Title	Dynamic Random Effects Models for Times between Repeated Events
Author(s)	Fong, DYT; Lam, KF; Lawless, JF; Lee, YW
Citation	Lifetime Data Analysis, 2001, v. 7 n. 4, p. 345-362
Issued Date	2001
URL	http://hdl.handle.net/10722/48490
Rights	Creative Commons: Attribution 3.0 Hong Kong License

Dynamic Random Effects Models for Times Between Repeated Events

Daniel Y.T. Fong

Clinical Trials Centre, Faculty of Medicine,

The University of Hong Kong, Pokfulam Road, Hong Kong

K. F. Lam

Department of Statistics & Actuarial Science,

The University of Hong Kong, Pokfulam Road, Hong Kong

J. F. Lawless

Department of Statistics & Actuarial Science,

University of Waterloo, Waterloo, Ont, Canada N2L 3G1

and

Y. W. Lee

Department of Statistics & Actuarial Science,

The University of Hong Kong, Pokfulam Road, Hong Kong.

February 14, 2000

Acknowledgement

The work described in this paper was substantially supported by grants from the Natural Sciences and Engineering Research Council of Canada, and the Research Grants Council of the Hong Kong Special Administrative Region, China (Project No. HKU 7140/98H) .

Dynamic Random Effects Models for Times Between Repeated Events

October 23, 2000

Abstract

We consider recurrent event data when the duration or gap times between successive event occurrences are of intrinsic interest. Subject heterogeneity not attributed to observed covariates is usually handled by random effects which result in an exchangeable correlation structure for the gap times of a subject. Recently, efforts have been put into relaxing this restriction to allow non-exchangeable correlation. Here we consider dynamic models where random effects can vary stochastically over the gap times. We extend the traditional Gaussian variance components models and evaluate a previously proposed proportional hazards model through a simulation study and some examples. Besides, semiparametric estimation of the proportional hazards models is considered. Both models are easily used. The Gaussian models are easily interpreted in terms of the variance structure. On the other hand, the proportional hazards models would be more appropriate in the context of survival analysis, particularly in the interpretation of the regression parameters. They can be sensitive to the choice of model for random effects but not to the choice of the baseline hazard function.

Keywords : Censoring; Dynamic Random Effects; Proportional Hazards Models; Recurrent Events.

1 Introduction

Equipment or subjects in clinical studies may experience recurrences of certain events. Examples include repeated failures in equipment (e.g. Follmann and Goldberg, 1988), the occurrence of tumors in animals or humans (e.g. Gail *et al.*, 1980; Byar, 1980), and cyclic behavior associated with human digestion, in which case the events correspond to the beginning of cycles (Aalen and Husebye, 1991). Various approaches are used to analyze recurrent events data including methods based on event intensity, rate and mean functions, times between events, and times to events. Lawless (1995) provides a review. The current paper follows the tradition of Gail *et al.* (1980), Prentice *et al.* (1981), Follmann and Goldberg (1988), and Aalen and Husebye (1991) who considered models based on the gaps times, or times between successive events. Lawless and Fong (1999) discussed models for sequences of gap times or duration times in a broader context.

Suppose that there are N subjects in a study and subject i is observed over the time interval $(0, \tau_i]$ and let $T_{i1} < T_{i2} < \cdots$ denote event occurrence times (for $i = 1, 2, \ldots, N$). The gap times are given by $S_{ij} = T_{ij} - T_{i,j-1}$ ($j = 1, 2, \ldots, n_i$), with T_{i0} defined to be 0 and n_i the number of occurrences of the event for the i^{th} subject during the observation period. We consider situations in which there is no pronounced trend in the occurrence of events. In the absence of explanatory variables, renewal processes are sometimes suitable in which gap times are independent and identically distributed. However, the gap times of an individual subject are generally found to be dependent. One approach is to model the dependence of S_{ij} on previous gap times via regression as in Gail *et al.* (1980) and Prentice *et al.* (1981). Another approach is to accommodate the within-subject dependence by the introduction of unobservable random effects. Aalen and Husebye (1991) and Follmann and Goldberg (1988) extended the renewal process model by introducing subject-specific random effects u_i, termed "frailty" such that, conditional on u_i, the gap times S_{ij} ($j = 1, 2, \ldots$) are independently distributed. Aalen and Husebye (1991) gave an excellent discussion on such models and advocate the use of two main approaches, (i) a Gaussian variance components model for appropriately transformed S_{ij}, and (ii) a proportional hazards frailty model for the S_{ij}.

The models of Follmann and Goldberg (1988) and Aalen and Husebye (1991) assume that the random effect for each unit or subject is constant across successive gap times, but in some situations it might be desirable to allow time-varying random effects. Models with dynamic (time-evolving) random effects have been studied in repeated measures problems (e.g. Jones, 1993) and in longitudinal data problems involving counts (e.g. Jørgensen *et al.*, 1996; Lambert, 1996b, 1996a). Yue and Chan (1997) considered gap times between successive events, as we do. The purpose of this paper is to consider dynamic random effects models for gap times between successive repeated events under the following framework. Let x_{ij} be a vector of covariates and $\underset{\sim}{u}_i = \{u_{i1}, u_{i2}, \ldots\}$ be a sequence of random effects of arbitrary length where u_{ij} can be regarded as a quantification of the missing covariates associated with S_{ij}. Denote $\{u_{i1}, \ldots, u_{ij}\}$ by u_i^j and $\{S_{i1}, \ldots, S_{ij}\}$ by S_i^j. All models considered in this paper satisfy the conditions

$$Pr\{S_{ij} \,|\, S_i^{j-1}, x_{ij}, u_i^j\} = Pr\{S_{ij} \,|\, x_{ij}, u_{ij}\} \quad (1)$$

$$Pr\{u_{ij} \,|\, S_i^{j-1}, x_{ij}, u_i^{j-1}\} = Pr\{u_{ij} \,|\, u_i^{j-1}\} \quad (2)$$

where "Pr" is used to represent a probability density or mass function. Conditions (1) and (2) specify the state-space models where, conditional on the u_{ij}'s and x_{ij}'s, the gap times $S_{i1}, S_{i2}, \ldots$ are independent. The models considered by Follmann and Goldberg (1988) and Aalen and Husebye (1991) are special cases with $u_{ij} = u_i$ ($j = 1, 2, \ldots$).

The state-space models considered here differ from many state-space settings, since each subject possesses an independent stochastic process. There has been extensive research in the area of general state-space models. For example, Carlin *et al.* (1992) and Shephard and Pitt (1997) considered non-linear and non-Gaussian state-space models respectively and developed estimation procedures using Gibbs sampling techniques. A recent paper by So (1999) introduced a class of state-space models when the unobservable state variable, i.e. u_{ij}, can be any Gaussian stochastic processes. Thus long memory processes that have non-finite state-space form can be handled. A new efficient recursive estimation procedure utilizing Markov Chain Monte Carlo techniques is also proposed in the paper. Also, the majority of the literature on state-space models has not dealt with censored observations. Exceptions

are Yue and Chan (1997), Harvey and Fernandes (1989), and Smith and Miller (1986) who considered conjugate filtering recursions for censored observations (also see Section 3).

We have two specific objectives. The first is to extend the families of models of Follmann and Goldberg (1988) and Aalen and Husebye (1991), and to consider the pros and cons of Gaussian versus proportional hazards formulations. It is observed that Gaussian models are easily used and interpreted in terms of the variance structure, but the proportional hazards models would sometimes be more appropriate in the context of survival analysis, particularly in the interpretation of the regression parameters. The second objective is to examine the dynamic random effects proportional hazards model proposed by Yue and Chan (1997). An EM-algorithm, similar to the method in Klein (1992), is suggested to estimate the parameters; this works in both the parametric and semiparametric settings. The organization of the paper is as follows. We consider the Gaussian variance components models and the proportional hazards models in Sections 2 and 3, respectively. The proposed estimation procedure for Yue and Chan's proportional hazards model (1997) will be discussed in Section 4. The models will be fitted to some real examples in Section 5 with additional remarks on the utility of various models given in Section 6.

2 Gaussian Models

Let $Y_{ij} = g(S_{ij})$ be a known transformation of the gap times. Following (1) and (2), we consider here the family of models

$$Y_{ij} = \mu_{ij} + u_{ij} + e_{ij}, \tag{3}$$

where μ_{ij} may depend on covariates x_{ij} that are assumed to remain constant over gaps, the e_{ij}'s are mutually independent $\mathcal{N}(0, \sigma_e^2)$ variables, and the u_{ij}'s ($j = 1, 2, \ldots$) follow a first order autoregressive (AR(1)) process with $u_{i1} \sim \mathcal{N}(0, \omega^2)$ and

$$u_{ij} = \phi u_{i,j-1} + \epsilon_{ij} \quad j = 2, 3, \ldots \tag{4}$$

where $|\phi| < 1$ and the ϵ_{ij}'s mutually independent $\mathcal{N}(0, \sigma_\epsilon^2)$ variables. This model is non-stationary, with

$$Cov(Y_{ij}, Y_{i,j+s}) = \phi^s \left[\phi^{2(j-1)} \omega^2 + \frac{1-\phi^{2(j-1)}}{1-\phi^2} \sigma_\epsilon^2 \right] + \sigma_e^2 I(s=0), \quad s \geq 0$$

where $I(A)$ is 1 when A is true, and 0 otherwise. The stationary version of the model arises when $\omega^2 = \sigma_\epsilon^2/(1-\phi^2)$, giving

$$Cov(Y_{ij}, Y_{i,j+s}) = \phi^s \omega^2 + I(s=0)\sigma_e^2, \quad s \geq 0.$$

This family of models contains some interesting sub-models. In particular, the Gaussian model of Aalen and Husebye (1991) is obtained when $\sigma_\epsilon = 0$ and $\phi \to 1$; an AR(1) model for the Y_{ij}'s is obtained when $\sigma_e = 0$; and a renewal process is given by $\omega = 0$, $\sigma_\epsilon = 0$, in which case ϕ drops out of the model. We also note that when $\phi = 0$, the model depends only on $\sigma_e^2 + \omega^2$ and $\sigma_e^2 + \sigma_\epsilon^2$.

Estimation of the parameters of the model specified in (3) and (4) can be handled easily by computing the conditional mean and variance of Y_{ij} given its past history S_i^{j-1}. Recall that $S_{i1}, S_{i2}, \ldots, S_{i,n_i}$ are the n_i complete duration times. Also denote $S^*_{i,n_i+1} = \tau_i - T_{n_i}$ as the last censored duration time and $Y^*_{i,n_i+1} = g(S^*_{i,n_i+1})$. With the conditional moments defined as

$$E(Y_{ij} \mid S_i^{j-1}) = y_{ij|j-1} \quad \text{and} \quad Var(Y_{ij} \mid S_i^{j-1}) = \sigma^2_{iy}(j|j-1), \tag{5}$$

the observed log-likelihood function can be decomposed as

$$\begin{aligned} l \;=\; & \sum_{i=1}^{N} \left\{ -\frac{n_i}{2}\log(2\pi) - \frac{1}{2}\sum_{j=1}^{n_i} \left[\log \sigma^2_{iy}(j|j-1) + \frac{(y_{ij} - y_{ij|j-1})^2}{\sigma^2_{iy}(j|j-1)} \right] \right. \\ & \left. + \log\left[1 - \Phi\left(\frac{y^*_{i,n_i+1} - y_{i,n_i+1|n_i}}{\sigma_{iy}(n_i+1|n_i)} \right) \right] \right\} \end{aligned} \tag{6}$$

where Φ is the cumulative distribution function of the standard Normal distribution. By fixing the parameters, the conditional moments in (5) can be evaluated recursively by the cel-

ebrated linear Kalman filter (Harvey, 1989). Estimates of the parameters and their standard errors are then easily obtained by maximizing (6) either via the EM algorithm (Dempster *et al.*, 1977) or some derivative-free optimization routine (e.g. the Nelder-Mead Simplex method in SAS/IML). This family of Gaussian models interprets the covariate effects marginally, since $\mu_{ij} = E(Y_{ij})$ and the association parameters are assumed to be independent of the covariates. Such assumptions are attractive and realistic for many situations.

3 Multiplicative Hazards Models

A common approach to analyze multivariate survival time data is through the introduction of unobservable random effects into the proportional or multiplicative hazards model (e.g. Aalen and Husebye, 1991; Klein, 1992; Clayton, 1994; Xue and Brookmeyer, 1996). A convenient feature of this class of models is the ability to deal with time-varying covariates, which is a distinct advantage over the Gaussian models in Section 2. In this case, the specification of (1) is equivalent to specifying the conditional hazard function of S_{ij}, $h_{ij}(s)$, by

$$h_{ij}(s) = v_{ij} h_0(s;\, \alpha) r(x_{ij}(s);\, \beta) \tag{7}$$

where $h_0(s;\, \alpha)$ is a baseline hazard function, $r(x(s);\, \beta)$ is a non-negative-valued function of $x(s)$ that can be time-varying, $v_{ij} = \exp(u_{ij})$ is the random effect associated with S_{ij}, and α and β are the unknown parameters. The frailty model considered by Aalen and Husebye (1991) is a special case of (7) with $v_{ij} = v_i$ ($j = 1, 2, \ldots$), that is, a constant random effect for each subject. Note that this class of models does not yield simple marginal interpretation of the effects of covariates on the S_{ij}'s. Further discussion on this point is given in Section 6.

It has proved difficult to develop computationally tractable dynamic models. One possibility is to express the random effects as linear combinations of gamma variables as in Petersen *et al.* (1996). On letting $Ga(a,\, b)$ denote a gamma distribution with mean ab^{-1} and variance ab^{-2}, we may take $v_{i1} \sim Ga(\omega^{-2},\, \omega^{-2})$ and

$$v_{ij} = \phi v_{i,j-1} + (1 - \phi) v_{ij}^*$$

where v_{i2}^*, v_{i3}^*, ... are independent $Ga(\frac{\omega^{-2}(1-\phi)^2}{1-\phi^2}, \frac{\omega^{-2}(1-\phi)^2}{1-\phi^2})$ variables. The v_{ij}'s are second order stationary with $E(v_{ij}) = 1$ and $Cov(v_{ij}, v_{ij'}) = \phi^{|j-j'|}\omega^{-2}$, though they do not have the same distribution. This model has the form of (1) and (2) and includes the Aalen and Husebye (1991) model as a special case when $\phi = 1$. As described in Petersen *et al.* (1996), the likelihood function for the data $s_{i1}, \ldots, s_{i,n_i}, s_{i,n_{i+1}}^*$ can be expressed in closed form, but it is the sum of many terms and is computationally forbidding. Further work on computation, perhaps aided by simulation, may enhance the applicability of such models.

We shall take a different approach and examine a family of models proposed by Yue and Chan (1997). The Yue and Chan (1997) model is based on ideas of Smith and Miller (1986) and Harvey and Fernandes (1989) in other contexts. It is non-stationary and assumes that the v_{ij}'s in (7) evolve according to a geometric independent increments model with

$$v_{i1} \sim Ga(\omega^{-2}, \omega^{-2}) \tag{8}$$

$$v_{i,j+1} = \psi^{-1} v_{ij} \eta_{ij} \qquad j = 1, 2, \ldots \tag{9}$$

where $0 < \psi < 1$ and the η_{ij}'s are independent $Beta(\psi a_{ij}, (1-\psi)a_{ij})$ variables with

$$a_{ij} = \psi^{j-1}\omega^{-2} + \psi^{j-1} + \cdots + \psi + I(j \leq n_i). \tag{10}$$

This model, a state-space model of the form (1) and (2), seems awkward but is analytically tractable since the conditional distributions of S_{ij} given S_i^{j-1}, which are needed to construct the likelihood function, are easy to compute; they will be described in the next section. The gamma shared frailty model in Aalen and Husebye (1991) is given by the special case $\psi \to 1$. Yue and Chan (1997) discussed maximum likelihood estimation and presented some simulation results.

4 Estimation for the Yue-Chan Model

A merit of the Yue-Chan model, specified by (7), (8) and (9), is the availability of the likelihood without the need for numerical integration or other approximations. The observed

data likelihood function for recurrent event data with independent right censoring can be constructed through the usual decomposition rule as

$$L_{obs} = \prod_{i=1}^{N} \left\{ \prod_{j=1}^{n_i} Pr\{s_{ij} \mid S_i^{j-1}\} \right\} Pr\{S_{i,n_i+1} \geq s^*_{i,n_i+1} \mid S_i^{n_i}\}$$

where each individual predictive density is computed, using (7), (8) and (9), by integrating over v_{ij} in $E_{v_{ij}|S_i^{j-1}}[Pr\{s_{ij} \mid v_{ij},\, S_i^{j-1}\}]$. Thus the observed data log-likelihood function, $\log(L_{obs}) = l_{obs}$ can be written as

$$\begin{aligned} l_{obs} \;=\; & \sum_{i=1}^{N} \left\{ \sum_{j=1}^{n_i} [\log h_0(s_{ij};\alpha) + \log r(x_{ij}(s_{ij});\beta) + \log(a_{ij}-1) - \log b_{ij}] \right. \\ & \left. + \sum_{j=1}^{n_i+1} (a_{ij}-1)[\log(b_{ij} - H(s_{ij}, x_{ij}(s_{ij});\alpha,\beta)) - \log b_{ij}] \right\} \end{aligned} \tag{11}$$

where

$$b_{ij} = \psi^{j-1}\omega^{-2} + \psi^{j-1} H(s_{i1}, x_{i1}(s_{i1});\alpha,\beta) + \cdots + H(s_{ij}; x_{ij}(s_{ij}),\alpha,\beta), \tag{12}$$

a_{ij} is given by (10), and $H(s; x(s),\alpha,\beta) = \int_0^s h_0(u;\alpha) r(x(u);\beta) du$ is the cumulative hazard function. The score function and Hessian matrix can be routinely evaluated by differentiating (11).

A popular choice of the baseline hazard function is the Weibull form with $h_0(s;\alpha) = cs^k$; $c > 0$, $k > -1$ and $\alpha = (c,k)'$. Assuming a piecewise constant baseline hazard function with the number of intervals and the jump points being known, Yue and Chan (1997) computed the likelihood by using a direct computation method based on individual predictive densities of the gap times. Common optimization algorithms such as the Downhill Simplex Method (which does not require first and second derivatives) or the Newton-Raphson Method are usually effective in obtaining the maximum likelihood estimates from (11).

We shall extend the Yue-Chan model by letting $h_0(s;\alpha) = h_0(s)$ be an unknown arbitrary positive function to provide a more flexible semiparametric analysis. An EM algorithm is proposed to estimate the parameters of interest, namely β, ψ, ω^2, in this case. Assume that

the v_{ij}'s are observable, based on (1), (2) and (7), so that the complete data likelihood, L_c, is proportional to

$$\begin{aligned} L_c &= \prod_{i=1}^{N}\prod_{j=1}^{n_i+1}[v_{ij}r(x_{ij});\beta)h_0(s_{ij};\alpha)]^{I(j\le n_i)}\exp\left[-v_{ij}H(s_{ij};x_{ij},\alpha,\beta)\right] \\ &\quad \cdot\frac{1}{\Gamma(\psi a_{i,j-1})}\exp(-\psi b_{i,j-1}v_{ij})(\psi b_{i,j-1})^{\psi a_{i,j-1}}\ v_{ij}^{\psi a_{i,j-1}-1}. \end{aligned} \tag{13}$$

For convenience, we let $\delta_{ij}=0$ if S_{ij} is censored and $\delta_{ij}=1$ otherwise, $r(x;\beta)=\exp(x'\beta)$, and $\Lambda_0(s;\alpha)=\int_0^s h_0(u;\alpha)du$ in the sequel. The complete data log-likelihood, $l_C=\log L_C$, can then be expressed as

$$l_C = l_1(\beta,\Lambda_0)+l_2(\psi,\omega^2) \tag{14}$$

where

$$\begin{aligned} l_1 &= \sum_{i=1}^{N}\sum_{j=1}^{n_i+1}\left\{\delta_{ij}[x_{ij}{}'\beta+\log h_0(s_{ij};\alpha)]-v_{ij}\exp(x_{ij}{}'\beta)\Lambda_0(s_{ij};\alpha)\right\}, \quad \text{and} \\ l_2 &= \sum_{i=1}^{N}\sum_{j=1}^{n_i+1}\left\{-\log\Gamma(\psi a_{i,j-1})-\psi b_{i,j-1}v_{ij}+\psi a_{i,j-1}\log(\psi b_{i,j-1})+(\psi a_{i,j-1}+\delta_{ij}-1)\log v_{ij}\right\}. \end{aligned}$$

The E-step is to take conditional expectations of l_1 and l_2 given the observed data. This can be done easily by replacing v_{ij} and $\log(v_{ij})$ by their conditional expectations $\frac{a_{ij}}{b_{ij}}$ and $\frac{\Gamma'(a_{ij})-\Gamma(a_{ij})\log b_{ij}}{\Gamma(a_{ij})}$, respectively. The M-step involves 2 steps. First, updated estimates of ψ and ω^2 are obtained by maximizing the conditional expectation of l_2, with other parameters held at their current values. To update the estimate of Λ_0, a profile likelihood approach as in Klein (1992) is adopted where $\Lambda_0(s)$ is estimated nonparametrically by

$$\hat{\Lambda}_0(s)=\sum_{S_{(i)}\le s}\frac{\delta_{(i)}}{\sum_{k\in R(S_{(i)})}\frac{a_k}{b_k}\exp(x_k{}'\beta)},$$

where $S_{(i)}$ is the i^{th} smallest gap time among all observations S_{ij}, and $R(S_{(i)})$ is the risk set just prior to time $S_{(i)}$. Substituting the nonparametric estimate of $\Lambda_0(t)$ into the conditional

expectation of l_1 gives a profile partial likelihood

$$E(l_1) = \sum_{i=1}^{N} \sum_{j=1}^{n_i+1} \delta_{ij}[x_{ij}{}'\beta - \log \sum_{k \in R(S_{ij})} \frac{a_k}{b_k} \exp(x_k{}'\beta)],$$

which is then maximized to estimate the regression parameter β. The E-step and M-step are applied repeatedly until convergence.

Standard errors of the estimates for the fully parametric models can be obtained from the inverse of the observed information matrix, $V(= \left[\frac{-\partial^2 l_{obs}(\theta)}{\partial \theta^2}\right]^{-1})$, based on the observed data log-likelihood (11) where $\theta = (\beta, \psi, \omega^2, \alpha)'$ is the vector of parameters of the model. In practice, V can be estimated by $\hat{V}$ by substituting $\theta = \hat{\theta}$, the maximum likelihood estimate of θ. For the semiparametric model, the method proposed by Andersen *et al.* (1997) can be adopted here to estimate the standard errors of the estimates by inverting the observed information matrix $V_1 = \left[\frac{-\partial^2 l_{obs}(\theta_1)}{\partial {\theta_1}^2}\right]^{-1}$ with $\theta_1 = (\beta, \psi, \omega^2, h_0)'$.

For assessment of estimation with this model, two sets of simulations are conducted. For both sets, the baseline hazard is taken from a Weibull distribution, i.e. $h_0(s; \alpha) = cs^k$; $c > 0$, $k > -1$ and $\alpha = (c, k)'$. Then, with some pre-specified values of $(\beta, c, k, \omega^2, \psi)$, we generate $s_{i1}, \ldots, s_{i,n_i}, s^*_{i,n_{i+1}}$ for subject i as follows $(i = 1, \ldots, N)$.

1. Fix the censoring time, τ_i, for subject i.
2. Generate v_{i1} from $Ga(\omega^{-2}, \omega^{-2})$.
3. Generate s_{i1} from its hazard function $v_{i1} h_0(t) r(x_{i1}; \beta)$.
4. Set $j = 1$.
5. Generate η_{ij} from $Beta(\psi a_{ij}, (1-\psi)a_{ij})$ where a_{ij} is computed from (10), and evaluate $v_{i,j+1} = \psi^{-1} v_{ij} \eta_{ij}$.
6. Generate $s_{i,j+1}$ from its hazard function $v_{i,j+1} h_0(t) r(x_{i,j+1}; \beta)$.
7. If the current total time span, $\sum_{k=1}^{j+1} s_{ik}$, is less than τ_i, set $j = j + 1$ and go back to step 5.

8. Put $n_i = j$ and the last censored duration time $s^*_{i,n_i+1} = \tau_i - \sum_{j=1}^{n_i} s_{ij}$.

Let M be the number of simulated data sets, $\hat{\theta}^{(m)}$ be the maximum likelihood estimate of θ, and $\hat{\Sigma}^{(m)}$ be the asymptotic covariance matrix of the estimates with the diagonal elements replaced by the asymptotic standard errors for the m-th ($m = 1, 2, \ldots, M$) data set. The efficiency of the estimator is assessed by computing the summary statistics, $\bar{\hat{\theta}} = \frac{1}{M}\sum_{m=1}^{M}\hat{\theta}^{(m)}$, $\bar{\hat{\Sigma}} = \frac{1}{M}\sum_{m=1}^{M}\hat{\Sigma}^{(m)}$, and Σ is the sample covariance matrix with diagonal elements replaced by the sample standard deviations.

In the first set of simulations we study estimation in the fully parametric Weibull model. We assumed no covariates and the same number of subjects ($N = 19$) and censoring times as in the set of small bowel motility data considered in Aalen and Husebye (1991). The data set has around 10 to 20 event recurrences per subject and is further analyzed in the next section. We generated $M = 1,000$ samples. The parameter to be estimated is $\theta = (c, k, \omega^2, \psi)$. Initial estimates for θ was obtained by fitting a gamma frailty model as in Aalen and Husebye (1991) and ψ was initially taken as 0.5. To avoid boundary value problems and highly correlated estimates, the set of parameters θ was transformed to $\theta_U = (u, \xi, \gamma, \tau)$ where

$$u = \frac{1}{k+1}\log(\frac{k+1}{c}),\ \ \xi = \log(k+1),\ \ \gamma = \log\omega^2 \text{ and } \tau = \gamma - \log(\frac{\psi}{1-\psi}) \qquad (15)$$

before estimation was performed. The nonlinear optimization subroutine NLPNMS (Nelder-Mead Simplex method) under SAS/IML was employed to maximize (11). The 95% coverage, which is the proportion of the 95% confidence intervals computed using the Normal assumption, i.e. $\hat{\theta}_U^{(m)} \pm 1.96 \times \text{standard error}(\hat{\theta}_U^{(m)})$, that includes the hypothesized value of θ_U, is also computed in this case. With obvious notation, $\bar{\hat{\theta}}_U$, $\bar{\hat{\Sigma}}_U$, Σ_U, and the 95% coverage for θ were also computed. The values of c and k are taken to be 0.00005 and 1.5 respectively. These are the approximate estimates obtained from fitting the small bowel motility data in the next section. The variability of the random effects, ω^2, is taken at two values: 0.1 and 1.5. In either case, ψ takes values in $\{0.1, 0.5, 0.9\}$. All summary statistics at different values of the parameters are tabulated in Table 1 for $\omega^2 = 0.1$ and Table 2 for $\omega^2 = 1.5$.

In Table 1 with $\omega^2 = 0.1$, except for τ, there is a fair agreement between the averages of the estimates and the true values. The same is observed between the standard errors

and the corresponding sample estimates. The coverage for θ_U agrees well with 0.95. The finite sample approximation by a Normal distribution for the sampling distribution of $\hat{\theta}_U$ can be reasonably assumed. It is intriguing to see the great discrepancy between the average standard error of $\hat{\tau}$ and its small finite sample standard deviation when $\psi = 0.9$ (Table 1(c)). This is due to the flatness of the likelihood when ψ is close to 1. A much better agreement can be seen for smaller values of ψ (e.g. Table 1(a) for $\psi = 0.1$). The empirical distribution of the estimates when ψ is close to 1 is studied in the small bowel motility example analyzed in the next section. In Table 2 with $\omega^2 = 1.5$, similar phenomena are observed.

Insert Tables 1 and 2.

In conclusion, with only a small number of subjects and around 10 to 20 duration times per subject, the asymptotic approximations behave reasonably well. Standard errors of the estimates of the baseline parameters (c, k) tend to be smaller as ψ gets larger, which corresponds to a more stable process, or when ω^2 gets smaller, which corresponds to smaller initial random effects. Interval estimates should be computed by a Normal approximation for $\hat{\theta}_U$, which gives closer to nominal coverage.

In the second simulation study, we considered the performance of the proposed EM algorithm in the semiparametric proportional hazards model. For simplicity, we consider the case with no censoring; $M = 200$ data sets, each with $N = 100$ subjects and $n_i = 5$ ($i = 1, \ldots, N$) successive gap times were generated. The covariate X was simulated from a uniform(0,1) distribution and assumed constant throughout the observation period (i.e., $x_{ij} = x_i$). The regression parameter β and the discounting parameter ψ were set at 1.0 and 0.6 respectively. The parameter ω^2, which characterizes the heterogeneity of the initial random effects, was set at $0.1, 0.2, 0.5, 1.0$. The parameter of interest is $\theta = (\beta, \psi, \omega^2)$. Estimator means $\bar{\hat{\theta}}$, the empirical standard deviations S, and the average asymptotic standard errors $\bar{se}$. are tabulated in Table 3.

The proposed method is efficient in the estimation of the parameters in the semiparametric dynamic random effects model of Yue and Chan (1997), even for moderate sizes of n_i and N. The performance of the regression estimator $\hat{\beta}$ is quite satisfactory in all cases. When ω^2 is small, all estimators are well-behaved, but for large value of ω^2, small discrepancy between

the mean estimates and the true values of the parameters is observed. Similar behavior was observed in the first set of simulation study. The average asymptotic standard errors of the estimators $\hat{\beta}$ and $\hat{\psi}$ agree with the empirical standard deviations reasonably well, but less well for $\hat{\omega}^2$.

Insert Table 3.

5 Examples

We consider two examples to assess the application of the Yue-Chan model. The first example, taken from Aalen and Husebye (1991), examines the empirical distribution of the estimates when random effects are not needed. The second example, taken from Gail *et al.* (1980), compares the estimates under different choices of the baseline hazard function: a Weibull hazard function, a piecewise constant function as used in Yue and Chan (1997) and an arbitrary, non-negative unknown function.

5.1 The Small Bowel Motility Data

Our small bowel has both absorptive and secretory functions and the muscular activity (motility) of it is vital for gastrointestinal function in humans. In a study described by Aalen and Husebye (1991), nineteen healthy individuals treated with a standardized mixed meal were monitored. An event occurrence is the detection of a certain phase characterizing a regular motility pattern in a fasting state. First detection of the phase was defined as the start of the fasting state and duration times for successive detection of the phase were continuously tracked until the experiment ended at a pre-specified time.

As in the simulation study, we assume a Weibull baseline hazard function, i.e. $h_0(s) = cs^k$; $c > 0$, $k > -1$. Initial estimates were obtained by fitting a gamma frailty model and ψ was initially taken as 0.5. With the same transformation in (15), the log-likelihood was maximized at -429.13 and maximum likelihood estimates together with their asymptotic standard errors and correlation coefficients are tabulated in Table 4. The sampling distribution of the estimates was examined by 500 parametric bootstrap samples. Figure 1 exhibits plots of

histograms for various estimates and shows a fairly symmetric empirical distribution for u and δ. The seemingly bi-modal behavior for the estimates of τ is due to the flatness of the likelihood as τ gets small when ψ is close to 1. Careful examination of the estimates shows that estimates for τ smaller than -9 usually have scores greater than -10^{-3} which keep increasing asymptotically to 0 as the estimates are pushed smaller. This is depicted in Figure 2 for a typical iterated estimate -13.31 for τ. Thus the left cluster of the estimates for τ should actually spread towards $-\infty$ and the empirical distributions of $\hat{\gamma}$ and $\hat{\tau}$ both have long left tail. Again, as observed in the simulation study, the bi-modal behavior disappears for small values of ψ.

Insert Table 4.

The log-likelihood ratio statistic for testing the null hypothesis $\omega^2 = 0$ is $R = 2.58$. However, since $\omega^2 = 0$ lies on the boundary of the parameter space, R is not distributed as a simple chi-square distribution. The empirical significance level of R is 0.09, which is the sample proportion of log-likelihood ratio statistics that are greater than R among $1,000$ bootstrapped samples with $(c, k, \omega^2) = (0.000044, 1.28, 0)$. Thus, we arrive at the same conclusion as in Aalen and Husebye (1991), that the data do not exhibit strong evidence of subject heterogeneity. Also, a graphical test (not shown here) of the Weibull model does not reveal a serious model departure. Consequently, the value of ψ becomes irrelevant. In fact, the log-likelihood ratio statistic for testing the null hypothesis $\omega^2 = 0$ against the alternative hypothesis $\omega^2 > 0$ but keeping $\psi = 1$ (i.e. the gamma frailty model) is only slightly smaller ($\approx 1.9 \times 10^{-4}$) than R. A renewal process model is therefore adequate for the small bowel motility data.

Insert Figures 1 and 2.

5.2 The Mammary Tumors Data

In this study, a total of 48 female rats under two treatments were followed until a prespecified time and the times to appearance of mammary tumors were recorded. The full data set is given in Gail *et al.* (1980). We define $x_{ij} = x_i = 1$ if the i-th rat was on the new

treatment and 0 otherwise. This covariate information is incorporated by taking $r(x_{ij};\beta) = e^{\beta x_{ij}}$ in (7). Three choices of $h_0(s;\alpha)$ were used, namely the Weibull baseline function, the three-step piecewise constant baseline function as described in Yue and Chan (1997), and an arbitrary but unknown baseline function. Some T_{ij} and $T_{i,j+1}$ are equal which leads to cases with $S_{ij} = 0$. For the cases with nonparametric baseline hazard function, the ties were broken at random while for parametric cases with Weibull and piecewise constant baseline hazard functions, we added a small positive quantity Δ to all the observations with $S_{ij} = 0$. Three values of Δ, namely 0.1, 0.3, and 0.5 were assumed to check on the sensitivity of the estimates. The estimates and the corresponding standard errors of the three models based on the proposed EM algorithm, and those obtained by Yue and Chan (1997) are tabulated in Table 5. The parameter estimates of the models with Weibull hazard function are influenced by Δ due to the large amount of zero observations in the data set. It is observed that a smaller value of Δ leads to a smaller value of k, indicating a more steeply declining estimated baseline hazard function. However the estimates of the models with piecewise constant hazard function are nearly identical indicating the parameter estimates are robust to the values of Δ. This is because only the first piece is affected, and its interval extends well beyond Δ. Except for the estimates for β, our results are somewhat contrary to that of Yue and Chan (1997), particularly in the estimation of the baseline hazard functions. The estimated baseline function of Yue and Chan is increasing while our estimated parametric baseline functions are all decreasing functions. It is possibly due to the different treatments for the cases with $S_{ij} = 0$ between the two approaches. Our estimates of ψ and ω^2 among the three models are very consistent. It is also observed that the two sets of estimated cumulative baseline hazard functions closely agree with each other, particularly at the jump points $s = 13$ and 30. The choice of baseline hazard functions does not seem to have a strong impact on the estimates of ω^2 and ψ, but may affect the estimates of β. However the significance of the treatment effect remains the same in this example.

Insert Table 5.

6 Additional Remarks

Random effects models are widely used in survival and event history analysis. They are perhaps best viewed as a modeling device for incorporating dependence or association among a set of response times. Gaussian models for appropriately transformed times are fairly easy to fit and provide simple marginal inference about means, variances, and covariate effects for observable quantities (i.e. the S_{ij}'s or transforms of them). They also provide a degree of robustness to distributional assumptions. Proportional hazards models, on the other hand, do not allow simple marginal interpretations on the covariate effects, and so require us to interpret the covariate effects conditional on unobservable random effects. They can be sensitive to the choice of model for the random effects. These issues are well known and widely discussed (e.g. Clayton, 1994; Pickles and Crouchley, 1994; Petersen *et al.*, 1996) as are advantages of proportional hazards, specifically, the ease of dealing with time-varying covariates and the availability of semi-parametric inference methods not requiring $h_0(s)$ in (7) to be modeled parametrically.

Dynamic random effects models seem attractive in some situations involving repeated times between recurrent events. Gaussian models as in Section 2 are easily used and interpreted. In view of the difficulties of marginal interpretation for proportional hazards models, it seems desirable to keep random effects structure as simple as possible. For (7) we might therefore prefer to consider models where $v_{ij} = v_i$ and x_{ij} is allowed to depend on S_i^{j-1} rather than to employ dynamic random effects v_{ij}. In some applications any of the repeated gap times S_{i1}, S_{i2}, ... may be censored, and not just the last one. For example, if the S_{ij}'s represent successive times from startup to failure of equipment as in Follmann and Goldberg (1988) then much of the time the equipment is shut off before failure occurs. This will not complicate the computation of the maximum likelihood estimates based on the Yue-Chan model using the proposed EM-algorithm. We simply replace (10) by

$$a_{ij} = \psi^{j-1}\omega^{-2} + \psi^{j-1}\delta_{i1} + \cdots + \psi\delta_{i,j-1} + \delta_{ij}.$$

We conclude with two additional remarks. First, there is a need to study diagnostic

procedures for proportional hazards models with random effects (e.g. Shih and Louis, 1996), which are not well understood relative to methods for Gaussian models. Second, there are situations where unobserved random effects vary continuously over time so that it is not feasible to associate a single v_{ij} with a given gap period. As far as we know this problem has not been studied for repeated events. One possible approach is to extend Fahrmeir's model (Fahrmeir, 1994) for discrete survival data to the multivariate case.

References

O. O. Aalen and E. Husebye, "Statistical analysis of repeated events forming renewal processes," *Statistics in Medicine* vol. 10, pp. 1227–1240, 1991.

P. K. Andersen, J. P. Klein, K. M. Knudsen, and R. Tabanera-Palacios, "Estimation of the variance in a Cox's regression model with shared frailties," *Biometrics* vol. 53, pp. 1475–1484, 1997.

D. P. Byar, "Why data bases should not replace randomized clinical trials," *Biometrics* vol. 36, pp. 337–342, 1980.

B. P. Carlin, N. G. Polson, and D. S. Stoffer, "A Monte Carlo approach to nonnormal and nonlinear state-space modeling," *Journal of American Statistical Association* vol. 87, pp. 493–500, 1992.

D. G. Clayton, "Some approaches to the analysis of recurrent event data," *Statistical Methods in Medical Research* vol. 3, pp. 244–262, 1994.

A. Dempster, N. Laird, and D. Rubin, "Maximum likelihood estimation with incomplete data via the EM algorithm (with discussion)," *Journal of the Royal Statistical Society* vol. B 39, pp. 1–38, 1977.

L. Fahrmeir, "Dynamic modelling and penalized likelihood estimation for discrete time survival data," *Biometrika* pp. 317–330, 1994.

D. A. Follmann and M. S. Goldberg, "Distinguishing heterogeneity from decreasing hazard rates," *Technometrics* vol. 30, pp. 389–396, 1988.

M. H. Gail, T. J. Santner, and C. C. Brown, "An analysis of comparative carcinogenesis experiments based on multiple times to tumor," *Biometrics* vol. 36, pp. 255–266, 1980.

A. C. Harvey, *Forecasting, structural time series models and the Kalman filter*, Cambridge University Press: New York, 1989.

A. C. Harvey and C. Fernandes, "Time series models for count or qualitative observations," *Journal of Business & Economic Statistics* vol. 7, pp. 407–423, 1989.

R. H. Jones, *Longitudinal data with serial correlation: A state-space approach*, Chapman and Hall: London, 1993.

B. Jørgensen, S. Lundbye-Christensen, X. Song, and L. Sun, "A longitudinal study of emergency room visits and air pollution for Prince George, Bristish Columbia," *Statistics in Medicine* vol. 15, pp. 823–836, 1996.

J. P. Klein, "Semiparametric estimation of random effects using the Cox model based on the EM algorithm," *Biometrics* vol. 48, pp. 795–806, 1992.

P. Lambert, "Modeling of nonlinear growth curve on series of correlated count data measured at unequally spaced times: A full likelihood based approach," *Biometrics* vol. 52, pp. 50–55, 1996a.

P. Lambert, "Modelling of repeated series of count data measured at unequally spaced times," *Applied Statistics* vol. 45, pp. 31–38, 1996b.

J. F. Lawless, "The analysis of recurrent events for multiple subjects," *Applied Statistics* vol. 44, pp. 487–498, 1995.

J. F. Lawless and D. Y. T. Fong, "State duration models in clinical and observational studies," *Statistics in Medicine* vol. 18, pp. 2365–2376, 1999.

J. H. Petersen, P. K. Andersen, and R. D. Gill, "Variance components models for survival data," *Statistica Neerlandica* vol. 50, pp. 193–211, 1996.

A. Pickles and R. Crouchley, "Generalization and applications of frailty models for survival and event data," *Statistical Methods in Medical Research* vol. 3, pp. 263–278, 1994.

R. L. Prentice, B. J. Williams, and A. V. Peterson, "On the regression analysis of multivariate failure time data," *Biometrika* vol. 68, pp. 373–379, 1981.

N. Shephard and M. K. Pitt, "Likelihood analysis of non-gaussian measurement time series," *Biometrika* vol. 84, pp. 653–667, 1997.

J. H. Shih and T. A. Louis, "Tests of independence for bivariate survival data," *Biometrics* vol. 52, pp. 1440–1449, 1996.

R. L. Smith and J. E. Miller, "A non-gaussian state space model and application to prediction of records," *Journal of Royal Statistical Society* vol. B 48, pp. 79–88, 1986.

M. K. P. So, "Time series with additive noise," *Biometrika* vol. 86, pp. 474–482, 1999.

X. Xue and R. Brookmeyer, "Bivariate frailty model for the analysis of multivariate failure time," *Lifetime Data Analysis* vol. 2, pp. 277–289, 1996.

H. Yue and K. S. Chan, "A dynamic frailty model for multivariate survival data," *Biometrics* vol. 53, pp. 785–793, 1997.

Table 1: Summary statistics from the dynamic proportional hazards model with $\omega^2 = 0.1$.

(a) $(c,\ k,\ w^2,\ \psi) = (0.00005,\ 1.5,\ 0.1,\ 0.1)$.

θ_U	$\overline{\hat{\theta}}_U$	$\overline{\hat{\Sigma}}_U$				Σ_U				95% coverage
4.3279	4.3250	0.0922	-0.1075	-0.3740	-0.3844	0.1051	-0.1956	-0.4154	-0.4533	0.92
0.9163	0.9459		0.1332	0.3550	0.5128		0.1446	0.3899	0.5360	0.94
-2.3026	-2.5601			0.8657	0.8940			0.9184	0.9057	0.96
-0.1054	-0.2356				0.5997				0.7289	0.95
θ	$\overline{\hat{\theta}}$	$\overline{\hat{\Sigma}}$				Σ				95% coverage
0.00005	0.00008	0.00009	-0.9841	-0.2912	0.0941	0.00023	-0.4339	-0.1294	0.3285	0.77
1.5	1.6023		0.3482	0.3550	-0.0621		0.3830	0.4150	-0.1066	0.95
0.1	0.1070			0.0743	0.7010			0.0851	0.4965	0.82
0.1	0.0942				0.0403				0.0353	0.91

(b) $(c,\ k,\ w^2,\ \psi) = (0.00005,\ 1.5,\ 0.1,\ 0.5)$.

θ_U	$\overline{\hat{\theta}}_U$	$\overline{\hat{\Sigma}}_U$				Σ_U				95% coverage
4.3279	4.3259	0.0684	-0.1097	-0.3465	-0.3841	0.0707	-0.1287	-0.3491	-0.3577	0.93
0.9163	0.9313		0.0909	0.3180	0.4796		0.0933	0.2781	0.4293	0.94
-2.3026	-2.5622			0.9178	0.8811			0.9604	0.7352	0.96
-2.3026	-2.5572				1.2650				0.9418	0.97
θ	$\overline{\hat{\theta}}$	$\overline{\hat{\Sigma}}$				Σ				95% coverage
0.00005	0.00006	0.00005	-0.9813	-0.2531	0.2010	0.00005	-0.8309	-0.1698	0.2227	0.84
1.5	1.5489		0.2321	0.3180	-0.1897		0.2409	0.2893	-0.2244	0.95
0.1	0.1072			0.0801	0.5779			0.0788	0.4907	0.86
0.5	0.4913				0.0937				0.1038	0.93

(c) $(c,\ k,\ w^2,\ \psi) = (0.00005,\ 1.5,\ 0.1,\ 0.9)$.

θ_U	$\overline{\hat{\theta}}_U$	$\overline{\hat{\Sigma}}_U$				Σ_U				95% coverage
4.3279	4.3239	0.0516	0.0204	-0.1915	-0.1477	0.0505	0.0503	-0.1189	-0.1109	0.95
0.9163	0.9364		0.0793	0.3283	0.2588		0.0786	0.2272	0.1807	0.95
-2.3026	-2.6262			1.4511	0.2112			1.1497	-0.0735	0.96
-4.4998	-9.001				396.87				5.81	0.95
θ	$\overline{\hat{\theta}}$	$\overline{\hat{\Sigma}}$				Σ				95% coverage
0.00005	0.00005	0.00004	-0.9865	-0.2918	0.1067	0.00004	-0.8520	-0.2312	0.1208	0.83
1.5	1.5588		0.2032	0.3283	-0.1128		0.2035	0.3212	-0.1475	0.96
0.1	0.1080			0.0877	0.3174			0.0816	0.3604	0.93
0.9	0.8649				0.1374				0.1573	0.53

Table 2: Summary statistics from the dynamic proportional hazards model with $\omega^2 = 1.5$.

(a) $(c,\ k,\ w^2,\ \psi) = (0.00005,\ 1.5,\ 1.5,\ 0.1)$.

θ_U	$\overline{\hat{\theta}}_U$	$\overline{\hat{\Sigma}}_U$				Σ_U				95% coverage
4.3279	4.3334	0.2217	-0.4514	-0.5451	-0.5189	0.2217	-0.4402	-0.5886	-0.5980	0.94
0.9163	0.9497		0.1975	0.6169	0.7334		0.2007	0.6450	0.7455	0.96
0.4055	0.5144			0.4881	0.8285			0.4822	0.8686	0.94
2.6027	2.4724				0.6368				0.6110	0.97
θ	$\overline{\hat{\theta}}$	$\overline{\hat{\Sigma}}$				Σ				95% coverage
0.00005	0.00010	0.00014	-0.9584	-0.5089	0.4549	0.00020	-0.5321	-0.2624	0.4 014	0.75
1.5	1.6390		0.5292	0.6169	-0.4643		0.5607	0.6856	-0.4347	0.97
1.5	1.8767			0.8890	-0.1110			0.9630	-0.1742	0.96
0.1	0.1275				0.0396				0.0357	0.99

(b) $(c,\ k,\ w^2,\ \psi) = (0.00005,\ 1.5,\ 1.5,\ 0.5)$.

θ_U	$\overline{\hat{\theta}}_U$	$\overline{\hat{\Sigma}}_U$				Σ_U				95% coverage
4.3279	4.3417	0.1716	-0.3521	-0.4239	-0.4457	0.1789	-0.4284	-0.4507	-0.3298	0.94
0.9163	0.9335		0.1265	0.4986	0.6811		0.1343	0.5026	0.5553	0.95
0.4055	0.3234			0.4232	0.7806			0.4515	0.5551	0.96
0.4055	0.2731				1.1875				0.9716	0.95
θ	$\overline{\hat{\theta}}$	$\overline{\hat{\Sigma}}$				Σ				95% coverage
0.00005	0.00007	0.00007	-0.9363	-0.3778	0.4933	0.00008	-0.7074	-0.2754	0.4156	0.81
1.5	1.5668		0.3264	0.4986	-0.5487		0.3521	0.5133	-0.5512	0.95
1.5	1.5207			0.6187	-0.1572			0.6613	-0.1640	0.90
0.5	0.5018				0.0947				0.1064	0.92

(c) $(c,\ k,\ w^2,\ \psi) = (0.00005,\ 1.5,\ 1.5,\ 0.9)$.

θ_U	$\overline{\hat{\theta}}_U$	$\overline{\hat{\Sigma}}_U$				Σ_U				95% coverage
4.3279	4.3379	0.1305	-0.1778	-0.1906	-0.1692	0.1395	-0.2153	-0.2160	-0.1734	0.93
0.9163	0.9344		0.0948	0.3780	0.3751		0.0988	0.3966	0.3030	0.93
0.4055	0.3192			0.3602	0.2359			0.3836	0.1547	0.94
-1.7918	-4.2859				33.1671				3.9415	0.93
θ	$\overline{\hat{\theta}}$	$\overline{\hat{\Sigma}}$				Σ				95% coverage
0.00005	0.00006	0.00005	-0.9370	-0.3132	0.2850	0.00005	-0.7787	-0.2846	0.2796	0.82
1.5	1.5582		0.2433	0.3780	-0.3253		0.2574	0.4035	-0.4143	0.94
1.5	1.4757			0.5202	-0.0406			0.5465	-0.0847	0.90
0.9	0.8939				0.1023				0.1121	0.62

Table 3: Estimation result for 200 datasets with $\beta = 1.0$ and $\psi = 0.6$ and unspecified baseline hazard function.

		β	ψ	ω^2
$\omega^2 = 0.1$	$\bar{\hat{\theta}}$	1.0161	0.6016	0.1024
	S	0.2290	0.0756	0.0407
	$s.\bar{e}.$	0.2317	0.0745	0.0667
$\omega^2 = 0.2$	$\bar{\hat{\theta}}$	0.9755	0.6094	0.1800
	S	0.2931	0.0582	0.0548
	$s.\bar{e}.$	0.2496	0.0640	0.0839
$\omega^2 = 0.5$	$\bar{\hat{\theta}}$	0.9639	0.6521	0.3873
	S	0.4376	0.0506	0.0744
	$s.\bar{e}.$	0.3096	0.0522	0.0944
$\omega^2 = 1.0$	$\bar{\hat{\theta}}$	0.9386	0.7300	0.8567
	S	0.5434	0.0584	0.2806
	$s.\bar{e}.$	0.3187	0.0499	0.3019

Table 4: Maximum likelihood estimates for the small bowel motility data.

Parameter	Estimate	Asymptotic correlation matrix[a]			
u	4.7525	0.0658	0.0789	-0.2031	0.1760
δ	0.8261		0.0991	0.3409	0.2237
γ	-1.9304			0.9084	-0.3668
τ	-6.2818				52.9457
c	0.000044	0.000044	-0.9887	-0.3026	0.2683
k	1.2844		0.2265	0.3409	-0.2165
w^2	0.1451			0.1318	0.3815
ψ	0.9873				0.6694

[a] The off-diagonal elements are the asymptotic correlation; the diagonal elements are the asymptotic standard errors.

Table 5: Maximum likelihood estimates for the mammary cancer tumors data.

		β	ψ	ω^2	c	k	λ_1	λ_2	λ_3
unspecified $h_0(t)$	$\hat{\bar{\theta}}$	-0.9820	0.8246	0.2948	-	-	-	-	-
	$\hat{se}(\hat{\theta})$	0.2390	0.0857	0.0821	-	-	-	-	-
$h_0(s) = cs^k$	$\hat{\bar{\theta}}$	-0.7614	0.9060	0.1027	0.0673	-0.2472	-	-	-
$(\Delta = 0.1)$	$\hat{se}(\hat{\theta})$	0.1895	0.0976	0.0449	0.0089	0.0451	-	-	-
$h_0(s) = cs^k$	$\hat{\bar{\theta}}$	-0.8218	0.8749	0.1494	0.0588	-0.1764	-	-	-
$(\Delta = 0.3)$	$\hat{se}(\hat{\theta})$	0.2134	0.0911	0.0588	0.0098	0.0504	-	-	-
$h_0(s) = cs^k$	$\hat{\bar{\theta}}$	-0.8618	0.8563	0.1806	0.0547	-0.1364	-	-	-
$(\Delta = 0.5)$	$\hat{se}(\hat{\theta})$	0.2365	0.0917	0.0695	0.0109	0.0544	-	-	-
Piecewise constant	$\hat{\bar{\theta}}$	-0.8701	0.7870	0.1903	-	-	0.0460	0.0342	0.0318
$(\Delta = 0.1)$	$\hat{se}(\hat{\theta})$	0.2395	0.0845	0.0717	-	-	0.0093	0.0076	0.0067
Piecewise constant	$\hat{\bar{\theta}}$	-0.8731	0.8038	0.1916	-	-	0.0463	0.0343	0.0318
$(\Delta = 0.3)$	$\hat{se}(\hat{\theta})$	0.2395	0.0850	0.0728	-	-	0.0093	0.0076	0.0068
Piecewise constant	$\hat{\bar{\theta}}$	-0.8696	0.8047	0.1912	-	-	0.0461	0.0343	0.0318
$(\Delta = 0.5)$	$\hat{se}(\hat{\theta})$	0.2386	0.0848	0.0725	-	-	0.0092	0.0076	0.0067
Yue-Chan	$\hat{\bar{\theta}}$	-0.9510	0.5120	0.6623	-	-	0.0180	0.0270	0.0380
	$\hat{se}(\hat{\theta})$	0.2900	0.1000	0.4238	-	-	0.0032	0.0045	0.0094

Figure 1: Histograms for estimates from a bootstrap sample of size 500.

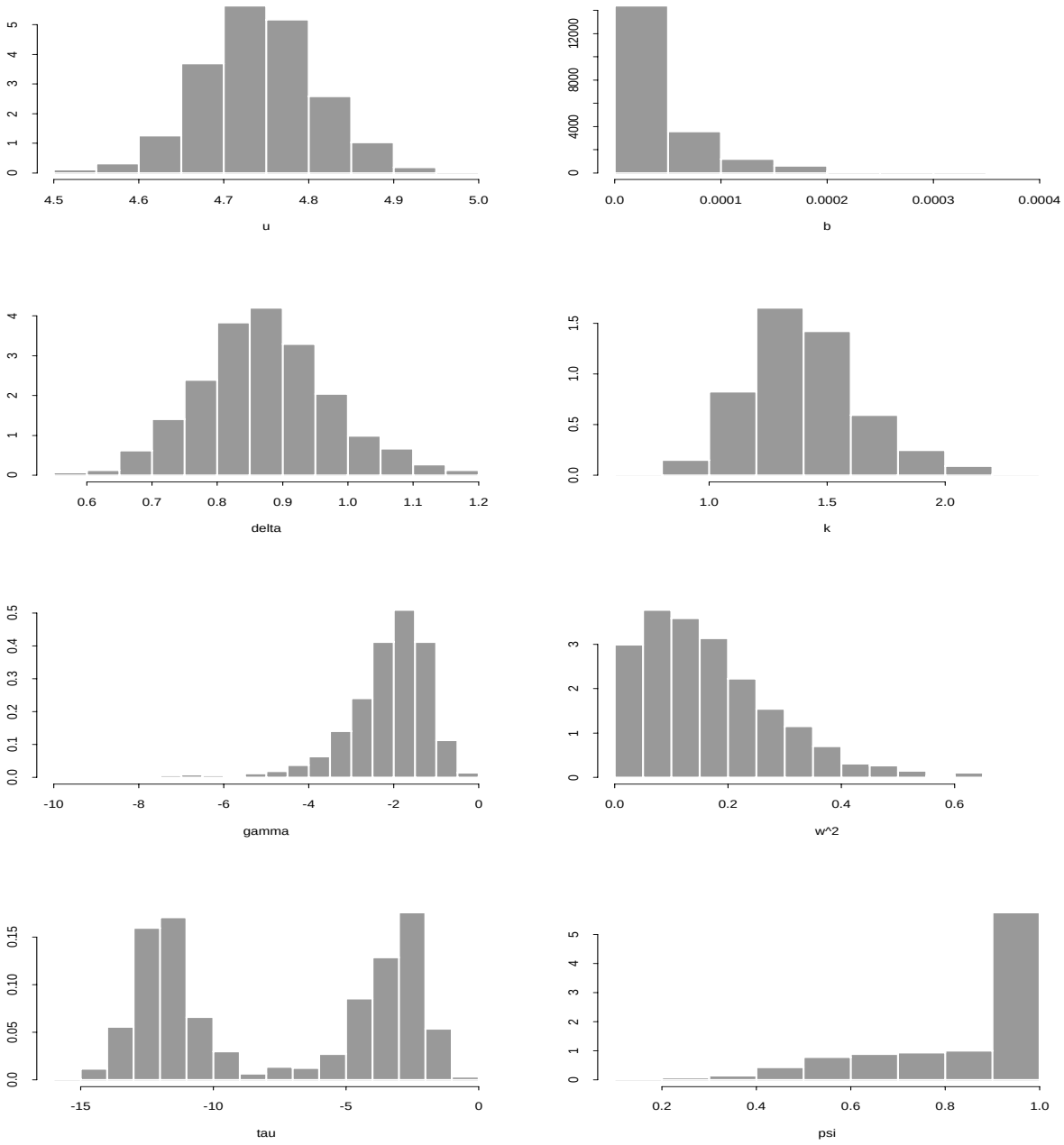

Figure 2: Plot of scores against τ for an iterated estimate of -13.31 for τ in a simulated data.