Analysis of Failure-Time Data

Edsel A. Peña

Department of Statistics

University of South Carolina, Columbia

Research support from NIH and NSF

Work joint with Prof. Elizabeth Slate and Juan Gonzalez

pena@stat.sc.edu

Goals and Outline of Talk

- Review of estimation methods for complete and censored single-event data.
- Discuss aspects of recurrent event modelling.
- Provide a general model for recurrent events.
- Present some properties of the estimation methods for general model for recurrent events.
- Provide an application to a biomedical data.
- Some concluding remarks.

Failure Times

- A positive-valued random variable, denoted by T, which denotes the time to the occurrence of an 'event' is called a failure-time.
- Events of interest might be, for example:
  - failure of a machine or computer software
  - death; relapse; onset of cancer
  - divorce
  - terrorist attack
  - sale of a house
- Relevant in a variety of settings; biomedical, engineering, economics, sociology, etc.

Distributions and Hazards

- Assume: T is a continuous failure-time.
- Distribution Function: $F(t) = \mathbf{P}\{T \leq t\}$
- Survivor Function: $\bar{F}(t) = 1 - F(t) = \mathbf{P}\{T > t\}$
- (Cumulative) Hazard Function: $\Lambda(t) = -\log \bar{F}(t)$
- Density Function: $f(t) = dF(t)/dt$

$$f(t)dt \approx \mathbf{P}\{t \leq T < t + dt\}$$

- Hazard Rate: $\lambda(t) = d\Lambda(t)/dt = f(t)/\bar{F}(t)$

$$\lambda(t)dt \approx \mathbf{P}\{t \leq T < t + dt | T \geq t\}$$

Relationships

$$\bar{F}(t) = \exp\{-\Lambda(t)\}$$

$$f(t) = \lambda(t)\exp\{-\Lambda(t)\}$$

A general representation is through product-integration:

$$\begin{aligned}\bar{F}(t) &= \lim_{\max|t_i - t_{i-1}|\to 0} \prod_{i=1}^{M} [1 - \{\Lambda(t_i) - \Lambda(t_{i-1})\}] \\ &\equiv \prod_{s=0}^{t} [1 - \Lambda(ds)]\end{aligned}$$

Single-Event Complete Data

- $T_1, T_2, \ldots, T_n$ IID $\bar{F}(t)$, where $\bar{F}(\cdot)$ is unknown.
- A nonparametric estimator of $\bar{F}(\cdot)$ is the empirical survivor function

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

  where $I(A) = 1$ iff A occurs, 0 otherwise.

- $E\{\hat{\bar{F}}(t)\} = \bar{F}(t)$, ie, it is unbiased.
- $Var\{\hat{\bar{F}}(t)\} = F(t)\bar{F}(t)/n$ since $n\hat{\bar{F}}(t) \sim BIN(n, \bar{F}(t))$

Asymptotic (as $n \longrightarrow \infty$) Properties of EDF

- Consistency property (Glivenko-Cantelli):

$$\sup_{s \geq 0} |\hat{\bar{F}}(t) - \bar{F}(t)| \xrightarrow{\text{pr}} 0$$

- Weak convergence property (Doob):

$$\sqrt{n}[\hat{\bar{F}}(\cdot) - \bar{F}(\cdot)] \Rightarrow W(\cdot) \quad \text{on} \quad \mathcal{D}[0, \infty)$$

where $W(\cdot)$ is a Gaussian process with mean function zero and variance function

$$Var\{W(t)\} = \bar{F}(t)F(t)$$

Single-Event Censored Data

- $T_1, T_2, \ldots, T_n$ IID from survivor function $\bar{F}(t)$ and hazard $\Lambda(t)$. T_is are failure times, and $\bar{F}$ and functionals of $\bar{F}$ such as the mean or median are the quantities of interest.
- $C_1, C_2, \ldots, C_n$ IID from a survivor function $\bar{G}$. C_is are the right-censoring variables.
- T_is are not completely observed, but what are observed are the censored data:

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

where $Z_i = \min(T_i, C_i)$ and $\delta_i = I\{T_i \leq C_i\}$.

Product-Limit Estimator

- Problem: Given the data $(Z_i, \delta_i), i = 1, 2, \ldots, n$, how to estimate $\bar{F}(\cdot)$ nonparametrically?
- Kaplan and Meier (1958) proposed and examined the product-limit estimator (PLE) of $\bar{F}$.
- The PLE is given by:

$$\hat{\bar{F}}(t) = \prod_{s=0}^{t} \left[1 - \frac{N(\Delta s)}{Y(s)} \right]$$

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

Properties of PLE (KM; Efron; BC; Gill)

- Asymptotic properties of the PLE are:

$$\sup_{t \geq 0} |\hat{\bar{F}}(t) - \bar{F}(t)| \xrightarrow{\text{pr}} 0$$

$$\sqrt{n}[\hat{\bar{F}}(\cdot) - \bar{F}(\cdot)] \Rightarrow W(\cdot) \quad \text{on} \quad \mathcal{D}[0, \tau]$$

$\{W(t) : 0 \leq t \leq \tau\}$ is a zero-mean Gaussian process with variance function

$$Var\{W(t)\} = \bar{F}^2(t) d(t)$$

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

Censored Data with Covariates

- In many practical situations, especially in biomedical settings, aside from the failure times, there usually is a set of covariates that affects the occurrence of the event of interest.
- It is imperative that these covariates be taken into account in the modelling and statistical inference. For instance, one of the covariates could be a treatment indicator, and it is desired to compare the treatment group with a control group.
- Observed censored data in this situation are of form:

$$(Z_1, \delta_1, \mathbf{X}_1), (Z_2, \delta_2, \mathbf{X}_2), \ldots, (Z_n, \delta_n, \mathbf{X}_n)$$

Cox (1972) PH Model

- D. Cox proposed a hazard-based model which incorporates covariates. For a unit with covariate vector $\mathbf{x}$, the (conditional) hazard-rate of failure is

$$\lambda(t|\mathbf{x}) = \lambda_0(t)\exp\{\mathbf{x}\beta\}$$

- $\lambda_0(\cdot) =$ an unknown baseline hazard rate.
- $\beta =$ is a regression parameter vector.
- $\exp(\beta_j)$ has the interpretation of being the change in hazard rate if the jth component of the covariate vector changes by one unit, others remaining the same.
- Goal: To estimate $\bar{F}_0(\cdot)$ and β based on $(Z_i, \delta_i, \mathbf{X}_i)$s.

Inference for Cox PH Model

- For estimating β, Cox introduced the partial likelihood (which is also a profile likelihood) given by:

$$L_P(\beta) = \prod_{i=1}^{n} \left[\frac{\exp(\mathbf{x}_i \beta)}{\sum_{\{j:\ t_j \geq t_i\}} \exp(\mathbf{x}_j \beta)} \right]^{\delta_i}$$

- Note that this partial likelihood is free of the unknown baseline hazard rate $\lambda_0(\cdot)$.
- $\mathbf{b}$, which maximizes this function, is called the partial likelihood MLE of β. Requires iterative methods to obtain $\mathbf{b}$.

Estimator of Λ_0 and $\bar{F}_0$

- Having obtained an estimator of β, the baseline hazard function $\Lambda_0(\cdot)$ is estimated by:

$$\hat{\Lambda}_0(t) = \int_0^t \frac{dN(s)}{\sum_{j=1}^n Y_j(s) \exp\{\mathbf{x}_j \mathbf{b}\}}$$

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

- Aalen-Breslow Estimator of $\bar{F}_0(\cdot)$:

$$\hat{\bar{F}}_0(t) = \prod_{s=0}^t \left[1 - d\hat{\Lambda}_0(t)\right]$$

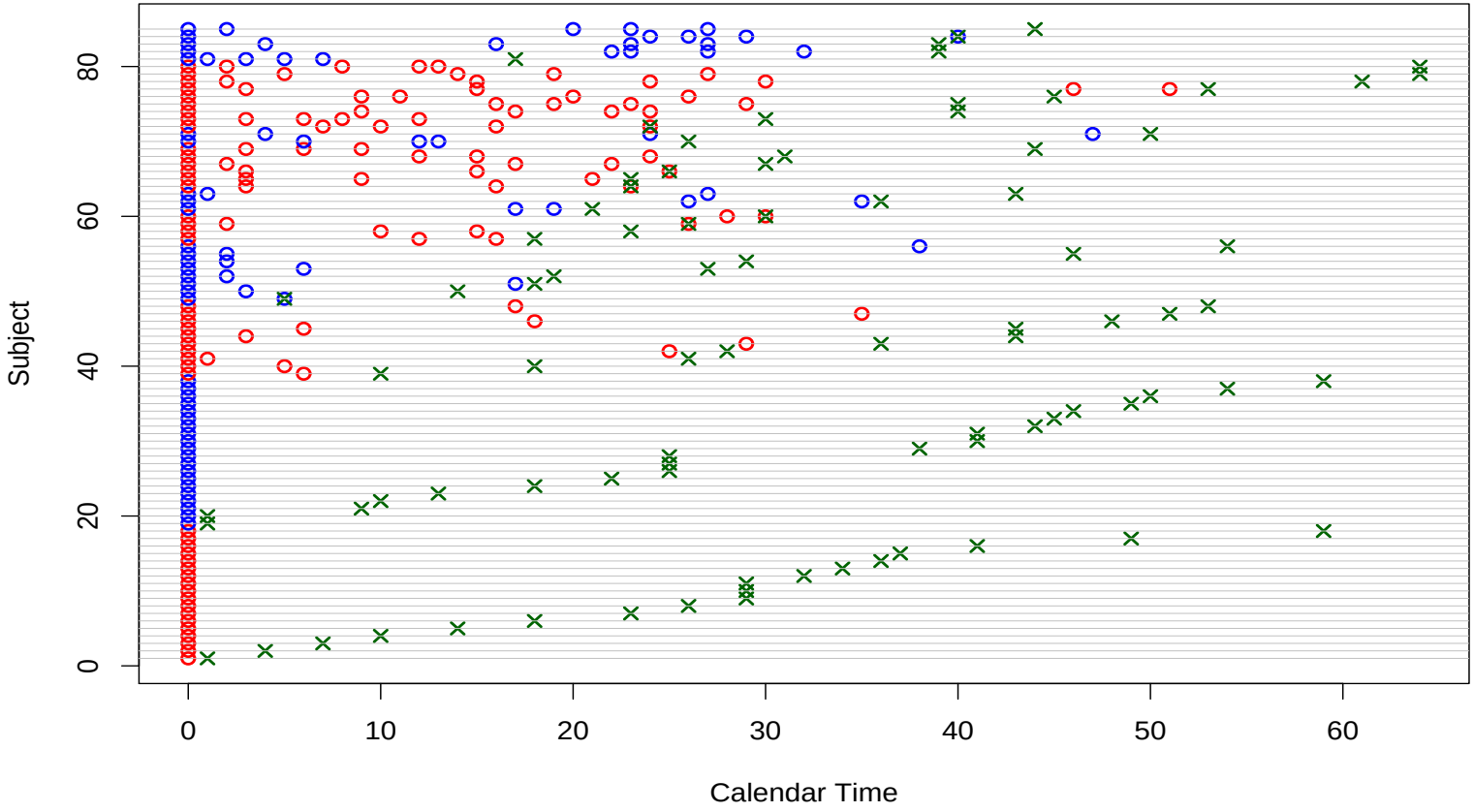
Properties of Estimators

- Andersen and Gill (1982), using counting process and martingale theory, obtained rigorously the properties of the estimators of β, Λ_0, and $\bar{F}_0$.
- In particular, they showed that under regularity conditions,$\mathbf{b}$ is asymptotically normal with mean β and some covariance matrix $\mathbf{\Sigma}$.
- This result is used to develop testing and confidence interval procedures for β, such as for example, in comparing control versus treatment groups.
- The estimator $\hat{\bar{F}}_0$ is also asymptotically Gaussian with mean $\bar{F}_0$, though for finite sample size, it is biased, but with the bias decreasing to zero at a geometric rate.

Recurrent Phenomena

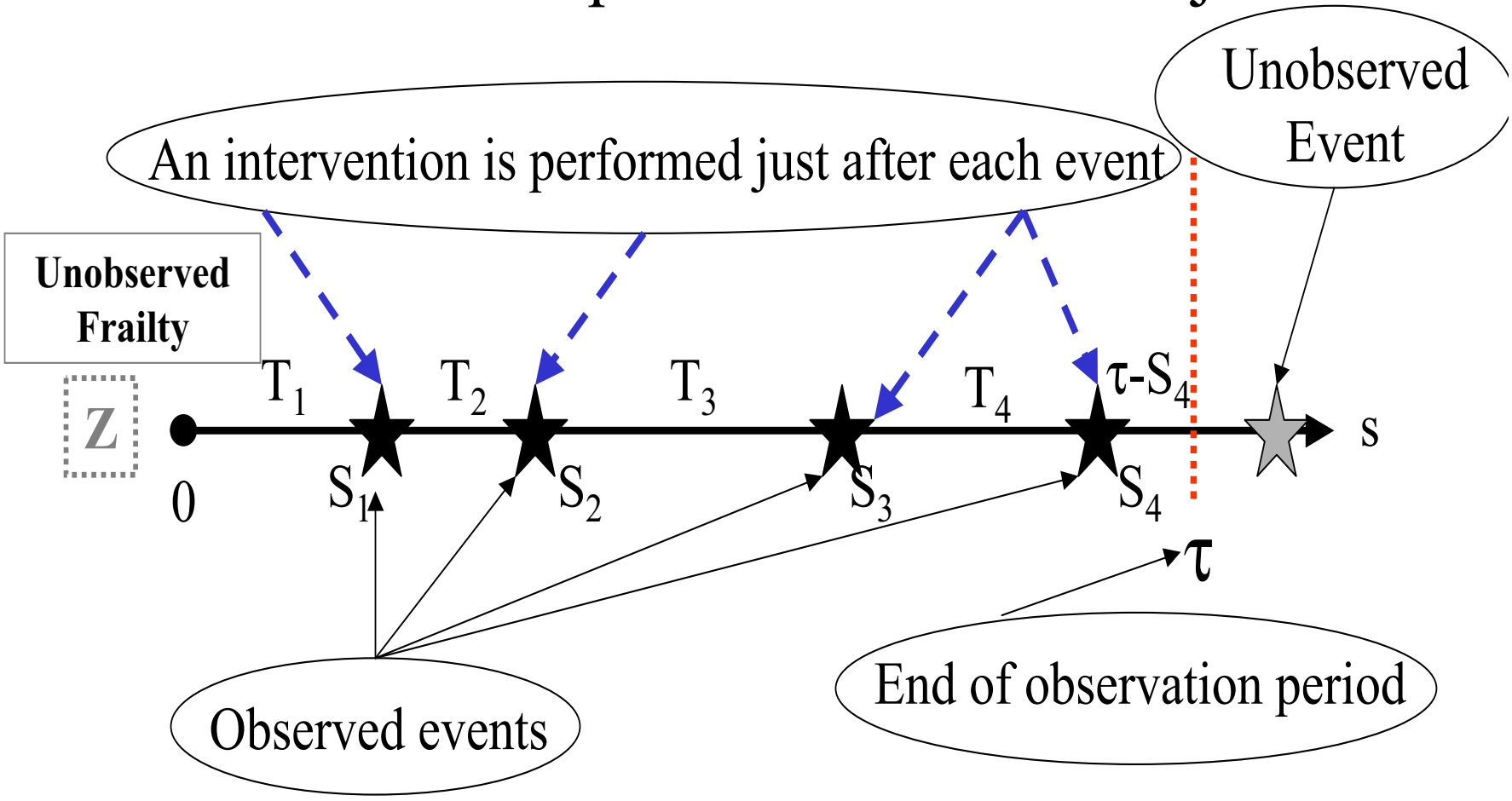
- In Health and Biomedical Settings
  - hospitalization due to a chronic disease
  - drug/alcohol abuse
  - occurrence of migraine headaches
  - onset of depression
  - episodes of epileptic seizures
  - asthma attacks
- In Engineering/Reliability and Other Settings
  - Software crashes and medical equipment failures

Bladder Cancer Data (WLW, ’89)



Questions: Difference in recurrence rates for placebo and thiotepa groups? Heterogeneity? Impact of more events?

A Pictorial Representation: One Subject



An observable covariate vector: $\mathbf{X}(s) = (X_1(s), X_2(s), \ldots, X_q(s))^t$

Random Entities: 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
- Accrued History: $\mathbf{F}^\dagger = \{\mathcal{F}_s^\dagger : s \geq 0\}$
- $Z =$ unobserved frailty variable
- $N^\dagger(s) =$ number of events in $[0, s]$
- $Y^\dagger(s) =$ at-risk indicator at time s

On Recurrent Event Modelling

- Intervention effects after each event occurrence.
- Effects of accumulating event occurrences. Could be weakening or strengthening effect.
- Effects of covariates.
- Associations of event occurrences per subject.
- Random observation monitoring period.
- Number of events observed informative about stochastic mechanism generating events.
- Informative right-censoring mechanism arising because of the sum-quota accrual scheme.

General Class of Models

- Peña and Hollander proposed a general class of models.
- $\{A^{\dagger}(s|Z) : s \geq 0\}$ is a predictable, nondecreasing process such that given Z and accrued information:

$$\{M^{\dagger}(s|Z) = N^{\dagger}(s) - A^{\dagger}(s|Z) : \ s \geq 0\}$$

is a zero-mean martingale (a fair game process). Assume multiplicative form:

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

Intensity Process

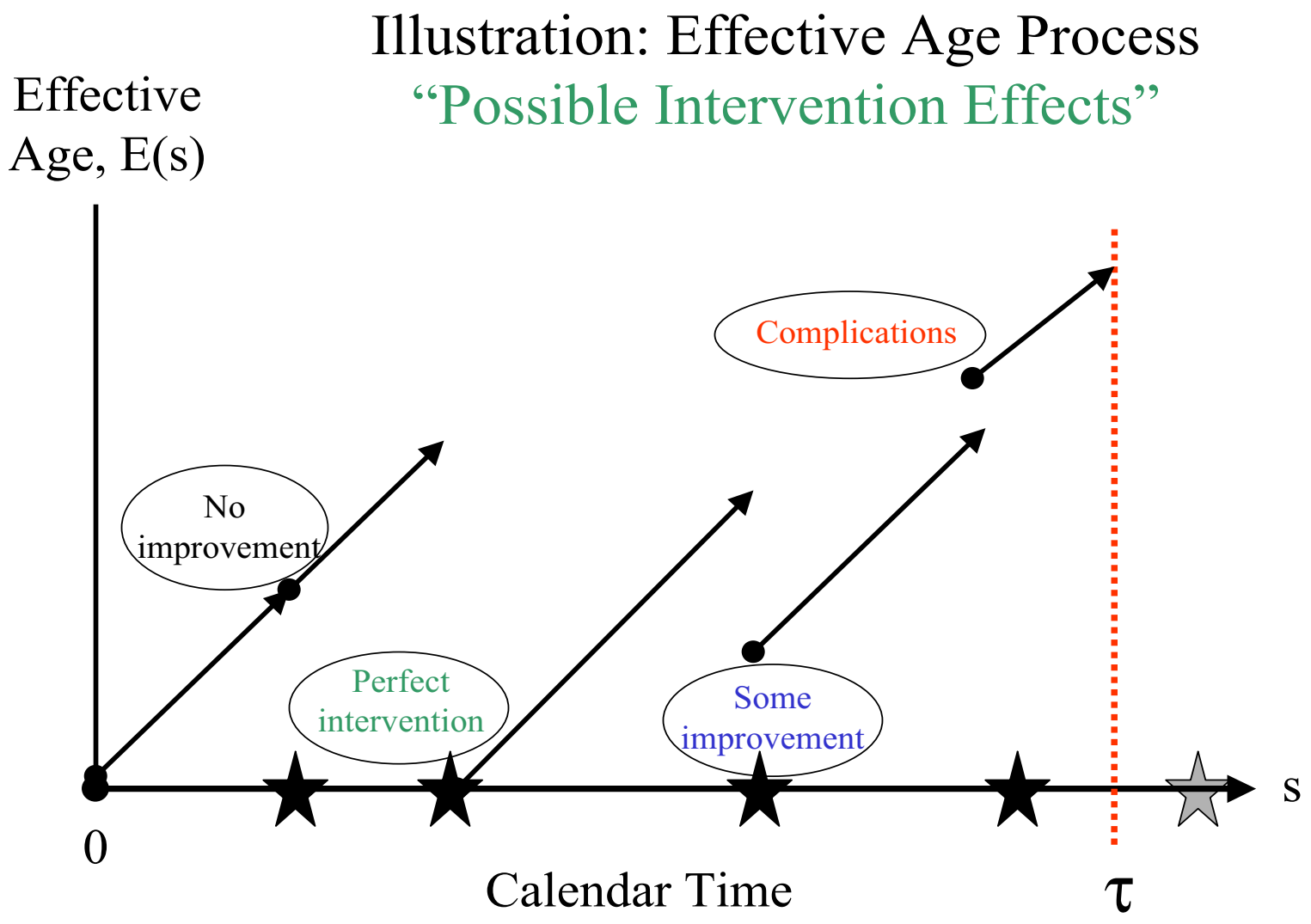
- Specify, possibly dynamically, a predictable, observable process $\{\mathcal{E}(s):\ 0 \le s \le \tau\}$ called the *effective age process*, satisfying
  - $\mathcal{E}(0) = e_0 \ge 0$;
  - $\mathcal{E}(s) \ge 0$ for every s;
  - On $[S_{k-1}, S_k)$, $\mathcal{E}(s)$ is monotone and differentiable with $\mathcal{E}'(s) \ge 0$.
- Specification:

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

Model Components

- $\lambda_0(\cdot) =$ unknown baseline hazard rate function.
- $\mathcal{E}(s) =$ effective age at calendar time s. **Rationale:** intervention changes effective age acting on baseline hazard.
- $\rho(\cdot;\alpha) =$ a positive function on $\mathcal{Z}_+$; known form; $\rho(0;\alpha) = 1$; unknown α. Encodes effect of accumulating events.
- $\psi(\cdot) =$ positive link function containing the effect of subject covariates. β is unknown.
- $Z =$ unobservable frailty variable. Induces associations among subject's inter-event times.

Effective Age Process



Flexibility

- IID renewal model: $\mathcal{E}(s) = s - S_{N^{\dagger}(s-)}$ (backward recurrence time), $\rho(k) = 1$, $\psi(x) = 1$.
- IID renewal model with frailties: same as above except for an unobserved frailty per subject/unit.
- Models dealt with in Peña, Strawderman and Hollander (*JASA*, 2001) and in Wang and Chang (*JASA*, 1999).
- Extended Cox (1972) PH model; Prentice, Williams and Peterson (1981) model; Lawless (1987):

$$\mathcal{E}(s) = s, \rho(k) = 1, \psi(x) = \exp(x)$$

Generality

- Gail, Santner and Brown (1980) carcinogenesis model and the Jelinski and Moranda (1972) software reliability model.

$$\rho(k;\alpha) = \max(0, \alpha - k + 1)$$

- Includes the Dorado, Hollander and Sethuraman (1997) general repair model.
- Also, reliability models of Kijima (1989); Baxter, Kijima and Tortorella (1996); Stadje and Zuckerman (1991); and Last and Szekli (1998).

‘Minimal Repair’ Models

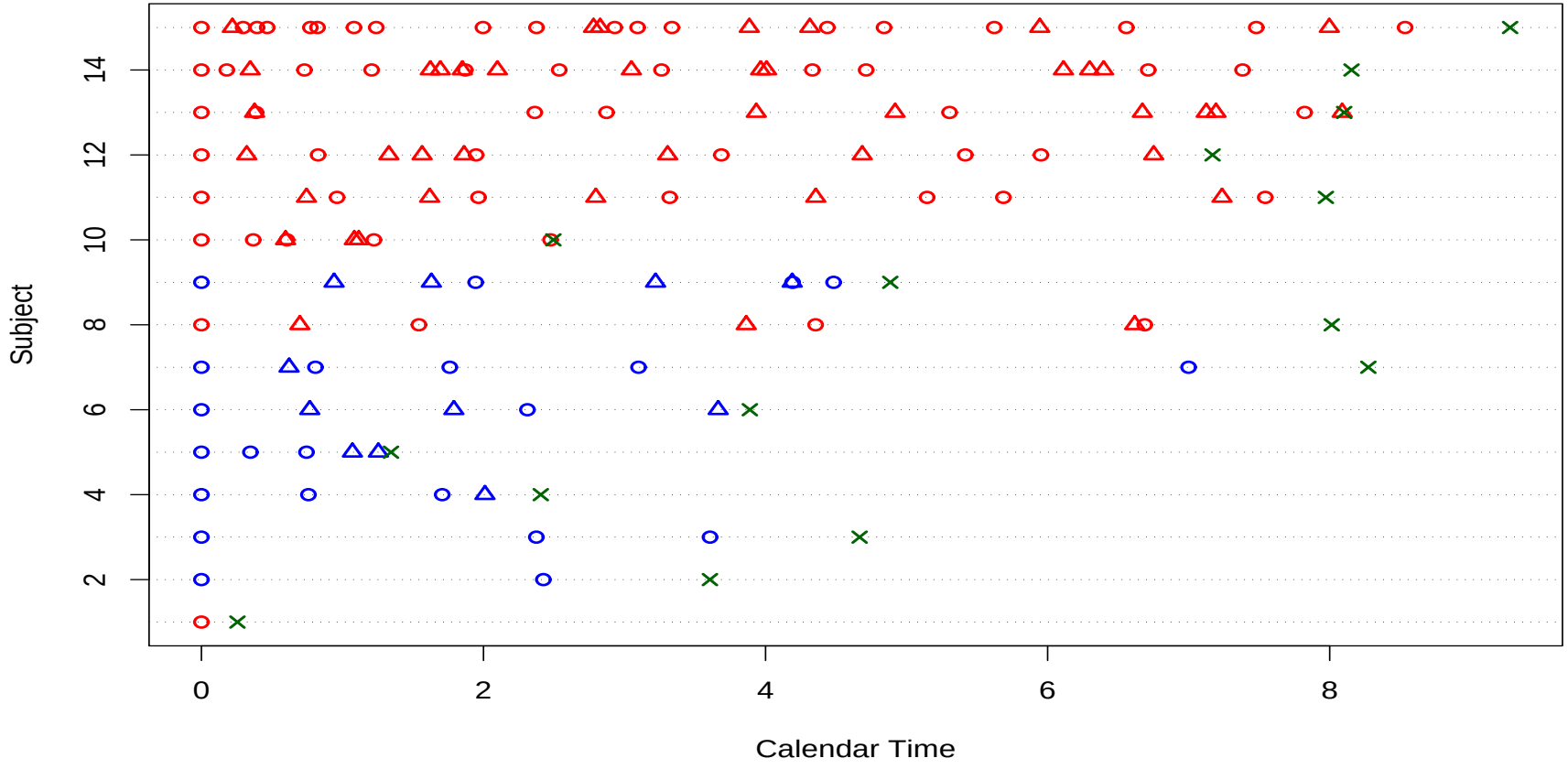
- (Generalized) Brown and Proschan (1983) minimal repair model: Let $I_1, I_2, \ldots$ IID Ber(p), p be the ‘perfect repair’ probability.
  - $\Gamma_k = \min\{j > \Gamma_{k-1} : \ I_j = 1\}$: index of kth perfect repair
  - $\eta(s) = \sum_{i=1}^{N^\dagger(s)} I_i$: # of perfect repairs till s
  - $\mathcal{E}(s) = s - S_{\Gamma_{\eta(s-)}}$: length since last perfect repair
- (Generalized) Block, Borges and Savits (1985): Perfect repair probability depends on s, so $p(s)$.

Estimation of Parameters

- Developed procedures for estimating the parameters of this general model, both with and without the frailty components.
- Estimator of baseline survivor function is also of product-limit type.
- Extended the idea of a partial likelihood.
- Expectation-Minimization (EM) algorithm utilized in estimating parameters for the model with frailties.
- Estimation procedure coded in R and Fortran.
- Lots of notation, so focus here instead on some empirical properties.

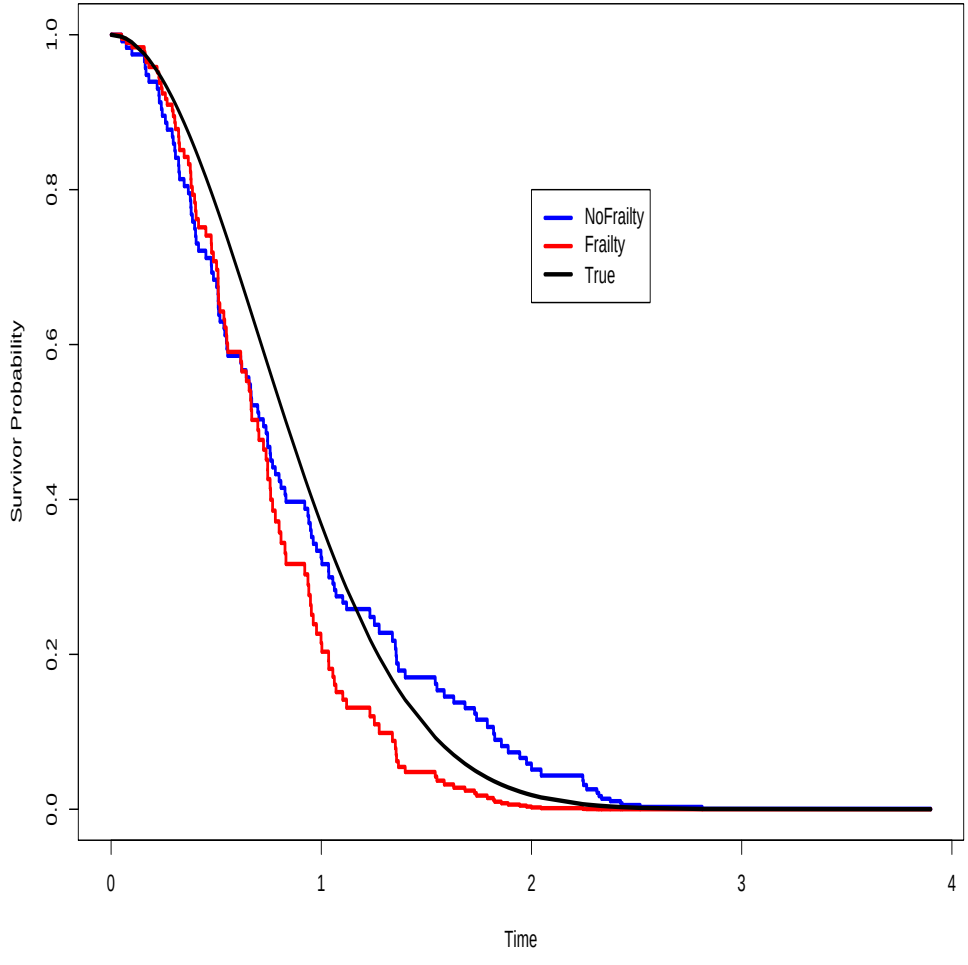
Simulated Data from the Model

True Model Parameters: $n = 15$; $\alpha = 0.90$; $\beta = (1.0, -1.0)$; $\xi = 2$; $X_1 \sim Ber(.5)$; $X_2 \sim N(0,1)$; $\tau \sim UNIF(0,10)$;
Minimal Repair with .6 prob; Baseline: Weibull(2,1)

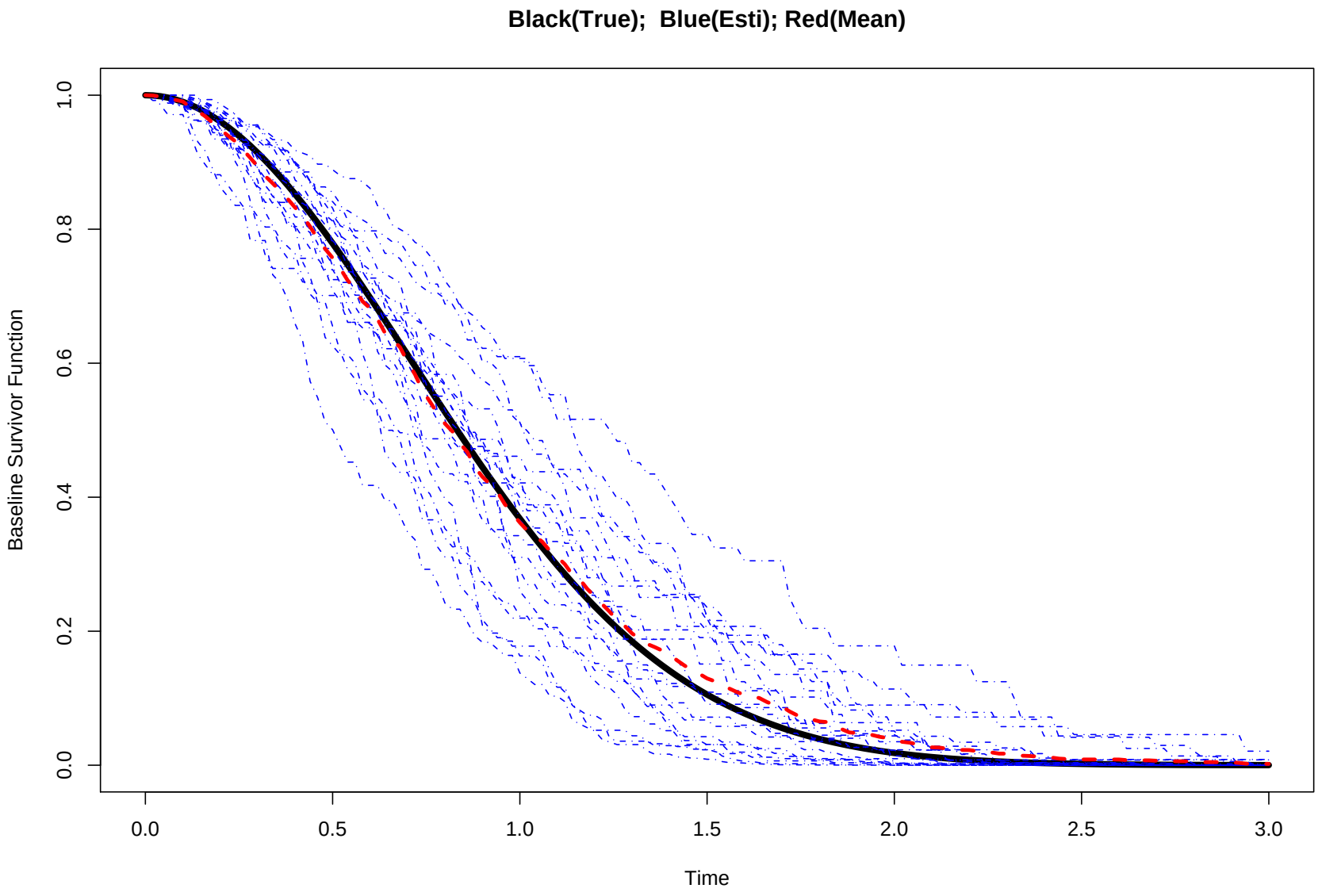


Estimates of Parameters

- **With Frailty Fit**
  - 102 iterations in EM
  - $\hat{\alpha} = .8748$
  - $\hat{\beta} = (1.099, -1.3986)$
  - $\hat{\xi} = 2.1831$
- **Without Frailty Fit**
  - $\hat{\alpha} = .963$
  - $\hat{\beta} = (0.590, -0.571)$



20 Simulated Estimates of BSF



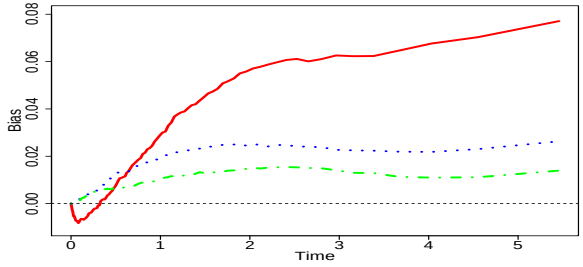
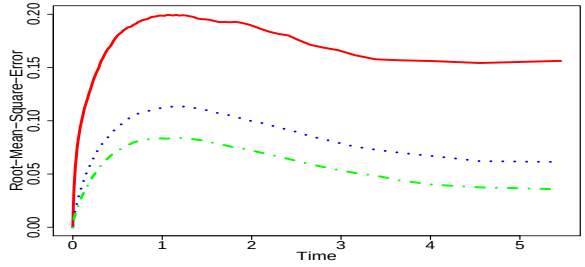
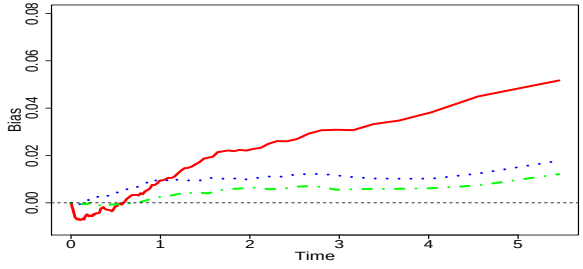
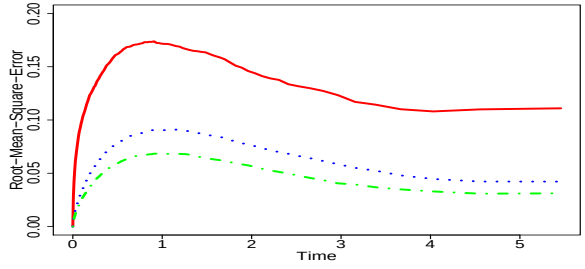
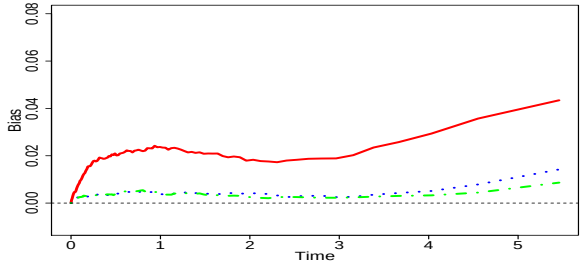
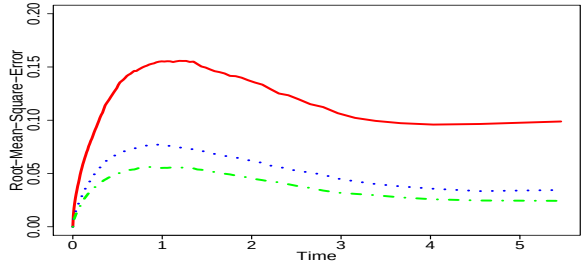
Properties: Simulated

- $\rho(k;\alpha) = \alpha^k$; $\alpha \in \{.9, 1.0, 1.05\}$
- $\psi(u) = \exp(u)$; $\beta = (1, -1)$; $X_1 \sim$ Ber$(.5)$; $X_2 \sim$ N$(0, 1)$
- Weibull baseline with shape $\gamma = .9$ (DFR) and $\gamma = 2$ (IFR)
- Gamma frailty parameter $\xi \in \{2, 6, \infty\}$
- Effective Age: Minimal repair model with $p = .6$
- Sample Size $n \in \{10, 30, 50\}$
- Censoring $\tau \sim$ Unif$(0, B)$ (approx 10 events/unit)
- 1000 replications per simulation combination

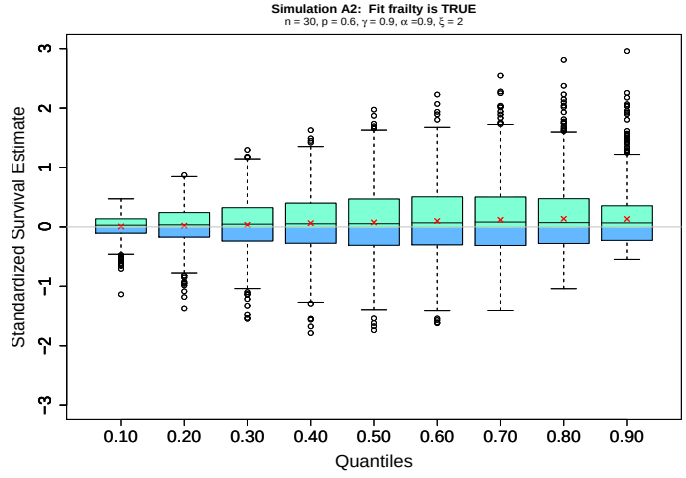
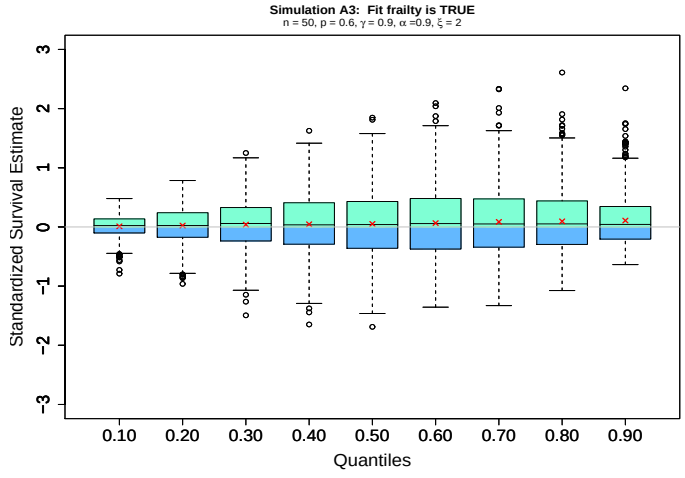
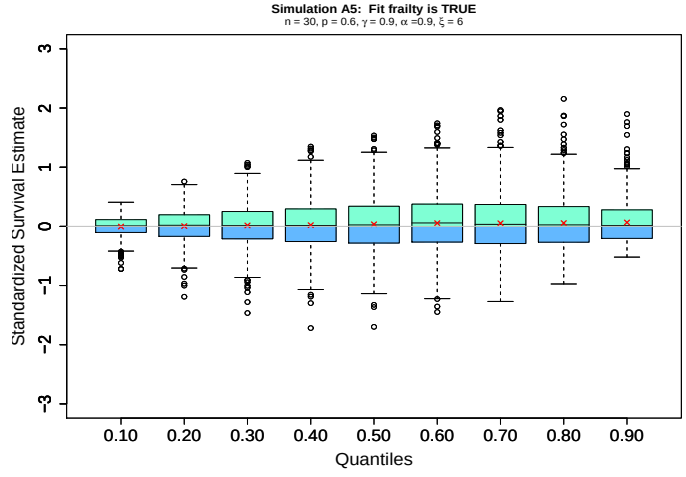
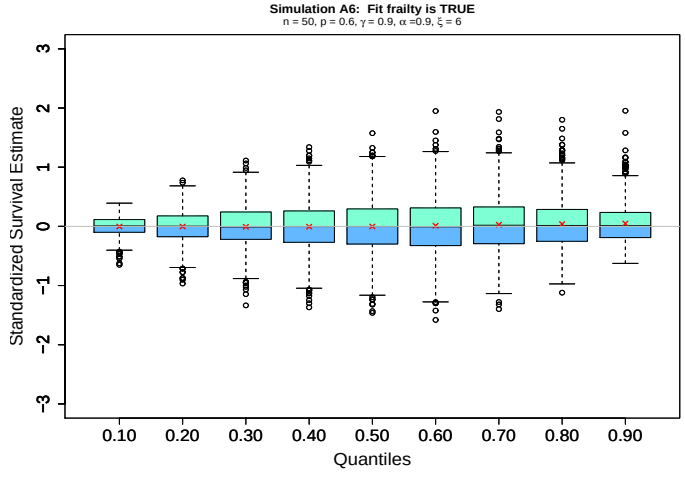
Finite-Dimensional Parameters

TableA	α	γ	ξ	η	n	$\hat{\mu}_{Ev}$	$\hat{\alpha}$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\eta}$
A2.me	0.9	0.9	2	0.67	30	4.1	0.898	1.01	-1.01	0.734
A2.sd							0.031	0.379	0.24	0.124
A3.me	0.9	0.9	2	0.67	50	5.2	0.899	1.02	-1	0.705
A3.sd							0.021	0.287	0.165	0.091
A5.me	0.9	0.9	6	0.86	30	4.3	0.9	0.988	-1.01	0.904
A5.sd							0.030	0.3	0.175	0.085
A6.me	0.9	0.9	6	0.86	50	5.3	0.899	0.998	-1	0.884
A6.sd							0.021	0.221	0.136	0.071
A8.me	0.9	0.9	∞	1	30	4.8	0.893	1.03	-1.03	
A8.sd							0.0247	0.222	0.135	
A9.me	0.9	0.9	∞	1	50	4.4	0.895	1.02	-1.02	
A9.sd							0.018	0.158	0.104	

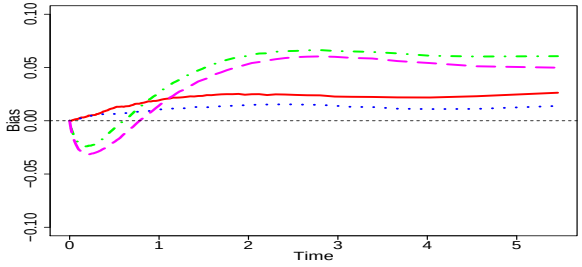
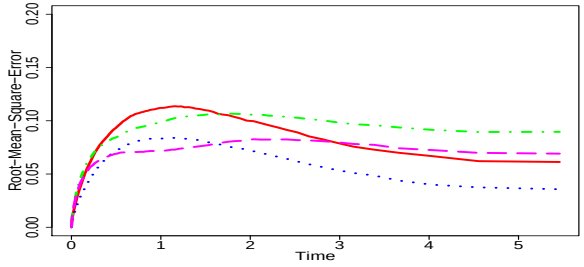
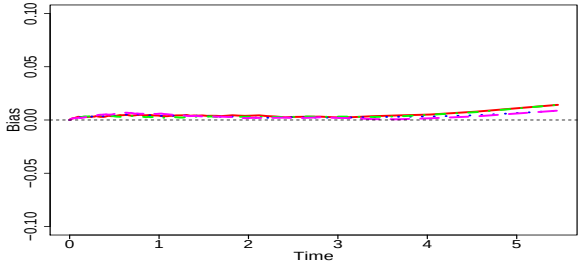
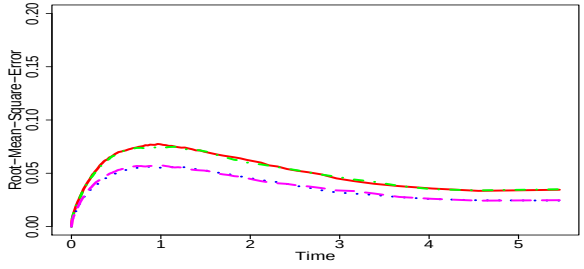
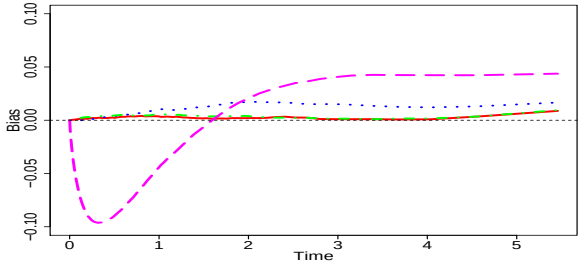
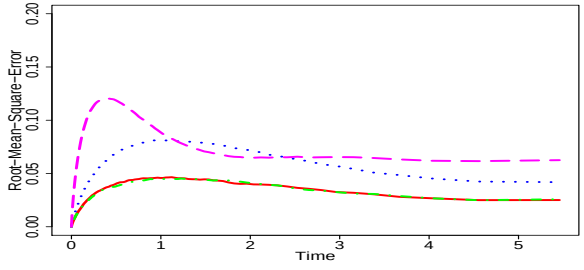
Baseline Survivor Function

ξ	Bias	RMSE
2		
6		
∞		

Peek Towards Asymptopia

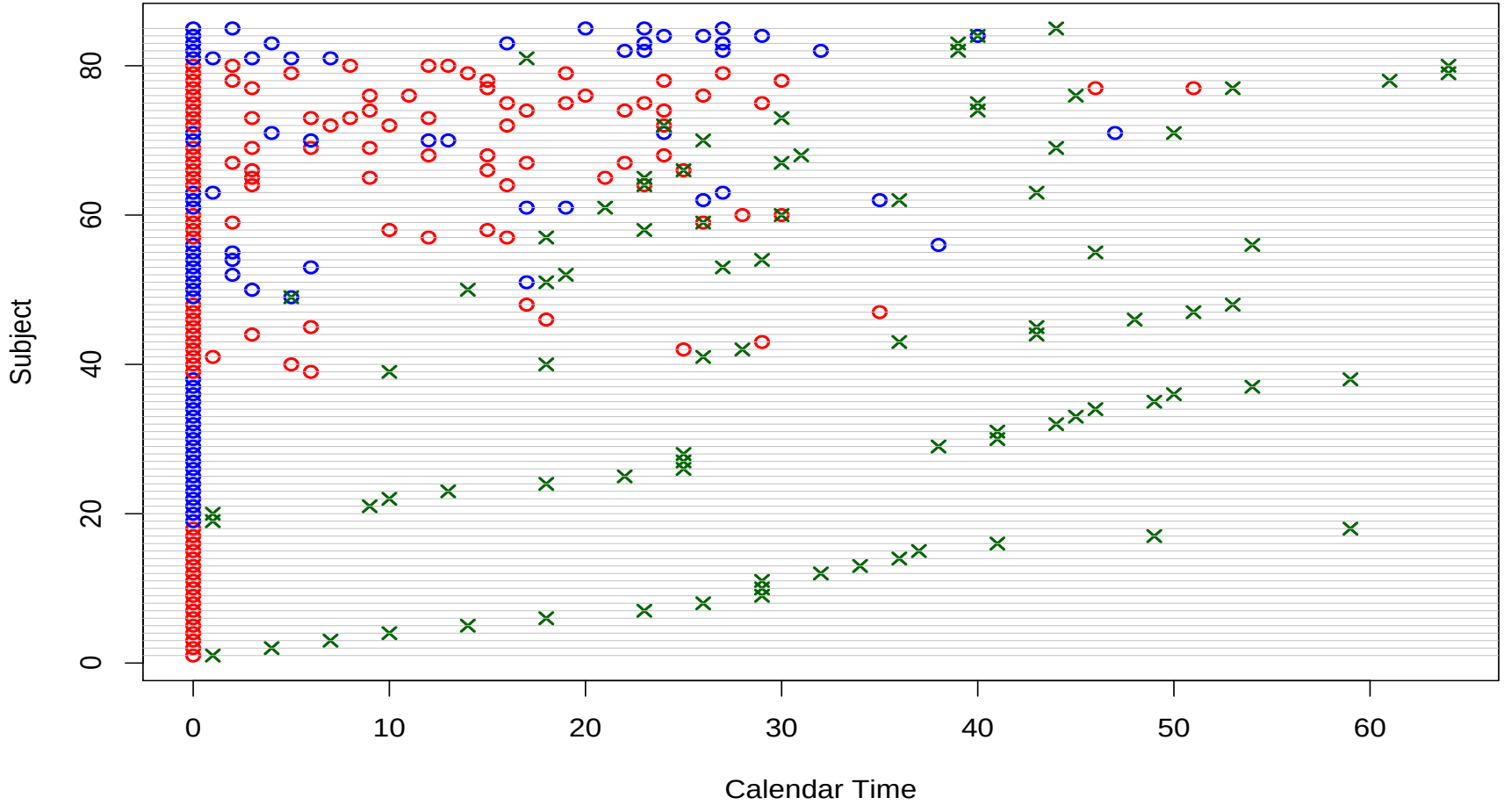
ξ	$n = 30$	$n = 50$
2		
6		

Effect of Mis-specifications

Type	Bias	RMSE
Under		
Over		
Comb.		

An Application: Bladder Data Set

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

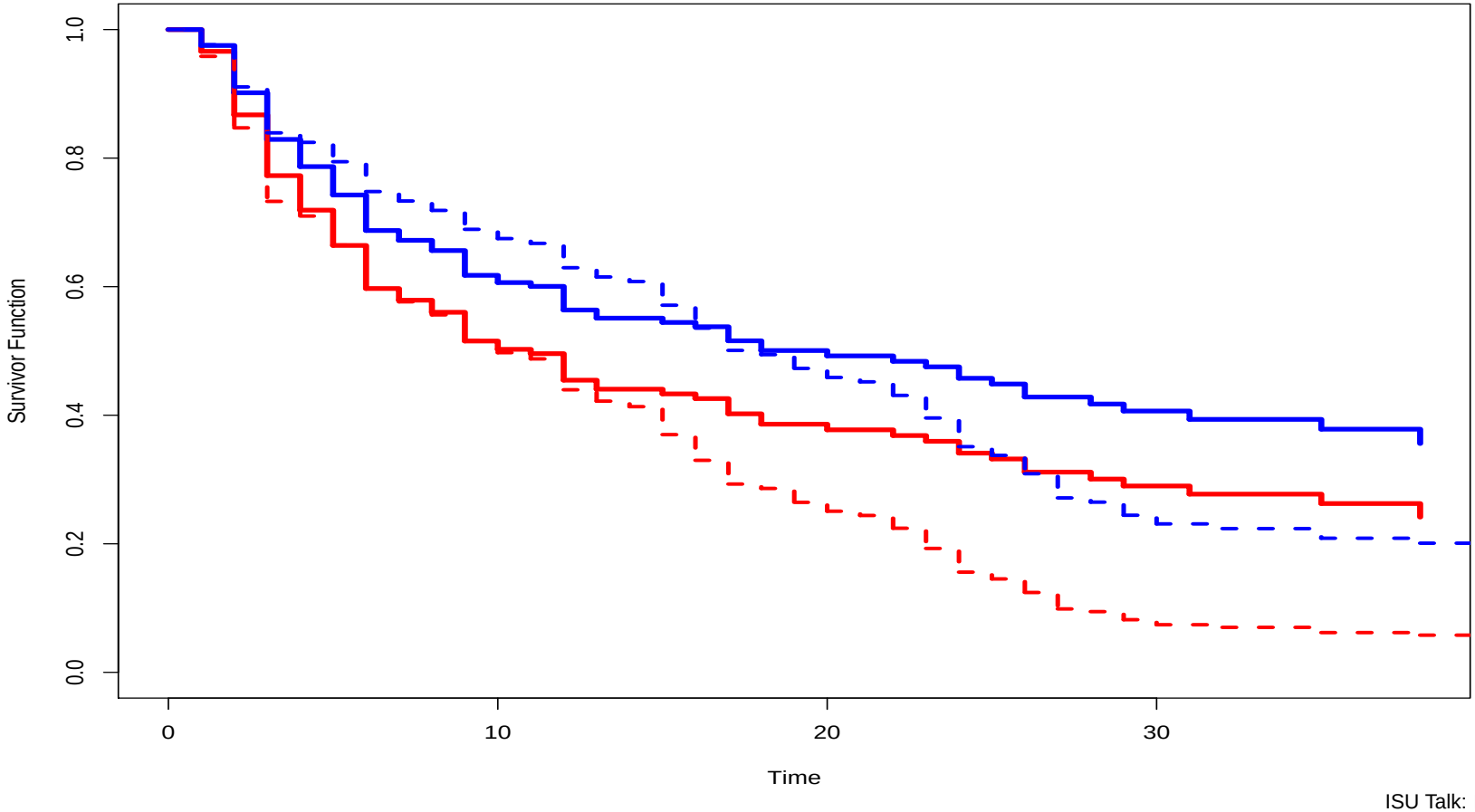


Estimates of Parameters

- X_1: (1 = placebo; 2 = thiotepa)
- X_2: size (cm) of largest initial tumor
- X_3: # of initial tumors
- Effective age: backward recurrence time (perfect repair) [also fitted with 'minimal' repair].
- Fitting model *without* frailties and 'perfect' repair:
  - $\hat{\alpha} = 0.98$ $(s.e. = 0.07)$;
  - $(\hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_3) = (-0.32, -0.02, 0.14)$;
  - s.e.s of $\hat{\beta} = (0.21, 0.07, 0.05)$.
- Fitting model *with* gamma frailties: 13 iterations in EM led to $\hat{\xi} = 5432999$ indicating absence of frailties.

Estimates of SFs for Two Groups

Blue: Thiotepa Group	Red: Placebo Group
Solid: Perfect Repair	Dashed: Minimal Repair



Concluding Remarks

- General and flexible model: incorporates aspects of recurrent event modelling.
- Allows a formal mathematical treatment which could enable reconciliation of different methods.
- Robust analysis of recurrent event data.
- Current deficiency: Effective age! Needed: paradigm shift in data gathering. Importance demonstrated in bladder data!
- Further studies: asymptotics; goodness of fit, and model validation aspects.
- Recurrent event model and longitudinal markers via latent classes. Research in progress with E. Slate.