Inference with Recurrent Event Data

by

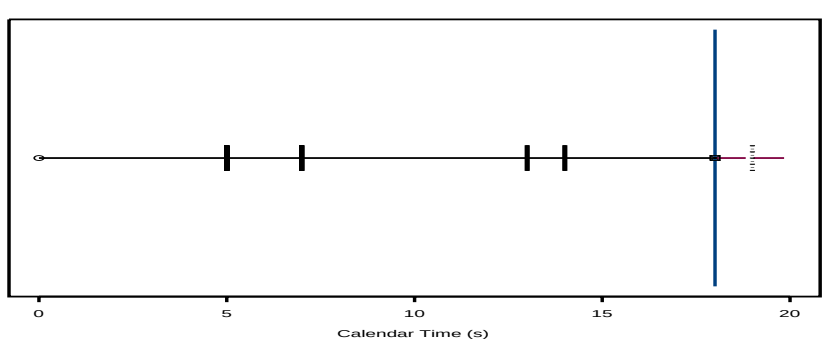
Edsel A. Peña
University of South Carolina, Columbia, SC
E-Mail: pena@stat.sc.edu

JSM 2000, Indianapolis, IN
August 13-17, 2000

Research supported by an US NIH/NIGMS Grant
Research joint with R. Strawderman and M. Hollander

Basic Data Accrual

- A reliability or engineering unit, or a biomedical subject.
- A recurrent event (such as stoppages of nuclear power plants, warranty claims for products, bouts with migraine headaches, recurrence of a tumor) is monitored for this unit over a random observation period $[0, \tau]$.
- τ has distribution $G(\cdot)$.
- $T_1, T_2, \ldots$ are the inter-occurrence times of the event, which are assumed to be IID from a distribution $F(\cdot)$.
- $S_1, S_2, \ldots$ are the calendar times of event occurrences, so $S_k = \sum_{j=1}^{k} S_j$.
- $K = \max\{k : \ S_k \leq \tau\} = \#$ of observed event occurrences in $[0, \tau]$.
- Observable Vector: $\mathbf{D} = (K, T_1, T_2, \ldots, T_K, \tau - S_K)$
- **Example:** $\tau = 18$ and $T_1 = 5, T_2 = 2, T_3 = 6, T_4 = 1, T_5 = 5$. Thus, $S_1 = 5, S_2 = 7, S_3 = 13, S_4 = 14, \tau - S_4 = 4$ and $K = 4$.



"Points to Ponder"

- K is informative about F.
- T_{K+1} is right-censored by $\tau - S_K$.
- Censoring mechanism is informative and dependent.
- Inter-occurrence time that gets censored tends to be larger (selection bias).

Problems Considered

- Suppose n units/subjects are available, with the ith unit observed over the period $[0, \tau_i]$.
- $\tau_1, \tau_2, \ldots, \tau_n$ are IID from G.
- $T_{i1}, T_{i2}, \ldots$ are IID F, $i = 1, 2, \ldots, n$.
- $S_{i1}, S_{i2}, \ldots$ are the calendar times of event occurrences.
- $K_i = \max\{k : \ S_{ik} \leq \tau_i\}$.

- Observables:

Unit #	Vector of Observables
1	$D_1 = (K_1, T_{11}, T_{12}, \ldots, T_{1K_1}, \tau_1 - S_{1K_1})$
2	$D_2 = (K_2, T_{21}, T_{22}, \ldots, T_{2K_1}, \tau_2 - S_{2K_2})$
$\vdots$	$\vdots$
n	$D_n = (K_n, T_{n1}, T_{n2}, \ldots, T_{nK_n}, \tau_n - S_{nK_n})$

- Given this data, how to estimate nonparametrically the event interoccurrence distribution F?

- Properties of the estimator?

- Comparison with existing estimator, for example with the Wang and Chang (JASA, 1999) estimator?

- Given this data, how do we perform goodness-of-fit tests?

Existing Approaches

- Consider the first, possibly right-censored, observation and use PLE.
- Ignore the right-censored last observation and use EDF.
- Wang and Chang (1999, JASA) proposed a PLE-type estimator.
- Aalen and Husebye's (1991) variance component model

$$g(T_{ij}) = \mu + U_i + E_{ij},$$

  and an intensity-based model incorporating a frailty component.
- Gill (80): very general paper dealing with a Markov renewal process.
- Vardi (82): an algorithm for the MLE when T_{ij}'s are arithmetic and the process starts at its stationary distribution.
- Sellke (88): nonparametric estimator for a single renewal process observed for a long time.
- Soon and Woodroofe (97): extended Vardi's results.
- Prentice, Williams and Peterson (1981); Andersen and Gill (1982), and Keiding, Andersen and Fledelius (1998).

Theoretical Developments

****Relevant Processes****

- Hazard rate function: $\lambda(\cdot) = f(\cdot)/\bar{F}(\cdot)$.

- Cumulative hazard function: $\Lambda(\cdot) = \int_0^{\cdot} \lambda(w)\mathrm{d}w$.

- *Calendar-time* processes

$$N_i^{\dagger}(s) = \sum_{j=1}^{\infty} I\{S_{ij} \leq s; S_{ij} \leq \tau_i\} = \#\text{ of events in } [0, s] \text{ for subj } i;$$

$$Y_i^{\dagger}(s) = I\{\tau_i \geq s\} = \text{at-risk indicator for subj } i;$$

$$\mathcal{F}_s^{\dagger} = \text{event history up to calendar-time } s.$$

- $A_i^{\dagger}(s) = \int_0^s Y_i^{\dagger}(v)\lambda\left(v - S_{iN_i^{\dagger}(v-)}\right)\mathrm{d}v, \quad i = 1, \ldots, n,$

- $M^{\dagger}(s) = (M_1^{\dagger}(s), \ldots, M_n^{\dagger}(s))$ with $M_i^{\dagger}(s) = N_i^{\dagger}(s) - A_i^{\dagger}(s)$ is a vector of square-integrable martingales with predictable covariance processes

$$\langle M_i^{\dagger}, M_{i'}^{\dagger}\rangle(s) = \begin{cases} A_i(s) & \text{if } i = i' \\ 0 & \text{if } i \neq i' \end{cases}.$$

- *Calendar-Duration Space* Processes: Define

$$Z_i(s, t) = I\{s - S_{iN_i^{\dagger}(s-)} \leq t\}.$$

  Indicates whether ith subject at calendar time s has been event-free for up to time t since last event occurrence.

- $Z_i(\cdot, t)$ is a bounded predictable process, nonincreasing in

$$s \in [S_{iN_i^{\dagger}(s-)}, S_{iN_i^{\dagger}(s-)+1}),$$

  for fixed t, and nondecreasing in t for fixed s.

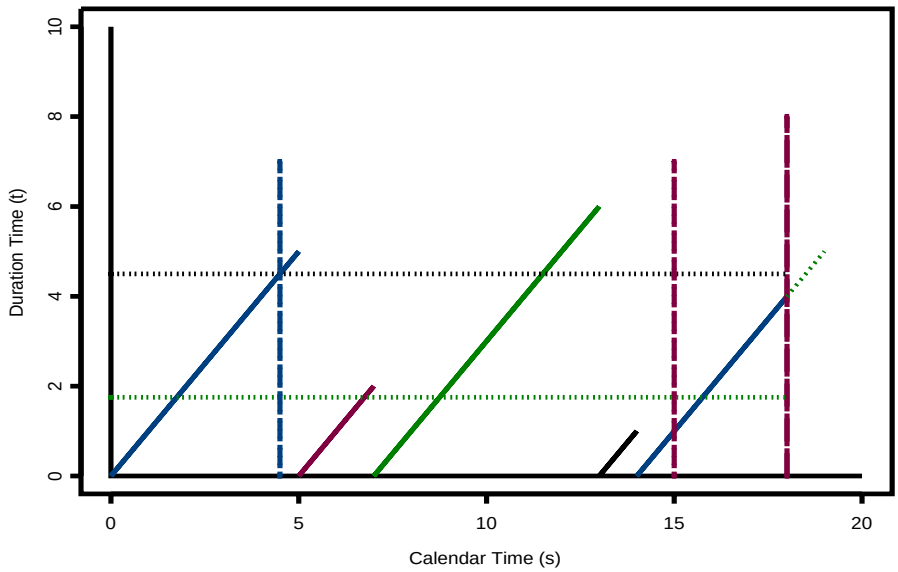


Figure 1: Picture of data $(K, T_1, T_2, T_3, T_4, \tau - S_4) = (4, 5, 2, 6, 1, 4)$. Inner horizontal lines: $t = 1.75, 4.5$. Inner vertical lines: $s = 4.5, 15$.

- From graph:

$$Z_1(4.5, 1.75) = 0 \quad \text{and} \quad Z_1(4.5, 4.5) = 1;$$

$$Z_1(15, 1.75) = 1 \quad \text{and} \quad Z_1(15, 4.5) = 1.$$

- Define the processes:

  - $N_i(s,t) = \int_0^s Z_i(v,t) N_i^\dagger(\mathrm{d}v)$
  - $A_i(s,t) = \int_0^s Z_i(v,t) A_i^\dagger(\mathrm{d}v)$
  - $M_i(s,t) = \int_0^s Z_i(v,t) M_i^\dagger(\mathrm{d}v) = N_i(s,t) - A_i(s,t)$
  - $Y_i(s,t) = \sum_{j=1}^{N_i^\dagger(s-)} I\{T_{ij} \geq t\} + I\{(s \wedge \tau_i) - S_{iN_i^\dagger(s-)} \geq t\}$

- $N_i(s,t)$: # of events in $[0,s]$ with gap times at most t.

- $Y_i(s,t)$: # of events in $[0,s]$ with gap times at least t.

- For data in Figure 1:

$$N_1(4.5, 1.75) = 0 \quad \text{and} \quad Y_1(4.5, 1.75) = 1;$$

$$N_1(4.5, 4.5) = 0 \quad \text{and} \quad Y_1(4.5, 4.5) = 1;$$

$$N_1(15, 1.75) = 1 \quad \text{and} \quad Y_1(15, 1.75) = 3;$$

$$N_1(15, 4.5) = 2 \quad \text{and} \quad Y_1(15, 4.5) = 2.$$

- Aggregated processes:

$$N(s,t) = \sum_{i=1}^{n} N_i(s,t);$$

$$A(s,t) = \sum_{i=1}^{n} A_i(s,t);$$

$$M(s,t) = \sum_{i=1}^{n} M_i(s,t).$$

****Intermediate Results****

- **Change-of-Integration Formulas**

  - ◇ **Identity #1:**

$$A(s,t) \;=\; \sum_{i=1}^{n} \int_0^s Z_i(v,t) A_i^{\dagger}(\mathrm{d}v) = \int_0^t Y(s,w)\lambda(w)\mathrm{d}w$$

  - ◇ **Identity #2:** For a predictable process $H_i(s,t)$,

$$\int_0^s H_i(s, v - S_{iN_i^{\dagger}(v-)}) M_i(\mathrm{d}v, t) = \int_0^t H_i(s,w) M_i(s, \mathrm{d}w)$$

- **Important Properties:**

  - ◇ **Property #1:** For fixed t, $M(\cdot,t)$ is a square-integrable martingale with predictable quadratic variation process

$$\langle M(\cdot,t), M(\cdot,t)\rangle(s) = \int_0^t Y(s,w)\lambda(w)\mathrm{d}w.$$

  - ◇ **Property #2:** With

$$G_s(w) = \begin{cases} G(w) & \text{if} \quad w < s \\ 1 & \text{if} \quad w \geq s \end{cases},$$

$$E\{Y_1(s,t)\} = y(s,t) \equiv \bar{F}(t)\left\{\bar{G}_s(t-) + \int_{[t,\infty)} R(w-t)\mathrm{d}G_s(w)\right\},$$

where

$$R(t) = \sum_{j=1}^{\infty} F^{\star j}(t) = \text{renewal function of } F;$$

$$F^{\star j} = j\text{th convolution.}$$

****A Weak Convergence Theorem****

Fix $s \in (0, \infty)$ and suppose for $t, t_1, t_2 \in [0, t^*]$ where $t^* \in (0, \infty)$:

(a) $\{\mathbf{H}_i(v, w) : 0 \leq v \leq s; 0 \leq w \leq t^*\}$ are left-continuous in (v, w); and there is a deterministic function $\mathbf{h}(v, w)$ on $[0, s] \times [0, t^*]$, continuous in (v, w) and bounded, with

$$\max_{1 \leq i \leq n} \sup_{0 \leq w \leq t^*} |\mathbf{H}_i(s, w) - \mathbf{h}(s, w)| \xrightarrow{\text{pr}} 0;$$

(b) For all $s \in (0, \infty)$, $\inf_{w \in [0, t^*]} y(s, w) > 0$;

(c) Matrix functions

$$\mathbf{V}^{(n)}(s, t) = \frac{1}{n} \sum_{i=1}^{n} \int_0^t \mathbf{H}_i(s, w)^{\otimes 2} Y_i(s, w) \lambda(w) \mathrm{d}w;$$

$$\mathbf{\Sigma}(s, t) = \int_0^t \mathbf{h}(s, w)^{\otimes 2} y(s, w) \lambda(w) \mathrm{d}w,$$

satisfies for each $t_1, t_2 \in (0, t^*]$ with $t_1 < t_2$,

$$0 < \det\{\mathbf{\Sigma}(s, t_2) - \mathbf{\Sigma}(s, t_1)\} < \infty \quad \text{and} \quad \left\| \mathbf{V}^{(n)}(s, t) - \mathbf{\Sigma}(s, t) \right\| \xrightarrow{\text{pr}} 0.$$

Then, the integral transforms,

$$\left\{ \mathbf{W}^{(n)}(s, t) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \int_0^t \mathbf{H}_i^{(n)}(s, w) \mathrm{d}M_i^{(n)}(s, \mathrm{d}w) : \quad t \in [0, t^*] \right\}$$

converges weakly to a zero-mean Gaussian process $\{\mathbf{W}^{(\infty)}(s, t) : t \in [0, t^*]\}$ with covariance matrix function

$$\mathbf{Cov}\{\mathbf{W}^{(\infty)}(s, t_1), \mathbf{W}^{(\infty)}(s, t_2)\} = \begin{bmatrix} \mathbf{\Sigma}(s, t_1) & \mathbf{\Sigma}(s, t_1) \\ \mathbf{\Sigma}(s, t_1) & \mathbf{\Sigma}(s, t_2) \end{bmatrix}, \quad \text{for} \quad t_1 \leq t_2.$$

Nonparametric Estimation of Λ and F

****Estimating Λ****

- For fixed t,

$$\left\{M(v,t) = N(v,t) - \int_0^t Y(v,w)\mathrm{d}\Lambda(w): \quad 0 \le v \le s\right\}$$

is a square-integrable martingale.

- Let $J(v,w) = I\{Y(v,w) > 0\}$.

- Then

$$\int_0^t \frac{J(s,w)}{Y(s,w)} M(s,\mathrm{d}w) = \int_0^t \frac{J(s,w)}{Y(s,w)} N(s,\mathrm{d}w) - \int_0^t J(s,w)\mathrm{d}\Lambda(w).$$

- By change-of-integration identity,

$$\int_0^t \frac{J(s,w)}{Y(s,w)} M(s,\mathrm{d}w) = \sum_{i=1}^n \int_0^s \frac{J(s, v - S_{iN_i^\dagger(v-)})}{Y(s, v - S_{iN_i^\dagger(v-)})} M_i(\mathrm{d}v, t).$$

- By stochastic integration theory,

$$\left\{\sum_{i=1}^n \int_0^s \frac{J(s, v - S_{iN_i^\dagger(v-)})}{Y(s, v - S_{iN_i^\dagger(v-)})} M_i(\mathrm{d}v, t) : 0 \le s < \infty\right\}$$

is a square-integrable zero-mean martingale.

- Consequently,

$$E\left\{\int_0^t \frac{J(s,w)}{Y(s,w)} N(s,\mathrm{d}w)\right\} = E\left\{\int_0^t J(s,w)\mathrm{d}\Lambda(w)\right\}.$$

- Method-of-moments estimator of $\Lambda(t)$:

$$\hat{\Lambda}(s,t) = \int_0^t \frac{J(s,w)}{Y(s,w)} N(s,\mathrm{d}w) = \int_0^t \frac{N(s,\mathrm{d}w)}{Y(s,w)}, \quad 0 \le t < \infty.$$

****Estimating $\bar{F}$****

- Product-integral representation of $\bar{F}(t)$: $\bar{F}(t) = \prod_{w \le t} [1 - \Lambda(\mathrm{d}w)]$.
- Substitution principle: Generalized PLE of $\bar{F}(t)$ is

$$\hat{\bar{F}}(s,t) = \prod_{w \le t} \left[1 - \hat{\Lambda}(s, \mathrm{d}w)\right] = \prod_{w \le t} \left[1 - \frac{N(s, \Delta w)}{Y(s,w)}\right].$$

- Case where $s \to \infty$ and under fixed (Type I) censoring considered by Gill (1980,1981).

Asymptotic Properties

- **Generalized NAE:** If $s \in (0, \infty)$ and $t^* \in (0, \infty)$ such that $y(s, t^*) > 0$ and if $\Lambda(t^*) < \infty$, then

$$\{V(s,t) = \sqrt{n}[\hat{\Lambda}(s,t) - \Lambda(t)] : t \in [0, t^*]\} \Rightarrow \{V^\infty(s,t) : t \in [0, t^*]\},$$

a zero-mean Gaussian process with covariance function

$$\mathrm{Cov}\{V^\infty(s,t_1), V^\infty(s,t_2)\} = d[s, \min(t_1, t_2)]$$

$$d(s,t) = \int_0^t \frac{\Lambda(\mathrm{d}w)}{y(s,w)},$$

$$y(s,t) \equiv \bar{F}(t) \left\{ \bar{G}_s(t-) + \int_{[t,\infty)} R(w-t) \mathrm{d}G_s(w) \right\}.$$

- **Generalized PLE:** Under same conditions,

$$\{W(s,t) = \sqrt{n}[\hat{\bar{F}}(s,t) - \bar{F}(t)] : t \in [0, t^*]\} \Rightarrow \{W^\infty(s,t) : t \in [0, t^*]\},$$

a zero-mean Gaussian process with covariance function

$$\mathrm{Cov}[W^\infty(s,t_1), W^\infty(s,t_2)] = \bar{F}(t_1)\bar{F}(t_2) d[s, \min(t_1, t_2)].$$

"A Sense of Closure"

Complete Data

- $T_1, T_2, \ldots, T_n$ IID positive with continuous SF $\bar{F}(t) = \mathbf{P}\{T > t\}$.

- EDF of $\bar{F}$:

$$\hat{\bar{F}}(t) = \frac{1}{n} \sum_{i=1}^{n} I\{T_i > t\}.$$

- Asymptotics:

$$\sqrt{n}\left[\hat{\bar{F}} - \bar{F}\right] \Rightarrow W_1$$

  W_1 a zero-mean Gaussian process with variance function

$$v_1(t) = \bar{F}(t)F(t) = \bar{F}(t)^2 \int_0^t \frac{\mathrm{d}\Lambda(w)}{\bar{F}(w)}.$$

Right-Censored Data

- Setting:

$$\begin{array}{rl} \text{Failure Times:} & T_1, T_2, \ldots, T_n \text{ IID } \bar{F}(t) = \mathbf{P}\{T > t\} \\ \text{Censoring Times:} & C_1, C_2, \ldots, C_n \text{ IID } \bar{G}(t) = \mathbf{P}\{C > t\} \end{array}$$

- Right-Censored Data:

$$(Z_1, \delta_1), (Z_2, \delta_2), \ldots, (Z_n, \delta_n)$$

$$Z_i = \min\{T_i, C_i\} \quad \text{and} \quad \delta_i = I\{T_i \leq C_i\}.$$

- PLE of $\bar{F}$:

$$\hat{\bar{F}}(t) = \prod_{\{i:\ Z_{(i)} \leq t\}} \left[1 - \frac{1}{n_{(i)}}\right]^{\delta_{(i)}}$$

$Z_{(1)} < Z_{(2)} < \ldots < Z_{(n)}$ the *ordered* values of the Z_i's, and $\delta_{(i)}$'s are the associated δ_i's; and

$$n_{(i)} = \sum_{j=1}^{n} I\{Z_j \geq Z_{(i)}\} = \# \text{ at risk at } Z_{(i)}.$$

- Asymptotics:

$$\sqrt{n}\left[\hat{\bar{F}} - \bar{F}\right] \Rightarrow W_2$$

W_2 a zero-mean Gaussian process with variance function

$$v_2(t) = \bar{F}(t)^2 \int_0^t \frac{\mathrm{d}\Lambda(w)}{\bar{F}(w)\bar{G}(w)}.$$

- $v_2(t)$ reduces to $v_1(t)$ when $\bar{G}(w) = 1$.

Recurrent Data Setting

- Generalized PLE of $\bar{F}$ (no tied values):

$$\hat{\bar{F}}(t) = \prod_{i=1}^{n} \prod_{\left\{j:\ \begin{array}{c} T_{ij} \leq t \\ S_{ij} \leq \tau_i \end{array}\right\}} \left[1 - \frac{1}{Y(T_{ij})}\right]$$

where

$$Y(t) = \sum_{i=1}^{n} \left\{ \sum_{j=1}^{K_i} I\{T_{ij} \geq t\} + I\{\tau_i - S_{iK_i} \geq t\} \right\}.$$

- Asymptotics:

$$\sqrt{n}\left[\hat{\bar{F}} - \bar{F}\right] \Rightarrow W_3,$$

a zero-mean Gaussian process with variance function

$$v_3(t) = \bar{F}(t)^2 \int_0^t \frac{\mathrm{d}\Lambda(w)}{\bar{F}(w)\bar{G}(w)\left\{1 + [\bar{G}(w)]^{-1} \int_w^\infty R(u-w)\mathrm{d}G(u)\right\}},$$

$$R(t) = \sum_{j=1}^{\infty} F^{\star j}(t) = \text{renewal function of } F;$$

$$F^{\star j} = j\text{th convolution of } F, j = 1, 2, \ldots.$$

- Effect in variance of sum-quota accrual scheme:

$$\left\{1 + \frac{1}{\bar{G}(w)} \int_w^{\infty} R(u-w)\mathrm{d}G(u)\right\}^{-1}.$$

- **An Approximation:** For large t, $R(t) \approx \frac{t}{\mu_F}$ where μ_F is the mean of F. So,

$$v_3(t) = \bar{F}(t)^2 \int_0^t \frac{\mathrm{d}\Lambda(w)}{\bar{F}(w)\bar{G}(w)\left\{1 + (\mu_F)^{-1}\mu_G(w)\right\}},$$

with

$$\begin{aligned} \mu_G(w) &= \mathbf{E}\left\{\tau - w | \tau \geq w\right\} \\ &= \text{MRL of } \tau \text{ given } \tau \geq w. \end{aligned}$$

A Concrete Example

- Assume: $F = \mathrm{EXP}(\theta), G = \mathrm{EXP}(\eta)$.
- Then: $v_3(t) = \left(\frac{\eta}{\theta+\eta}\right) v_2(t)$.
- Therefore, if $\frac{1}{\theta} << \frac{1}{\eta}$, a considerable gain in efficiency accrues by using all the data compared to just using the first, possibly right-censored, observation.
- Furthermore,

$$d(s,t) = I\{t \leq s\} \times \\ \theta \int_0^t \frac{\exp\{(\theta+\eta)w\}}{1 + \frac{\theta}{\eta}\left[1 - \exp\{-\eta(s-w)\} - \eta(s-w)\exp\{-\eta(s-w)\}\right]} \mathrm{d}w.$$

- If $s \to \infty$:

$$d(\infty, t) = \frac{\theta\eta}{(\theta+\eta)^2}\left\{\exp\{(\theta+\eta)t\} - 1\right\}.$$

- Thus:

$$\mathrm{Avar}\left(\sqrt{n}\hat{\bar{F}}(\infty,t)\right) = \frac{\theta\eta}{(\theta+\eta)^2} \times \\ \exp\{-(\theta-\eta)t\}\left[1 - \exp\{-(\theta+\eta)t\}\right].$$

- How adequate are these asymptotic results for moderate sample sizes?
- Results of simulation under the exponential model follow.

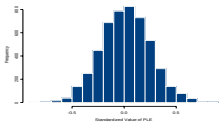
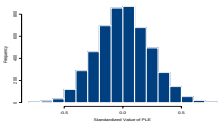
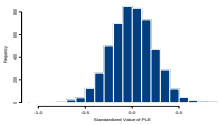
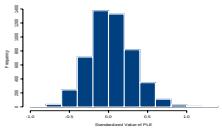
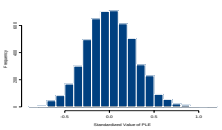
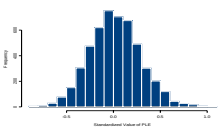
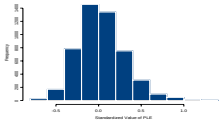
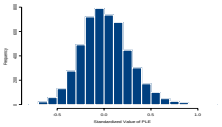
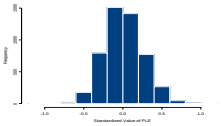
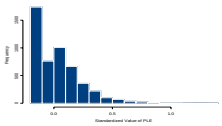
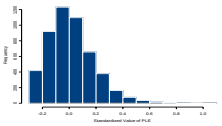
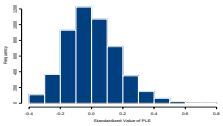
t ($\bar{F}(t)$)	Simulated Property of the Sampling Distribution	Sample Size (n)		
		10	30	50
.10 (.7408)	Mean	.0163	.0053	.0010
	Std. Error (Theoretical)	.2350 (.2250)	.2261 (.2250)	.2274 (.2250)
	Histogram			
.25 (.4723)	Mean	.0329	.0138	.0189
	Std. Error (Theoretical)	.2818 (.2681)	.2723 (.2681)	.2697 (.2681)
	Histogram			
.50 (.2231)	Mean	.0330	.0158	.0112
	Std. Error (Theoretical)	.2790 (.2442)	.2549 (.2442)	.2469 (.2442)
	Histogram			
1.0 (.0497)	Mean	.0312	.0123	.0033
	Std. Error (Theoretical)	.2007 (.1578)	.1738 (.1578)	.1668 (.1578)
	Histogram			

Table 1: Simulated properties of $\sqrt{n}[\hat{\bar{F}}(\infty, t) - \bar{F}(t)]$ under exponential model with $\theta = 3$ and $\eta = 1$.

Comparison with Wang and Chang Estimator

- Wang and Chang Estimator (JASA, 1999) of $\bar{F}$:

- $$K_i^* = \begin{cases} 1 & \text{if} \quad K_i = 0 \\ K_i & \text{if} \quad K_i > 0 \end{cases}$$

- $$d^*(t) = \sum_{i=1}^{n} \left\{ \frac{I\{K_i > 0\}}{K_i^*} \sum_{j=1}^{K_i} I\{T_{ij} = t\} \right\}$$

- $$R^*(t) = \sum_{i=1}^{n} \frac{1}{K_i^*} \left[\sum_{j=1}^{K_i} I\{T_{ij} \geq t\} + I\{\tau_i - S_{iK_i} \geq t\} I\{K_i = 0\} \right]$$

- $$\hat{S}(t) = \prod_{i=1}^{n} \prod_{\{j:\ T_{ij} \leq t\}} \left[1 - \frac{d^*(T_{ij})}{R^*(T_{ij})} \right]$$

- Estimator developed for a model with correlated data; so comparison a bit unfair to their estimator.

- Simulated comparison of $\hat{S}(t)$ and $\hat{\bar{F}}(\infty, t)$ for the exponential model.

- Simulated biases and root-mean-square errors (RMSE) obtained for several time points.

- Results of the simulation presented in Figure 2.

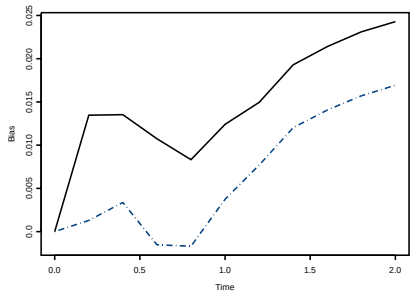
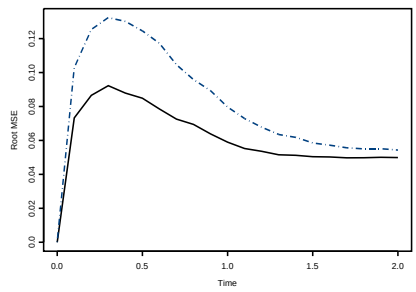
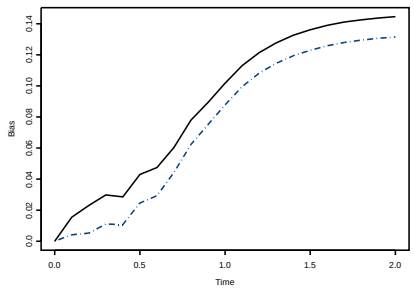
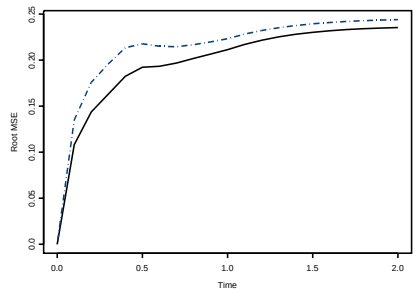
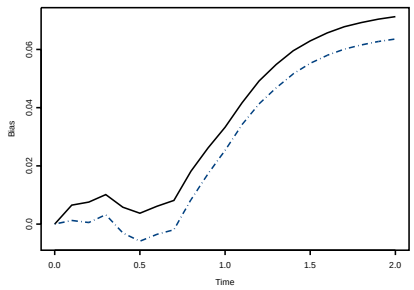
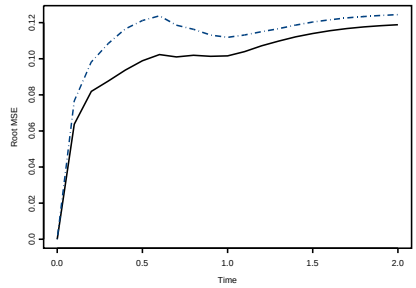
(n, θ, η)	Simulated Biases	Simulated RMSE
$(10, 3, 1)$		
$(10, 3, 3)$		
$(30, 3, 1)$		

Figure 2: Simulated biases and root-mean-squared errors of the estimator $\hat{\bar{F}}(\infty, t)$ (SOLID) and the WC estimator.

Application to a Real Data

- Aalen and Husebye (1991, *Stat in Medicine*): study concerning small bowel motility during the fasting state.
- Consecutive migrating motor complex (MMC) periods for 19 healthy subjects (15 men and 4 women).

Unit # (i)	# of Complete Obs. (K_i)	Complete Observed Periods (T_{ij}'s)	Censored Obs. ($\tau_i - S_{iK_i}$)
1	8	112 145 39 52 21 34 33 51	54
2	2	206 147	30
3	3	284 59 186	4
4	3	94 98 84	87
5	1	67	131
6	9	124 34 87 75 43 38 58 142 75	23
7	5	116 71 83 68 125	111
8	4	111 59 47 95	110
9	4	98 161 154 55	44
10	2	166 56	122
11	5	63 90 63 103 51	85
12	4	47 86 68 144	72
13	3	120 106 176	6
14	4	112 25 57 166	85
15	3	132 267 89	86
16	5	120 47 165 64 113	12
17	4	162 141 107 69	39
18	6	106 56 158 41 41 168	13
19	5	147 134 78 66 100	4

Table 2: Gastroenterology data set from Aalen and Husebye (1991) consisting of the migrating motor complex (MMC) periods (in minutes) for 19 individuals.

- Estimate of the mean MMC period is: 104 minutes (s.e. = 5.87 minutes); while from Aalen and Husebye's variance component model, they obtained the mean estimate of 106.8 minutes (s.e. = 6.9 minutes).

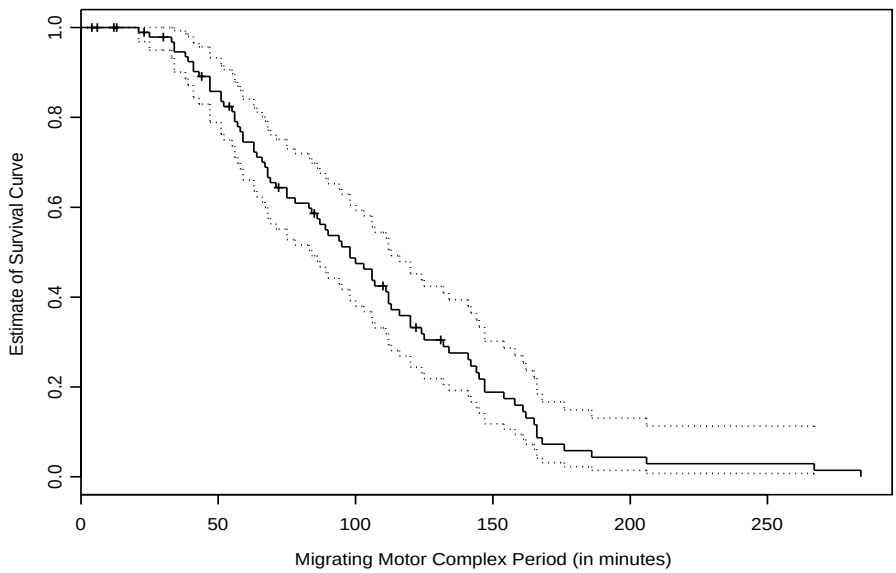


Figure 3: Graph of the estimate of the MMC period survival function together with its asymptotic 95% pointwise confidence interval.

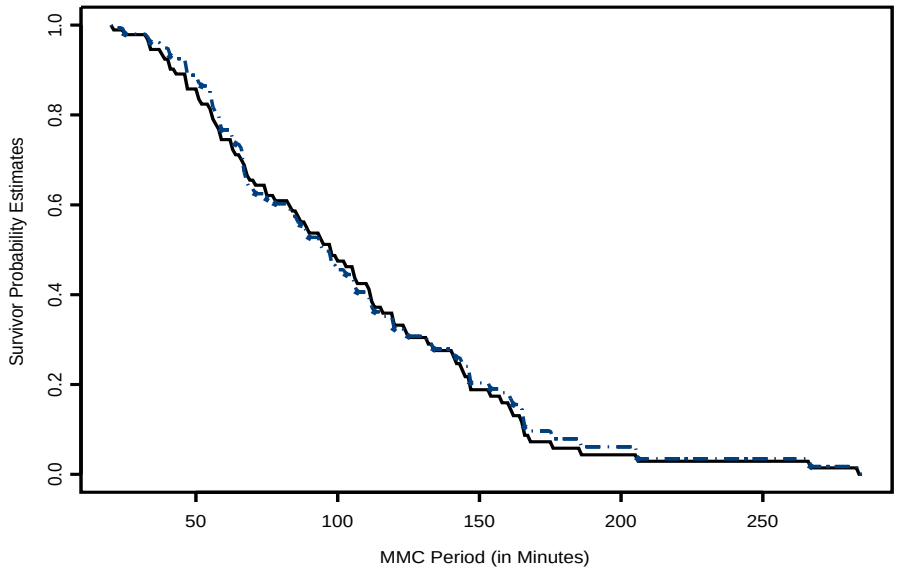


Figure 4: Graphs of the generalized PLE (solid line) and the Wang and Chang PLE (dashed line) for the MMC Data.

Related Problems

- **Goodness-of-Fit Problem:** $F = F_0$? (the simple case), or $F \in \{F(\cdot;\theta) : \theta \in \Theta\}$? (the composite case). The idea introduced in Peña (1998) and Agustin and Peña (2000) for hazard-based smooth goodness-of-fit tests apply to this situation. Still being worked out!
- Frailty or random mixing component in the model.
- Model validation and generalized residuals?
- Testing that the renewal assumption is valid.