Modeling and Analysis of Recurrent Events

Edsel A. Peña

Department of Statistics

University of South Carolina, Columbia, SC 29208

Talk at the

Conference on Frontiers in Applied and Computational Mathematics

New Jersey Institute of Technology

May 20, 2008

Some (Recurrent) Events

- Submission of a manuscript for publication.
- Occurrence of tumor.
- Onset of depression.
- Patient hospitalization.
- Machine/system failure.
- Occurrence of a natural disaster.
- Non-life insurance claim.
- Change in job.
- Onset of economic recession.
- At least a 200 points decrease in the DJIA.
- Marital disagreement.

Event Times and Distributions

- T : the time to the occurrence of an event of interest.
- $F(t) = \Pr\{T \leq t\}$: the distribution function of T.
- $S(t) = \bar{F}(t) = 1 - F(t)$: survivor/reliability function.
- Hazard rate/ probability and Cumulative Hazards:

$$\text{Cont:} \quad \lambda(t)dt \approx \Pr\{T \leq t + dt | T \geq t\} = \frac{f(t)}{S(t-)}dt$$

$$\text{Disc:} \quad \lambda(t_j) = \Pr\{T = t_j | T \geq t_j\} = \frac{f(t_j)}{S(t_j-)}$$

$$\text{Cumulative:} \quad \Lambda(t) = \int_0^t \lambda(w)dw \quad \text{or} \quad \Lambda(t) = \sum_{t_j \leq t} \lambda(t_j)$$

Representation/Relationships

- $0 < t_1 < \ldots < t_M = t, \mathcal{M}(t) = \max |t_i - t_{i-1}| = o(1)$,

$$\begin{aligned} S(t) = \Pr\{T > t\} &= \prod_{i=1}^{M} \Pr\{T > t_i | T \geq t_{i-1}\} \\ &\approx \prod_{i=1}^{M} \left[1 - \{\Lambda(t_i) - \Lambda(t_{i-1})\}\right]. \end{aligned}$$

- Identities:

$$S(t) = \prod_{w \leq t} [1 - \Lambda(dw)]$$

$$\Lambda(t) = \int_0^t \frac{dF(w)}{1 - F(w-)}$$

Estimating F

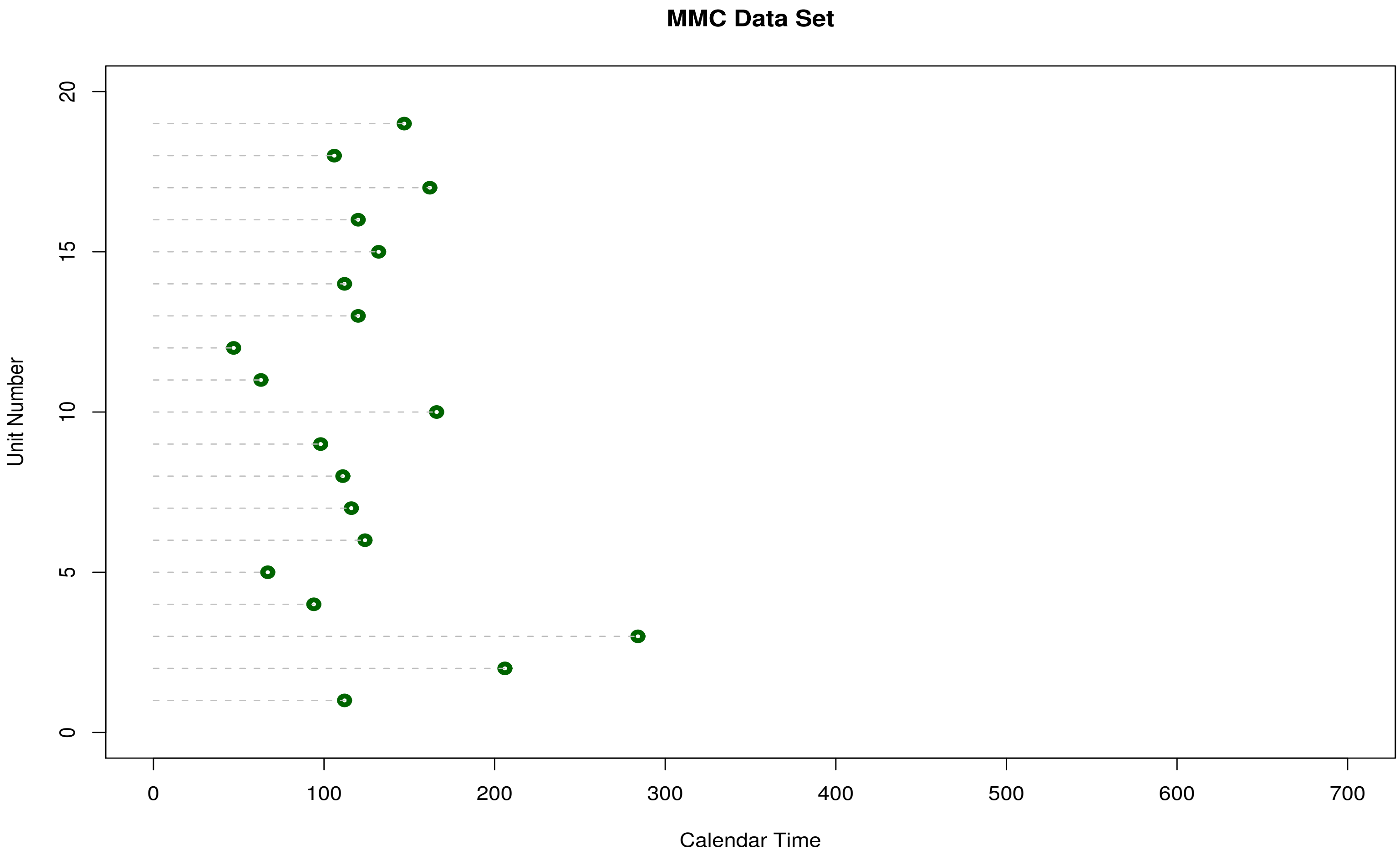
- A Classic Problem: Given $T_1, T_2, \ldots, T_n$ IID F, obtain an estimator $\hat{F}$ of F.
- Importance?
  - $\theta(F)$ of F (e.g., mean, median, variance) estimated via $\hat{\theta} = \theta(\hat{F})$.
  - Prediction of time-to-event for new units.
  - Comparing groups, e.g., thru a statistic

$$Q = \int W(t) d\left[\hat{F}_1(t) - \hat{F}_2(t)\right]$$

where $W(t)$ is some weight function.

Gastroenterology Data: Aalen and Husebye ('91)

Migratory Motor Complex (MMC) Times for 19 Subjects



Question: How to estimate the MMC period dist, F?

Parametric Approach

- Assume a Model: $\mathcal{F} = \{F(t;\theta) : \theta \in \Theta \subset \Re^p\}$
- Given $t_1, t_2, \ldots, t_n$, θ estimated by $\hat{\theta}$ such as ML.
- $L(\theta) = \prod_{i=1}^n f(t_i;\theta) = \prod_{i=1}^n \lambda(t_i;\theta)\exp\{-\Lambda(t_i;\theta)\}$
- $\hat{\theta} = \arg\max_\theta L(\theta)$
- $\hat{F}_{pa}(t) = F(t;\hat{\theta})$
- $\mathcal{I}(\theta) = \text{Var}\{\frac{\partial}{\partial\theta}\log f(T_1;\theta)\}$
- $\hat{\theta} \sim \text{AN}\left(\theta, \frac{1}{n}\mathcal{I}(\theta)^{-1}\right)$
- $\overset{\bullet}{F}(t;\theta) = \frac{\partial}{\partial\theta}F(t;\theta)$
- $\hat{F}_{pa}(t) \sim \text{AN}\left(F(t;\theta), \frac{1}{n}\overset{\bullet}{F}(t;\theta)'\mathcal{I}(\theta)^{-1}\overset{\bullet}{F}(t;\theta)\right)$

Nonparametric Approach

- No assumptions are made regarding the family of distributions to which the unknown df F belongs.
- Empirical Distribution Function (EDF):

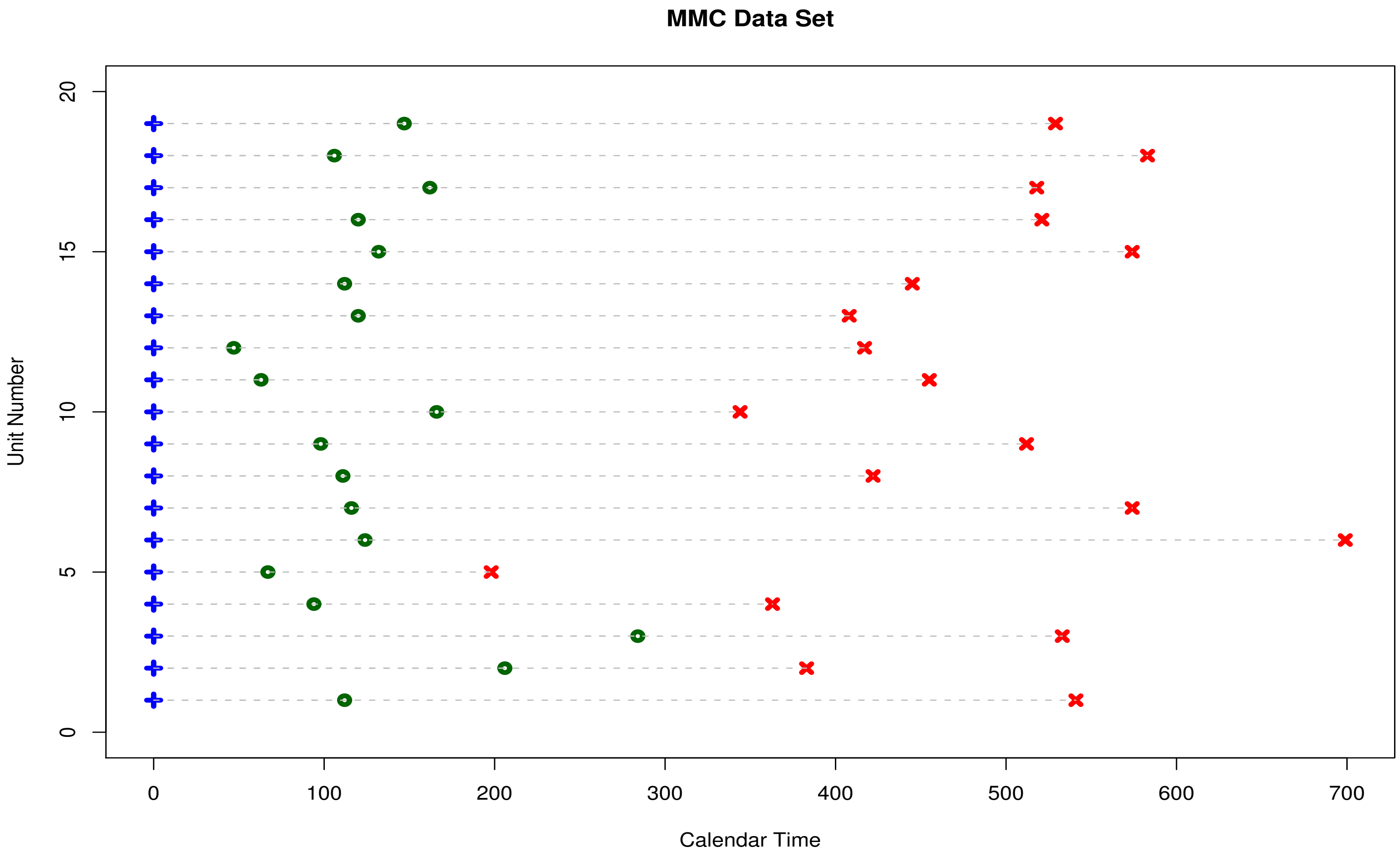
$$\hat{F}_{np}(t) = \frac{1}{n}\sum_{i=1}^{n} I\{T_i \le t\}$$

- $\hat{F}_{np}(\cdot)$ is a *nonparametric* MLE of F.
- Since $I\{T_i \le t\}, i = 1, 2, \ldots, n,$ are IID Ber$(F(t))$, by Central Limit Theorem,

$$\hat{F}_{np}(t) \sim AN\left(F(t), \frac{1}{n}F(t)[1 - F(t)]\right).$$

MMC Data: Censoring Aspect

For each unit, red mark is the potential termination time.



Remark: All 19 MMC times completely observed.

Estimation of F: With Censoring

- Right-censoring variables: $C_1, C_2, \ldots, C_n$ IID G.
- Observables: $(Z_i, \delta_i), i = 1, 2, \ldots, n$ with $Z_i = \min\{T_i, C_i\}$ and $\delta_i = I\{T_i \leq C_i\}$.
- Problem: Given (Z_i, δ_i)s, estimate df F or hazard function Λ of the T_is.
- Nonparametric Approaches:
  - Nonparametric MLE (Kaplan-Meier).
  - Martingale and method-of-moments.
- Pioneers: Kaplan & Meier; Efron; Nelson; Breslow; Breslow & Crowley; Aalen; Gill.

Product-Limit Estimator

- Counting and At-Risk Processes:

$$N(t) = \sum_{i=1}^{n} I\{Z_i \le t; \delta_i = 1\};$$
$$Y(t) = \sum_{i=1}^{n} I\{Z_i \ge t\}$$

- Hazard probability estimate at t:

$$\hat{\Lambda}(dt) = \frac{\Delta N(t)}{Y(t)} = \frac{\text{\# of Observed Failures at } t}{\text{\# at-risk at } t}$$

- Product-Limit Estimator (PLE):

$$1 - \hat{F}(t) = \hat{S}(t) = \prod_{w \le t} \left[1 - \frac{\Delta N(t)}{Y(t)}\right]$$

Stochastic Process Approach

- A martingale M is a zero-mean process which models a fair game. With $\mathcal{H}_t =$ history up to t:

$$E\{M(s+t)|\mathcal{H}_t\} = M(t).$$

- $M(t) = N(t) - \int_0^t Y(w)\Lambda(dw)$ is a martingale, so with $J(t) = I\{Y(t) > 0\}$ and stochastic integration,

$$E\left\{\int_0^t \frac{J(w)}{Y(w)} dN(w)\right\} = E\left\{\int_0^t J(w)\Lambda(dw)\right\}.$$

- Nelson-Aalen estimator of Λ, and PLE:

$$\hat{\Lambda}(t) = \int_0^t \frac{dN(w)}{Y(w)}, \quad \text{so} \quad \hat{S}(t) = \prod_{w \le t}[1 - \hat{\Lambda}(dw)].$$

Asymptotic Properties

- NAE: $\sqrt{n}[\hat{\Lambda}(t) - \Lambda(t)] \Rightarrow Z_1(t)$ with $\{Z_1(t) : t \geq 0\}$ a zero-mean *Gaussian process* with

$$d_1(t) = \mathrm{Var}(Z_1(t)) = \int_0^t \frac{\Lambda(dw)}{S(w)\bar{G}(w-)}.$$

- PLE: $\sqrt{n}[\hat{F}(t) - F(t)] \Rightarrow Z_2(t) \stackrel{st}{=} S(t)Z_1(t)$ so

$$d_2(t) = \mathrm{Var}(Z_2(t)) = S(t)^2 \int_0^t \frac{\Lambda(dw)}{S(w)\bar{G}(w-)}.$$

- If $\bar{G}(w) \equiv 1$ (no censoring), $d_2(t) = F(t)S(t)$!

Regression Models

- **Covariates:** temperature, degree of usage, stress level, age, blood pressure, race, etc.
- How to account of covariates to improve knowledge of time-to-event.
- Modelling approaches:
  - Log-linear models:

$$\log(T) = \beta' \mathbf{x} + \sigma\epsilon.$$

    The accelerated failure-time model. Error distribution to use? Normal errors not appropriate.
  - Hazard-based models: Cox proportional hazards (PH) model; Aalen's additive hazards model.

Cox ('72) PH Model: Single Event

- Conditional on x, hazard rate of T is:

$$\lambda(t|\mathbf{x}) = \lambda_0(t)\exp\{\beta'\mathbf{x}\}.$$

- $\hat{\beta}$ maximizes partial likelihood function of β:

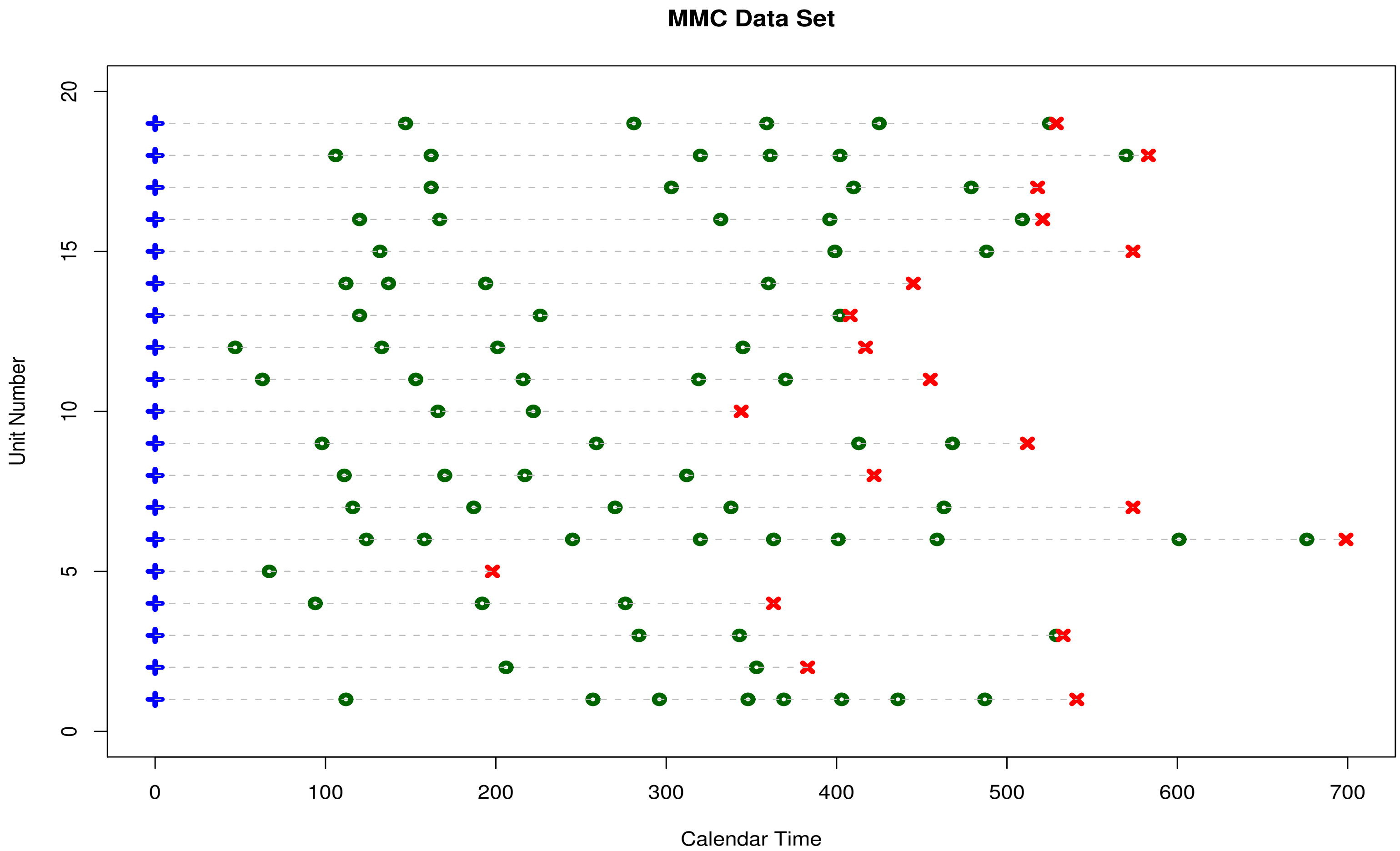
$$L_P(\beta) \equiv \prod_{i=1}^{n} \prod_{t<\infty} \left[\frac{\exp(\beta'\mathbf{x}_i)}{\sum_{j=1}^{n} Y_j(t)\exp(\beta'\mathbf{x}_j)} \right]^{\Delta N_i(t)}.$$

- Aalen-Breslow semiparametric estimator of $\Lambda_0(\cdot)$:

$$\hat{\Lambda}_0(t) = \int_0^t \frac{\sum_{i=1}^{n} dN_i(w)}{\sum_{i=1}^{n} Y_i(w)\exp(\hat{\beta}'\mathbf{x}_i)}.$$

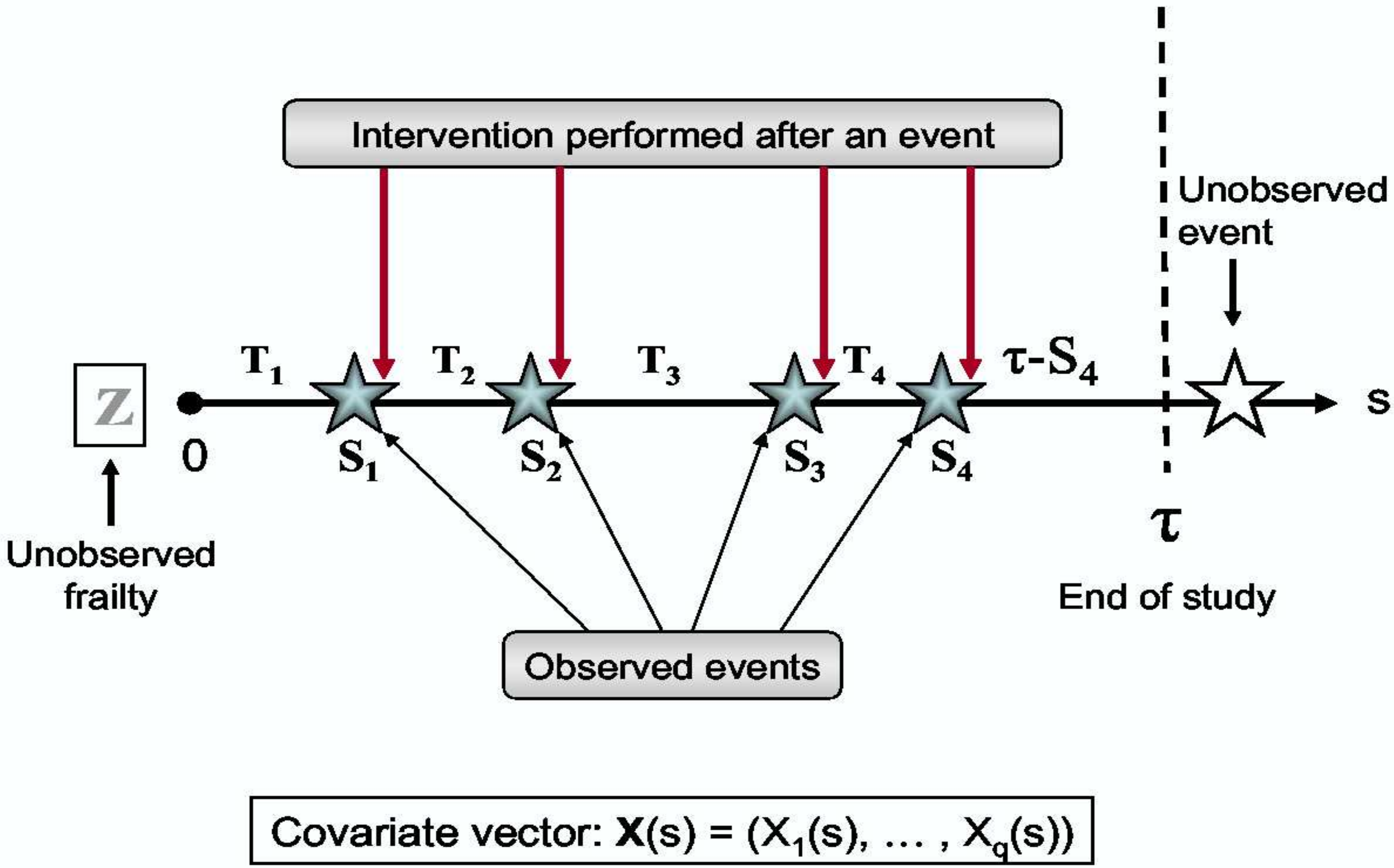
MMC Data: Recurrent Aspect

Aalen and Husebye ('91) Full Data



Problem: Estimate inter-event time distribution.

Representation: One Subject



Observables: One Subject

- $\mathbf{X}(s) =$ covariate vector, possibly time-dependent
- $T_1, T_2, T_3, \ldots =$ inter-event or gap times
- $S_1, S_2, S_3, \ldots =$ calendar times of event occurrences
- $\tau =$ end of observation period: Assume $\tau \sim G$
- $K = \max\{k : S_k \leq \tau\} =$ number of events in $[0, \tau]$
- $Z =$ unobserved frailty variable
- $N^\dagger(s) =$ number of events in $[0, s]$
- $Y^\dagger(s) = I\{\tau \geq s\} =$ at-risk indicator at time s
- $\mathbf{F}^\dagger = \{\mathcal{F}_s^\dagger : s \geq 0\} =$ filtration: information that includes interventions, covariates, etc.

Aspect of Sum-Quota Accrual

- Observed Number of Events:

$$K = \max\left\{k : \sum_{j=1}^{k} T_j \leq \tau\right\}$$

- Induced Constraint:

$$(T_1, T_2, \ldots, T_K) \text{ satisfies } \sum_{j=1}^{K} T_j \leq \tau < \sum_{j=1}^{K+1} T_j.$$

- $(K, T_1, T_2, \ldots, T_K)$ are all random and dependent, and K is informative about F.

Recurrent Event Models: IID Case

- Parametric Models:
  - HPP: $T_{i1}, T_{i2}, T_{i3}, \ldots$ IID EXP(λ).
  - IID Renewal Model: $T_{i1}, T_{i2}, T_{i3}, \ldots$ IID F where

$$F \in \mathcal{F} = \{F(\cdot;\theta) : \theta \in \Theta \subset \Re^p\};$$

    e.g., Weibull family; gamma family; etc.

- Non-Parametric Model: $T_{i1}, T_{i2}, T_{i3}, \ldots$ IID F which is some df.
- With Frailty: For each unit i, there is an *unobservable* Z_i from some distribution $H(\cdot;\xi)$ and $(T_{i1}, T_{i2}, T_{i3}, \ldots)$, given Z_i, are IID with survivor function

$$[1 - F(t)]^{Z_i}.$$

A General Class of Full Models

- Peña and Hollander (2004) model.

$$N^\dagger(s) = A^\dagger(s|Z) + M^\dagger(s|Z)$$

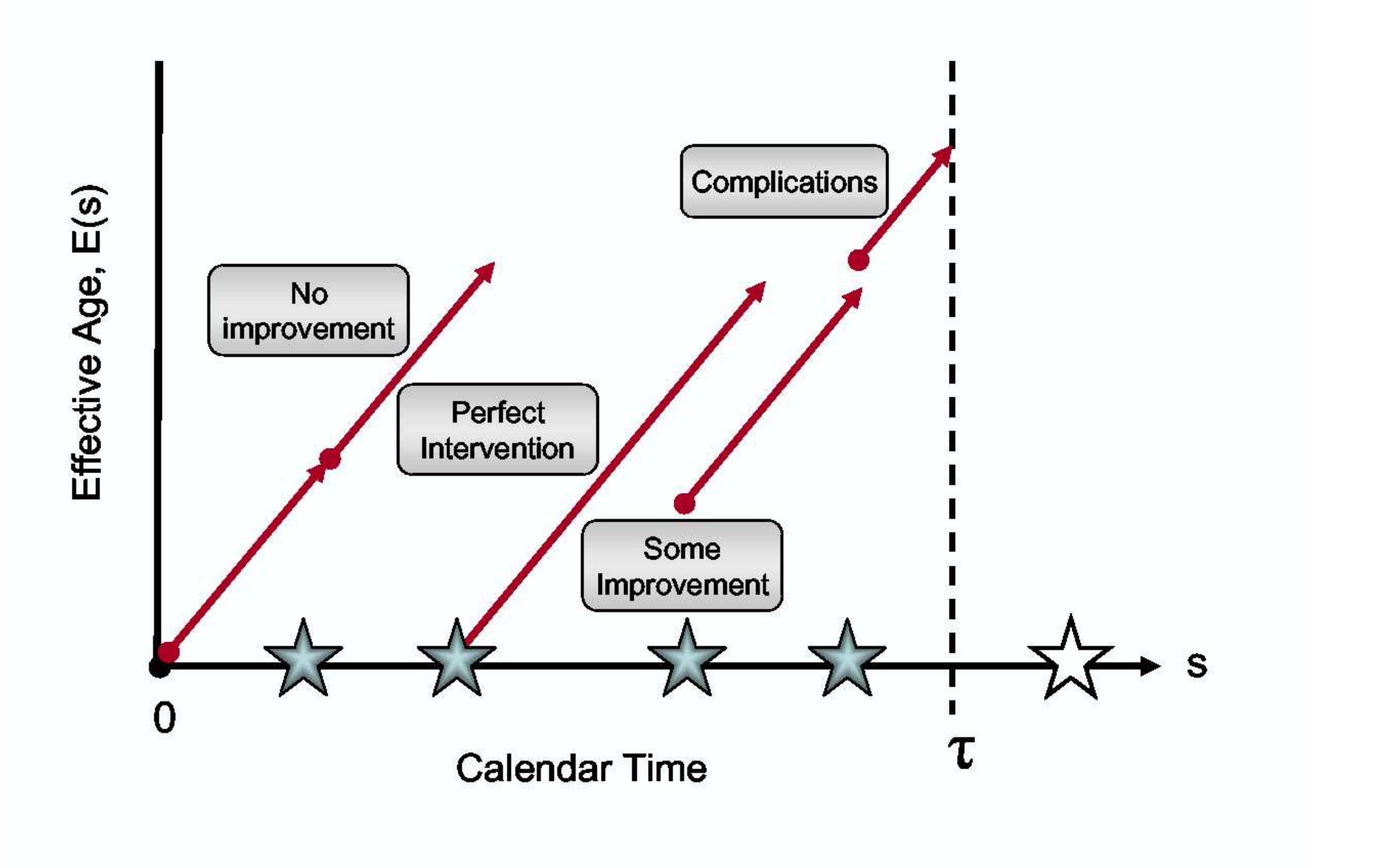
$$M^\dagger(s|Z) \in \mathcal{M}_0^2 = \text{sq-int martingales}$$

$$A^\dagger(s|Z) = \int_0^s Y^\dagger(w)\lambda(w|Z)dw$$

- Intensity Process:

$$\lambda(s|Z) = Z\,\lambda_0[\mathcal{E}(s)]\,\rho[N^\dagger(s-);\alpha]\,\psi[\beta^{\mathtt{t}}X(s)]$$

Effective Age Process: $\mathcal{E}(s)$



Effective Age Process, $\mathcal{E}(s)$

- PERFECT Intervention: $\mathcal{E}(s) = s - S_{N^\dagger(s-)}$.
- IMPERFECT Intervention: $\mathcal{E}(s) = s$.
- MINIMAL Intervention (BP '83; BBS '85):

$$\mathcal{E}(s) = s - S_{\Gamma_{\eta(s-)}}$$

where, with $I_1, I_2, \ldots$ IID BER(p),

$$\eta(s) = \sum_{i=1}^{N^\dagger(s)} I_i \quad \text{and} \quad \Gamma_k = \min\{j > \Gamma_{k-1} : I_j = 1\}.$$

Semi-Parametric Estimation: No Frailty

Observed Data for n Subjects:

$$\{(\mathbf{X}_i(s), N_i^\dagger(s), Y_i^\dagger(s), \mathcal{E}_i(s)) : \ 0 \le s \le s^*\}, i = 1, \ldots, n$$

$$N_i^\dagger(s) = \text{\# of events in } [0, s] \text{ for } i\text{th unit}$$

$$Y_i^\dagger(s) = \text{at-risk indicator at } s \text{ for } i\text{th unit}$$

with the model for the 'signal' being

$$A_i^\dagger(s) = \int_0^s Y_i^\dagger(v)\, \rho[N_i^\dagger(v-); \alpha]\, \psi[\beta^{\mathsf{t}} \mathbf{X}_i(v)]\, \lambda_0[\mathcal{E}_i(v)] dv$$

where $\lambda_0(\cdot)$ is an unspecified baseline hazard rate function.

Processes and Notations

Calendar/Gap Time Processes:

$$N_i(s,t) = \int_0^s I\{\mathcal{E}_i(v) \le t\} N_i^{\dagger}(dv)$$

$$A_i(s,t) = \int_0^s I\{\mathcal{E}_i(v) \le t\} A_i^{\dagger}(dv)$$

Notational Reductions:

$$\mathcal{E}_{ij-1}(v) \equiv \mathcal{E}_i(v) I_{(S_{ij-1}, S_{ij}]}(v) I\{Y_i^{\dagger}(v) > 0\}$$

$$\varphi_{ij-1}(w|\alpha,\beta) \equiv \frac{\rho(j-1;\alpha)\psi\{\beta^{\mathsf{t}}\mathbf{X}_i[\mathcal{E}_{ij-1}^{-1}(w)]\}}{\mathcal{E}'_{ij-1}[\mathcal{E}_{ij-1}^{-1}(w)]}$$

Generalized At-Risk Process

$$Y_i(s, w|\alpha, \beta) \equiv$$

$$\sum_{j=1}^{N_i^\dagger(s-)} I_{(\mathcal{E}_{ij-1}(S_{ij-1}),\ \mathcal{E}_{ij-1}(S_{ij})]}(w)\ \varphi_{ij-1}(w|\alpha, \beta) +$$

$$I_{(\mathcal{E}_{iN_i^\dagger(s-)}(S_{iN_i^\dagger(s-)}),\ \mathcal{E}_{iN_i^\dagger(s-)}((s \wedge \tau_i))]}(w)\ \varphi_{iN_i^\dagger(s-)}(w|\alpha, \beta)$$

For IID Renewal Model (PSH, 01) this simplifies to:

$$Y_i(s, w) = \sum_{j=1}^{N_i^\dagger(s-)} I\{T_{ij} \geq w\} + I\{(s \wedge \tau_i) - S_{iN_i^\dagger(s-)} \geq w\}$$

Estimation of Λ_0

$$A_i(s,t|\alpha,\beta) = \int_0^t Y_i(s,w|\alpha,\beta)\Lambda_0(dw)$$

$$S_0(s,t|\alpha,\beta) = \sum_{i=1}^n Y_i(s,t|\alpha,\beta)$$

$$J(s,t|\alpha,\beta) = I\{S_0(s,t|\alpha,\beta) > 0\}$$

Generalized Nelson-Aalen 'Estimator':

$$\hat{\Lambda}_0(s,t|\alpha,\beta) = \int_0^t \left\{\frac{J(s,w|\alpha,\beta)}{S_0(s,w|\alpha,\beta)}\right\}\left\{\sum_{i=1}^n N_i(s,dw)\right\}$$

Estimation of α and β

- Partial Likelihood (PL) Process:

$$L_P(s^*|\alpha,\beta) = \prod_{i=1}^{n} \prod_{j=1}^{N_i^\dagger(s^*)} \left[\frac{\rho(j-1;\alpha)\psi[\beta^{\mathtt{t}}\mathbf{X}_i(S_{ij})]}{S_0[s^*,\mathcal{E}_i(S_{ij})|\alpha,\beta]}\right]^{\Delta N_i^\dagger(S_{ij})}$$

- PL-MLE: $\hat{\alpha}$ and $\hat{\beta}$ are maximizers of the mapping

$$(\alpha,\beta) \mapsto L_P(s^*|\alpha,\beta)$$

- Iterative procedures. Implemented in an `R` package called `gcmrec` (Gonzaléz, Slate, Peña '04).

Estimation of $\bar{F}_0$

- G-NAE of $\Lambda_0(\cdot)$: $\hat{\Lambda}_0(s^*, t) \equiv \hat{\Lambda}_0(s^*, t|\hat{\alpha}, \hat{\beta})$
- G-PLE of $\bar{F}_0(t)$:

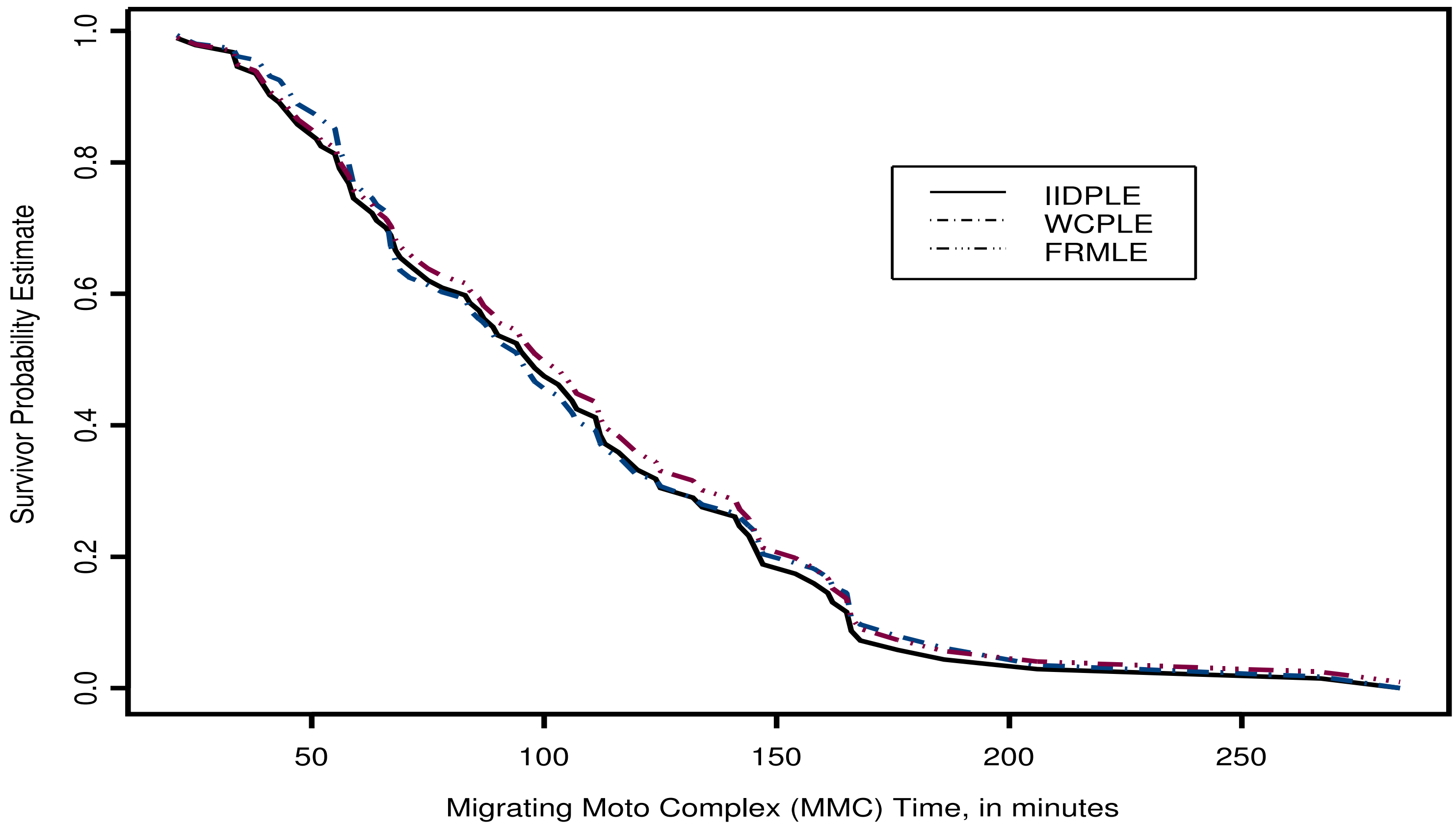
$$\hat{\bar{F}}_0(s^*, t) \;=\; \prod_{w=0}^{t}\left[1 - \frac{\sum_{i=1}^{n} N_i(s^*, dw)}{S_0(s^*, w|\hat{\alpha}, \hat{\beta})}\right]$$

- For IID renewal model with $\mathcal{E}_i(s) = s - S_{iN_i^{\dagger}(s-)}$, $\rho(k;\alpha) = 1$, and $\psi(w) = 1$, the generalized product-limit estimator in PSH (2001, JASA) obtains.

First Application: MMC Data Set

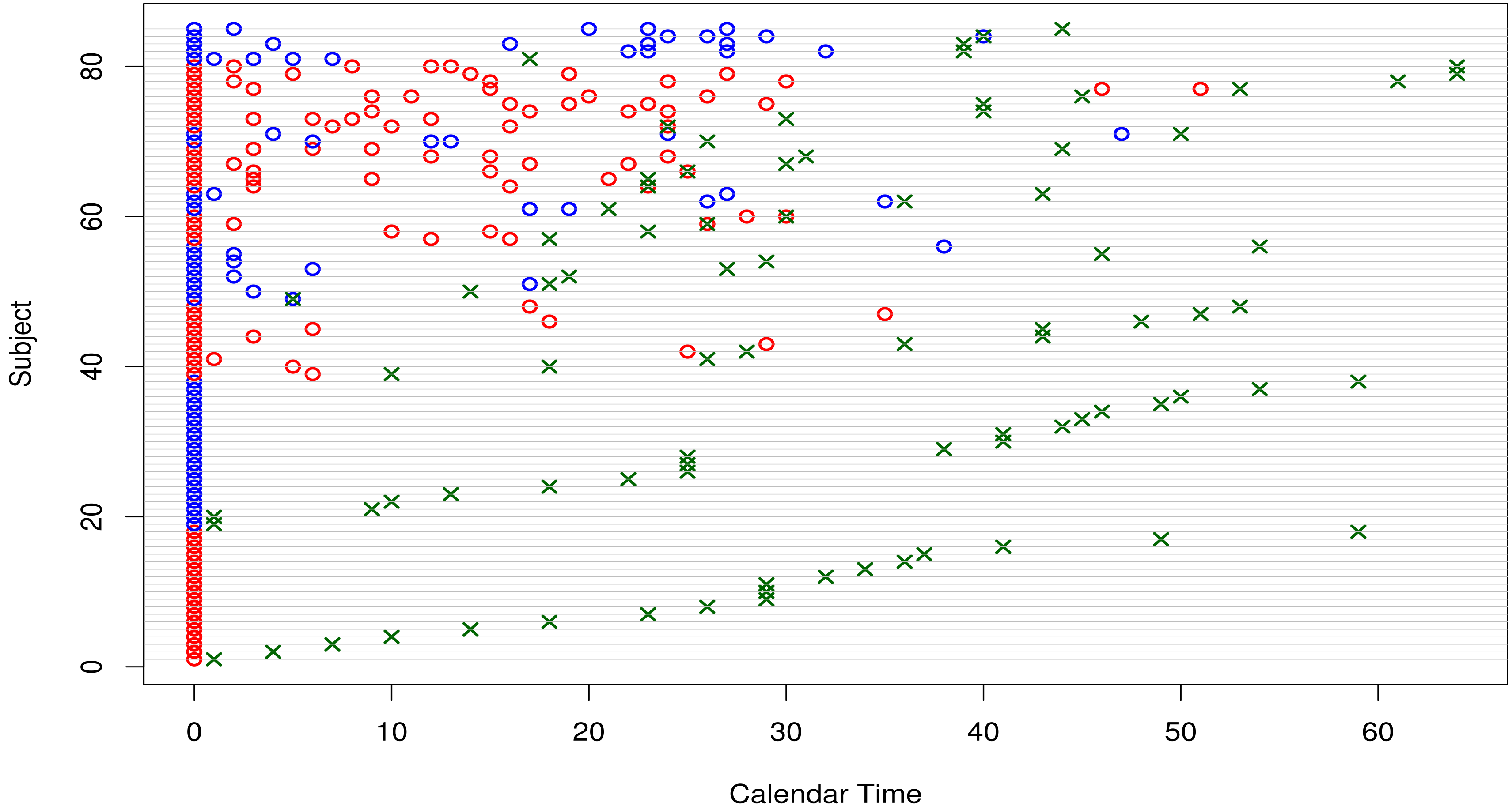
Aalen and Husebye (1991) Data

Estimates of distribution of MMC period



Second Application: Bladder Data Set

Bladder cancer data pertaining to times to recurrence for $n = 85$ subjects studied in Wei, Lin and Weissfeld ('89).



Results and Comparisons

- Estimates from Different Methods for Bladder Data

Cova	Para	AG	WLW	PWP	General Model	
			Marginal	Cond*nal	Perfect[a]	Minimal[b]
$\log N(t-)$	α	-	-	-	.98 (.07)	.79
Frailty	ξ	-	-	-	∞	.97
rx	β_1	-.47 (.20)	**-.58 (.20)**	***-.33 (.21)***	***-.32 (.21)***	**-.57**
Size	β_2	-.04 (.07)	**-.05 (.07)**	***-.01 (.07)***	***-.02 (.07)***	**-.03**
Number	β_3	.18 (.05)	**.21 (.05)**	***.12 (.05)***	***.14 (.05)***	**.22**

[a] Effective Age is backward recurrence time ($\mathcal{E}(s) = s - S_{N^\dagger(s-)}$).
[b] Effective Age is calendar time ($\mathcal{E}(s) = s$).

Details: Peña, Slate, and Gonzalez (2007). JSPI.

Sum-Quota Effect: IID Renewal

- Generalized product-limit estimator $\hat{F}$ of common gap-time df F presented in PSH (2001, JASA).

$$\sqrt{n}(\hat{\bar{F}}(\cdot) - \bar{F}(\cdot)) \Longrightarrow \text{GP}(0, \sigma^2(\cdot))$$

$$\sigma^2(t) = \bar{F}(t)^2 \int_0^t \frac{d\Lambda(w)}{\bar{F}(w)\bar{G}(w-)\left[1 + \nu(w)\right]}$$

$$\nu(w) = \frac{1}{\bar{G}(w-)} \int_w^\infty \rho^*(v - w) dG(v)$$

$$\rho^*(\cdot) = \sum_{j=1}^{\infty} F^{\star j}(\cdot) = \text{renewal function}$$

Efficiency: Some Questions

- Is it worth using the additional event recurrences in the analysis? How much do we gain in efficiency?
- Impact of G, the distribution of the τ_is, when G is related to inter-event time distribution? Loss if this informative monitoring structure is ignored?
- (In)Efficiency of GPLE relative to estimator which exploits informative monitoring structure?
- What is a reasonable informative monitoring model for examining these questions?
- Could we extend similar studies that were performed for the PLE using the so-called Koziol-Green Model?

Koziol-Green Model

- Koziol & Green (1976); Chen Hollander & Langberg (1982); Cheng & Lin (1989)
- $T \sim F$ and $C \sim G$ with T failure time and C right censoring time.
- Assumption: $1 - G = (1 - F)^{\beta}$ for some $\beta \geq 0$.
- $Z = \min(T, C)$ and $\delta = I\{T \leq C\}$ are independent; $Z \sim \bar{H} = \bar{F}^{\beta+1}$; $\delta \sim Ber(1/(1+\beta))$; $\beta =$ censoring parameter.
- CHL: Exact properties of the PLE: mean, variance, mean-squared error.
- CL: Efficiency of PLE relative to estimator that exploits KG assumption.

Generalized KG: Recurrent Events

- For a unit or subject,
- Inter-event times T_js are IID F;
- End-of-monitoring time τ has distribution G.
- Assumption: $1 - G = (1 - F)^{\beta}$ for some $\beta \geq 0$.
- Remark: Independence property that allowed exact derivations in right-censored single-event settings does not play a role in this recurrent event setting.
- Efficiency comparisons performed via asymptotic analysis and through computer simulations.
- Two Cases: (i) F is exponential, and (ii) F is two-parameter Weibull.

Ignoring Informative Structure

- When F is exponential, so model is HPP.
- $\hat{\theta}_n$: estimator that exploits informative monitoring.
- $\tilde{\theta}_n$: estimator that ignores informative monitoring.
- Efficiency Result: $\Delta ARE(\hat{\theta}_n : \tilde{\theta}_n) = 0$.
- Surprising result!? It turns out that in the exponential setting, these two estimators are identical.
- The estimators are:

$$\hat{\theta}_n = \tilde{\theta}_n = \frac{\sum_{i=1}^n K_i}{\sum_{i=1}^n \tau_i}$$

Single-Event Analysis

- Single-Event Analysis: Only the first, possibly right-censored, observations are used in the statistical analysis?
- $\check{\theta}$: depends only on the first event times.
- When F is exponential, we have

$$\Delta ARE(\hat{\theta}_n : \check{\theta}_n) = \frac{1}{\beta}.$$

- $1/\beta$: (approximate) expected number of events per unit.

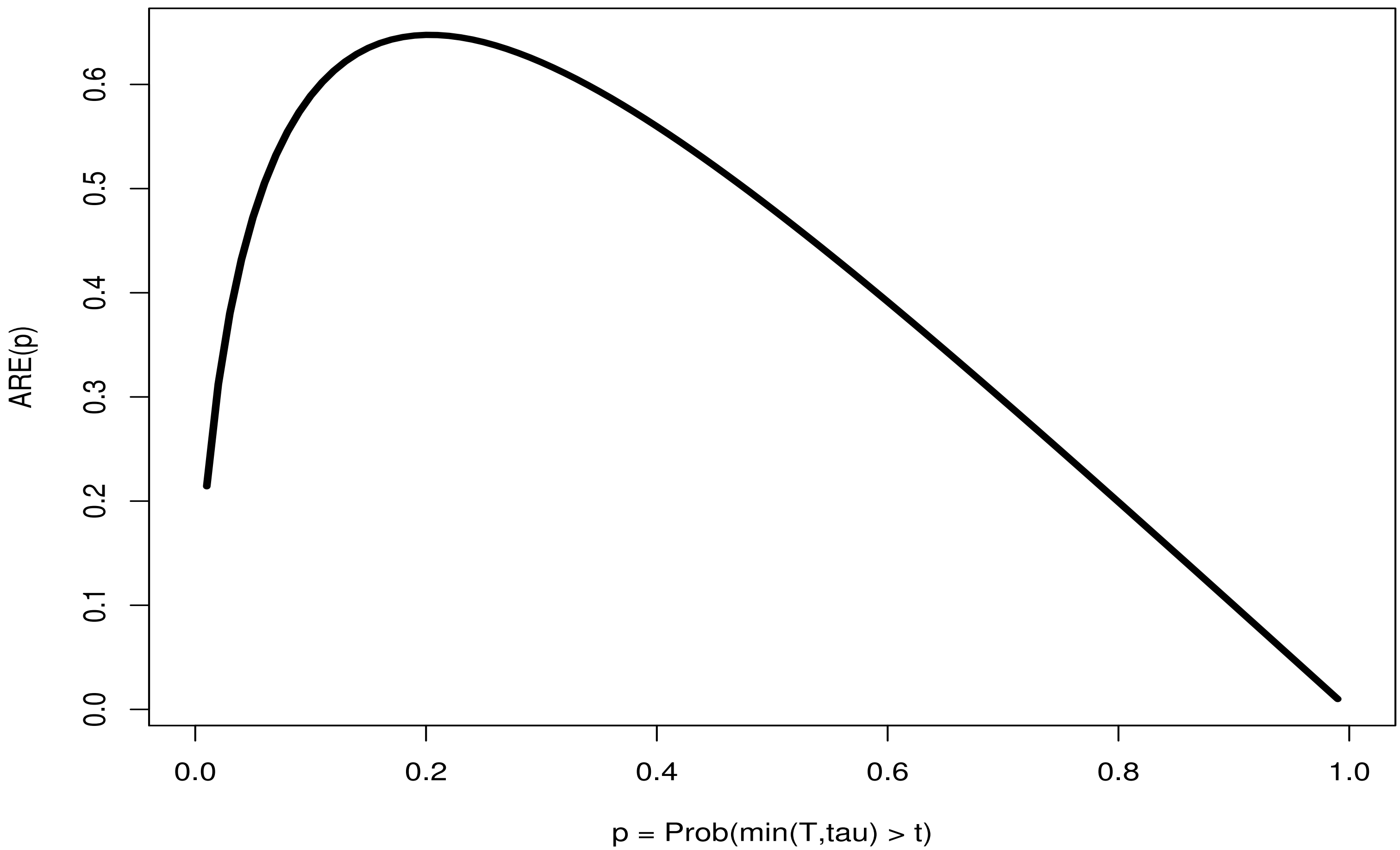
Inefficiency of GPLE

- How inefficient is the generalized PLE of F compared to the parametric estimator that exploits informative monitoring structure?
- $\hat{F}_n$: parametric estimator and exploits informative structure.
- $\tilde{F}_n$: generalized PLE in PSH (JASA, 01).
- When F is exponential,

$$ARE(\tilde{F}_n(t) : \hat{F}_n(t)) = \frac{[(1+\beta)t]^2}{\exp[(1+\beta)t] - 1}, t \geq 0.$$

Efficiency Plot: Exponential F

GPLE ($\tilde{F}$) versus Parametric Estimator ($\hat{F}$)

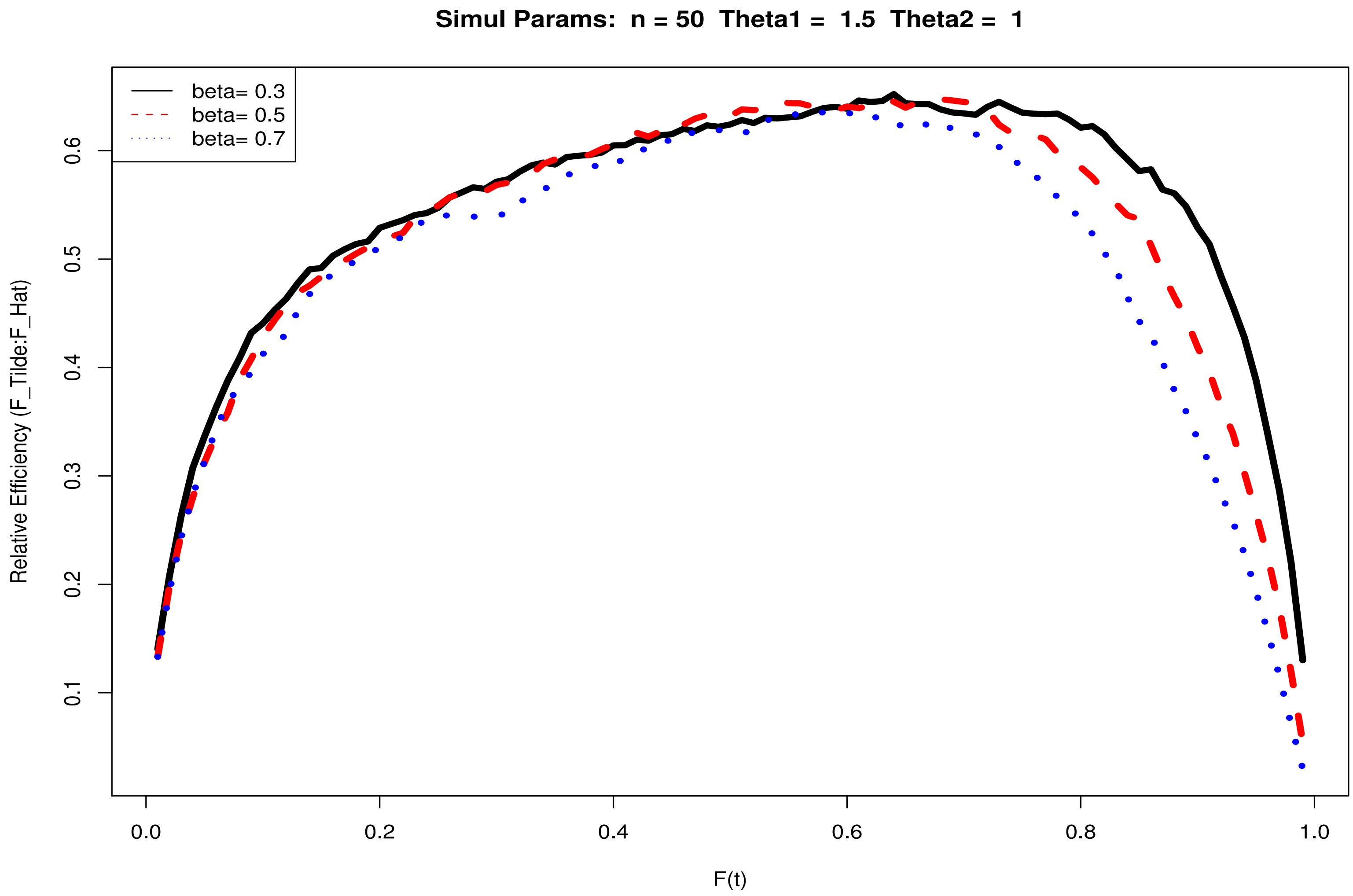


Efficiencies: Weibull F

n	θ_1	θ_2	β	MeanEvs	Eff$(\hat{\theta} : \tilde{\theta})$	Eff$(\hat{\theta} : \check{\theta})$
50	0.9	1	0.3	3.90	1.26	30.27
50	0.9	1	0.5	2.25	1.41	13.48
50	0.9	1	0.7	1.57	1.66	7.96
50	1.0	1	0.3	3.34	1.30	23.29
50	1.0	1	0.5	2.00	1.53	9.91
50	1.0	1	0.7	1.42	1.71	6.69
50	1.5	1	0.3	1.97	1.49	9.53
50	1.5	1	0.5	1.34	1.75	5.55
50	1.5	1	0.7	1.02	2.11	3.82

Efficiency Plot: Weibull F

GPLE ($\tilde{F}$) versus Parametric Estimator ($\hat{F}$)



On Marginal Modeling: WLW and PWP

- k_0 specified (usually the maximum value of the observed Ks).
- Assume a Cox PH-type model for each S_k, $k = 1, \ldots, k_0$.
- Counting Processes ($k = 1, 2, \ldots, k_0$):

$$N_k(s) = I\{S_k \leq s; S_k \leq \tau\}$$

- At-Risk Processes ($k = 1, 2, \ldots, k_0$):

$$Y_k^{WLW}(s) = I\{S_k \geq s; \tau \geq s\}$$

$$Y_k^{PWP}(s) = I\{S_{k-1} < s \leq S_k; \tau \geq s\}$$

Working Model Specifications

- **WLW Model**

$$\left\{ N_k(s) - \int_0^s Y_k^{WLW}(v)\lambda_{0k}^{WLW}(v)\exp\{\beta_k^{WLW}X(v)\}dv \right\}$$

- **PWP Model**

$$\left\{ N_k(s) - \int_0^s Y_k^{PWP}(v)\lambda_{0k}^{PWP}(v)\exp\{\beta_k^{PWP}X(v)\}dv \right\}$$

- are *assumed* to be zero-mean martingales (in s).

Parameter Estimation

- See Therneau & Grambsch's book *Modeling Survival Data: Extending the Cox Model*.
- $\hat{\beta}_k^{WLW}$ and $\hat{\beta}_k^{PWP}$ obtained via partial likelihood (Cox (72) and Andersen and Gill (82)).
- Overall β-estimate:

$$\hat{\beta}^{WLW} = \sum_{k=1}^{k_0} \hat{c}_k \hat{\beta}_k^{WLW};$$

  c_ks being 'optimal' weights. See WLW paper.
- $\hat{\Lambda}_{0k}^{WLW}(\cdot)$ and $\hat{\Lambda}_{0k}^{PWP}(\cdot)$: Aalen-Breslow-Nelson type estimators.

Two Relevant Questions

- **Question 1:** When one assumes marginal models for S_ks that are of the Cox PH-type, does there exist a *full model* that *actually induces* such PH-type marginal models?

  Answer: YES, by a very nice paper by Nang and Ying (Biometrika:2001). BUT, the joint model obtained is rather 'limited'.

- **Question 2:** If one *assumes* Cox PH-type marginal models for the S_ks (or T_ks), but the *true* full model does not induce such PH-type marginal models *[which may usually be the case in practice]*, what are the *consequences*?

Case of the HPP Model

- *True Full Model:* for a unit with covariate $X = x$, events occur according to an HPP model with rate:

$$\lambda(t|x) = \theta \exp(\beta x).$$

- For this unit, inter-event times $T_k, k = 1, 2, \ldots$ are IID exponential with mean time $1/\lambda(t|x)$.
- Assume also that $X \sim BER(p)$ and $\mu_\tau = E(\tau)$.
- Main goal is to infer about the regression coefficient β which relates the covariate X to the event occurrences.

Full Model Analysis

- $\hat{\beta}$ solves

$$\frac{\sum X_i K_i}{\sum K_i} = \frac{\sum \tau_i X_i \exp(\beta X_i)}{\sum \tau_i \exp(\beta X_i)}.$$

- $\hat{\beta}$ does not directly depend on the S_{ij}s. Why?
- **Sufficiency:** (K_i, τ_i)s contain all information on (θ, β).

$$(S_{i1}, S_{i2}, \ldots, S_{iK_i}) | (K_i, \tau_i) \stackrel{d}{=} \tau_i (U_{(1)}, U_{(2)}, \ldots, U_{(K_i)}).$$

- Asymptotics:

$$\hat{\beta} \sim AN\left(\beta, \frac{1}{n} \frac{(1-p) + pe^{\beta}}{\mu_\tau \theta [(1-p) + pe^{\beta}]}\right).$$

Some Questions

- Under WLW or the PWP: how are β_k^{WLW} and β_k^{PWP} related to θ and β?
- Impact of event position k?
- Are we *ignoring* that K_is are informative? Why not also a marginal model on the K_is?
- Are we violating the *Sufficiency Principle*?
- Results simulation-based: Therneau & Grambsch book ('01) and Metcalfe & Thompson (SMMR, '07).
- Comment by D. Oakes that PWP estimates *less biased* than WLW estimates.

Properties of $\hat{\beta}_k^{WLW}$

- Let $\hat{\beta}_k^{WLW}$ be the partial likelihood MLE of β based on at-risk process $Y_k^{WLW}(v)$.
- **Question:** Does $\hat{\beta}_k^{WLW}$ converge to β?
- $g_k(w) = w^{k-1}e^{-w}/\Gamma(k)$: standard gamma pdf.
- $\bar{\mathcal{G}}_k(v) = \int_v^\infty g_k(w)dw$: standard gamma survivor function.
- $\bar{G}(\cdot)$: survivor function of τ.
- $E(\cdot)$: denotes expectation wrt X.

Limit Value (LV) of $\hat{\beta}_k^{WLW}$

- **Limit Value** $\beta_k^* = \beta_k^*(\theta, \beta)$ of $\hat{\beta}_k^{WLW}$: solution in β^* of

$$\int_0^\infty E(X\theta e^{\beta X} g_k(v\theta e^{\beta X}))\bar{G}(v)dv =$$

$$\int_0^\infty e_k^{WLW}(v;\theta,\beta,\beta^*)E(\theta e^{\beta X} g_k(v\theta e^{\beta X}))\bar{G}(v)dv$$

where

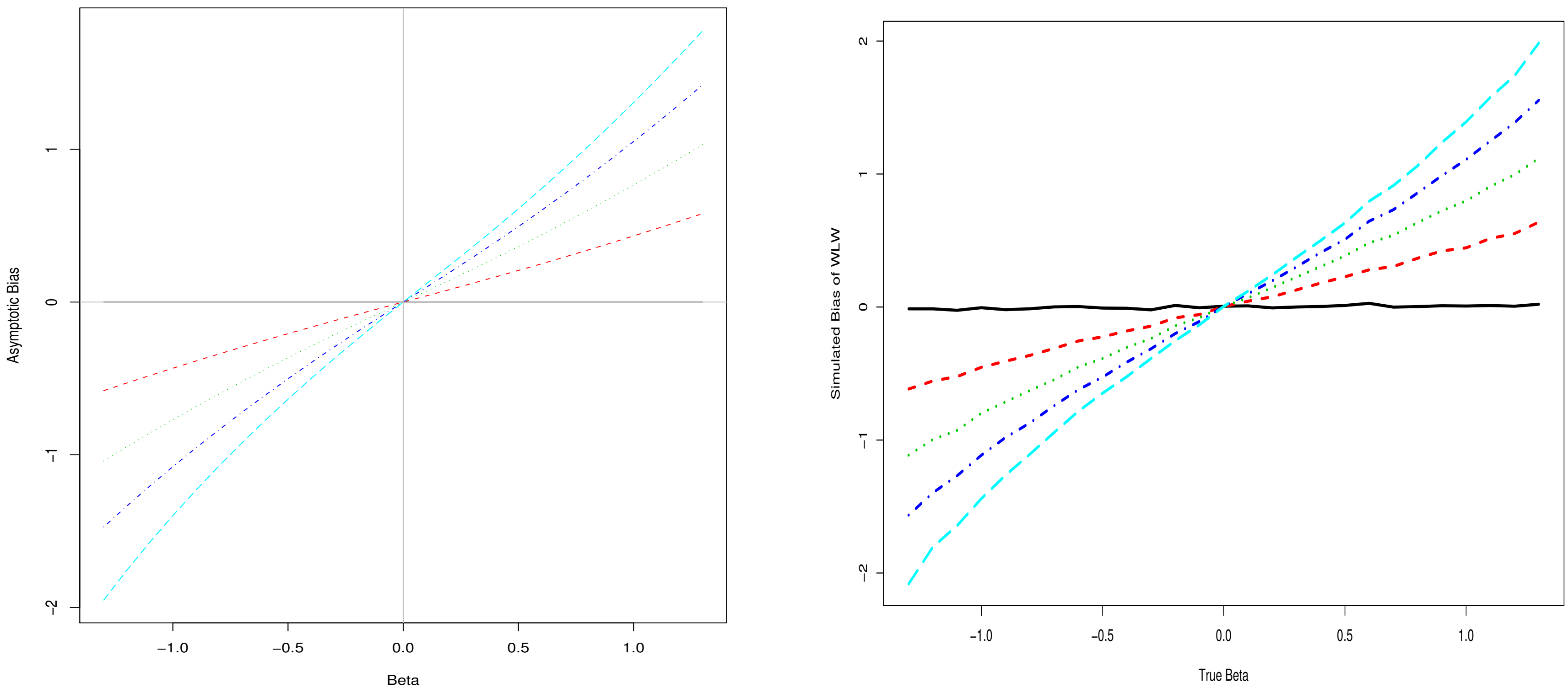
$$e_k^{WLW}(v;\theta,\beta,\beta^*) = \frac{E(Xe^{\beta^* X}\bar{\mathcal{G}}_k(v\theta e^{\beta X}))}{E(e^{\beta^* X}\bar{\mathcal{G}}_k(v\theta e^{\beta X}))}$$

- Asymptotic Bias of $\hat{\beta}_k^{WLW} = \beta_k^* - \beta$

Bias Plots for WLW Estimator

Colors pertain to value of k, the Event Position
$k = 1$: Black; $k = 2$: Red; $k = 3$: Green; $k = 4$: DarkBlue;
$k = 5$: LightBlue

Theoretical Simulated



On PWP Estimators

- Main Difference Between WLW and PWP:

$$E(Y_k^{WLW}(v)|X) = \bar{G}(v)\bar{\mathcal{G}}_k(v\theta\exp(\beta X));$$

$$E(Y_k^{PWP}(v)|X) = \bar{G}(v)\frac{g_k(v\theta\exp(\beta X))}{\theta\exp(\beta X)}.$$

- Leads to: $u_k^{PWP}(s;\theta,\beta) = 0$ for $k = 1, 2, \ldots$.
- $\hat{\beta}_k^{PWP}$ are **asymptotically unbiased** for β for each k (at least in this HPP model)!
- Theoretical result consistent with observed results from simulation studies and D. Oakes' observation.

Concluding Remarks

- Recurrent events prevalent in many areas.
- Dynamic models: accommodate unique aspects.
- More research in inference for dynamic models.
- *Current limitation:* tracking effective age.
- Efficiency gains with recurrences.
- Caution: informative aspects of model.
- Caution: marginal modeling approaches.
- *Dynamic recurrent event modeling*: challenging and still fertile!

Acknowledgements

- Thanks to NIH grants that partially support research.
- Research collaborators: Myles Hollander, Rob Strawderman, Elizabeth Slate, Juan Gonzalez, Russ Stocker, Akim Adekpedjou, and Jonathan Quiton.
- Thanks to current students: Alex McLain, Laura Taylor, Josh Habiger, and Wensong Wu.
- Thanks to all of you!