Nonparametric Estimation for Current Status Data with Competing Risks

Marloes Maathuis
University of Washington

joint work with

Piet Groeneboom and Jon Wellner

Motivating example: HIV/AIDS vaccine trials

- Clinical trial to assess vaccine efficacy
  - Does the vaccine work?
  - Does vaccine efficacy depend on the subtype?

Motivating example: HIV/AIDS vaccine trials

- Clinical trial to assess vaccine efficacy
  - Does the vaccine work?
  - Does vaccine efficacy depend on the subtype?
- Variables of interest:
  - X = time to HIV infection
  - Y = subtype of the infecting virus

Motivating example: HIV/AIDS vaccine trials

- Clinical trial to assess vaccine efficacy
  - Does the vaccine work?
  - Does vaccine efficacy depend on the subtype?
- Variables of interest:
  - X = time to HIV infection
  - Y = subtype of the infecting virus
- People are tested for HIV at follow-up times $\Rightarrow$ interval censoring:
  - X is never observed
  - Y is sometimes observed

Motivating example: HIV/AIDS vaccine trials

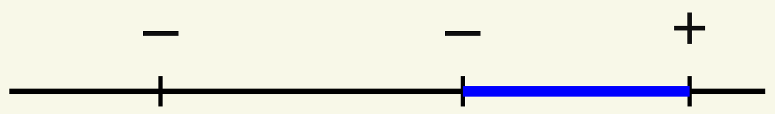
- Clinical trial to assess vaccine efficacy
  - Does the vaccine work?
  - Does vaccine efficacy depend on the subtype?
- Variables of interest:
  - X = time to HIV infection
  - Y = subtype of the infecting virus
- People are tested for HIV at follow-up times $\Rightarrow$ interval censoring:
  - X is never observed
  - Y is sometimes observed

Motivating example: HIV/AIDS vaccine trials

- Clinical trial to assess vaccine efficacy
  - Does the vaccine work?
  - Does vaccine efficacy depend on the subtype?
- Variables of interest:
  - X = time to HIV infection
  - Y = subtype of the infecting virus
- People are tested for HIV at follow-up times $\Rightarrow$ interval censoring:
  - X is never observed
  - Y is sometimes observed

Motivating example: HIV/AIDS vaccine trials

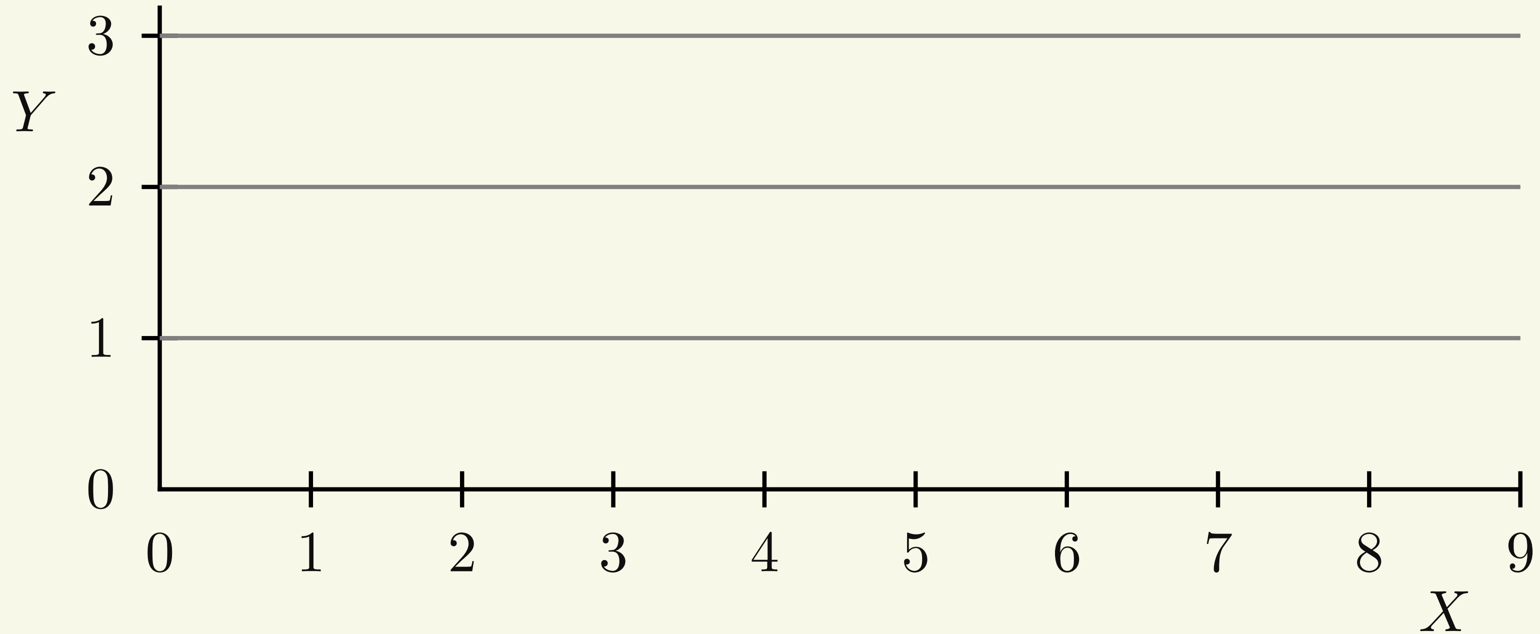
- Clinical trial to assess vaccine efficacy
  - Does the vaccine work?
  - Does vaccine efficacy depend on the subtype?
- Variables of interest:
  - X = time to HIV infection
  - Y = subtype of the infecting virus
- People are tested for HIV at follow-up times $\Rightarrow$ interval censoring:
  - X is never observed
  - Y is sometimes observed



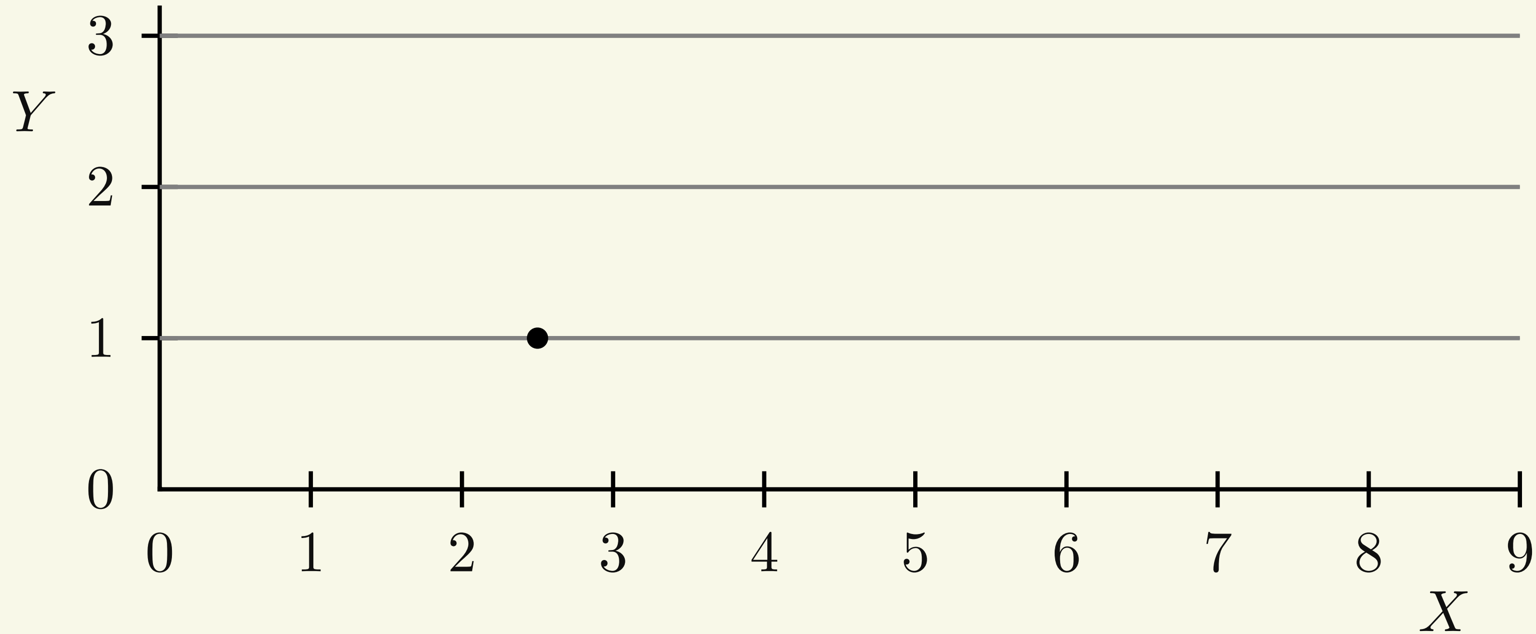
- Current status censoring: one observation time per person
  - Cross-sectional studies with several failure causes
  - First step towards interval censored data with competing risks

Graphical representation

Example, $K = 3$ competing risks

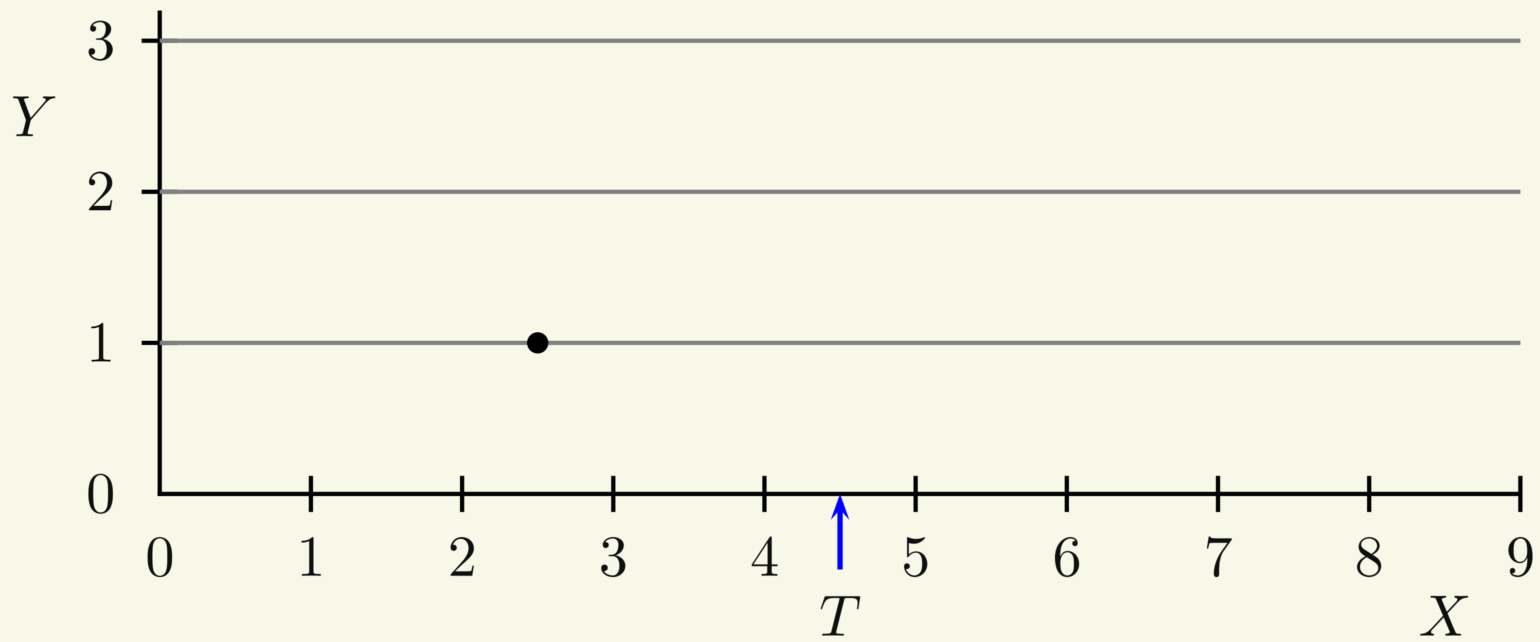


Example, $K = 3$ competing risks



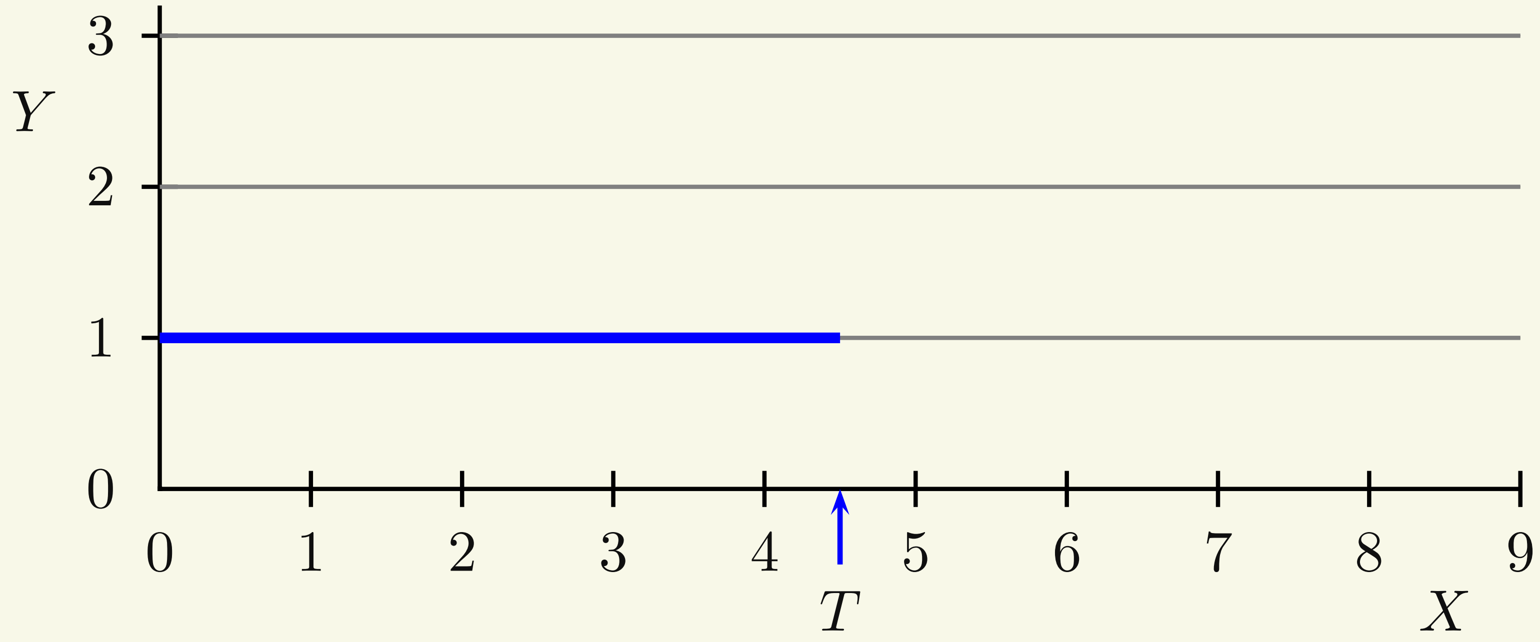
Graphical representation

Example, $K = 3$ competing risks



Graphical representation

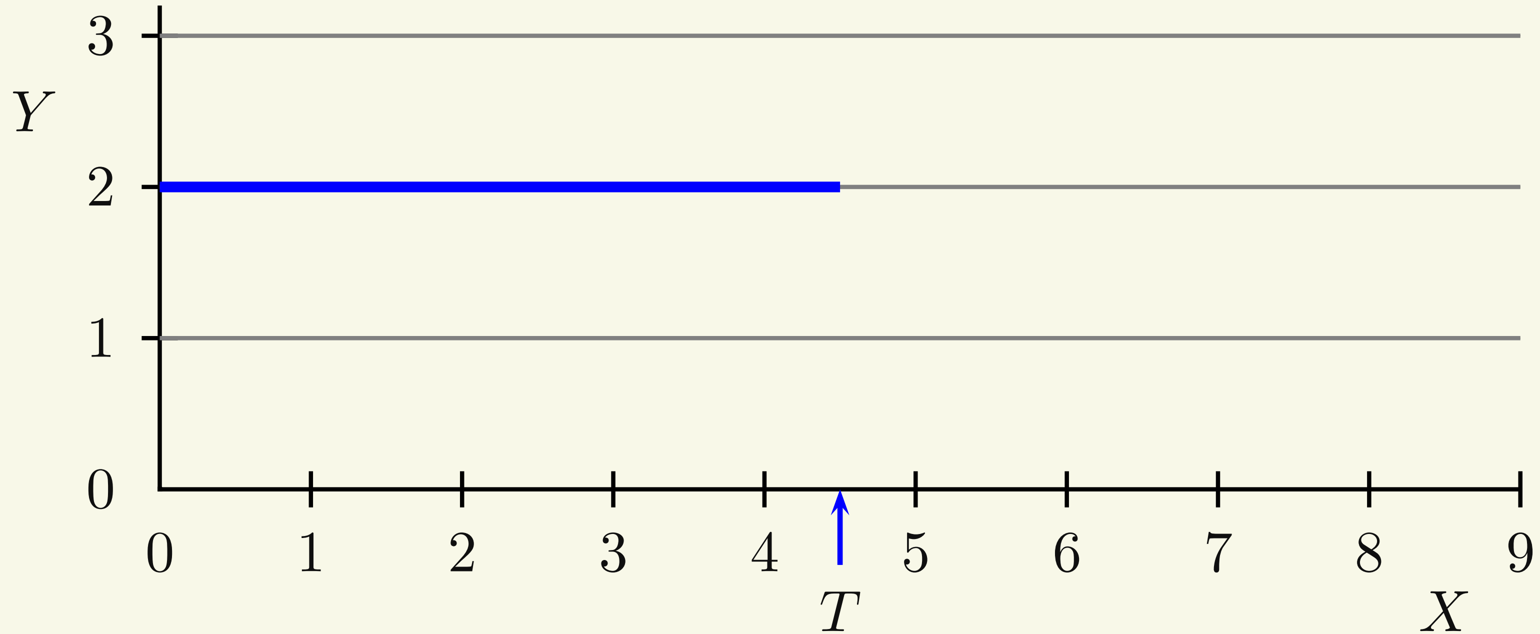
Example, $K = 3$ competing risks



$$\Delta = (\Delta_1, \Delta_2, \Delta_3, \Delta_4) = (1, 0, 0, 0)$$

Graphical representation

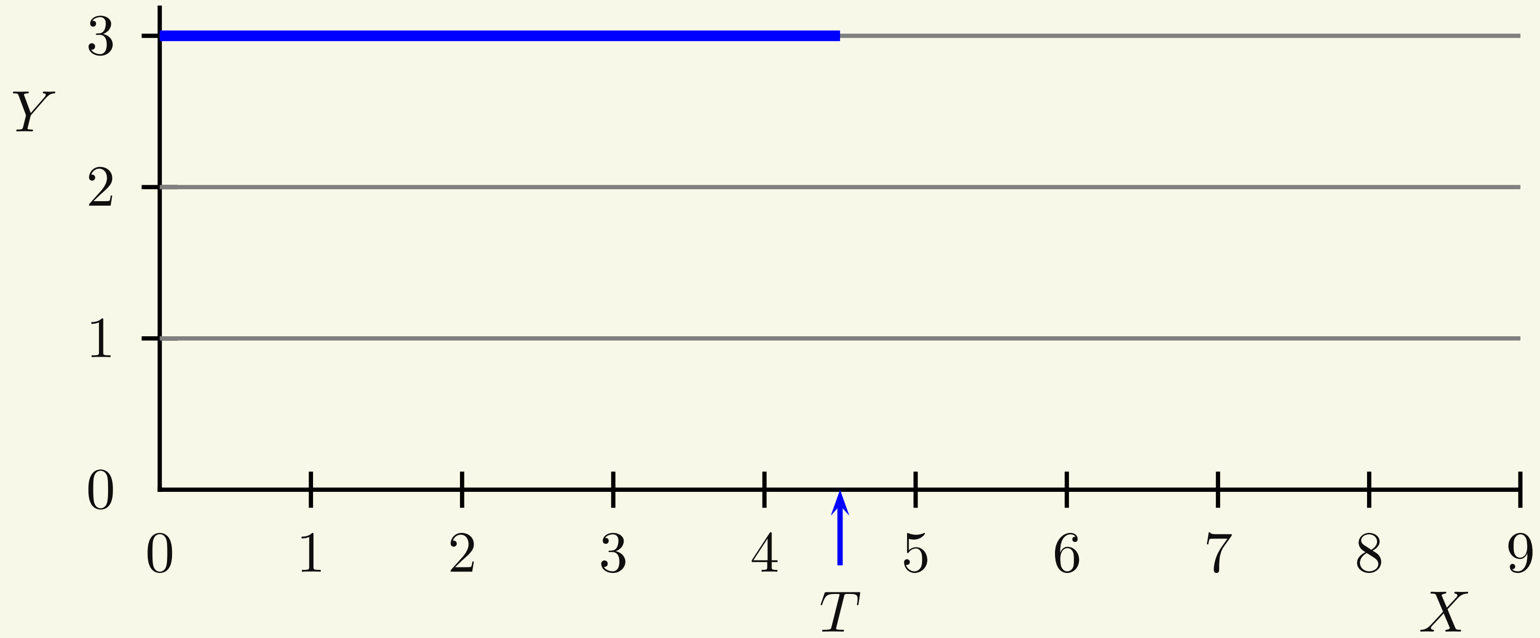
Example, $K = 3$ competing risks



$$\Delta = (\Delta_1, \Delta_2, \Delta_3, \Delta_4) = (0, 1, 0, 0)$$

Graphical representation

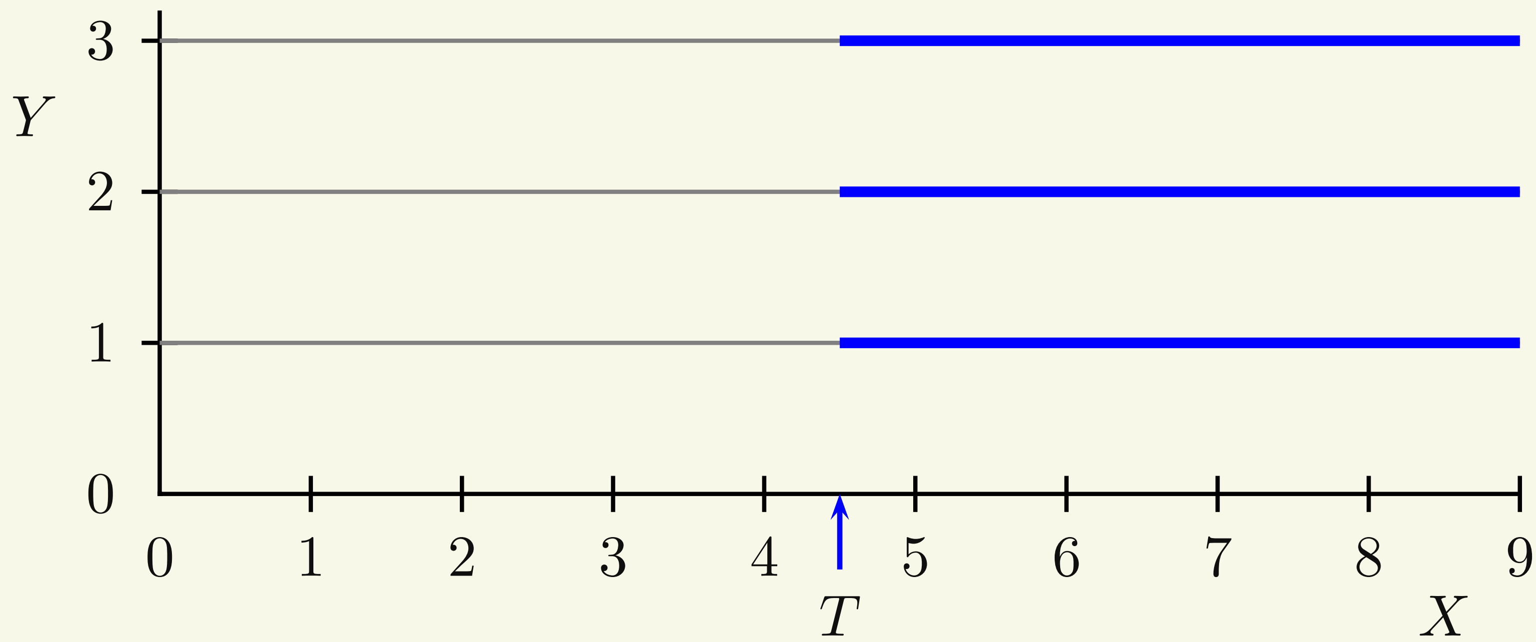
Example, $K = 3$ competing risks



$$\Delta = (\Delta_1, \Delta_2, \Delta_3, \Delta_4) = (0, 0, 1, 0)$$

Graphical representation

Example, $K = 3$ competing risks



$$\Delta = (\Delta_1, \Delta_2, \Delta_3, \Delta_4) = (0, 0, 0, 1)$$

General framework

- Study failure time of a system that can fail from K different causes (competing risks)

General framework

- Study failure time of a system that can fail from K different causes (competing risks)
- Variables of interest are (X, Y):
  - $X \in \mathbb{R}$ is the failure time
  - $Y \in \{1, \ldots, K\}$ is the failure cause

General framework

- Study failure time of a system that can fail from K different causes (competing risks)
- Variables of interest are (X, Y):
  - $X \in \mathbb{R}$ is the failure time
  - $Y \in \{1, \ldots, K\}$ is the failure cause
- Current status censoring:
  - n i.i.d. observations of (T, Δ)
  - T is the observation time
  - $\Delta = (\Delta_1, \ldots, \Delta_{K+1})$
    $\Delta_j = 1\{X \leq T, Y = j\}, \qquad j = 1, \ldots, K$
    $\Delta_{K+1} = 1\{X > T\}$

General framework

- Study failure time of a system that can fail from K different causes (competing risks)
- Variables of interest are (X, Y):
  - $X \in \mathbb{R}$ is the failure time
  - $Y \in \{1, \dots, K\}$ is the failure cause
- Current status censoring:
  - n i.i.d. observations of (T, Δ)
  - T is the observation time
  - $\Delta = (\Delta_1, \dots, \Delta_{K+1})$
    $\Delta_j = 1\{X \leq T, Y = j\}, \qquad j = 1, \dots, K$
    $\Delta_{K+1} = 1\{X > T\}$
- T is independent of (X, Y)

Estimation

- Goal: nonparametric estimation of the sub-distribution functions $F_{0j} : \mathbb{R} \to [0,1]$:

$$F_{0j}(t) = P(X \leq t, Y = j), \quad j = 1, \ldots, K$$

Estimation

- Goal: nonparametric estimation of the sub-distribution functions $F_{0j} : \mathbb{R} \to [0,1]$:

$$F_{0j}(t) = P(X \le t, Y = j), \quad j = 1, \ldots, K$$

- These sub-distribution functions are related to each other, since their sum

$$\sum_{j=1}^{K} F_{0j}(t) \;=\; P(X \le t)$$

is the overall failure time distribution.

Estimation

- Goal: nonparametric estimation of the sub-distribution functions $F_{0j} : \mathbb{R} \to [0,1]$:

$$F_{0j}(t) = P(X \leq t, Y = j), \quad j = 1, \ldots, K$$

- These sub-distribution functions are related to each other, since their sum

$$\sum_{j=1}^{K} F_{0j}(t) \;=\; P(X \leq t)$$

is the overall failure time distribution.

- In particular, $\sum_{j=1}^{K} F_{0j}(t) \leq 1$.

Previous work

- Key papers:
  - Hudgens, Satten and Longini - Biometrics 2001
  - Jewell, van der Laan and Henneman - Biometrika 2003
  - Jewell and Kalbfleisch - Biostatistics 2004
- Results:
  - Various nonparametric estimators
  - Computational algorithms
  - Simulation studies
- Open problem:
  - Large sample properties of the estimators

Our work

- We consider two estimators:
  - Nonparametric maximum likelihood estimator (MLE):
    - Natural estimator
    - Often has good behavior
  - ‘Naive estimator’ of Jewell, Van der Laan and Henneman
    - Simple: computation and theory
- Main focus:
  - Large sample properties

Outline

- Definition of the estimators
- Example
- Characterization
- Asymptotic properties:
  - Consistency
  - Rate of convergence
  - Limiting distribution
- Simulation results

Definition of the estimators

- Notation:
  - $\Delta_+ = \sum_{j=1}^K \Delta_j, \qquad F_+(t) = \sum_{j=1}^K F_j(t)$
  - $\int h(t,\delta) d\mathbb{P}_n(t,\delta) = \frac{1}{n}\sum_{i=1}^n h(T_i, \Delta_i)$
  - $\widehat{F}_n$ = MLE, $\widetilde{F}_n$ = naive estimator

Definition of the estimators

- Notation:
  - $\Delta_+ = \sum_{j=1}^K \Delta_j, \qquad F_+(t) = \sum_{j=1}^K F_j(t)$
  - $\int h(t,\delta) d\mathbb{P}_n(t,\delta) = \frac{1}{n}\sum_{i=1}^n h(T_i,\Delta_i)$
  - $\widehat{F}_n$ = MLE, $\widetilde{F}_n$ = naive estimator
- The MLE $\widehat{F}_n = (\widehat{F}_{n1}, \dots, \widehat{F}_{nK})$ maximizes

$$\int \left[\sum_{j=1}^K \delta_j \log F_j(t) + (1-\delta_+) \log\{1 - F_+(t)\} \right] d\mathbb{P}_n(t,\delta)$$

over all K-tuples $F = (F_1, \dots, F_K)$ of sub-distribution functions with $F_+ \leq 1$.
- The naive estimator $\widetilde{F}_n = (\widetilde{F}_{n1}, \dots, \widetilde{F}_{nK})$ maximizes

$$\int \sum_{j=1}^K \Big[\delta_j \log F_j(t) + (1-\delta_j) \log\{1 - F_j(t)\} \Big] d\mathbb{P}_n(t,\delta)$$

over all K-tuples $F = (F_1, \dots, F_K)$ of sub-distribution functions.

Definition of the estimators

- Notation:
  - $\Delta_+ = \sum_{j=1}^K \Delta_j, \qquad F_+(t) = \sum_{j=1}^K F_j(t)$
  - $\int h(t,\delta) d\mathbb{P}_n(t,\delta) = \frac{1}{n}\sum_{i=1}^n h(T_i, \Delta_i)$
  - $\widehat{F}_n$ = MLE, $\widetilde{F}_n$ = naive estimator
- The MLE $\widehat{F}_n = (\widehat{F}_{n1}, \ldots, \widehat{F}_{nK})$ maximizes

$$\int \left[\sum_{j=1}^K \delta_j \log F_j(t) + (1-\delta_+) \log\{1 - F_+(t)\}\right] d\mathbb{P}_n(t,\delta)$$

over all K-tuples $F = (F_1, \ldots, F_K)$ of sub-distribution functions with $F_+ \le 1$.

- The naive estimator $\widetilde{F}_n = (\widetilde{F}_{n1}, \ldots, \widetilde{F}_{nK})$ maximizes

$$\int \sum_{j=1}^K \left[\delta_j \log F_j(t) + (1-\delta_j) \log\{1 - F_j(t)\}\right] d\mathbb{P}_n(t,\delta)$$

over all K-tuples $F = (F_1, \ldots, F_K)$ of sub-distribution functions.

Alternative definition of the naive estimators $\widetilde{F}_{nj}$

- For all $j = 1, \ldots, K$, $\widetilde{F}_{nj}$ maximizes
$$\int \Big[\delta_j \log F_j(t) + (1-\delta_j) \log\{1 - F_j(t)\}\Big] d\mathbb{P}_n(t,\delta)$$
over all sub-distribution functions.

Alternative definition of the naive estimators $\widetilde{F}_{nj}$

- For all $j = 1, \ldots, K$, $\widetilde{F}_{nj}$ maximizes
$$\int \Big[\delta_j \log F_j(t) + (1 - \delta_j) \log\{1 - F_j(t)\} \Big] d\mathbb{P}_n(t, \delta)$$
over all sub-distribution functions.
- $\widetilde{F}_{nj}$ is MLE for reduced current status data (T, Δ_j)

Alternative definition of the naive estimators $\widetilde{F}_{nj}$

- For all $j = 1, \ldots, K$, $\widetilde{F}_{nj}$ maximizes
$$\int \Big[\delta_j \log F_j(t) + (1-\delta_j)\log\{1 - F_j(t)\}\Big] d\mathbb{P}_n(t,\delta)$$
over all sub-distribution functions.
- $\widetilde{F}_{nj}$ is MLE for reduced current status data (T, Δ_j)
- Computation and asymptotic properties of the naive estimator follow easily from known results
- But how much do we lose?

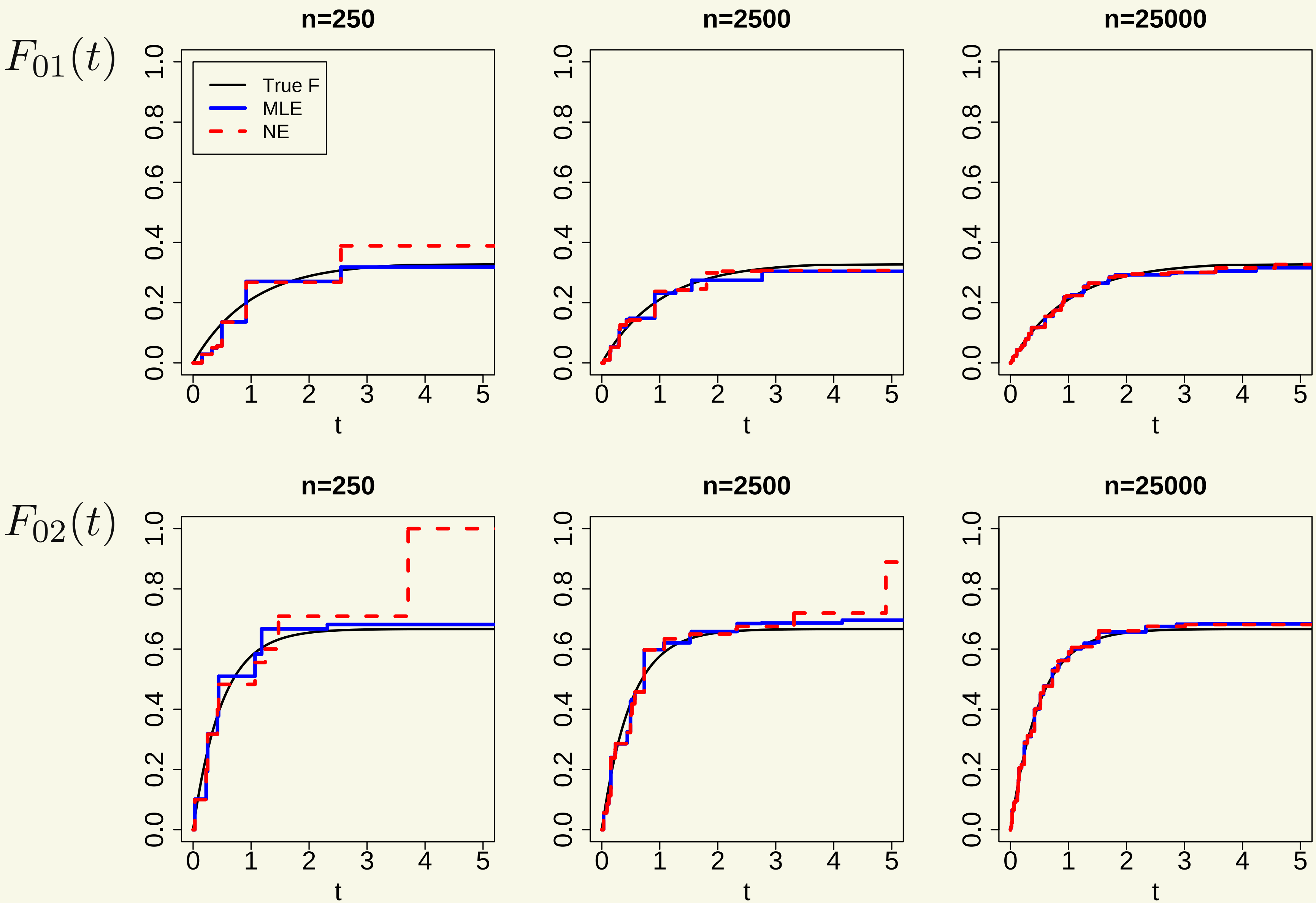
Example: $K = 2$ competing risks

- Notation:
  - G is the distribution of T
- The following example is used throughout:

$$
\begin{aligned}
&G(t) = 1 - \exp(-t) \\
&P(Y = j) = j/3, && j = 1, 2 \\
&P(X \leq t | Y = j) = 1 - \exp(-jt), && j = 1, 2
\end{aligned}
$$

  so that $F_{0j}(t) = (j/3)\{1 - \exp(-jt)\}$ for $j = 1, 2$.
- Note:
  - T is independent of (X, Y)
  - X and Y are dependent

Example: $K = 2$ competing risks



Characterization of the estimators

- Notation: τ_{nj} is a jump point of $\widetilde{F}_{nj}$ or $\widehat{F}_{nj}$.

Characterization of the estimators

- Notation: τ_{nj} is a jump point of $\widetilde{F}_{nj}$ or $\widehat{F}_{nj}$.
- The naive estimator $(\widetilde{F}_{n1}, \ldots, \widetilde{F}_{nK})$ satisfies for all j and $s \in (\tau_{nj}, \infty)$

$$0 \leq \int_{[\tau_{nj}, s)} \{\delta_j - \widetilde{F}_{nj}(t)\} d\mathbb{P}_n(t, \delta)$$

and equality holds if s is a jump point of $\widetilde{F}_{nj}$.

Characterization of the estimators

- Notation: τ_{nj} is a jump point of $\widetilde{F}_{nj}$ or $\widehat{F}_{nj}$.
- The naive estimator $(\widetilde{F}_{n1}, \ldots, \widetilde{F}_{nK})$ satisfies for all j and $s \in (\tau_{nj}, \infty)$

$$0 \leq \int_{[\tau_{nj}, s)} \{\delta_j - \widetilde{F}_{nj}(t)\} d\mathbb{P}_n(t, \delta)$$

  and equality holds if s is a jump point of $\widetilde{F}_{nj}$.
- The MLE $(\widehat{F}_{n1}, \ldots, \widehat{F}_{nK})$ satisfies for all j and $s \in (\tau_{nj}, T_{(n)})$

$$0 \leq \int_{[\tau_{nj}, s)} \left\{ \delta_j - \widehat{F}_{nj}(t) + \{\delta_+ - \widehat{F}_{n+}(t)\} \frac{\widehat{F}_{nj}(t)}{1 - \widehat{F}_{n+}(t)} \right\} d\mathbb{P}_n(t, \delta)$$

  and equality holds if s is a jump point of $\widehat{F}_{nj}$.

Characterization of the estimators

- Notation: τ_{nj} is a jump point of $\widetilde{F}_{nj}$ or $\widehat{F}_{nj}$.
- The naive estimator $(\widetilde{F}_{n1}, \ldots, \widetilde{F}_{nK})$ satisfies for all j and $s \in (\tau_{nj}, \infty)$

$$0 \leq \int_{[\tau_{nj}, s)} \{\delta_j - \widetilde{F}_{nj}(t)\} d\mathbb{P}_n(t, \delta)$$

  and equality holds if s is a jump point of $\widetilde{F}_{nj}$.
- The MLE $(\widehat{F}_{n1}, \ldots, \widehat{F}_{nK})$ satisfies for all j and $s \in (\tau_{nj}, T_{(n)})$

$$\begin{aligned} 0 &\leq \int_{[\tau_{nj}, s)} \left\{ \delta_j - \widehat{F}_{nj}(t) + \{\delta_+ - \widehat{F}_{n+}(t)\} \frac{\widehat{F}_{nj}(t)}{1 - \widehat{F}_{n+}(t)} \right\} d\mathbb{P}_n(t, \delta) \\ &= \int_{[\tau_{nj}, s)} \left\{ \delta_j - \widehat{F}_{nj}(t) + \{\delta_+ - \widehat{F}_{n+}(t)\} \frac{F_{0j}(s)}{1 - F_{0+}(s)} + R \right\} d\mathbb{P}_n(t, \delta) \end{aligned}$$

  and equality holds if s is a jump point of $\widehat{F}_{nj}$.

Characterization of the estimators

- Notation: τ_{nj} is a jump point of $\widetilde{F}_{nj}$ or $\widehat{F}_{nj}$.
- The naive estimator $(\widetilde{F}_{n1}, \ldots, \widetilde{F}_{nK})$ satisfies for all j and $s \in (\tau_{nj}, \infty)$

$$0 \leq \int_{[\tau_{nj}, s)} \{\delta_j - \widetilde{F}_{nj}(t)\} d\mathbb{P}_n(t, \delta)$$

and equality holds if s is a jump point of $\widetilde{F}_{nj}$.

- The MLE $(\widehat{F}_{n1}, \ldots, \widehat{F}_{nK})$ satisfies for all j and $s \in (\tau_{nj}, T_{(n)})$

$$\begin{aligned} 0 &\leq \int_{[\tau_{nj}, s)} \left\{ \delta_j - \widehat{F}_{nj}(t) + \{\delta_+ - \widehat{F}_{n+}(t)\} \frac{\widehat{F}_{nj}(t)}{1 - \widehat{F}_{n+}(t)} \right\} d\mathbb{P}_n(t, \delta) \\ &= \int_{[\tau_{nj}, s)} \left\{ \delta_j - \widehat{F}_{nj}(t) + \{\delta_+ - \widehat{F}_{n+}(t)\} \frac{F_{0j}(s)}{1 - F_{0+}(s)} + R \right\} d\mathbb{P}_n(t, \delta) \end{aligned}$$

and equality holds if s is a jump point of $\widehat{F}_{nj}$.

Characterization of the estimators

- Notation: τ_{nj} is a jump point of $\widetilde{F}_{nj}$ or $\widehat{F}_{nj}$.
- The naive estimator $(\widetilde{F}_{n1}, \ldots, \widetilde{F}_{nK})$ satisfies for all j and $s \in (\tau_{nj}, \infty)$

$$0 \leq \int_{[\tau_{nj}, s)} \{\delta_j - \widetilde{F}_{nj}(t)\} d\mathbb{P}_n(t, \delta)$$

  and equality holds if s is a jump point of $\widetilde{F}_{nj}$.
- The MLE $(\widehat{F}_{n1}, \ldots, \widehat{F}_{nK})$ satisfies for all j and $s \in (\tau_{nj}, T_{(n)})$

$$\begin{aligned} 0 &\leq \int_{[\tau_{nj}, s)} \left\{ \delta_j - \widehat{F}_{nj}(t) + \{\delta_+ - \widehat{F}_{n+}(t)\} \frac{\widehat{F}_{nj}(t)}{1 - \widehat{F}_{n+}(t)} \right\} d\mathbb{P}_n(t, \delta) \\ &= \int_{[\tau_{nj}, s)} \left\{ \delta_j - \widehat{F}_{nj}(t) + \{\delta_+ - \widehat{F}_{n+}(t)\} \frac{F_{0j}(s)}{1 - F_{0+}(s)} + R \right\} d\mathbb{P}_n(t, \delta) \end{aligned}$$

  and equality holds if s is a jump point of $\widehat{F}_{nj}$.

Local asymptotic properties

- Three steps: consistency, rate of convergence, and limiting distribution

Local asymptotic properties

- Three steps: consistency, rate of convergence, and limiting distribution
- Estimate population mean μ using sample mean $\bar{X}_n$:
  - Consistency: $\bar{X}_n \rightarrow_{\text{a.s.}} \mu$
  - Rate of convergence: $\sqrt{n}(\bar{X}_n - \mu) = O_p(1)$
  - Limiting distribution: $\sqrt{n}(\bar{X}_n - \mu) \rightarrow_{\text{d}} N(0, \sigma^2)$
  - Limiting distribution is used for inference, hypothesis testing

Local asymptotic properties

- Three steps: consistency, rate of convergence, and limiting distribution
- Estimate population mean μ using sample mean $\bar{X}_n$:
  - Consistency: $\bar{X}_n \to_{\text{a.s.}} \mu$
  - Rate of convergence: $\sqrt{n}(\bar{X}_n - \mu) = O_p(1)$
  - Limiting distribution: $\sqrt{n}(\bar{X}_n - \mu) \to_{\text{d}} N(0, \sigma^2)$
  - Limiting distribution is used for inference, hypothesis testing
- Current status data with competing risks:
  - Consistency: $\widehat{F}_{nj}(t_0) \to_{\text{a.s.}} F_{0j}(t_0)$
  - Rate of convergence: $n^{1/3}\{\widehat{F}_{nj}(t_0) - F_{0j}(t_0)\} = O_p(1)$
  - Limiting distribution: $n^{1/3}\{\widehat{F}_{nj}(t_0) - F_{0j}(t_0)\} \to_{\text{d}} ?$
  - Finding the limiting distribution is the first step towards making inference, drawing conclusions.

Consistency

- Notation:
  - $\overline{F}_{nj}$ stands for $\widetilde{F}_{nj}$ or $\widehat{F}_{nj}$
  - G is the distribution of T
- Global consistency:

$$\int \left| \overline{F}_{nj}(t) - F_{0j}(t) \right| dG(t) \to_{\text{a.s.}} 0, \qquad \forall j$$

  Proof: Empirical process theory

- Under some regularity conditions, there exists an $\epsilon > 0$ such that

$$\sup_{t \in [t_0 - \epsilon, t_0 + \epsilon]} \left| \overline{F}_{nj}(t) - F_{0j}(t) \right| \to_{\text{a.s.}} 0, \qquad \forall j$$

  Proof: Follows from global consistency

Rate of convergence

- Global rate of convergence:

$$n^{1/3} \int \left|\overline{F}_{nj}(t) - F_{0j}(t)\right| dG(t) = O_p(1), \qquad \forall j$$

  Proof: Empirical process theory

- Under some regularity conditions:

$$n^{1/3} \sup_{t \in [-M,M]} \left|\overline{F}_{nj}(t_0 + n^{-1/3}t) - F_{0j}(t_0)\right| = O_p(1), \qquad \forall j$$

  Proof: Difficult, new method

- Local minimax lower bound: $n^{1/3}$
- Both estimators converge locally at the optimal rate, in a minimax sense

Proof of local rate of convergence of the MLE

- Result is intuitively clear
  - The local rate for the naive estimator is $n^{1/3}$
  - The MLE should be at least as good as the naive estimator
  - No estimator can have a better rate than $n^{1/3}$
- Difficulties:
  - No explicit form for the MLE
  - No standard methods
  - System of sub-distribution functions
- Proof:
  - Use characterization of the MLE
  - First derive rate result for $\widehat{F}_{n+}$ that holds on a fixed neighborhood of t_0

Limiting distribution

- Question:

$$n^{1/3} \begin{pmatrix} \widehat{F}_{n1}(t_0) - F_{01}(t_0) \\ \vdots \\ \widehat{F}_{nK}(t_0) - F_{0K}(t_0) \end{pmatrix} \to_{\mathrm{d}} \; ?$$

Limiting distribution

- Question:

$$n^{1/3} \begin{pmatrix} \widehat{F}_{n1}(t_0) - F_{01}(t_0) \\ \vdots \\ \widehat{F}_{nK}(t_0) - F_{0K}(t_0) \end{pmatrix} \to_{\mathrm{d}} \, ?$$

- Outline:
  - Intermezzo: current status data without competing risks
  - Current status data with competing risks
    - Naive estimator
    - MLE

Intermezzo: current status data without competing risks

- Set-up:
  - $X \sim F_0, \quad T \sim G, \quad f_0 = F_0', \quad g = G'$
  - Observed data (T, Δ) where $\Delta = 1\{X \leq T\}$

Intermezzo: current status data without competing risks

- Set-up:
  - $X \sim F_0, \quad T \sim G, \quad f_0 = F_0', \quad g = G'$
  - Observed data (T, Δ) where $\Delta = 1\{X \leq T\}$
- Let W be a two-sided Brownian motion with mean 0 and variance

$$E\{W(s)W(t)\} = (|s| \wedge |t|)1\{st > 0\}F_0(t_0)\{1 - F_0(t_0)\}/g(t_0)$$

Intermezzo: current status data without competing risks

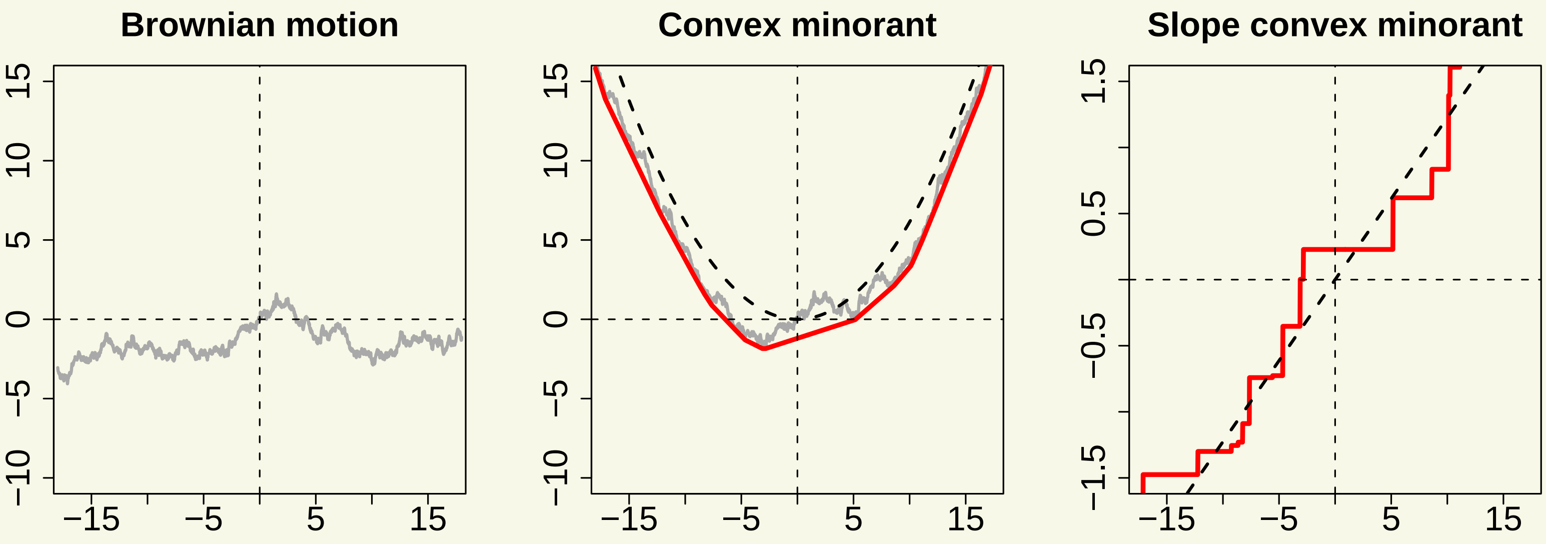
- Set-up:
  - $X \sim F_0, \quad T \sim G, \quad f_0 = F_0', \quad g = G'$
  - Observed data (T, Δ) where $\Delta = 1\{X \leq T\}$
- Let W be a two-sided Brownian motion with mean 0 and variance

$$E\{W(s)W(t)\} = (|s| \wedge |t|)1\{st > 0\}F_0(t_0)\{1 - F_0(t_0)\}/g(t_0)$$

- W has a Bernoulli variance structure, since $\Delta|T \sim \text{Bern}(F_0(T))$

- Let $V(t) = W(t) + \frac{1}{2} f_0(t_0) t^2$
- Let H be the convex minorant of V
- Groeneboom and Wellner (1992): Then, under some regularity conditions

$$n^{1/3}\{\widehat{F}_n(t_0) - F_0(t_0)\} \to_{\mathrm{d}} H'(0)$$



Competing risks gives a K-tuple of Brownian motions

- Let $g = G'$ and $W = (W_1, \ldots, W_K)$ be a K-tuple of two-sided Brownian motions with mean 0 and covariances

$$E\{W_j(s)W_k(t)\} = (|s| \wedge |t|)1\{st > 0\}\Sigma_{jk}$$

$$\Sigma_{jk} = \{1\{j = k\}F_{0k}(t_0) - F_{0j}(t_0)F_{0k}(t_0)\}/g(t_0)$$

Competing risks gives a K-tuple of Brownian motions

- Let $g = G'$ and $W = (W_1, \ldots, W_K)$ be a K-tuple of two-sided Brownian motions with mean 0 and covariances

$$E\{W_j(s)W_k(t)\} = (|s| \wedge |t|)1\{st > 0\}\Sigma_{jk}$$

$$\Sigma_{jk} = \{1\{j = k\}F_{0k}(t_0) - F_{0j}(t_0)F_{0k}(t_0)\}/g(t_0)$$

- W has multinomial covariance structure since $\Delta | T \sim \text{Multinomial}_1(F_{01}(T), \ldots, F_{0K}(T), 1 - F_{0+}(T))$

Competing risks gives a K-tuple of Brownian motions

- Let $g = G'$ and $W = (W_1, \ldots, W_K)$ be a K-tuple of two-sided Brownian motions with mean 0 and covariances

$$E\{W_j(s)W_k(t)\} = (|s| \wedge |t|)1\{st > 0\}\Sigma_{jk}$$

$$\Sigma_{jk} = \{1\{j = k\}F_{0k}(t_0) - F_{0j}(t_0)F_{0k}(t_0)\}/g(t_0)$$

- W has multinomial covariance structure since $\Delta|T \sim \text{Multinomial}_1(F_{01}(T), \ldots, F_{0K}(T), 1 - F_{0+}(T))$
- Correlation

$$r_{jk} = \frac{\Sigma_{jk}}{\sqrt{\Sigma_{jj}\Sigma_{kk}}} = -\frac{\sqrt{F_{0j}(t_0)F_{0k}(t_0)}}{\sqrt{\{1 - F_{0j}(t_0)\}\{1 - F_{0k}(t_0)\}}}$$

is negative and decreasing in t_0.

Limiting distribution of the naive estimator

- Let $f_{0j} = F'_{0j}$ and $V_j(t) = W_j(t) + \frac{1}{2} f_{0j}(t_0) t^2$, $j = 1, \ldots, K$

Limiting distribution of the naive estimator

- Let $f_{0j} = F'_{0j}$ and $V_j(t) = W_j(t) + \frac{1}{2} f_{0j}(t_0) t^2$, $j = 1, \ldots, K$
- Let $\widetilde{H}_j$ be the convex minorant of V_j, $j = 1, \ldots, K$

Limiting distribution of the naive estimator

- Let $f_{0j} = F'_{0j}$ and $V_j(t) = W_j(t) + \frac{1}{2} f_{0j}(t_0) t^2$, $j = 1, \ldots, K$
- Let $\widetilde{H}_j$ be the convex minorant of V_j, $j = 1, \ldots, K$
- Then, under some regularity conditions

$$n^{1/3} \begin{pmatrix} \widetilde{F}_{n1}(t_0) - F_{01}(t_0) \\ \vdots \\ \widetilde{F}_{nK}(t_0) - F_{0K}(t_0) \end{pmatrix} \to_{\mathrm{d}} \begin{pmatrix} \widetilde{H}'_1(0) \\ \vdots \\ \widetilde{H}'_K(0) \end{pmatrix}$$

Limiting distribution of the naive estimator

- Let $f_{0j} = F'_{0j}$ and $V_j(t) = W_j(t) + \frac{1}{2} f_{0j}(t_0) t^2$, $j = 1, \ldots, K$
- Let $\widetilde{H}_j$ be the convex minorant of V_j, $j = 1, \ldots, K$
- Then, under some regularity conditions

$$n^{1/3} \begin{pmatrix} \widetilde{F}_{n1}(t_0) - F_{01}(t_0) \\ \vdots \\ \widetilde{F}_{nK}(t_0) - F_{0K}(t_0) \end{pmatrix} \to_{\mathrm{d}} \begin{pmatrix} \widetilde{H}'_1(0) \\ \vdots \\ \widetilde{H}'_K(0) \end{pmatrix}$$

- This is a K-dimensional version of the limit for current status data

Limiting distribution of the naive estimator

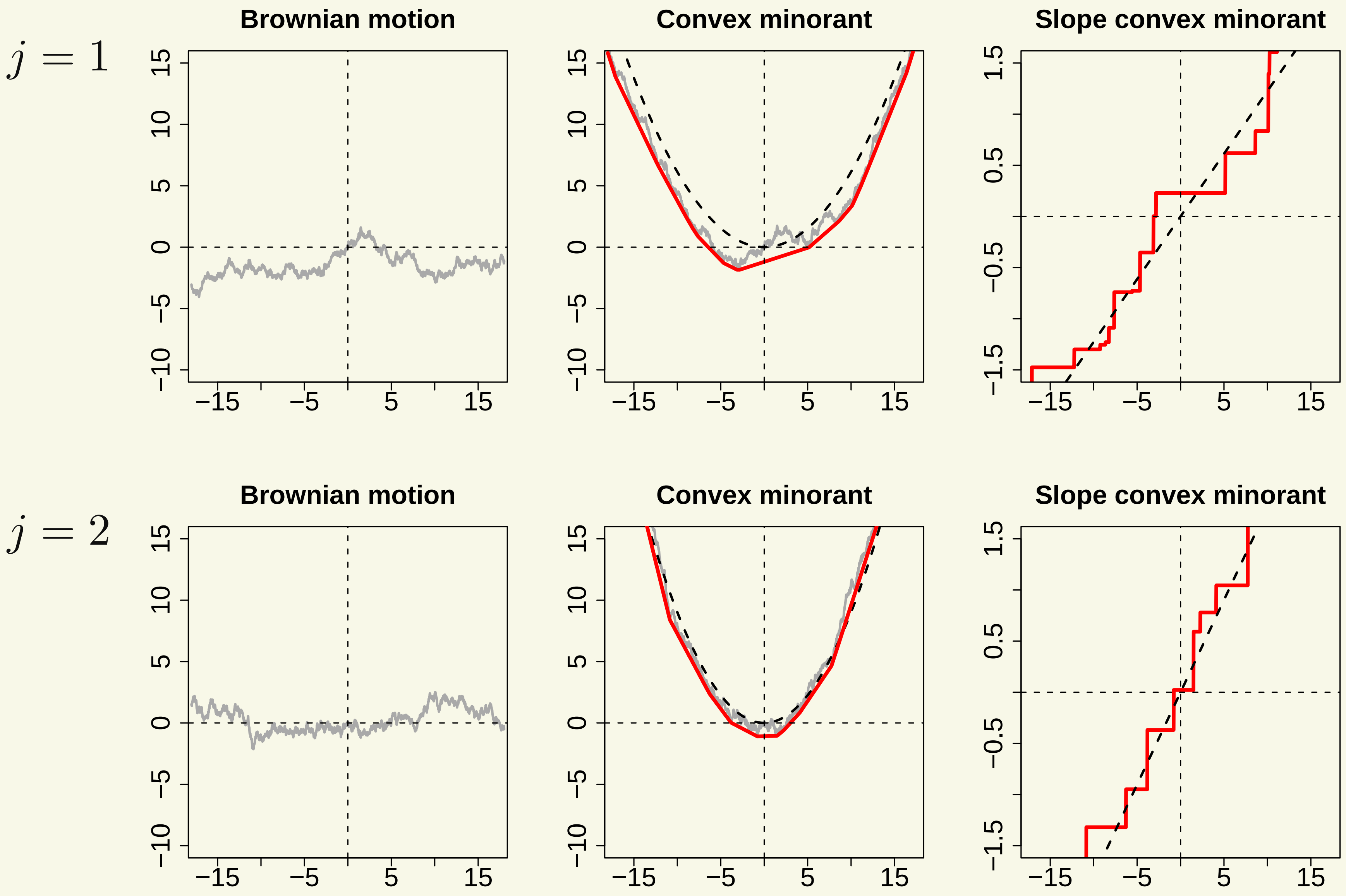
- Let $f_{0j} = F'_{0j}$ and $V_j(t) = W_j(t) + \frac{1}{2} f_{0j}(t_0) t^2$, $j = 1, \ldots, K$
- Let $\widetilde{H}_j$ be the convex minorant of V_j, $j = 1, \ldots, K$
- Then, under some regularity conditions

$$n^{1/3} \begin{pmatrix} \widetilde{F}_{n1}(t_0) - F_{01}(t_0) \\ \vdots \\ \widetilde{F}_{nK}(t_0) - F_{0K}(t_0) \end{pmatrix} \to_d \begin{pmatrix} \widetilde{H}'_1(0) \\ \vdots \\ \widetilde{H}'_K(0) \end{pmatrix}$$

- This is a K-dimensional version of the limit for current status data
- In the example with $K = 2$ competing risks:

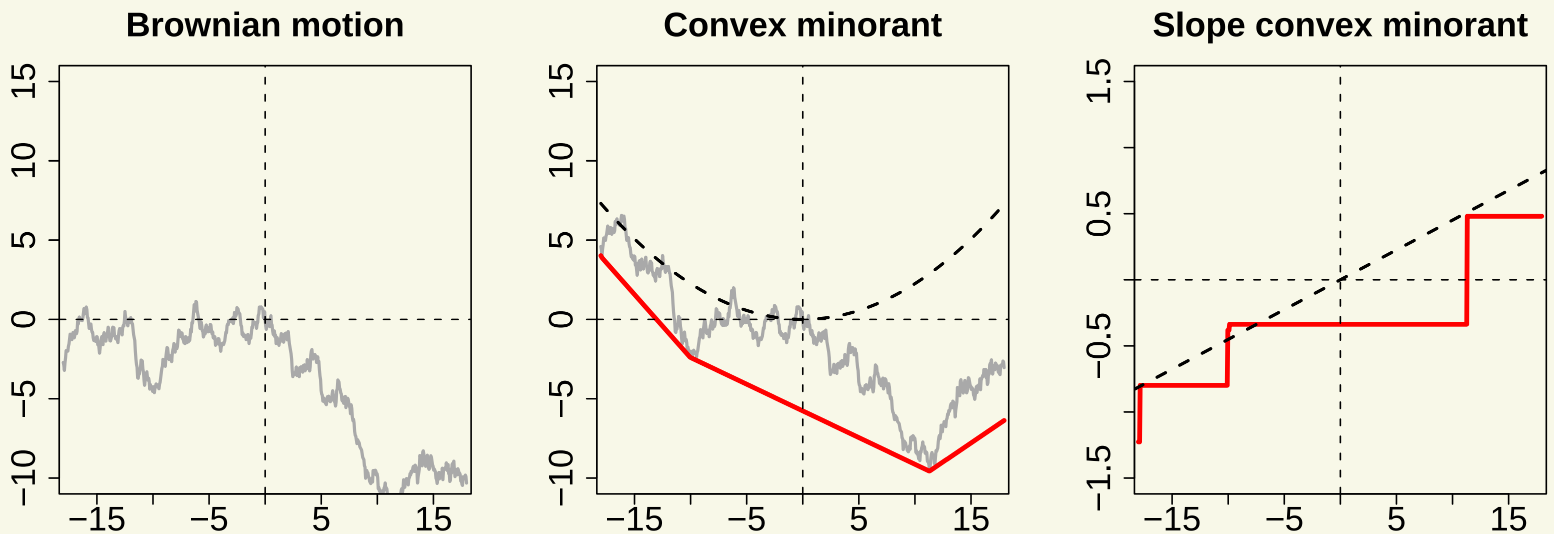
	$f_{01}(t_0)$	$f_{02}(t_0)$	Σ_{11}	Σ_{22}	r_{12}
$t_0 = 1$	0.12	0.18	0.5	0.7	-0.6
$t_0 = 2$	0.05	0.02	1.5	1.7	-0.9

Limiting processes of the naive estimator, $t_0 = 1$

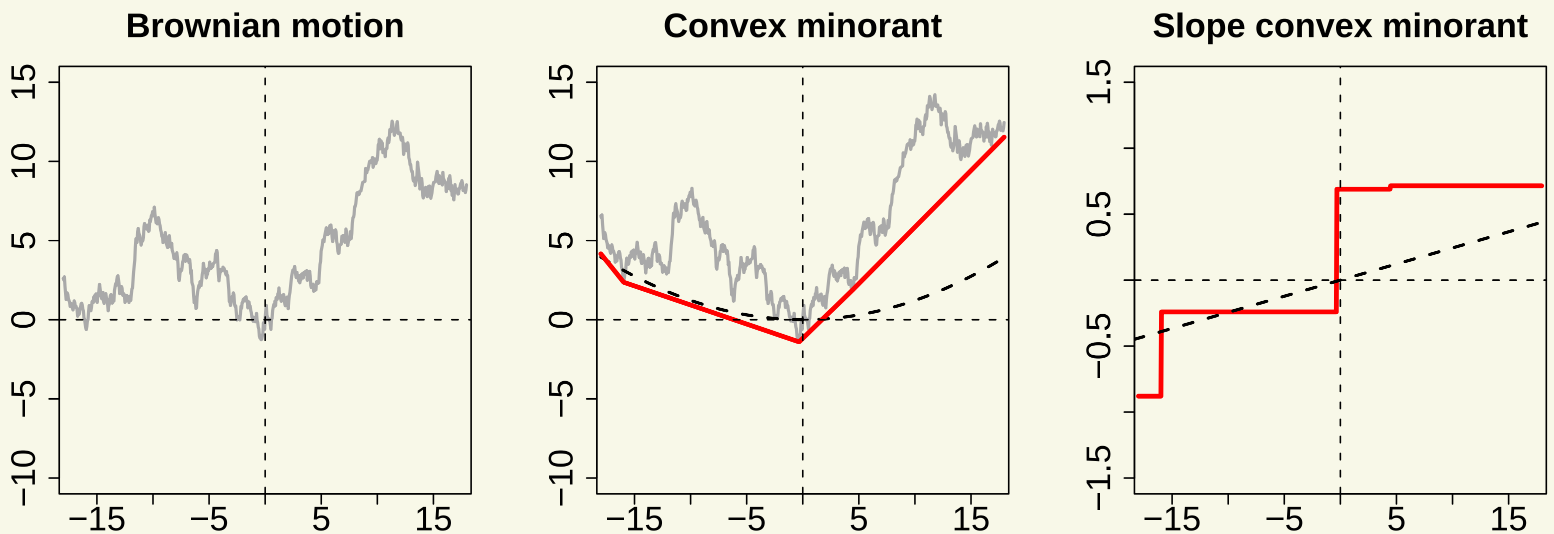


Limiting processes for the naive estimator, $t_0 = 2$

$j = 1$



$j = 2$



Limiting distribution of the MLE

- Let $V_j(t) = W_j(t) + \frac{1}{2} f_{0j}(t_0) t^2$, $j = 1, \ldots, K$

Limiting distribution of the MLE

- Let $V_j(t) = W_j(t) + \frac{1}{2} f_{0j}(t_0) t^2$, $j = 1, \ldots, K$
- For each $j = 1, \ldots, K$, let $\widehat{H}_j$ be the convex minorant of

$$V_j(t) + \frac{F_{0j}(t_0)}{1 - F_{0+}(t_0)} \{V_+(t) - \widehat{H}_+(t)\}$$

Limiting distribution of the MLE

- Let $V_j(t) = W_j(t) + \frac{1}{2} f_{0j}(t_0) t^2$, $j = 1, \ldots, K$
- For each $j = 1, \ldots, K$, let $\widehat{H}_j$ be the convex minorant of

$$V_j(t) + \frac{F_{0j}(t_0)}{1 - F_{0+}(t_0)} \{V_+(t) - \widehat{H}_+(t)\}$$

- Then, under some regularity conditions

$$n^{1/3} \begin{pmatrix} \widehat{F}_{n1}(t_0) - F_{01}(t_0) \\ \vdots \\ \widehat{F}_{nK}(t_0) - F_{0K}(t_0) \end{pmatrix} \to_{\text{d}} \begin{pmatrix} \widehat{H}_1'(0) \\ \vdots \\ \widehat{H}_K'(0) \end{pmatrix}$$

Limiting distribution of the MLE

- Let $V_j(t) = W_j(t) + \frac{1}{2} f_{0j}(t_0) t^2$, $j = 1, \ldots, K$
- For each $j = 1, \ldots, K$, let $\widehat{H}_j$ be the convex minorant of

$$V_j(t) + \frac{F_{0j}(t_0)}{1 - F_{0+}(t_0)} \{V_+(t) - \widehat{H}_+(t)\}$$

- Then, under some regularity conditions

$$n^{1/3} \begin{pmatrix} \widehat{F}_{n1}(t_0) - F_{01}(t_0) \\ \vdots \\ \widehat{F}_{nK}(t_0) - F_{0K}(t_0) \end{pmatrix} \to_{\mathrm{d}} \begin{pmatrix} \widehat{H}'_1(0) \\ \vdots \\ \widehat{H}'_K(0) \end{pmatrix}$$

- Note that $\widehat{H}$ is a new self-induced limiting process, and hence existence and uniqueness of this process are not automatic.

The process $\widehat{H} = (\widehat{H}_1, \ldots, \widehat{H}_K)$

- For each $j = 1, \ldots, K$, let $\widehat{H}_j$ be the convex minorant of

$$V_j(t) + \frac{F_{0j}(t_0)}{1 - F_{0+}(t_0)}\{V_+(t) - \widehat{H}_+(t)\}$$

The process $\widehat{H} = (\widehat{H}_1, \ldots, \widehat{H}_K)$

- For each $j = 1, \ldots, K$, let $\widehat{H}_j$ be the convex minorant of

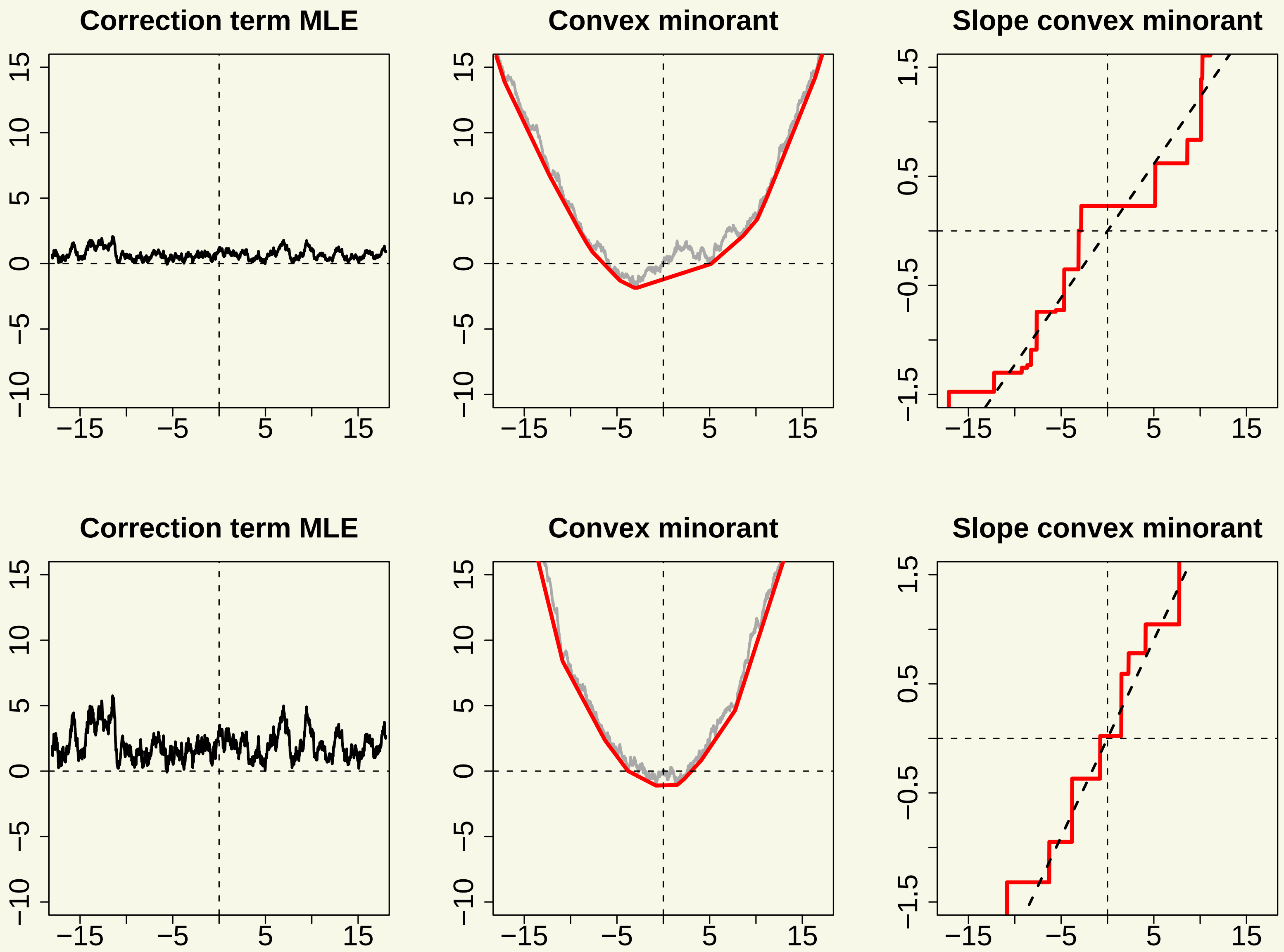
$$V_j(t) + \frac{F_{0j}(t_0)}{1 - F_{0+}(t_0)} \{V_+(t) - \widehat{H}_+(t)\}$$

- Note:
  - Correction term comes from $(1 - \delta_+) \log(1 - F_+)$ in log likelihood
  - Correction term involves $\widehat{H}_+ \Rightarrow$ self-induced process
  - Correction term same for all components, except for scaling
  - Scaling factor $F_{0j}(t_0)/(1 - F_{0+}(t_0))$ is increasing in t_0
  - One can prove that the correction term is nonnegative
  - $\widetilde{H}_j(t) \leq \widehat{H}_j(t)$

Comparison MLE and naive estimator, $t_0 = 1$

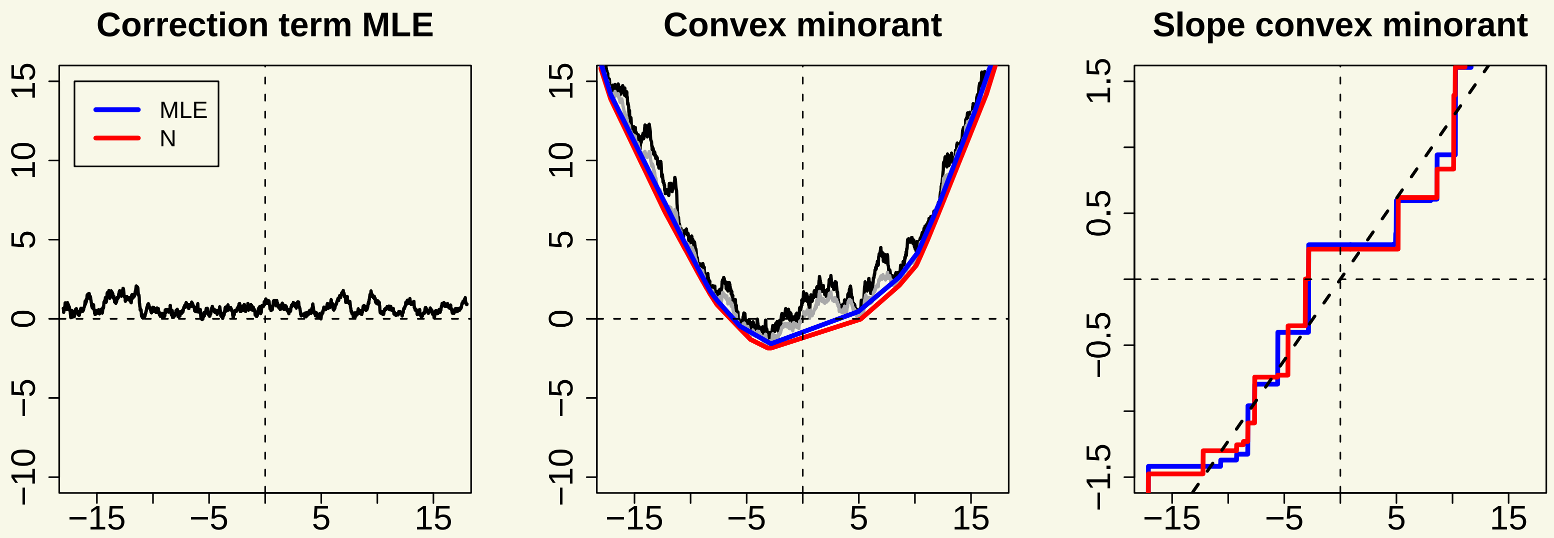
$j = 1$

$j = 2$

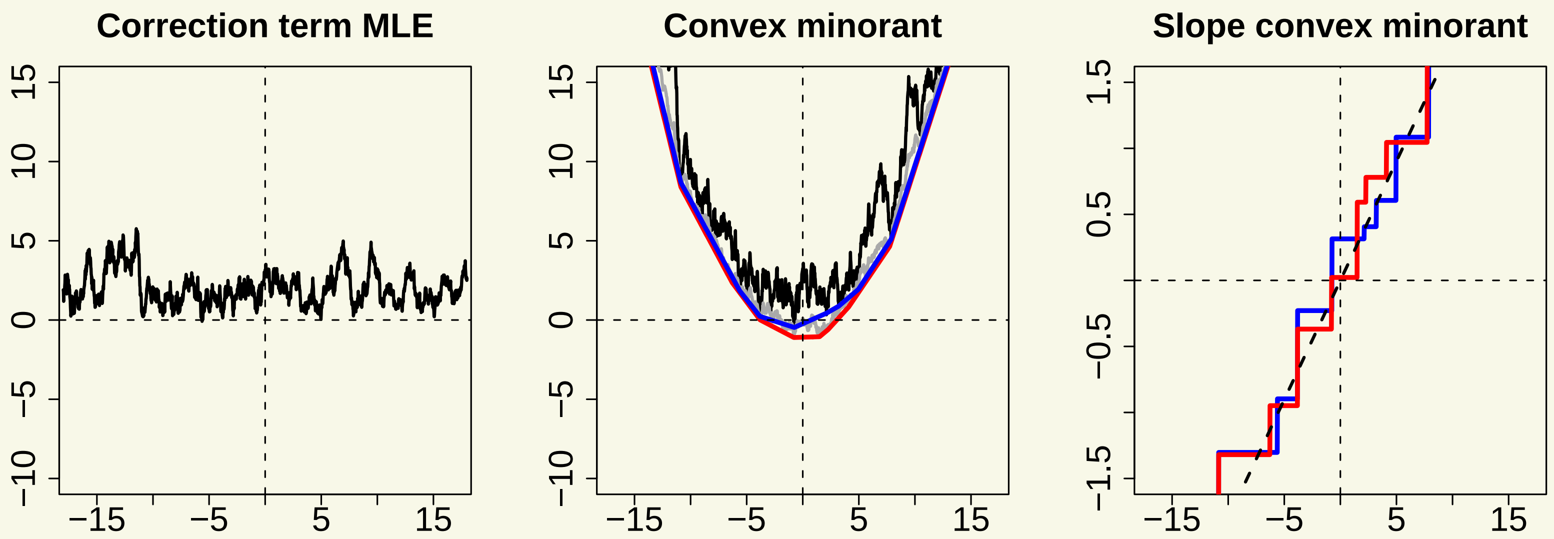


Comparison MLE and naive estimator, $t_0 = 1$

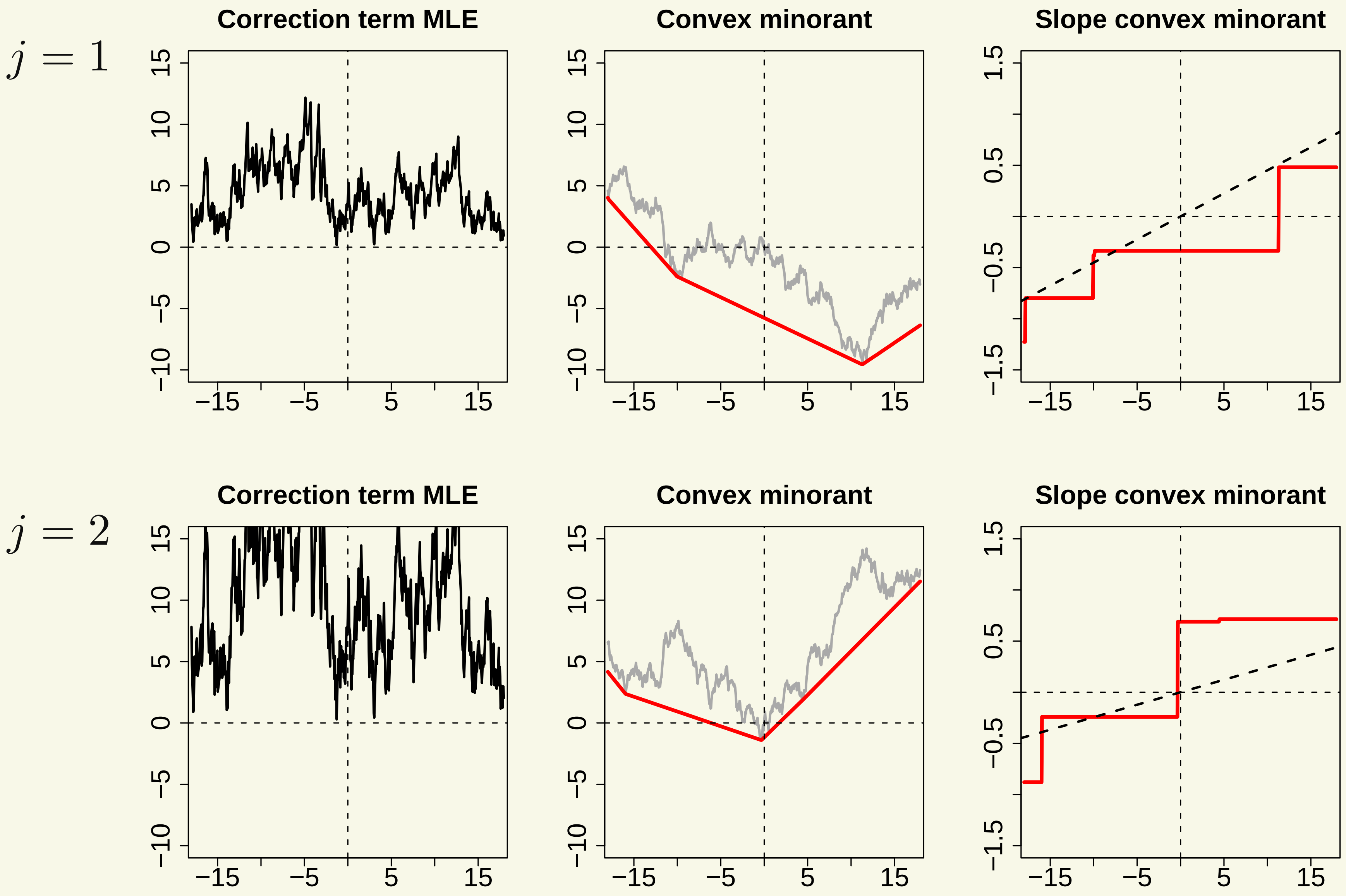
$j = 1$



$j = 2$



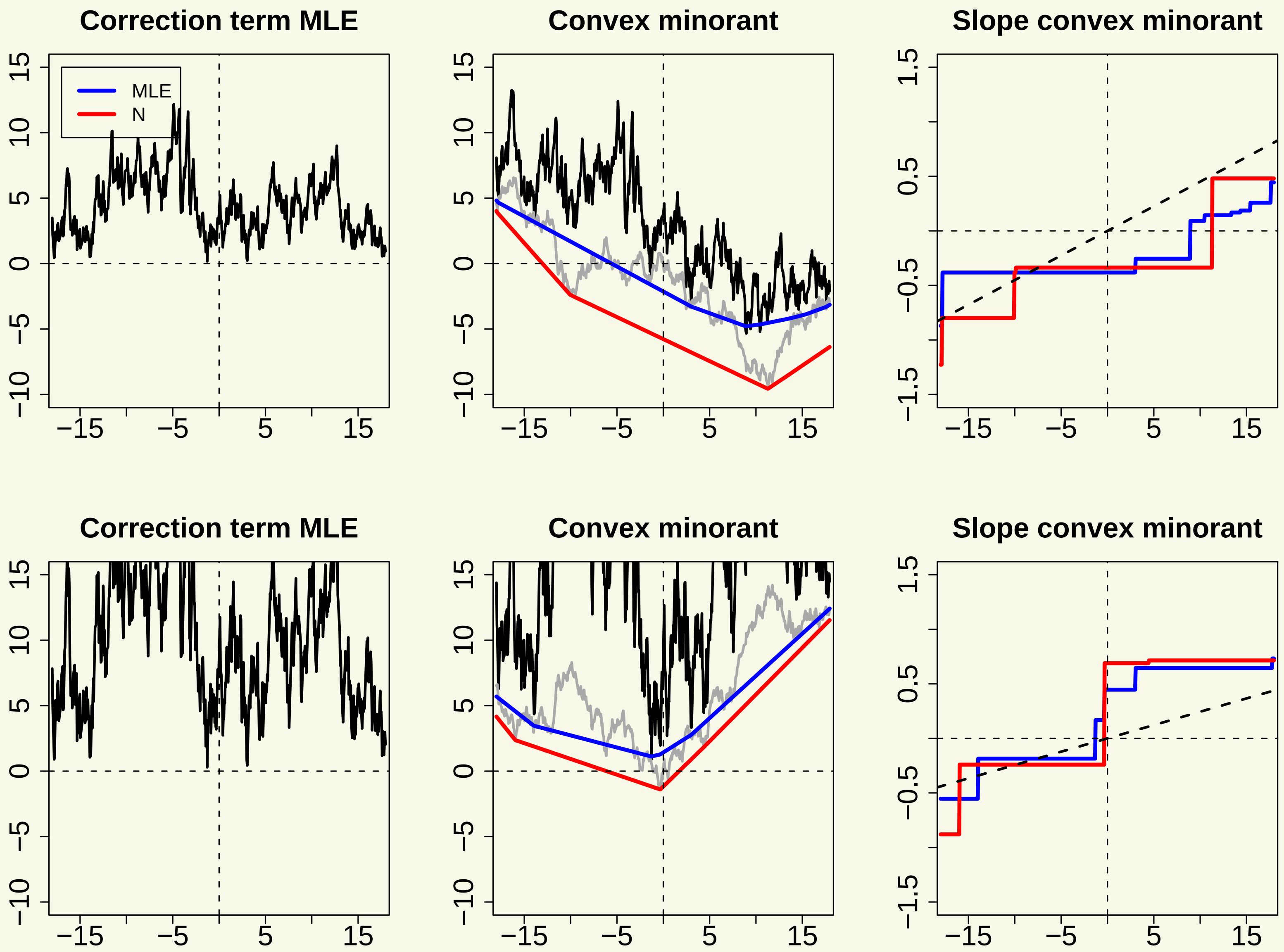
Comparison MLE and naive estimator, $t_0 = 2$



Comparison MLE and naive estimator, $t_0 = 2$

$j = 1$

$j = 2$



Simulation study

- Same example
- 1000 simulations for sample sizes 250, 2500 and 25000

Simulation study

- Same example
- 1000 simulations for sample sizes 250, 2500 and 25000
- Estimators:
  - MLE
  - Naive estimator (NE)
  - Truncated naive estimator (TNE):
    - truncate as soon as $\widetilde{F}_{n+}(t) > 1$
  - Scaled naive estimator (SNE):
    - fix a large (we took $a = 3.5$)
    - if $\widetilde{F}_{n+}(a) > 1$, divide all estimates by $\widetilde{F}_{n+}(a)$

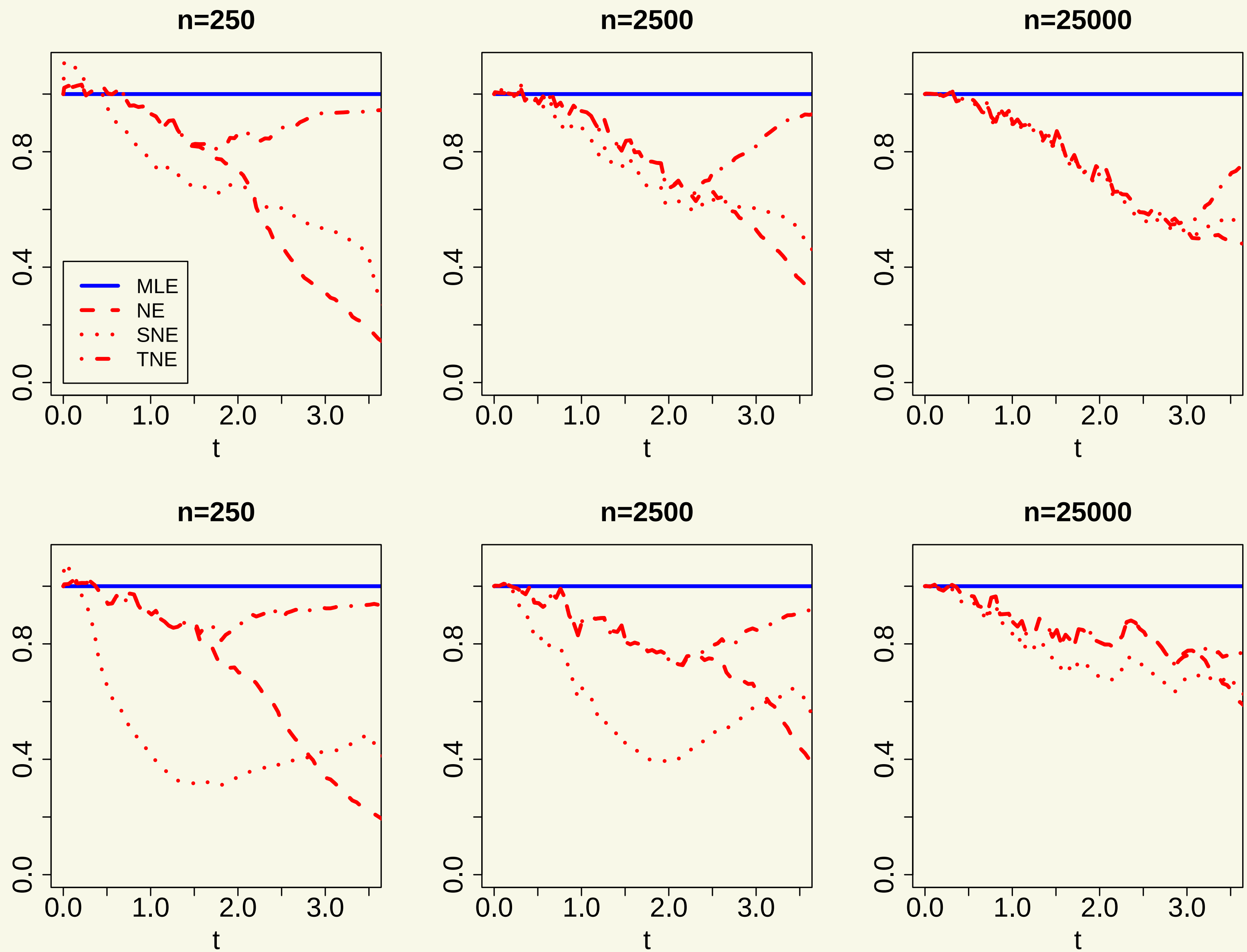
Simulation study

- Same example
- 1000 simulations for sample sizes 250, 2500 and 25000
- Estimators:
  - MLE
  - Naive estimator (NE)
  - Truncated naive estimator (TNE):
    - truncate as soon as $\widetilde{F}_{n+}(t) > 1$
  - Scaled naive estimator (SNE):
    - fix a large (we took $a = 3.5$)
    - if $\widetilde{F}_{n+}(a) > 1$, divide all estimates by $\widetilde{F}_{n+}(a)$
- Computed mean squared error on a grid
- Compute relative efficiency: (MSE MLE) / (MSE estimator)

Relative efficiency

$j = 1$

$j = 2$



Summary and future work

- Summary:
  - Current status data with competing risks
  - Limiting distribution of MLE and naive estimator
  - New self-induced limiting process for the MLE

Summary and future work

- Summary:
  - Current status data with competing risks
  - Limiting distribution of MLE and naive estimator
  - New self-induced limiting process for the MLE
- Immediate implication:
  - Pointwise confidence intervals via sub-sampling (Politis et al. (1999))

Summary and future work

- Summary:
  - Current status data with competing risks
  - Limiting distribution of MLE and naive estimator
  - New self-induced limiting process for the MLE
- Immediate implication:
  - Pointwise confidence intervals via sub-sampling (Politis et al. (1999))
- Future work:
  - Better understand structure of new limiting distribution
  - Incorporate covariate information, two-sample tests
  - Generalize to interval censored data

References

- P. Groeneboom, M.H. Maathuis, J.A. Wellner (2006a), *"Current status data with competing risks: consistency and rates of convergence of the MLE"*, submitted to the Annals of Statistics.
- P. Groeneboom, M.H. Maathuis, J.A. Wellner (2006b), *"Current status data with competing risks: limiting distribution of the MLE"*, submitted to the Annals of Statistics.

Available at http://www.stat.washington.edu/marloes

Thanks!