

An Appetizer

The important thing is not to stop questioning. Curiosity has its own reason for existing. One cannot help but be in awe when he contemplates the mysteries of eternity, of life, of the marvelous structure of reality. It is enough if one tries merely to comprehend a little of this mystery every day. Never lose a holy curiosity. Albert Einstein

Time-to-Event Modeling and Statistical Analysis

Edsel A. Peña

(Portions joint work with J. Gonzalez,
M. Hollander, E. Slate, R. Strawderman)

Department of Statistics

University of South Carolina, Columbia, SC

Research support from NIH Grants (RO1, COBRE)

Some Events of Interest

- Death.
- First publication after PhD graduation.
- Occurrence of tumor.
- Onset of depression.
- Machine/system failure.
- Occurrence of a natural disaster.
- Hospitalization.
- Non-life insurance claim.
- Accident or terrorist attack.
- Onset of economic recession.
- Divorce.

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:

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

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

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

Representation/Relationships

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

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

- S as a product-integral of Λ : When $\mathcal{M}(t) \rightarrow 0$,

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

- In general, Λ in terms of F : $\Lambda(t) = \int_0^t \frac{dF(w)}{1-F(w-)}.$

Estimation of F and Why?

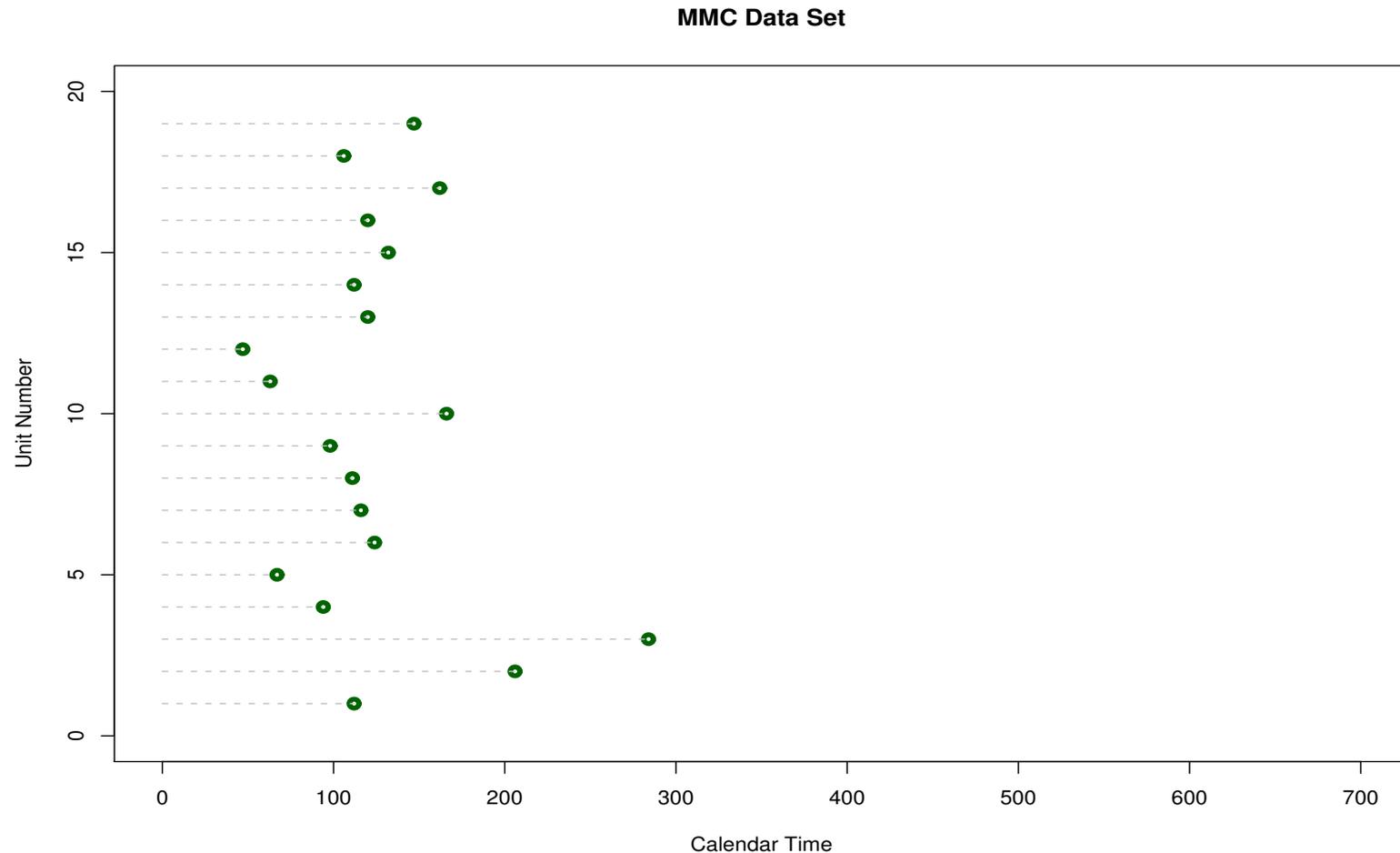
- **Most Basic Problem:** Given a sample T_1, T_2, \dots, T_n from an **unknown** distribution F , to obtain an estimator \hat{F} of F .
- **Why is it important to know how to estimate F ?**
 - Functionals/parameters $\theta(F)$ of F (e.g., mean, median, variance) can be estimated via $\hat{\theta} = \theta(\hat{F})$.
 - Prediction of time-to-event for new units.
 - Knowledge of population of units or event times.
 - For 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

- Unknown df F is assumed to belong to some parametric family (e.g., exponential, gamma, Weibull)

$$\mathcal{F} = \{F(t; \theta) : \theta \in \Theta \subset \mathbb{R}^p\}$$

with functional form of $F(\cdot; \cdot)$ known; θ is unknown.

- Based on data t_1, t_2, \dots, t_n , θ is estimated by $\hat{\theta}$, say, via **maximum likelihood (ML)**. $\hat{\theta}$ maximizes likelihood

$$L(\theta) = \prod_{i=1}^n f(t_i; \theta) = \prod_{i=1}^n \lambda(t_i; \theta) \exp\{-\Lambda(t_i; \theta)\}.$$

- The distribution function F is estimated by

$$\hat{F}_{pa}(t) = F(t; \hat{\theta}).$$

Parametric Estimation: Asymptotics

- When \mathcal{F} holds, MLE of θ satisfies

$$\hat{\theta} \sim \text{AN} \left(\theta, \frac{1}{n} \mathcal{I}(\theta)^{-1} \right);$$

$\mathcal{I}(\theta) = \text{Var} \left\{ \frac{\partial}{\partial \theta} \log f(T_1; \theta) \right\} =$ Fisher information.

- Therefore, when \mathcal{F} holds, by δ -method, with

$$\dot{F}(t; \theta) = \frac{\partial}{\partial \theta} F(t; \theta)$$

then

$$\hat{F}_{pa}(t) \sim \text{AN} \left(F(t; \theta), \frac{1}{n} \dot{F}(t; \theta)' \mathcal{I}(\theta)^{-1} \dot{F}(t; \theta) \right).$$

Nonparametric Approach

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

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

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

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

An Efficiency Comparison

- Assume that family $\mathcal{F} = \{F(t; \theta) : \theta \in \Theta\}$ holds. Both \hat{F}_{pa} and \hat{F}_{np} are **asymptotically unbiased**.
- To compare under \mathcal{F} , we take ratio of asymptotic variances to give the **efficiency** of parametric estimator over nonparametric estimator.

$$\text{Eff}(\hat{F}_{pa}(t) : \hat{F}_{np}(t)) = \frac{F(t; \theta)[1 - F(t; \theta)]}{\dot{F}(t; \theta)' \mathcal{I}(\theta)^{-1} \dot{F}(t; \theta)}.$$

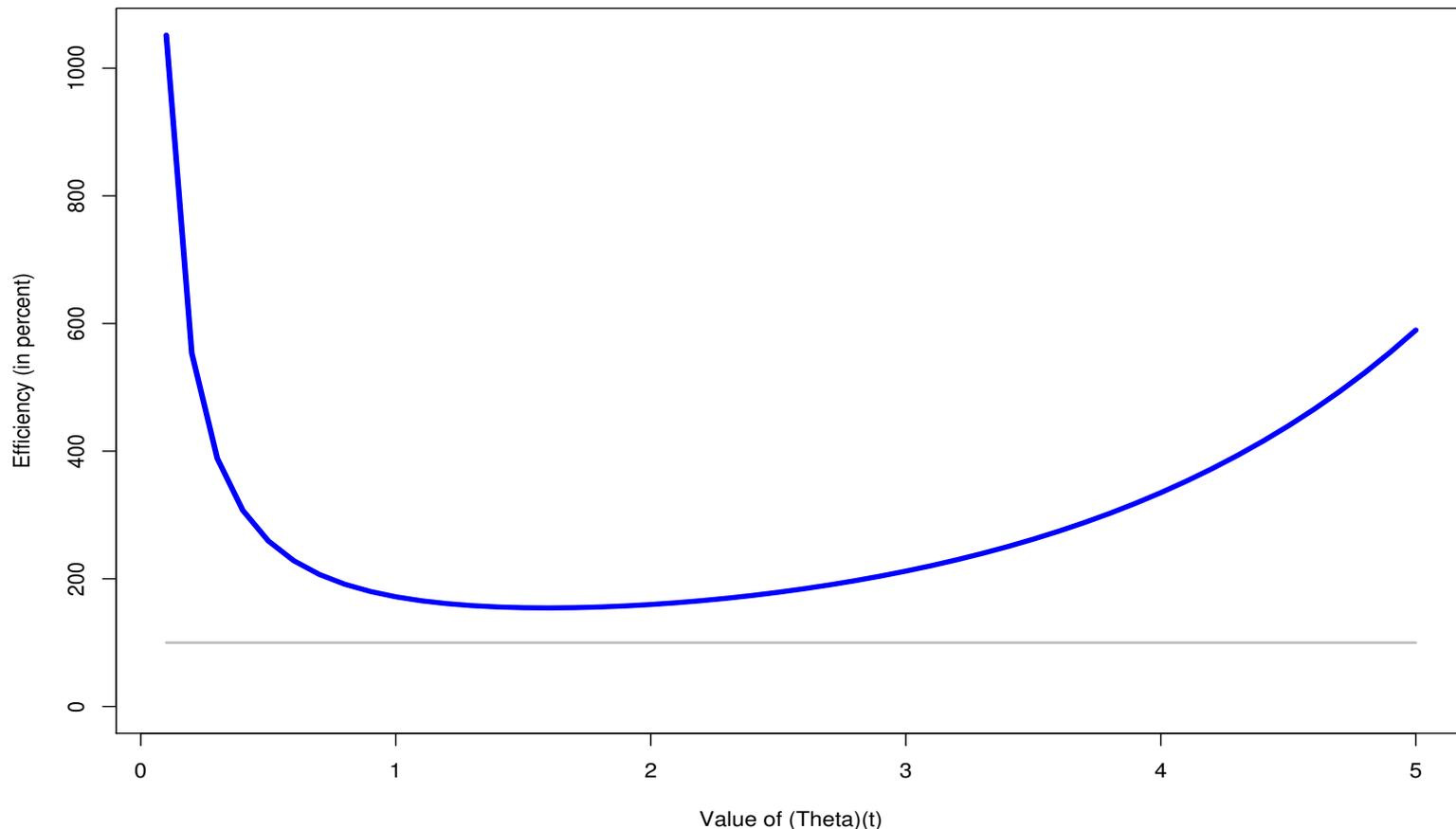
- When $\mathcal{F} = \{F(t; \theta) = 1 - \exp\{-\theta t\} : \theta > 0\}$, then

$$\text{Eff}(\hat{F}_{pa}(t) : \hat{F}_{np}(t)) = \frac{\exp\{\theta t\} - 1}{(\theta t)^2}$$

Efficiency: Parametric/Nonparametric

Asymptotic efficiency of parametric versus nonparametric estimators under a *correct* negative exponential family model.

Effi of Para Relative to NonPara in Expo Case



Whither Nonparametrics?

- Consider however the case where the negative exponential family is fitted, **but it is actually not the correct model**. Let us suppose that the gamma family of distributions is the *correct* model.
- *Under wrong model*, with $\bar{T} = \frac{1}{n} \sum_{i=1}^n T_i$ the sample mean, the parametric estimator of F is

$$\hat{F}_{pa}(t) = 1 - \exp\{-t/\bar{T}\}.$$

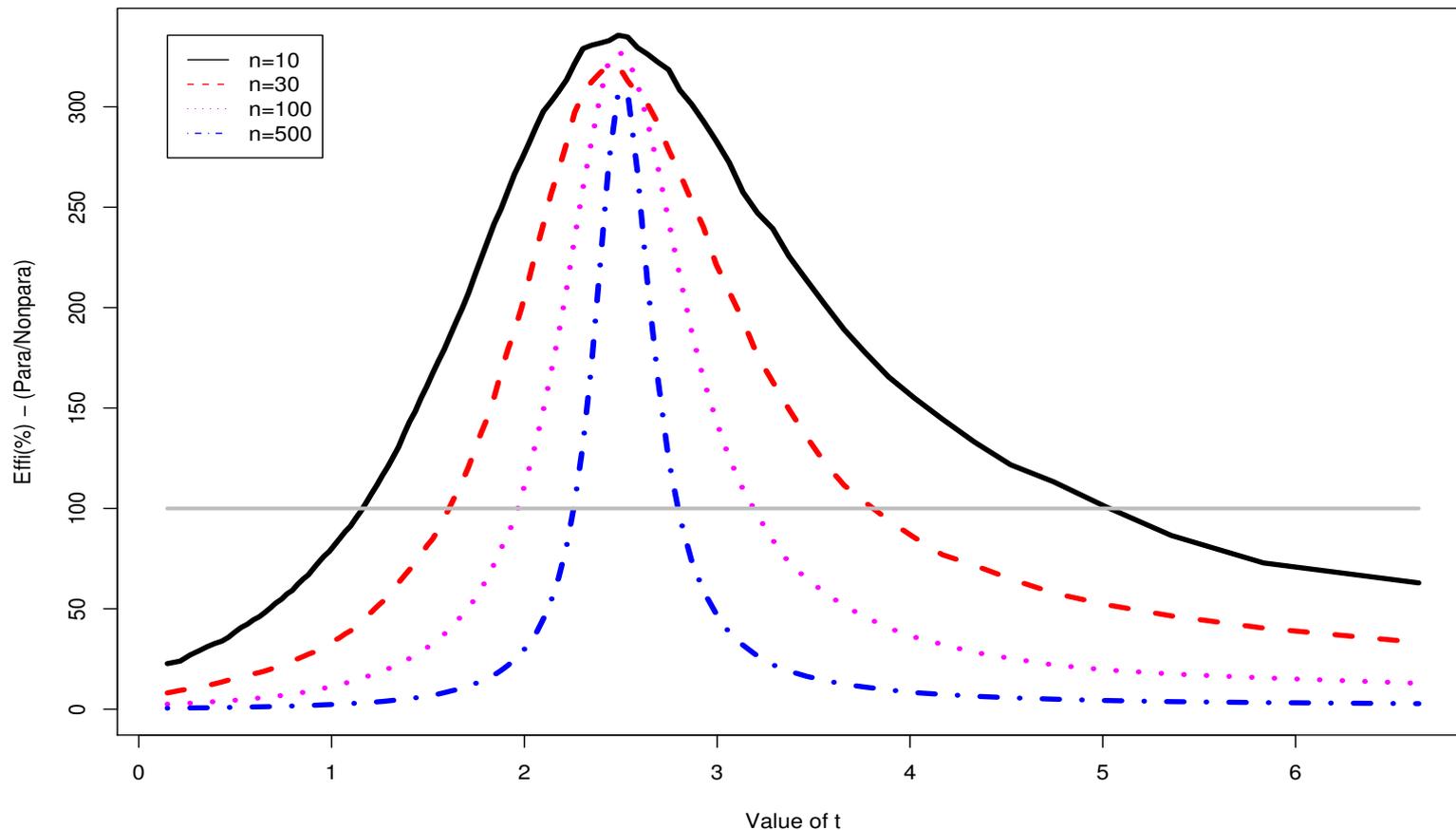
- Under gamma with shape α and scale θ , and since $\bar{T} \sim AN(\alpha/\theta, \alpha/(n\theta^2))$, by δ -method

$$\hat{F}_{pa}(t) \sim AN \left(1 - \exp \left\{ -\frac{\theta t}{\alpha} \right\}, \frac{1}{n} \frac{(\theta t)^2}{\alpha^3} \exp \left\{ -\frac{2(\theta t)}{\alpha} \right\} \right).$$

Efficiency: Under a Mis-specified Model

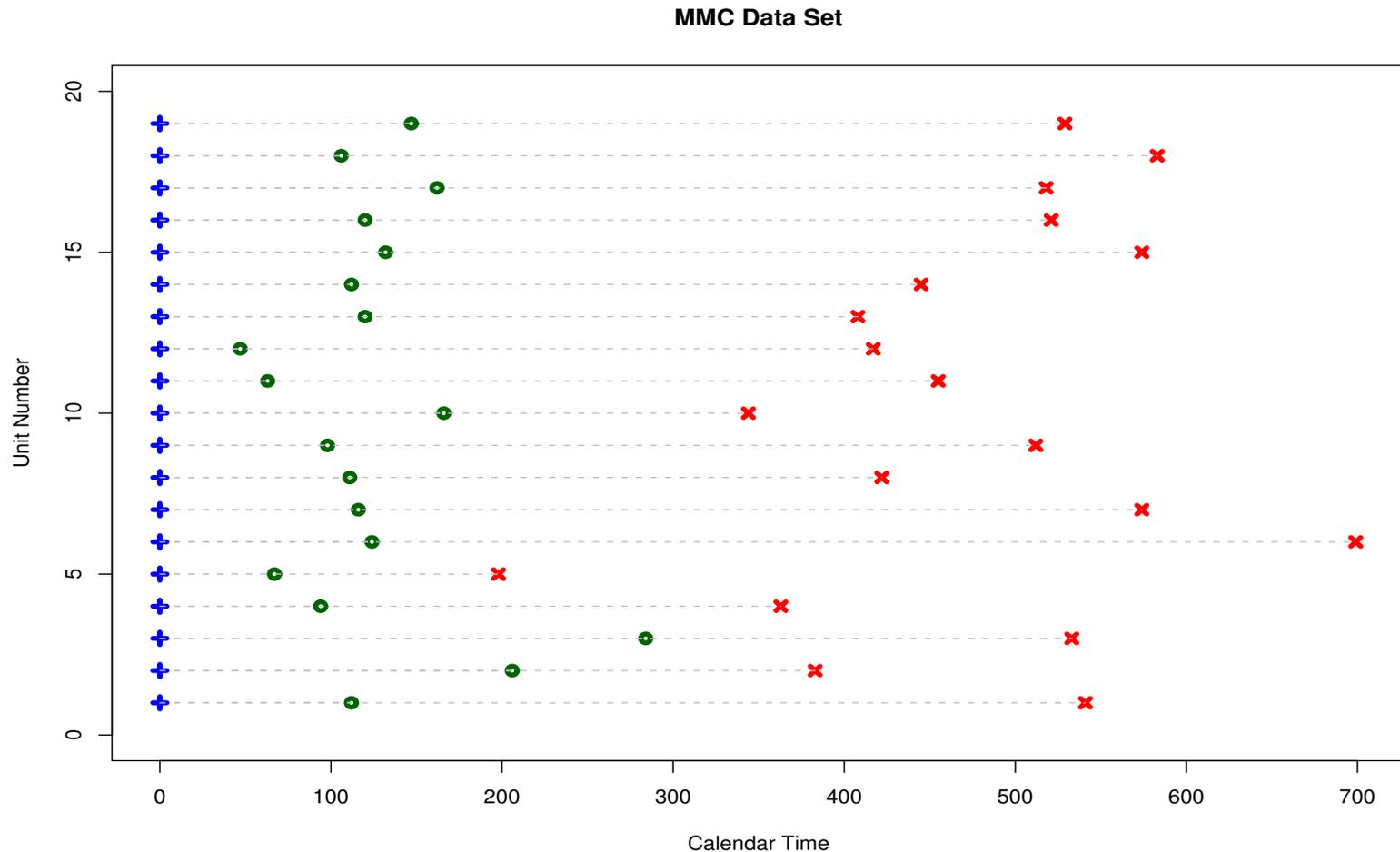
Simulated Effi: $\text{MSE}(\text{NonParametric})/\text{MSE}(\text{Parametric})$
under a *mis-specified* exponential family model.
True Family of Model: Gamma Family

Effi of Para vs NonPara under Misspecified Model



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

- For i th unit, a right-censoring variable C_i with C_1, C_2, \dots, C_n IID df G .
- Observables are $(Z_i, \delta_i), i = 1, 2, \dots, n$ with $Z_i = \min\{T_i, C_i\}$ and $\delta_i = I\{T_i \leq C_i\}$.
- **Problem:** For observed (Z_i, δ_i) s, to estimate df F or hazard function Λ of the T_i s.
- **Nonparametric Approaches:**
 - ‘Naive’ (product-limit)!
 - 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:

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

$$Y(t) = \sum_{i=1}^n I\{Z_i \geq t\}$$

- 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):

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

Some Properties of PLE

- Nonparametric MLE of F (Kaplan-Meier, '58).
- PLE is a step-function which jumps only at observed failure times.
- With censored data, unequal jumps.
- Efron ('67): Possesses self-consistency property.
- Biased for finite n .
- When no censoring and no tied values: $\Delta N(t_{(i)}) = 1$ and $Y(t_{(i)}) = n - i + 1$, so

$$\hat{S}(t_{(i)}) = \prod_{j=1}^i \left[1 - \frac{1}{n - j + 1} \right] = 1 - \frac{i}{n}.$$

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 $J(t) = I\{Y(w) > 0\}$ 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:

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

Likelihood Process: Hazard-Based

- J. Jacod's likelihood:

$$L_t(\Lambda(\cdot)) = \prod_{w \leq t} [Y(w)\Lambda(dw)]^{N(dw)} [1 - Y(w)\Lambda(dw)]^{1-N(dw)}.$$

- When $\Lambda(\cdot)$ is **continuous**,

$$L_t(\Lambda(\cdot)) = \left\{ \prod_{w \leq t} [Y(w)\Lambda(dw)]^{N(dw)} \right\} e^{-\int_0^t Y(w)\Lambda(dw)}.$$

- With $\mathcal{T}(t) = \int_0^t Y(w)dw = \text{TTOT}(t)$, for $\lambda(t) = \theta$,

$$L_t(\theta) = \theta^{N(t)} \exp\{-\theta\mathcal{T}(t)\}.$$

Asymptotic Properties

- Proofs uses martingale central limit theorem.
- 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

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

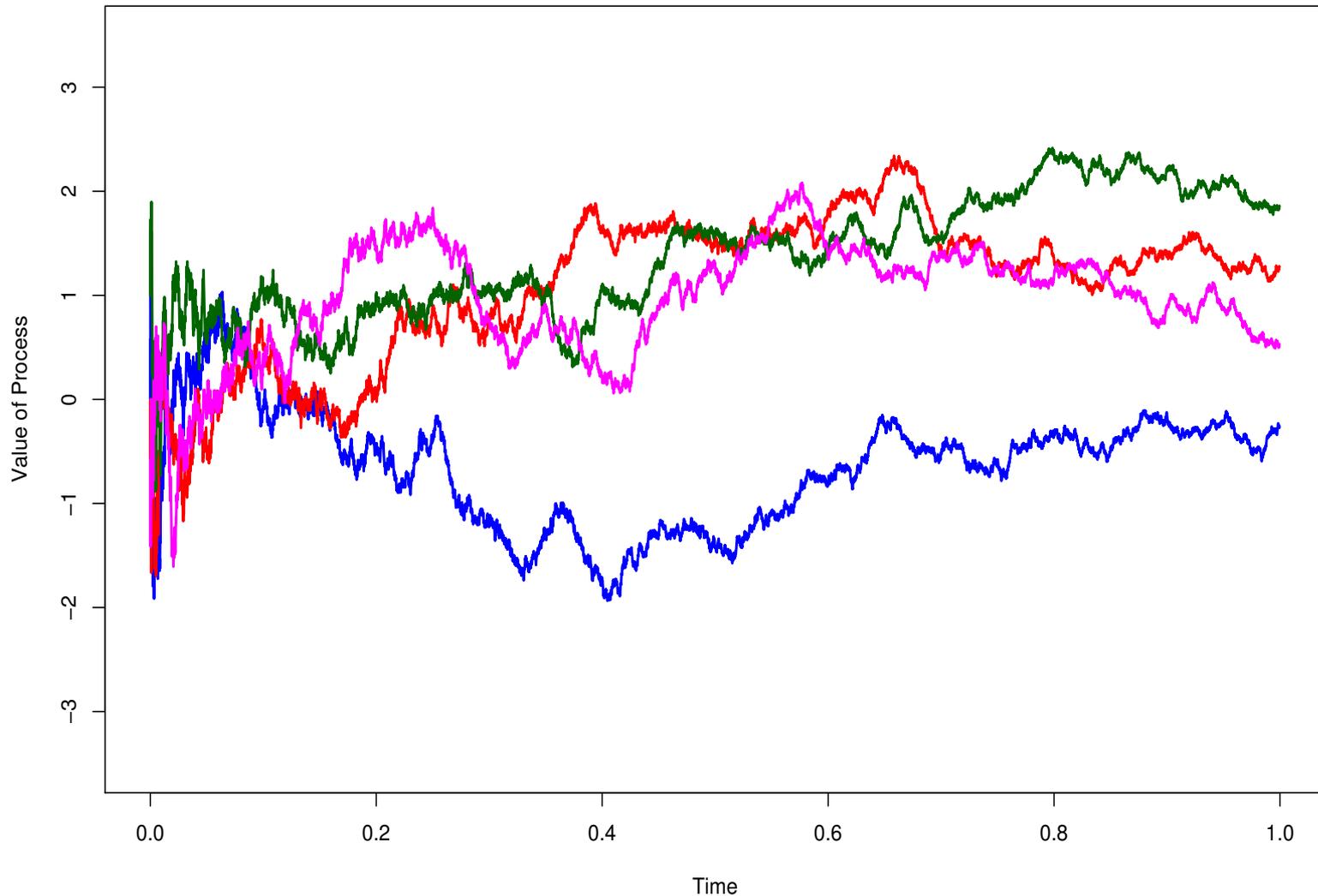
- PLE: $\sqrt{n}[\hat{F}(t) - F(t)] \Rightarrow Z_2(t) \stackrel{st}{=} S(t)Z_1(t)$ so

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

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

Gaussian Process: Sample Paths

Brownian Motion Paths



Regression Models

- In many situations we observe covariates: temperature, degree of usage, stress level, age, blood pressure, race, etc. How to account of them 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 \mathbf{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 :

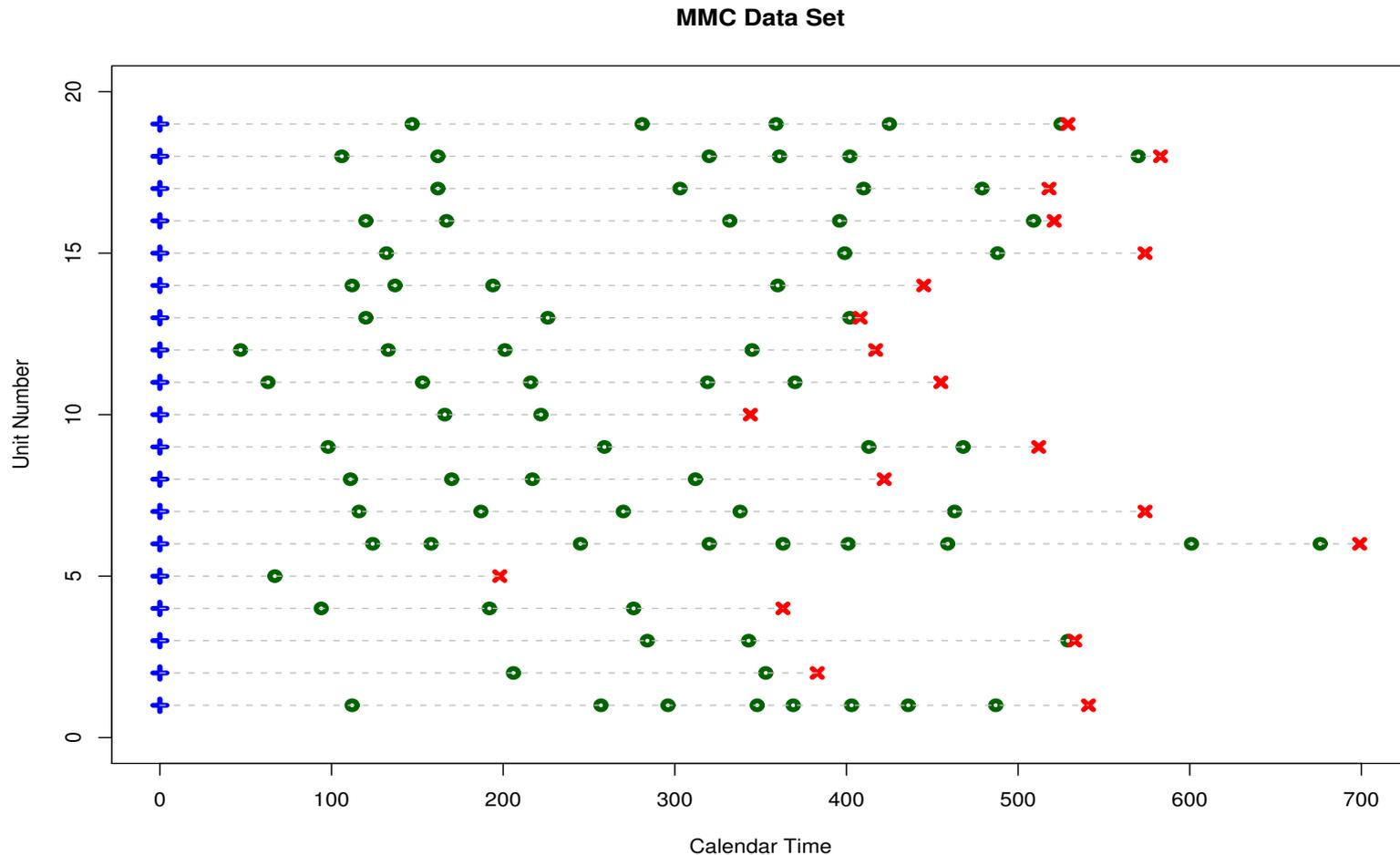
$$L_P(\beta) \equiv \prod_{i=1}^n \prod_{t < \infty} \left[\frac{\exp(\beta' \mathbf{x}_i)}{\sum_{j=1}^n Y_j(t) \exp(\beta' \mathbf{x}_j)} \right]^{\Delta N_i(t)}.$$

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

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

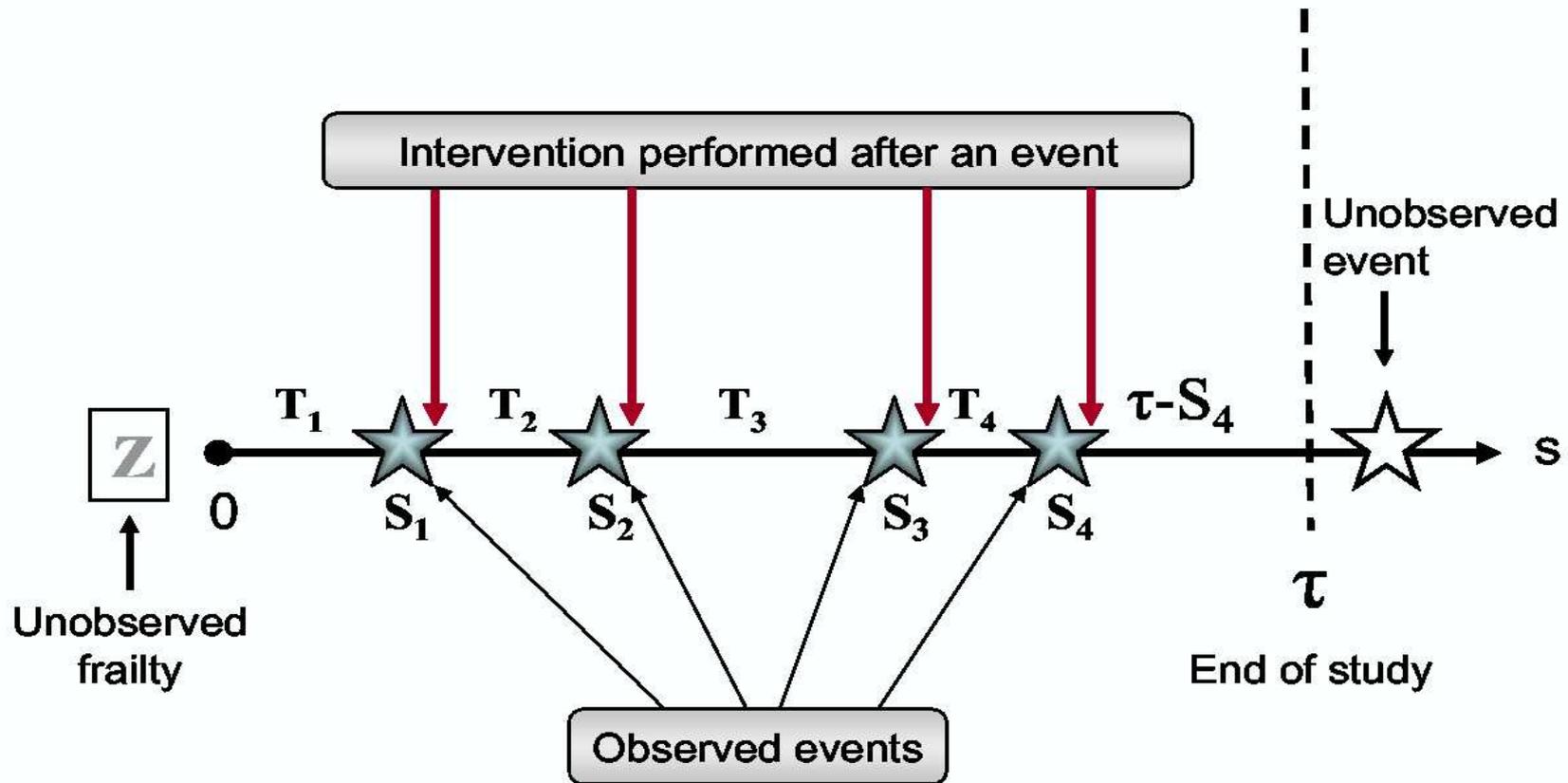
MMC Data: Recurrent Aspect

Aalen and Husebye ('91) Full Data



Problem: Estimate inter-event time distribution.

Representation: One Subject



Covariate vector: $\mathbf{X}(s) = (X_1(s), \dots, X_q(s))$

Observables: One Subject

- $\mathbf{X}(s)$ = covariate vector, possibly time-dependent
- T_1, T_2, T_3, \dots = inter-event or gap times
- S_1, S_2, S_3, \dots = calendar times of event occurrences
- τ = 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

Remark: A unique feature of recurrent event modeling is the **sum-quota constraint** that arises due to a fixed or random observation window. Failure to recognize this in the statistical analysis leads to erroneous conclusions.

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

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

Recurrent Event Models: IID Case

- Parametric Models:

- HPP: $T_{i1}, T_{i2}, T_{i3}, \dots$ IID EXP(λ).

- IID Renewal Model: $T_{i1}, T_{i2}, T_{i3}, \dots$ IID F where

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

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

- Non-Parametric Model: $T_{i1}, T_{i2}, T_{i3}, \dots$ 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 $(T_{i1}, T_{i2}, T_{i3}, \dots)$, given Z_i , are IID with survivor function

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

A General Class of Models

- Peña and Hollander (2004) model.

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

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

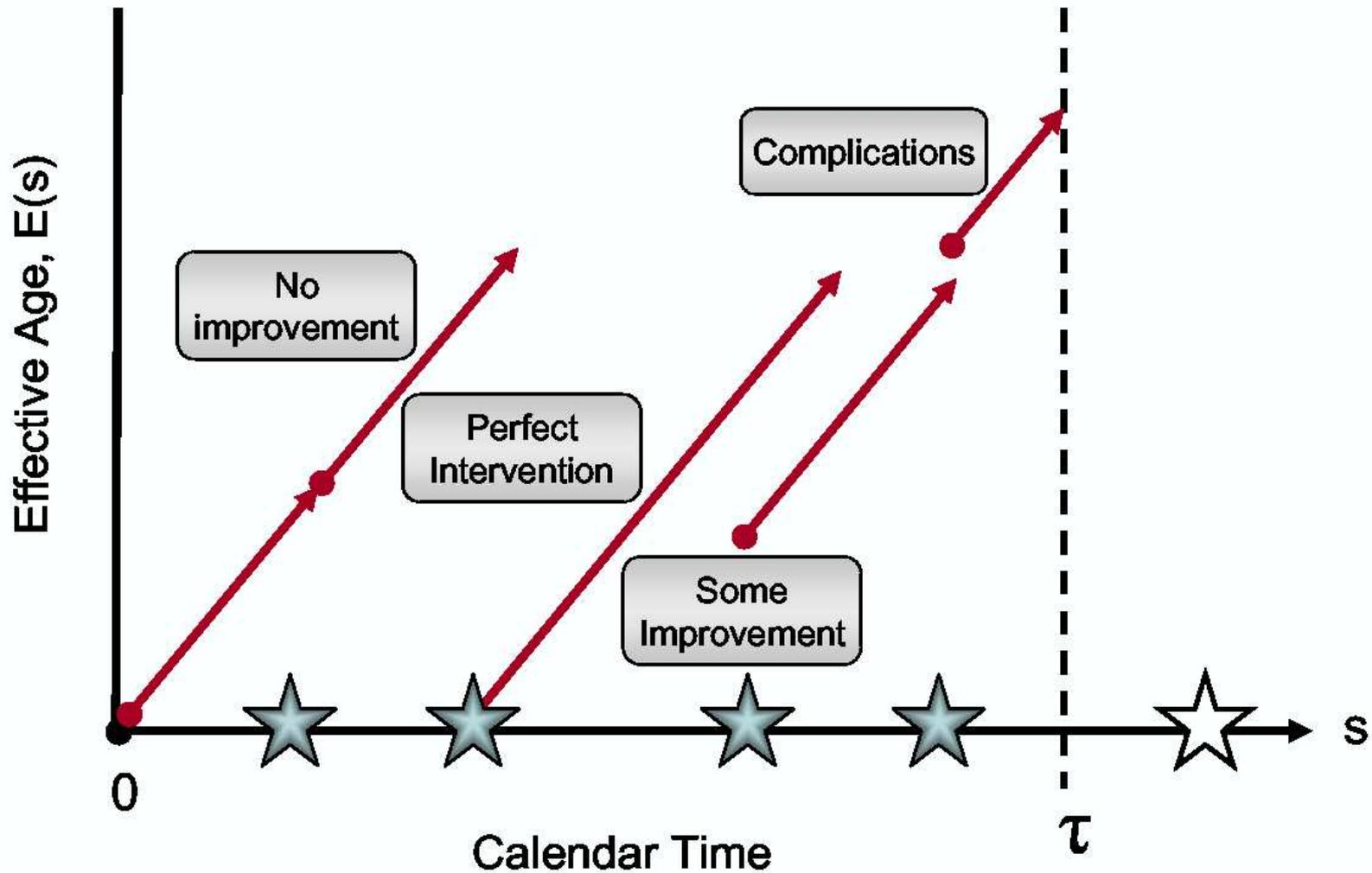
$$A^\dagger(s|Z) = \int_0^s Y^\dagger(w) \lambda(w|Z) dw$$

- Intensity:

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

- This class of models includes as special cases many models in reliability and survival analysis.

Effective Age Process



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

- Predictable, observable, nonnegative, dynamically specified, monotone, and differentiable on $[S_{k-1}, S_k)$, $\mathcal{E}(s)$ with $\mathcal{E}'(s) \geq 0$.
- **Perfect** Intervention: $\mathcal{E}(s) = s - S_{N^+(s-)}$.
- **Imperfect** Intervention: $\mathcal{E}(s) = s$.
- **Minimal** Intervention (Brown & Proschan, '83; Block, Borges & Savits, '85):

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

where, with I_1, I_2, \dots IID BER(p),

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

Semi-Parametric Estimation: No Frailty

Observed Data for n Subjects:

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

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

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

with the model for the ‘signal’ being

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

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

Processes and Notations

Calendar/Gap Time Processes:

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

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

Notational Reductions:

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

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

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:

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

- 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` (González, 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 estimator in PSH (2001) obtains.

Sum-Quota Effect: IID Renewal

- Generalized product-limit estimator \hat{F}_0 of common gap-time df F_0 presented in PSH (2001, JASA).

$$\sqrt{n}(\hat{F}_0(\cdot) - \bar{F}_0(\cdot)) \implies \mathbf{GP}(0, \sigma^2(\cdot))$$

$$\sigma^2(t) = \bar{F}_0(t)^2 \int_0^t \frac{d\Lambda_0(w)}{\bar{F}_0(w)\bar{G}(w-)[1 + \nu(w)]}$$

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

$$\rho^*(\cdot) = \sum_{j=1}^{\infty} F_0^{*j}(\cdot) = \text{renewal function}$$

Semi-Parametric Estimation: With Frailty

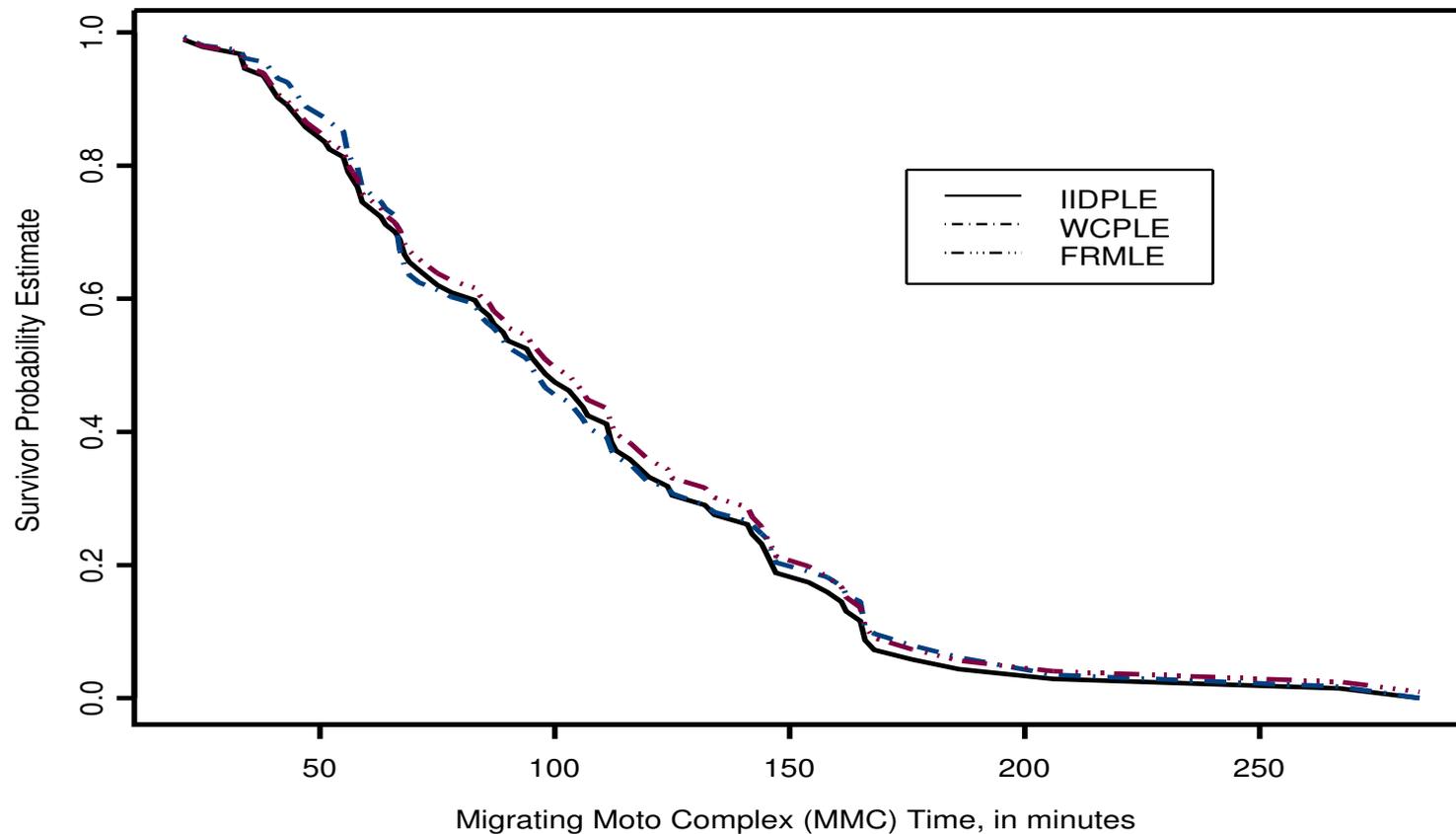
- Recall the intensity rate:

$$\lambda_i(s|Z_i, \mathbf{X}_i) = Z_i \lambda_0[\mathcal{E}_i(s)] \rho[N_i^\dagger(s-); \alpha] \psi(\beta^\top \mathbf{X}_i(s))$$

- Frailties Z_1, Z_2, \dots, Z_n are **unobserved** and assumed to be **IID Gamma**(ξ, ξ)
- Unknown parameters: $(\xi, \alpha, \beta, \lambda_0(\cdot))$
- Use of the **EM algorithm** (Dempster, et al; Nielsen, et al), with frailties as missing observations.
- Estimator of baseline hazard function under no-frailty model plays an important role.
- Details are in Peña, Slate & Gonzalez (To appear in *JSPI*, 2006).

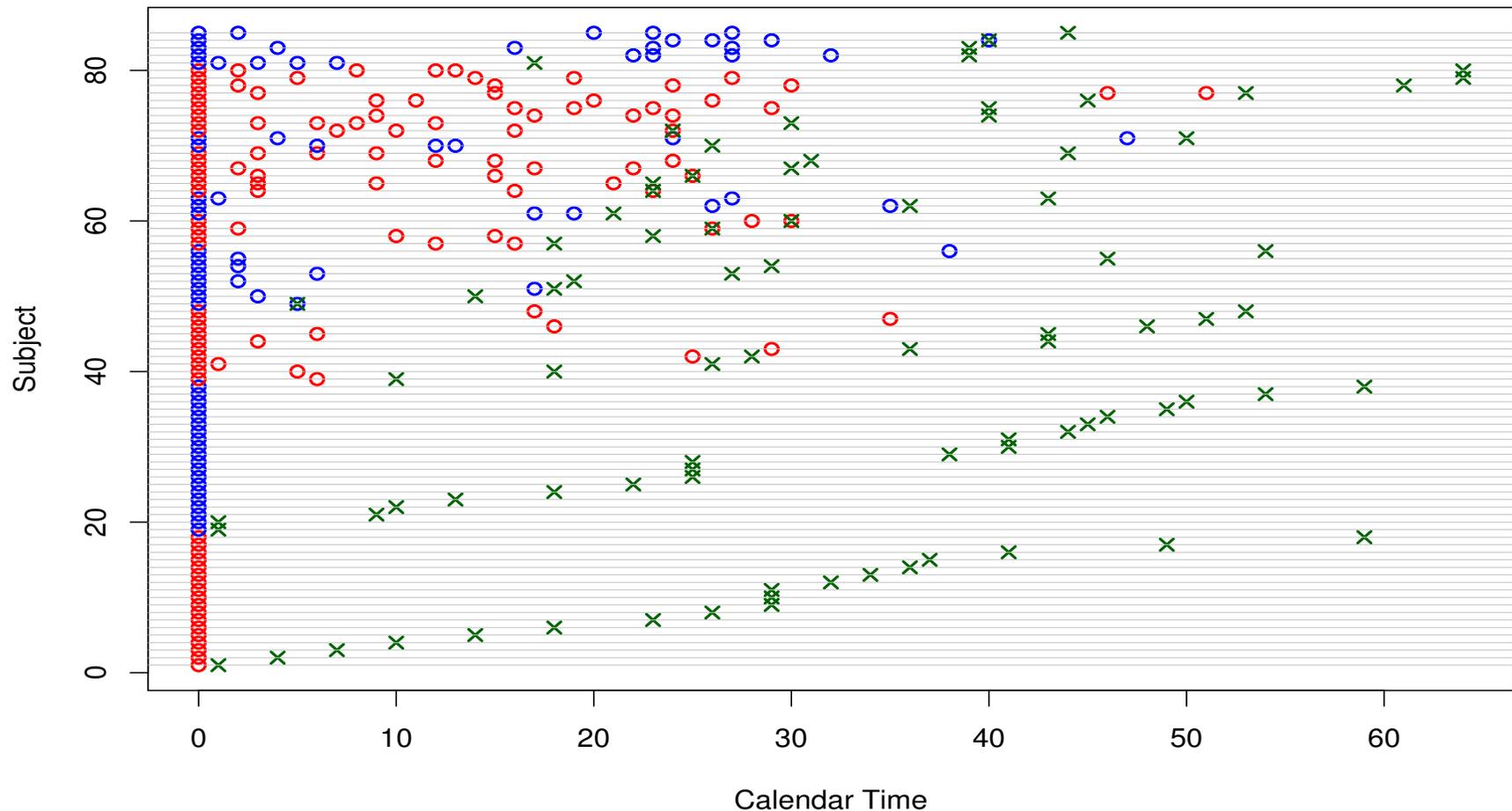
First Application: MMC Data Set

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 Marginal	PWP Cond*nal	General Model	
					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

^aEffective Age is backward recurrence time ($\mathcal{E}(s) = s - S_{N^+(s-)}$).

^bEffective Age is calendar time ($\mathcal{E}(s) = s$).

On Asymptotic Properties

- Asymptotics under the **no-frailty models**.
- **Difficulty:** $\lambda_0(\cdot)$ has $\mathcal{E}(s)$ as argument in the model; **whereas**, interest is usually on $\Lambda_0(t)$.
- **No** martingale structure in gap-time axis. MCLT not **directly** applicable.
- Under regularity conditions: **consistency** and **joint weak convergence** to Gaussian processes of standardized $(\hat{\alpha}, \hat{\beta})$ and $\hat{\Lambda}_0(s^*, \cdot)$.
- Results **extend** those in Andersen and Gill (AoS 82) regarding Cox PHM, though proofs different.
- Research on the asymptotics for the model **with frailty** *in progress*.

Asymptotics: Master Theorem

- $\{\mathbf{H}_i\}$ a sequence defined on $[0, s^*] \times [0, t^*]$.
- $M_i(s, t) = \int_0^s I\{\mathcal{E}_i(v) \leq t\} M_i^\dagger(dv)$.
- $Y_i(s, t)$ - generalized at-risk process.
- Under some regularity conditions, and if

$$\frac{1}{n} \sum_{i=1}^n \mathbf{H}_i^{\otimes 2}(s^*, \cdot) Y_i(s^*, \cdot) \xrightarrow{upr} \mathbf{v}(s^*, \cdot),$$

- then, with $\Sigma(s^*, t) = \int_0^t \mathbf{v}(s^*, w) \Lambda_0(dw)$,

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n \int_0^\cdot \mathbf{H}_i(s^*, w) M_i(s^*, dw) \implies \mathbf{GP}(0, \Sigma(s^*, \cdot)).$$

Relevant Empirical Measures

- Simplified model (one unit):

$$\Pr\{dN_i^\dagger(v) = 1 | \mathcal{F}_{s-}\} = Y_i^\dagger(v) \lambda_0[\mathcal{E}_i(v)] \Xi_i(v; \eta) dv.$$

- *Conditional PM* $Q(s^*, w; \eta)$ on $\{1, 2, \dots, N^\dagger(s-) + 1\}$:

$$Q(\{j\}; s^*, w; \eta) = \frac{\varphi_{j-1}(w; \eta) I\{\mathcal{E}(S_{j-1}) < w \leq \mathcal{E}(S_j)\}}{Y(s^*, w)}$$

with $S_{N^\dagger(s-)+1} = \min(s, \tau)$.

- *Conditional PM* $P(s^*, w; \eta)$ on $\{1, 2, \dots, n\}$:

$$P(\{i\}; s^*, w; \eta) = \frac{Y_i(s^*, w; \eta)}{nPY(s^*, w; \eta)}.$$

Empirical Means & Variances

$$\mathbb{P}f(\mathbf{D}) = \frac{1}{n} \sum_{i=1}^n f(\mathbf{D}_i)$$

$$\mathbb{E}_{Q(s^*, w; \eta)} g(J) = \sum_{j=1}^{N^\dagger(s^* -) + 1} g(j) Q(\{j\}; s^*, w; \eta)$$

$$\mathbb{V}_{Q(s^*, w; \eta)} g(J) = \mathbb{E}_{Q(s^*, w; \eta)} [g^2(J)] - (\mathbb{E}_{Q(s^*, w; \eta)} g(J))^2$$

$$\mathbb{E}_{P(s^*, w; \eta)} g(I) = \sum_{i=1}^n g(i) P(\{i\}; s^*, w; \eta)$$

$$\mathbb{V}_{P(s^*, w; \eta)} g(I) = \mathbb{E}_{Q(s^*, w; \eta)} [g^2(I)] - (\mathbb{E}_{Q(s^*, w; \eta)} g(I))^2$$

Relevant Limit Functions

- $s_0(s^*, w; \eta, \Lambda_0) = \text{plim } \mathbb{P}Y(s^*, w; \eta)$.
- Partial Likelihood Information Limit:

$$\mathcal{I}_p(s^*, t; \eta, \Lambda_0) = \text{plim}$$

$$\int_0^t \left\{ \left[\mathbb{E}_{P(s^*, w; \eta)} \mathbb{V}_{Q(s^*, w; \eta)} \left(\nabla_{\eta} \log \Xi_I(\mathcal{E}_{IJ-1}^{-1}(w); \eta) \right) + \mathbb{V}_{P(s^*, w; \eta)} \mathbb{E}_{Q(s^*, w; \eta)} \left(\nabla_{\eta} \log \Xi_I(\mathcal{E}_{IJ-1}^{-1}(w); \eta) \right) \right] \right\} \times s_0(s^*, w; \eta, \Lambda_0) \Lambda_0(dw).$$

- With $\mathbf{e}(s^*, w; \eta, \Lambda_0) = \text{plim } \frac{\mathbb{P}\nabla_{\eta} Y(s^*, w; \eta)}{\mathbb{P}Y(s^*, w; \eta)}$, let

$$A(s^*, t; \eta, \Lambda_0) = \int_0^t \mathbf{e}(s^*, w; \eta, \Lambda_0) \Lambda_0(dw).$$

Weak Convergence Results

As $n \rightarrow \infty$ and under certain regularity conditions:

$$\sqrt{n}(\hat{\eta}(s^*, t^*) - \eta) \Rightarrow N(0, \mathcal{I}_p(s^*, t^*; \eta, \Lambda_0)^{-1})$$

$$\sqrt{n}(\hat{\Lambda}_0(s^*, \cdot) - \Lambda_0(\cdot)) \Rightarrow GP(0, \Gamma(s^*, \cdot; \eta, \Lambda_0))$$

where the limiting variance function is given by

$$\begin{aligned} \Gamma(s^*, t; \eta, \Lambda_0) &= \int_0^t \frac{\Lambda_0(dw)}{s_0(s^*, w; \eta)} \\ &+ A(s^*, t; \eta, \Lambda_0) \mathcal{I}_p(s^*, t^*; \eta, \Lambda_0)^{-1} A(s^*, t; \eta, \Lambda_0)^\dagger. \end{aligned}$$

Concluding Remarks

- Many aspects of the general dynamic recurrent event model still under investigation.
- Asymptotics for the model with frailty.
- Testing hypothesis procedures.
- Goodness-of-fit and residual analysis.
- Its practical relevance still needs exploring, e.g., could the effective age process be determined appropriately in practice.
- Comparisons with marginal-based models (PWP, WLW).
- *Dynamic recurrent event modeling* remains a challenge and is a fertile area for research.