( i \) such that \( N_i(t) \) counts the number of events they experience over the interval \( (0, t], i = 1, \ldots, m \). In some settings a \( p \times 1 \) covariate vector \( x_i \) may be associated with individual \( i \), and interest may lie in assessing the relationship between \( x_i \) and event occurrence. If \( \mathcal{H}_i(t) = \{N_i(u), 0 \leq u < t; x_i\} \) denotes the history of the process at \( t \), including the covariate value, then the event intensity function is

\[
\lambda_i(t|\mathcal{H}_i(t)) \equiv \lim_{\Delta t \to 0} \frac{P(N_i(t + \Delta t^-) - N_i(t^-) = 1|\mathcal{H}_i(t))}{\Delta t}, \quad (2.1)
\]

which completely defines the probability model for the events (Cook and Lawless, 2007). There is a large number of ways in which the event intensity can depend on the history but for the canonical Poisson model

\[
\lambda_i(t|\mathcal{H}_i(t)) = \rho_i(t), \quad (2.2)
\]

where \( \rho_i(t) \) is a rate function and \( \mu_i(t) = E\{N_i(t)|x_i\} \). Covariate effects are often expressed through multiplicative models in which \( \rho_i(t) = \rho_0(t) \exp(x_i^T \beta) \) where \( \rho_0(t) \) is a baseline rate function giving a baseline mean function \( \mu_0(t) = \int_0^t \rho(s) \, ds \) corresponding to a subject with \( x_i = 0 \), and \( \beta \) is a \( p \times 1 \) vector of regression coefficients. We consider the problem in which it is of interest to estimate the rate function \( \rho(t) \) and mean function \( \mu(t) = E\{N_i(t)\} = \int_0^t \rho(u) \, du \), and possibly regression coefficients.
2.2 Local Likelihood Methods for Right-Censored Count Data

Suppose events for different individuals are all generated according to the same Poisson process. Let \( t_{i1}, \ldots, t_{in_i} \) denote the times of the \( N_i(\tau_i) = n_i \) events experienced by individual \( i \) over \( (0, \tau_i] \). The likelihood contribution from individual \( i \) is then

\[
L_i = \left\{ \prod_{j=1}^{n_i} \rho(T_{ij}) \right\} \exp \left( -\int_{0}^{\tau_i} \rho(u) \, du \right).
\] (2.3)

With local likelihood, we approximate the rate function in a neighbourhood around \( t \) by a function indexed by a parameter vector \( \alpha_t = (\alpha_{t0}, \ldots, \alpha_{tp})' \). For example, we might use a \( p \)-th order polynomial function and set

\[
\log \rho(u) = \alpha_{t0} + \alpha_{t1}(u-t) + \cdots + \alpha_{tp}(u-t)^p.
\]

The local log-likelihood at \( t \) then becomes

\[
\ell(\alpha_t; t) = \sum_{i=1}^{m} \left\{ \sum_{j=1}^{n_i} [K_b(T_{ij} - t) \log(\rho(T_{ij}; \alpha_t))] - \int_{0}^{\tau_i} K_b(u-t)\rho(u; \alpha_t) \, du \right\}
\] (2.4)

where \( K_b(t) = K(t/b)/b \) and \( K(u) \) is a kernel function. Typical examples of kernel functions are the rectangular kernel where \( K(u) = 1/2 \) for \( |u| < 1 \), the Gaussian kernel where \( K(u) = 1/\sqrt{2\pi} \exp(-u^2/2) \), and the Epanechnikov kernel where \( K(u) = 3/4(1-u^2) \) for \( |u| < 1 \); we use the latter in this paper. For a given kernel function, estimates of \( \alpha_t \) are obtained by solving the \( p+1 \) estimating equations,

\[
U_{tq}(\alpha_t) = \frac{\partial \ell(\alpha_t; t)}{\partial \alpha_{tq}} = 0,
\] (2.5)

where \( U_{tq} \) is obtained by differentiating (2.4) with respect to the \( q \)-th component of \( \alpha_t \).

Multiplicative models of the form \( \rho_t(t) = \rho_0(t) \exp(\mathbf{x}_t'\beta) \) are perhaps the most common way of assessing covariate effects. Here we propose estimation of \( \beta \) via profile likelihood. An estimate of \( \rho_0(t) \) for fixed \( \beta \), denoted \( \tilde{\rho}_0(\cdot; \beta) \), is obtained by maximizing the local log-likelihood

\[
\ell(\alpha_t, \beta; t) = \sum_{i=1}^{m} \left\{ \sum_{j=1}^{n_i} [K_b(T_{ij} - t) [\log(\rho(T_{ij}; \alpha_t) + \mathbf{x}_t'\beta]]
\right. \left. - \exp(\mathbf{x}_t'\beta) \int_{0}^{\tau_i} K_b(u-t)\rho(u; \alpha_t) \, du \right\}
\] (2.6)
with respect to $\alpha_t$ to get $\tilde{\alpha}_t(\beta)$. One can then insert the estimate $\hat{\rho}_0(\cdot; \beta) = \rho_0(\cdot; \tilde{\alpha}_t(\beta))$ into the log-likelihood (2.6) to get

$$
\ell(\hat{\rho}_0(\cdot; \beta), \beta) = \sum_{i=1}^{m} \left\{ \sum_{j=1}^{n_i} [\log \hat{\rho}_0(T_{ij}; \beta) + \mathbf{x}'_i \beta] - \exp(\mathbf{x}'_i \beta) \hat{\mu}_i(\tau_i; \beta) \right\}.
$$

(2.7)

The pseudo-profile log-likelihood, $\ell(\hat{\rho}_0(\cdot; \beta), \beta)$ is then viewed as a function of $\beta$ and maximized to obtain $\hat{\beta}$. In the case of a single covariate, confidence intervals for $\beta$ can be obtained by finding the $\beta$ that satisfy

$$
2[\ell(\hat{\rho}_0(\cdot; \hat{\beta}), \hat{\beta}) - \ell(\hat{\rho}_0(\cdot; \beta), \beta)] \leq \chi^2_{1, \alpha}.
$$

We give empirical evidence regarding the validity of the chi-squared assumption for interval-censored data in Section 4.

2.3 Local Likelihood Methods for Data from Overdispersed Poisson Processes

The Poisson model is restrictive in the sense that $\text{var}\{N_i(t)\} = \mu_i(t)$. If there is evidence of extra-Poisson variation, the alternative variance function $\text{var}\{N_i(t)\} = \nu \mu_i(t)$ is often used in generalized linear models. In this case, the Poisson log-likelihood is multiplied by $\nu^{-1}$ so the estimate of $\beta$ is unaffected but the asymptotic variance of $\hat{\beta}$ is inflated by $\nu$. The variance inflation factor $\nu$ can be estimated by the method of moments (McCullagh and Nelder, 1989) as

$$
\hat{\nu} = \frac{1}{n} \sum_{i=1}^{m} \frac{(n_i - \hat{\mu}_i(\tau_i))^2}{\hat{\mu}_i(\tau_i)},
$$

(2.8)

and the adjusted profile likelihood interval then becomes the set of $\beta$ which satisfy

$$
2[\ell(\hat{\rho}_0(\cdot; \hat{\beta}), \hat{\beta}) - \ell(\hat{\rho}_0(\cdot; \beta), \beta)] \leq \hat{\nu} \cdot \chi^2_{1, \alpha}
$$

where $\ell(\hat{\rho}_0(\cdot; \beta), \beta)$ is (2.7).

Mixed-Poisson models offer an alternative strategy for dealing with extra-Poisson variation. Let $u_i$ denote an individual-specific random effect with $E(u_i) = 1$ and $\text{var}(u_i) = \phi$ and $\lambda_i(t|H_i(t), u_i) = u_i \rho_i(t)$ denote the conditional (given $u_i$) Poisson intensity function. Unconditionally $E\{N_i(t)\} = \mu_i(t)$ and $\text{var}\{N_i(t)\} = \mu_i(t) + \mu_i^2(t)\phi$, so $\phi > 0$ corresponds to the setting with extra-Poisson variation. Taking $u_i$ as independent and identically distributed gamma random variables is most common, in which case the events are generated according to a negative binomial process.
Inference can proceed as follows. If an estimate of $\rho_0(t)$ for given $\beta$ is obtained using the Poisson model of Section 2.2, an estimate of $\phi$ can be obtained by the method of moments as the solution to

$$\sum_{i=1}^{m} \left\{ \frac{(n_i - \hat{\mu}_i)^2 - \hat{\mu}_i(1 + \phi \hat{\mu}_i)}{(1 + \phi \hat{\mu}_i)^2} \right\} = 0,$$

where $\hat{\mu}_i = \hat{\mu}_i(\tau_i; \beta)$ (Dean, 1991), and we denote this estimate as $\hat{\phi} = \hat{\phi}(\hat{\mu}, \beta_0)$ where $\hat{\mu} = (\hat{\mu}_1, \ldots, \hat{\mu}_n)'$.

We let

$$\ell(\beta) = \sum_{i=1}^{m} \left\{ \sum_{j=1}^{n_i} \log \hat{\rho}_0(T_{ij}) + n_i (\log \hat{\phi} + \mathbf{x}_i' \beta) - (n_i + \hat{\phi}^{-1}) \log (1 + \hat{\phi} \hat{\mu}_i(\tau_i)) \right\} \log \Gamma(n_i + \hat{\phi}^{-1}) - \log \Gamma(\hat{\phi}^{-1}) \right\},$$

and refer to this as a pseudo-profile likelihood obtained by substituting $\hat{\mu}$ and the moment estimate $\hat{\phi}$ into the negative binomial log-likelihood. An estimate of $\beta$ is then obtained by maximizing (2.10).

3. Local Likelihood for Rate Function Estimation with Interval Count Data

As in Section 2 we begin by considering the case of no covariates. Suppose individuals are only seen at periodic inspection times which vary from subject to subject. Let $0 = b_{i1} < \cdots < b_{imi} = \tau_i$ denote the $m_i$ inspection times for individual $i$, $i = 1, \ldots, m$. We assume here that the inspection process is ignorable in the sense of Grüger et al. (1991). Here the aim is to develop a local likelihood EM algorithm to estimate the rate under a Poisson model in a one-sample problem. This is in the spirit to the EM approach of Lawless and Zhan (1998), but the advantage here is that one can obtain a smooth estimate of the rate function $\rho(t)$.

Let $T_{i1}, \ldots, T_{imi}$ denote the times of the $n_i$ events experienced by individual $i$ over the period of observation from 0 to $\tau_i$. Let $(L_{ij}, R_{ij}]$ be the censoring interval for $T_{ij}$. These censoring intervals are created as a result of the recurrent event process only being observed at the assessment times. The complete data likelihood contribution from individual $i$ is given by (2.3). Let the baseline rate function around $t$ be approximated by a function dependent on a parameter vector $\alpha_i = (\alpha_{i0}, \ldots, \alpha_{ip})'$ so the local log-likelihood at $t$ is given by (2.4). The parameters can now be estimated by an EM-type algorithm which integrates smoothing
at the E-step as proposed by Braun et al. (2005) for the case of density estimation and Betensky et al. (1999) for interval-censored survival data.

A grid of equally spaced points at which the baseline hazard will be estimated, \( t_1, \ldots, t_M \), must be defined. Let \( \hat{\rho} \) denote the estimate of \( \rho(t) \) at each of the grid points, so the \( g \)th element of \( \hat{\rho} \) depends on \( \alpha_{tg} \). Define \( Q^{(r)}(\alpha_t; \hat{\rho}^{(r-1)}) \) as

\[
\sum_{i=1}^{m} \left\{ \sum_{j=1}^{n_i} \left[ E[K_b(T_{ij} - t)|\hat{\rho}^{(r-1)}, L_{ij}, R_{ij}] \log \rho(T_{ij}; \alpha_t) \right] - \int_0^{\tau_i} K_b(u - t) \rho(u; \alpha_t) \, du \right\} \tag{3.1}
\]

where parameters with a superscript \((r-1)\) indicate the parameter estimate at the \((r-1)\)st iteration. The expectations involving the \( T_{ij} \) must be evaluated numerically, for example using the trapezoid rule. To obtain the conditional density of the \( T_{ij} \) given the data, it is useful to transform the times by defining \( S_{ij} = \mu(T_{ij}) \). The \( S_{ij} \) then follow a homogeneous Poisson process with rate 1. Due to the independent increments property of Poisson processes the density of \( S_{ij} \) given it occurred in the interval \((\mu(b_{ik-1}), \mu(b_{ik})]\) does not depend on \( \mu(b_{ik-1}) \) and \( b_{ik} \). The M-step is to maximize \( Q^{(r)}(\alpha_t; \hat{\rho}^{(r-1)}) \) to obtain \( \alpha_{ti}^{(r)} \) at each grid point using, for example, a Newton-Raphson algorithm. The EM algorithm is continued until the difference between successive parameter estimates becomes negligible.

When interest lies in the effect of fixed covariates, multiplicative models can again be used. Estimation of \( \rho_0(\cdot) \) and \( \beta \) can be done using an approach analogous to the profile likelihood as described for the right-censored case. With interval censoring, the log-likelihood for the total event counts is

\[
\ell(\rho_0(\cdot; \beta), \beta) = \sum_{i=1}^{m} \sum_{j=1}^{n_i} \left\{ n_{ij}(\log \mu_{ij} + x_i'\beta) - \exp(x_i'\beta) \mu_{ij} \right\} \tag{3.2}
\]

where \( n_{ij} \) is the number of events in \((b_{ij-1}, b_{ij}]\) and \( \mu_{ij} = \int_{b_{ij-1}}^{b_{ij}} \rho_0(u) \, du \). In order to obtain \( \hat{\rho}(\cdot; \beta) \) an EM algorithm can be used with \( \beta \) at a fixed value. The complete data likelihood contribution from individual \( i \) is

\[
\left\{ \prod_{j=1}^{n_i} \rho_0(T_{ij}) \exp(x_i'\beta) \right\} \exp \left( -\int_0^{\tau_i} \rho_0(u) \exp(x_i'\beta) \, du \right) \tag{3.3}
\]
and the log-likelihood is,

$$\ell_C(\rho_0(\cdot), \beta) = \sum_{i=1}^{m} \left[ \sum_{j=1}^{n_i} \left[ \log \rho_0(T_{ij}) + x'_i\beta \right] - \exp(x'_i\beta) \int_{0}^{T_i} \rho_0(u) \, du \right].$$

(3.4)

Again, we let the baseline rate function around $t$ be approximated by a function dependent on a parameter vector $\alpha_t$, then the local log-likelihood at $t$ becomes,

$$\ell_C(\alpha_t, \beta) = \sum_{i=1}^{m} \left( \sum_{j=1}^{n_i} K_b(T_{ij} - t) \left[ \log \rho_0(T_{ij}; \alpha_t) + x'_i\beta \right] - \exp(x'_i\beta) \int_{0}^{T_i} K_b(u - t) \rho_0(u; \alpha_t) \, du \right)$$

Let $\hat{\rho}_0$ denote the estimate of $\rho_0$ at each of the grid points, so the $g$th element of $\hat{\rho}_0$ depends on $\hat{\alpha}_{tg}$. Define $Q^{(t)}(\alpha_t, \beta; \hat{\rho}_0^{(r-1)}, \hat{\beta}^{(r-1)})$ as

$$\sum_{i=1}^{m} \left( \sum_{j=1}^{n_i} E[K_b(T_{ij} - t)|\hat{\rho}_0^{(r-1)}, \hat{\beta}^{(r-1)}, L_{ij}, R_{ij}] \left[ \log \rho_0(T_{ij}; \alpha_t) + x'_i\beta \right] - \exp(x'_i\beta) \int_{0}^{T_i} K_b(u - t) \rho_0(u; \alpha_t) \, du \right)$$

(3.5)

where parameters with a superscript $(r-1)$ indicate the parameter estimate at the $(r-1)$st iteration. Again, the expectations involving the $T_{ij}$ must be evaluated numerically. The M-step is to maximize $Q^{(t)}(\alpha_t, \beta; \hat{\rho}_0^{(r-1)}, \hat{\beta}^{(r-1)})$ to obtain $\hat{\alpha}_t^{(r)}$ at each grid point. The EM algorithm is continued until the difference between successive parameter estimates becomes negligible.

The pseudo-profile log-likelihood, $\ell(\hat{\rho}_0(\cdot; \beta), \beta)$ is then maximized to obtain $\hat{\beta}$. In the one-parameter case, confidence intervals for $\beta$ can be easily constructed by finding the $\beta$ values that satisfy $2[\ell(\hat{\rho}_0(\cdot; \hat{\beta})), \hat{\beta}) - \ell(\hat{\rho}_0(\cdot; \beta), \beta)] \leq \chi^2_{1, \alpha}$. To accommodate extra-Poisson variation, the methods of Section 2.3 can be employed with variance functions of the form $\text{var}(N_i(t)) = \nu\mu_i(t)$ or $\text{var}(N_i(t)) = \mu_i(t) + \mu_i^2(t)\phi$. Alternatively, resampling techniques such as the bootstrap may be considered.

4. Simulation Studies

4.1 Simulations Assessing the Distribution of Profile Local Likelihood

Here we report on a simulation study to assess the validity of the chi-square approximation for the distribution of the profile likelihood ratio statistic under interval-censored case. Here we consider the case of a single binary covariate with takes the value of 1 with probability
0.5 and is zero otherwise; the value of $\beta$ was set to log 0.75 for a strong covariate effect. Recurrent events were generated according to a Poisson process with $\rho_i(t) = \rho_0(t) \exp(\beta x_i)$ where $\rho_0(t) = \gamma \theta(\theta t)^{-1}$. We set $b_{i1} = 0$, but generated the followup assessment times $b_{i2}, \ldots, b_{im_i}$ according to a time-homogeneous Poisson process over the interval $(0,1]$ with mean 4. One thousand datasets composed of 500 subjects were simulated. Poisson local likelihood methods with nearest-neighbours bandwidths of $b = 0.2$ and $b = 0.6$ were used for estimation. Figure 1 shows QQ-plots of the simulated quantiles of the profile likelihood ratio statistic based on (3.2) versus the theoretical quantiles of the $\chi^2_1$ distribution along with 95% pointwise confidence intervals for a range of trend parameters ($\gamma$) and the two bandwidths. The QQ-plots were obtained using the `qq.plot` function in the R library `car` (Fox, 2007). There appears to be good agreement between the empirical and chi-square (1 df) quantiles for most configurations since the empirical quantiles are within the “confidence envelope” (Fox, 1997) in all cases except $\gamma = 1.2$, $b = 0.2$. Omnibus Kolmogorov-Smirnov tests of the null hypothesis that the profile likelihood ratio statistic

$$2[\ell(\hat{\rho}_0(\cdot; \hat{\beta}), \hat{\beta}) - \ell(\hat{\rho}_0(\cdot; \beta), \beta)]$$

follows a $\chi^2_1$ distribution is given in Table 1 and as can be seen there is insufficient evidence to reject this null hypothesis for any of the parameter configurations considered here. The validity of this distributional assumption is further substantiated empirically by the good empirical coverage probability of the 95% profile likelihood ratio confidence intervals. This empirical investigation has limited scope and so in practical applications it may be worthwhile to use bootstrap methods to corroborate confidence intervals obtained by profile likelihood.

4.2 Empirical Studies of Local Likelihood Estimators

Simulation studies were conducted to assess the performance of the Poisson local likelihood estimator. Data was generated over the interval $(0,1]$ such that the mean number of events experienced by an individual was 4. The underlying processes considered were Poisson, mixed-Poisson and renewal processes. For the Poisson processes, the mean function took the form $(\theta t)^\gamma$ as in Section 4.1. The values of $\gamma$ were taken to be 1, 0.75 and 1.2 in order to examine the effects of a trend in the intensity. The same form was used as the baseline mean function conditional on the random effect in the mixed-Poisson case, while the random
effect was taken to be gamma distributed with mean 1 and variance 0.25. The interarrival
distribution for the renewal process was taken to be gamma with shape parameter 2. The
scale parameters in all cases were chosen such that $\mu(1) = 4$. The assessment times were
generated according to a homogeneous Poisson process with mean number of visits equal to
5 or 10.

For comparison, the Poisson local likelihood estimator was compared with estimates
obtained using a piecewise constant rate function with 3 or 6 pieces. In the case of mean
function estimation, comparisons can also be made with a simple estimator of the mean
function based on state occupancy probabilities in a multistate model. If we consider a
recurrent event process as a multistate model with states defined by the number of events
an individual has experienced, then estimates of the state occupancy probabilities can be
obtained as in Pepe et al. (1991). If $Z(t)$ indicates the state an individual is in at time
$t$ then $P(Z(t) = k) = P(T_{k+1} > t) - P(T_k > t)$ where $T_k$ denotes the time of the $k$th
event and $T_0 = 0$. The marginal survivor functions can be estimated separately using local
likelihood as in Betensky et al. (1999) and an estimate of the state occupancy probabilities
can be obtained by taking the appropriate difference. An estimate of the mean function is
obtained by $\tilde{\mu}(t) = \sum_{j=1}^{J} j \tilde{p}_j(t)$ where $\tilde{p}_j(t)$ is the prevalence function estimate for state $j,$ $j = 1, \ldots, J.$

A locally constant polynomial and 20% “nearest neighbors” bandwidth Loader (1999)
was used in all cases. The estimators were examined in terms of integrated mean squared
error and integrated absolute bias and the results of the simulations are displayed in Table 2.
Figures 2 and 3 graphically display the performance of the estimators over the relevant
range.

Examination of Table 2 shows that the local likelihood estimator outperforms the Pepe
estimator in all situations. It can also be seen that the MSE and bias of the local likelihood
estimator are close to that of the piecewise constant estimators. The local likelihood estima-
tor tends to perform slightly better than the 3-piece model, while the 6-piece model tends to
perform slightly better than the local likelihood estimator. Figure 2 displays the bias of the
estimators. In all cases, the bias of the Pepe estimator is largest. All methods suffer early
bias when the model is not time homogeneous. The piecewise constant methods have large
bias early on, which gets smaller with time, while the local likelihood estimator has smaller
initial bias, but there is a slight amount of bias that persists over time. Figure 3 displays the
performance of the local likelihood estimator versus the piecewise constant estimators of the rate function. The local likelihood estimate tracks the true rate function very closely except very early on near the boundary at zero. It is clear from this figure that the local likelihood method provides a smooth estimate as compared with the piecewise constant estimates.

Simulations were also conducted to assess the performance of the estimators of a regression coefficient. This was done for the right-censored case. The true value of the regression coefficient was set to $\beta = \log 0.75 = -0.288$ with covariate values taking the values 1 and 0 with equal probabilities. Baseline rate functions considered were the same as for the previous case. Poisson and mixed-Poisson models were used to generate the data with an average of two or four events over the interval $(0, 1]$. For the local likelihood methods, the mixed-Poisson procedure involving estimation of $\phi$ was used when the data were generated according to a mixed-Poisson process. The results are displayed in Table 2. For comparison, results obtained using a piecewise constant rate function (with four or ten pieces) and robust variance estimates are also displayed.

The biases are all small relative to their standard errors. The coverage probabilities are all close to the nominal level of 0.95. When the Poisson local likelihood approach is used for data generated according to a mixed-Poisson process the bias is still small, but the coverage probability drops significantly (results not shown in table). It is also worth noting that using a variance function of the form $\text{var}(N_i(t)) = \nu \mu_i(t)$ results in estimators with little bias and coverage close to nominal (results not shown in table).

Insert Table 2 and Figures 1, 2 and 3 about here.

5. Applications

5.1 Data on the Reliability of Nuclear Plant Safety

Sun and Kalbfleisch (1995) consider an example in which 30 nuclear plants are examined for losses of feedwater flow. The cumulative number of episodes of losses in flow were recorded at a single inspection time for each plant, creating current status data (Sun, 2006). The inspection times ranged from 1 to 15 years (with quantiles of 2.0, 3.5 and 5.0 years), and the cumulative number of episodes of losses in feedwater flow ranged from 0 to a maximum of 58. The raw data are plotted in Figure 4 with the total number of losses on the vertical
axis and the duration of operating time on the horizontal axis, using open circles. Figure 4 also contains the consistent nonparametric estimate of the mean function derived by Sun and Kalbfleisch (1995) obtained using isotonic regression methodology (Barlow et al., 1972) to ensure the estimate is non-decreasing. A limitation of this estimate is that it is only defined at the times of inspections; the estimate at these times is denoted by the filled in circle.

Two local likelihood estimate of the mean function are superimposed using the Epanechnikov kernel with 100 grid points. The dashed line corresponds to a bandwidth of 1.0 while the solid line corresponds to a bandwidth of 0.33 chosen by cross validation. This bandwidth was obtained by finding the $b$ which maximizes the likelihood criterion given by

$$\sum_{i=1}^{n} \left\{ n_i \log \hat{\mu}_i(\tau_i; b) - \hat{\mu}_i(\tau_i; b) \right\}$$

where $\hat{\mu}_i(t; b)$ is the local likelihood estimate of $\mu_i(t)$ with bandwidth $b$ and $\tau_i$ is the assessment time for the $i$th plant. We considered the use of other kernels but the resulting estimates were also comparable.

The local likelihood estimate can be seen to be non-decreasing by construction, and it also tracks the nonparametric estimate of Sun and Kalbfleisch (1995) very well. This estimate indicates that the average number of feedwater flow losses per nuclear plant at each time point is increasing fairly steadily over time; specifically the plot suggests that the rate of occurrence of loss of feedwater flow begins to plateau at just after five years and increases more rapidly after about 7 years.

5.2 Psoriatic Arthritis

The University of Toronto Psoriatic Arthritis Clinic was established in 1978 to provide detailed prospective data on the progression of psoriatic arthritis in patients attending this tertiary referral center (Gladman et al., 1995; Gladman et al., 1998). According to the protocol of this clinic, participating patients are scheduled to be examined at annual clinic visits. In addition to completing a detailed questionnaire at these visits, patients underwent a detailed clinical examination of their joints, which were classified according to the Sharp scoring system (Sharp et al., 1971). The event of interest here is the development of damage in joints and interest lies in the rate of occurrence of damage and the expected cumulative
number of damaged joints over time. Moreover, identification of important covariate effects is also of interest to help characterize risk among patients presenting in the clinic. Here we analyse data from this clinic as of March 21, 2005, which involves 623 patients. The mean time from clinic entry to closure of the database is 8.8 years (ranging from 0.10 to 31.01).

While assessments are scheduled annually, in reality there is considerable variability in the frequency and timing of clinic visits. Figure 5 shows the times of assessments for a sample of patients, along with counts of the number of damaged joints occurring between assessments. As can be seen, some patients adhere to a regular schedule of clinic visits and others attend less frequently and more irregularly.

Estimates of the expected cumulative number of damaged joints over time are given in Figure 6 based on Poisson models via the local likelihood approach of Section 3 as well as based on piecewise constant models of Lawless and Zhan (1998) with 3 and 6 pieces; 95% bootstrap confidence intervals are also provided for the local likelihood approach. The local likelihood estimate, based on an Epanechnikov kernel, a bandwidth of 3, and a locally constant function, agrees very well with the piecewise estimates over the majority of the followup. There is a slight divergence in these estimates towards the end of followup due to boundary problems commonly observed in local likelihood methods (Loader, 1999).

A multiplicative mixed Poisson model was fit using (2.11) for estimation and bootstrap resampling for estimation of standard errors and confidence intervals. The covariates included duration of arthritis (years), sex (1=male, 0=female), and family history of psoriasis (1=yes, 0=no). The findings show that the rate of damaged joints increases with each additional year of disease (RR=1.026; 95% CI (1.014, 1.038)), and that males have a lower rate of damage than females (RR=0.698; 95% CI (0.521, 0.934)); there is no significant effect of family history of disease.

Insert Figures 5 and 6 about here.

6. Discussion
This paper explores the use of local likelihood methods for the analysis of right-censored and interval-censored recurrent event data. For right-censored data, we have found (results not reported on here) that the estimates of rate and mean functions agree very well with weakly parametric methods based on piecewise constant models, and the regression coefficients are in close agreement with those from an Andersen-Gill regression model (Andersen and Gill,
The greatest appeal in using local likelihood methods is for the setting with interval censored data, where smooth estimates of rate and mean functions are readily obtained and estimates of regression coefficients are again in close agreement with those from piecewise constant models.

Variance estimation remains a challenge in this setting as empirical studies have shown that sandwich variance estimates do not behave well in the settings we considered. Pseudo-profile likelihood based confidence interval estimation performs well, however, and is reasonably computationally efficient for both Poisson and mixed Poisson settings. Moment estimation of the variance parameter appears preferable to profiling it out, and simulations confirm this approach performs well. Local likelihood methods could also be adapted to deal with multi-type interval-censored recurrent event as discussed in Chen et al. (2005). In this case, the moment estimates of the variance parameters can be supplemented with moment estimates of covariance parameters; see Cook and Lawless (2007) for related expressions.

In many settings recurrent event data are both censored and left truncated. This occurs when subjects must have experienced one or more events, for example, to enter into a particular database. The algorithm of Sections 2 and 3 can be readily adapted to deal with truncation in the spirit of Turnbull (1976). The use of local likelihood would be particularly appealing in this case as it would furnish smooth estimates over the positive real line. Another common complication is the presence of dependent termination (Cook and Lawless, 1997; Strawderman, 2000). Relatively little work has been done in this area for interval censored recurrent event data, and this warrants methodologic development; local likelihood represents one promising area of investigation.

Recurrent event data can also be viewed from a multistate perspective in which the states correspond to the number of events experienced over time. In this case, one can think of transition intensities rather than marginal event rates, and estimate these under Markov assumptions using local likelihood methods for right or interval censored event time data. This approach is taken in Cook et al. (2008) for right censored data, where the robustness of the Aalen-Johansen estimator of the state occupancy probabilities (Aalen et al., 2001; Datta and Satten, 2001) yield consistent estimates of the mean function. Again, however, further work is warranted in the context of interval censored data.
ACKNOWLEDGEMENTS

The authors thank Dr. Dafna Gladman for permission to use the data from the University of Toronto Psoriatic Arthritis Clinic and for helpful comments. This research was supported by grants from the Canadian Institutes for Health Research and the Natural Sciences and Engineering Research Council of Canada. R.J. Cook is a Canada Research Chair in Statistical Methods for Health Research.

REFERENCES


Figure 1. QQ-plots comparing the simulated quantiles of the profile likelihood ratio statistic with the quantiles of the $\chi^2_1$ distribution based on 1000 simulated datasets of 500 subjects; Poisson data with mean function $(\theta t)^{\gamma}$.
Figure 2. Bias of the Poisson local likelihood and piecewise constant estimates, and Pepe estimates of the mean function for interval-censored count data; visits were distributed according to a Poisson process with a mean of 5 visits over (0, 1].
Figure 3. Poisson local likelihood and piecewise constant estimates of the rate function for interval-censored count data; visits were distributed according to a Poisson process with a mean of 5 visits over (0, 1].
Figure 4. Estimates of the mean number of episodes of feedwater loss based on nonparametric estimator of Sun and Kalbfleisch (1995) and local likelihood (dashed lines).
Figure 5. Timeline diagrams for followup assessments for a sample of patients from the University of Toronto Psoriatic Arthritis Clinic; vertical lines indicate times of assessments and numbers indicate the number of new damaged joints detected over the corresponding interval.
Figure 6. Poisson local likelihood and piecewise constant estimates (3 and 6 pieces) of the mean cumulative number of damaged joints among patients in the University of Toronto Psoriatic Arthritis Clinic.
Table 1

Kolmogorov-Smirnov test statistics (KS) and p-values for assessing the validity of the $\chi^2_1$ approximation to the profile likelihood ratio statistic based on 1000 simulated datasets of 500 subjects; Poisson data with mean function $(\theta t)^\gamma$.

<table>
<thead>
<tr>
<th>$\gamma$</th>
<th>$b$</th>
<th>KS</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.75</td>
<td>0.2</td>
<td>0.029</td>
<td>0.390</td>
</tr>
<tr>
<td></td>
<td>0.6</td>
<td>0.019</td>
<td>0.872</td>
</tr>
<tr>
<td>1.00</td>
<td>0.2</td>
<td>0.020</td>
<td>0.818</td>
</tr>
<tr>
<td></td>
<td>0.6</td>
<td>0.014</td>
<td>0.991</td>
</tr>
<tr>
<td>1.20</td>
<td>0.2</td>
<td>0.038</td>
<td>0.106</td>
</tr>
<tr>
<td></td>
<td>0.6</td>
<td>0.016</td>
<td>0.952</td>
</tr>
</tbody>
</table>
**Table 2**

Integrated mean squared error and integrated absolute bias for Poisson-based local likelihood, Pepe local likelihood, and piecewise (PW) constant methods with three and six pieces.

<table>
<thead>
<tr>
<th></th>
<th>Five Visits</th>
<th></th>
<th></th>
<th>Ten Visits</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Local Likelihood</td>
<td></td>
<td></td>
<td>Local Likelihood</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Poisson</td>
<td>Pepe</td>
<td>PW3</td>
<td>PW6</td>
<td>Poisson</td>
</tr>
<tr>
<td>IMSE γ = 0.75</td>
<td>0.00619</td>
<td>0.00863</td>
<td>0.00901</td>
<td>0.00575</td>
<td>0.00600</td>
<td>0.01317</td>
</tr>
<tr>
<td>IMSE γ = 1.0</td>
<td>0.00481</td>
<td>0.00834</td>
<td>0.00443</td>
<td>0.00491</td>
<td>0.00418</td>
<td>0.01095</td>
</tr>
<tr>
<td>IMSE γ = 1.2</td>
<td>0.00453</td>
<td>0.00838</td>
<td>0.00476</td>
<td>0.00466</td>
<td>0.00398</td>
<td>0.01013</td>
</tr>
<tr>
<td>Mixed Poisson Processes</td>
<td>0.01007</td>
<td>0.01340</td>
<td>0.01283</td>
<td>0.00956</td>
<td>0.00945</td>
<td>0.02183</td>
</tr>
<tr>
<td>Renewal Processes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IMSE</td>
<td>0.00334</td>
<td>0.00739</td>
<td>0.00446</td>
<td>0.00304</td>
<td>0.00340</td>
<td>0.00784</td>
</tr>
<tr>
<td>IAE γ = 0.75</td>
<td>0.03431</td>
<td>0.05340</td>
<td>0.04132</td>
<td>0.01366</td>
<td>0.03013</td>
<td>0.08439</td>
</tr>
<tr>
<td>IAE γ = 1.0</td>
<td>0.01153</td>
<td>0.04602</td>
<td>0.00107</td>
<td>0.00093</td>
<td>0.01027</td>
<td>0.06899</td>
</tr>
<tr>
<td>IAE γ = 1.2</td>
<td>0.00800</td>
<td>0.04668</td>
<td>0.01653</td>
<td>0.00495</td>
<td>0.01073</td>
<td>0.06492</td>
</tr>
<tr>
<td>Mixed Poisson Processes</td>
<td>0.03408</td>
<td>0.06171</td>
<td>0.04126</td>
<td>0.01382</td>
<td>0.03879</td>
<td>0.10949</td>
</tr>
<tr>
<td>Renewal Processes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IAE</td>
<td>0.02218</td>
<td>0.05301</td>
<td>0.03002</td>
<td>0.01046</td>
<td>0.03082</td>
<td>0.06293</td>
</tr>
</tbody>
</table>