Current status data with competing risks:

Nonparametric estimation

Marloes Maathuis

Joint work with Piet Groeneboom and Jon Wellner

Department of Statistics, University of Washington

marloes@stat.washington.edu

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

Current status data with competing risks

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

Current status data with competing risks

- We study the failure time of a system that can fail from K different causes (competing risks)
- Variables of interest are (X, Y), where
  - $X > 0$ is the failure time
  - $Y \in \{1, \ldots, K\}$ is the failure cause

Current status data with competing risks

- We study the failure time of a system that can fail from K different causes (competing risks)
- Variables of interest are (X, Y), where
  - $X > 0$ is the failure time
  - $Y \in \{1, \ldots, K\}$ is the failure cause
- Current status censoring: we observe the ‘current status’ of the system at a random time T

Current status data with competing risks

- We study the failure time of a system that can fail from K different causes (competing risks)
- Variables of interest are (X, Y), where
  - $X > 0$ is the failure time
  - $Y \in \{1, \ldots, K\}$ is the failure cause
- Current status censoring: we observe the 'current status' of the system at a random time T
- We observe (T, Δ_+, Z), where
  - T is the observation time
  - $\Delta_+ = 1\{X \leq T\}$
  - $Z = \Delta_+ Y$

  We have n i.i.d. observations (T_i, Δ_{+i}, Z_i), $i = 1, \ldots, n$

Current status data with competing risks

- We study the failure time of a system that can fail from K different causes (competing risks)
- Variables of interest are (X, Y), where
  - $X > 0$ is the failure time
  - $Y \in \{1, \ldots, K\}$ is the failure cause
- Current status censoring: we observe the 'current status' of the system at a random time T
- We observe (T, Δ_+, Z), where
  - T is the observation time
  - $\Delta_+ = 1\{X \leq T\}$
  - $Z = \Delta_+ Y$

  We have n i.i.d. observations (T_i, Δ_{+i}, Z_i), $i = 1, \ldots, n$
- We assume T is independent of (X, Y)

Where does this type of data arise?

- Cross-sectional studies

Where does this type of data arise?

- Cross-sectional studies
- HIV Example:
  - HIV testing & HIV subtype determination

Where does this type of data arise?

- Cross-sectional studies
- HIV Example:
  - HIV testing & HIV subtype determination
  - T = age at time of HIV test
  - X = age at time of HIV infection
  - Δ_+ = result of the HIV test
  - Y = HIV subtype

Where does this type of data arise?

- Cross-sectional studies
- HIV Example:
  - HIV testing & HIV subtype determination
  - T = age at time of HIV test
  - X = age at time of HIV infection
  - Δ_+ = result of the HIV test
  - Y = HIV subtype
- See Hudgens, Satten, Longini (2001) for such a study with a more complicated censoring scheme.

Estimation of the sub-distribution functions

We study nonparametric estimation of the sub-distribution functions $F_k : \mathbb{R}^+ \to [0,1]$, where:

$$F_k(t) = P(X \leq t, Y = k), \quad k = 1, \ldots, K$$

Estimation of the sub-distribution functions

We study nonparametric estimation of the sub-distribution functions $F_k : \mathbb{R}^+ \to [0, 1]$, where:

$$F_k(t) = P(X \leq t, Y = k), \quad k = 1, \ldots, K$$

The sub-distribution functions are related to each other in the sense that $F_+(t) \equiv \sum_{k=1}^K F_k(t) = P(X \leq t) \leq 1$.

Overview of previous work in this area

Key papers:

- Hudgens, Satten and Longini - Biometrics 2001
- Jewell, van der Laan and Henneman - Biometrika 2003
- Jewell and Kalbfleisch - Biostatistics 2004

Overview of previous work in this area

Key papers:

- Hudgens, Satten and Longini - Biometrics 2001
- Jewell, van der Laan and Henneman - Biometrika 2003
- Jewell and Kalbfleisch - Biostatistics 2004

Some of the main results in these papers:

- Development of different nonparametric estimators
- Algorithms to compute the estimators
- Simulation studies comparing the estimators

Overview of previous work in this area

Key papers:

- Hudgens, Satten and Longini - Biometrics 2001
- Jewell, van der Laan and Henneman - Biometrika 2003
- Jewell and Kalbfleisch - Biostatistics 2004

Some of the main results in these papers:

- Development of different nonparametric estimators
- Algorithms to compute the estimators
- Simulation studies comparing the estimators

In this talk the focus is on the MLE and a ‘naive’ estimator:

- Convex minorant characterizations
- Simulation results

MLE and naive estimator

- Notation: $\Delta_{ki} = 1\{Z_i = k\}$ and $\Delta_{+i} = \sum_{k=1}^{K} \Delta_{ki}$

MLE and naive estimator

- Notation: $\Delta_{ki} = 1\{Z_i = k\}$ and $\Delta_{+i} = \sum_{k=1}^{K} \Delta_{ki}$
- The MLE $\hat{F}_n = (\hat{F}_{n1}, \ldots, \hat{F}_{nK})$ maximizes

$$\sum_{i=1}^{n} \left[\sum_{k=1}^{K} \Delta_{ki} \log F_k(T_i) + (1 - \Delta_{+i}) \log\{1 - F_+(T_i)\} \right]$$

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

MLE and naive estimator

- Notation: $\Delta_{ki} = 1\{Z_i = k\}$ and $\Delta_{+i} = \sum_{k=1}^K \Delta_{ki}$
- The MLE $\hat{F}_n = (\hat{F}_{n1}, \ldots, \hat{F}_{nK})$ maximizes

$$\sum_{i=1}^n \left[\sum_{k=1}^K \Delta_{ki} \log F_k(T_i) + (1 - \Delta_{+i}) \log\{1 - F_+(T_i)\} \right]$$

  over all K-tuples $F = (F_1, \ldots, F_K)$ of sub-distribution functions summing to at most one.
- The naive estimator $\tilde{F}_{nk}$ maximizes

$$\sum_{i=1}^n [\Delta_{ki} \log F_k(T_i) + (1 - \Delta_{ki}) \log\{1 - F_k(T_i)\}]$$

  over all sub-distribution functions F_k, for $k = 1, \ldots, K$.

Convex minorant characterizations (1)

- Assume for notational simplicity that the T_i are ordered and have no ties: $T_1 < T_2 < \cdots < T_n$

Convex minorant characterizations (1)

- Assume for notational simplicity that the T_i are ordered and have no ties: $T_1 < T_2 < \cdots < T_n$
- The naive estimator $\tilde{F}_{nk}$ can be characterized as the left derivative of the greatest convex minorant of the points

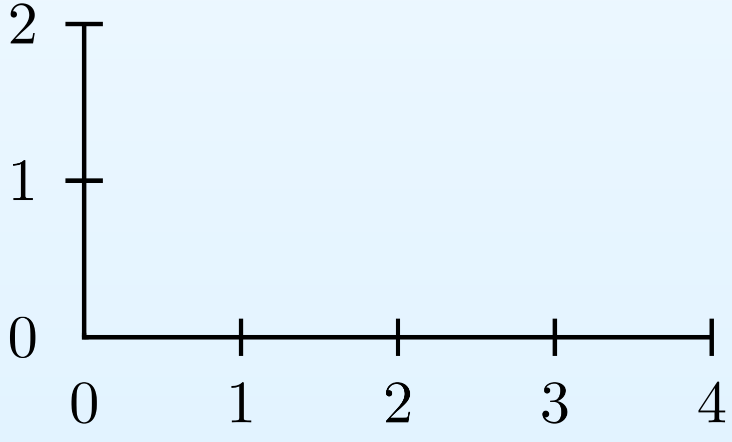
$$\left(i, \sum_{j=1}^{i} \Delta_{kj}\right), \quad i = 0, \ldots, n$$

Convex minorant characterizations (1)

- Assume for notational simplicity that the T_i are ordered and have no ties: $T_1 < T_2 < \cdots < T_n$
- The naive estimator $\tilde{F}_{nk}$ can be characterized as the left derivative of the greatest convex minorant of the points

$$\left(i, \sum_{j=1}^{i} \Delta_{kj}\right), \quad i = 0, \ldots, n$$

Example: $n = 4$ and $(\Delta_{k1}, \Delta_{k2}, \Delta_{k3}, \Delta_{k4}) = (0, 1, 1, 0)$

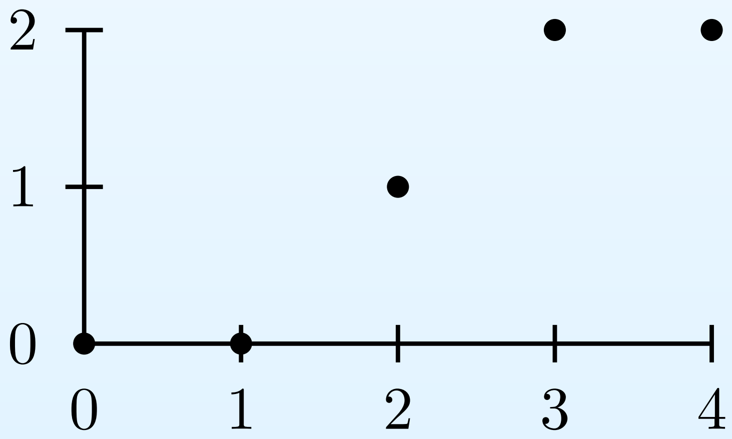


Convex minorant characterizations (1)

- Assume for notational simplicity that the T_i are ordered and have no ties: $T_1 < T_2 < \cdots < T_n$
- The naive estimator $\tilde{F}_{nk}$ can be characterized as the left derivative of the greatest convex minorant of the points

$$\left(i, \sum_{j=1}^{i} \Delta_{kj}\right), \quad i = 0, \ldots, n$$

Example: $n = 4$ and $(\Delta_{k1}, \Delta_{k2}, \Delta_{k3}, \Delta_{k4}) = (0, 1, 1, 0)$

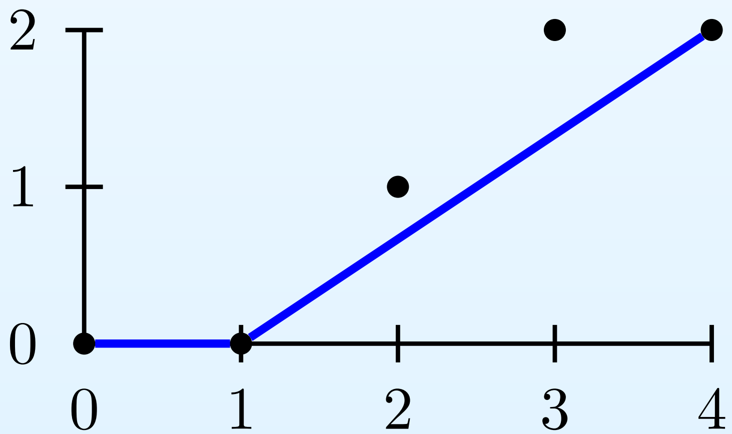


Convex minorant characterizations (1)

- Assume for notational simplicity that the T_i are ordered and have no ties: $T_1 < T_2 < \cdots < T_n$
- The naive estimator $\tilde{F}_{nk}$ can be characterized as the left derivative of the greatest convex minorant of the points

$$\left(i, \sum_{j=1}^{i} \Delta_{kj}\right), \quad i = 0, \ldots, n$$

Example: $n = 4$ and $(\Delta_{k1}, \Delta_{k2}, \Delta_{k3}, \Delta_{k4}) = (0, 1, 1, 0)$

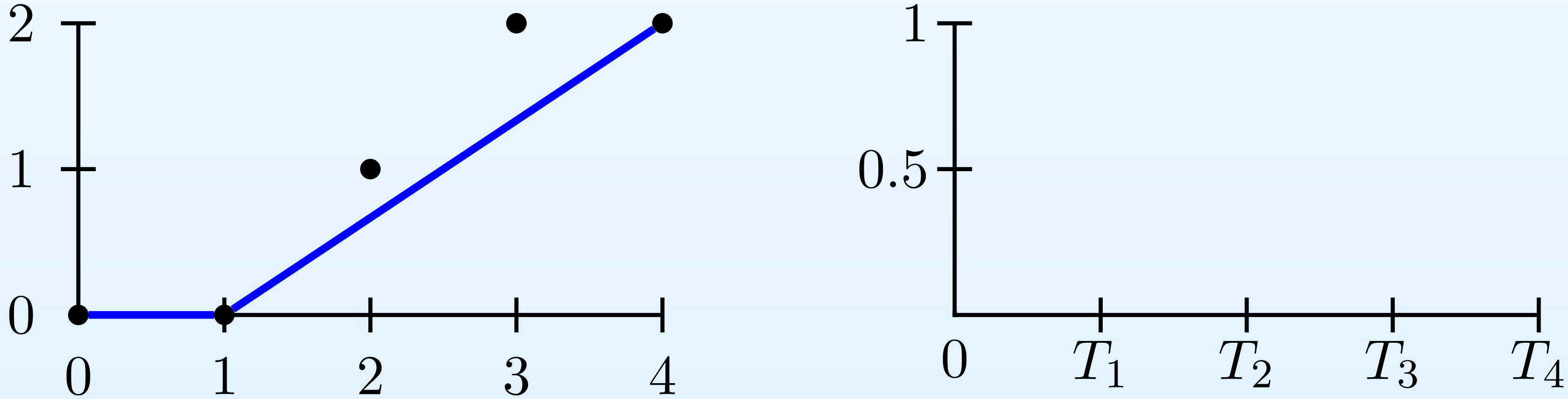


Convex minorant characterizations (1)

- Assume for notational simplicity that the T_i are ordered and have no ties: $T_1 < T_2 < \cdots < T_n$
- The naive estimator $\tilde{F}_{nk}$ can be characterized as the left derivative of the greatest convex minorant of the points

$$\left(i, \sum_{j=1}^{i} \Delta_{kj}\right), \quad i = 0, \ldots, n$$

Example: $n = 4$ and $(\Delta_{k1}, \Delta_{k2}, \Delta_{k3}, \Delta_{k4}) = (0, 1, 1, 0)$

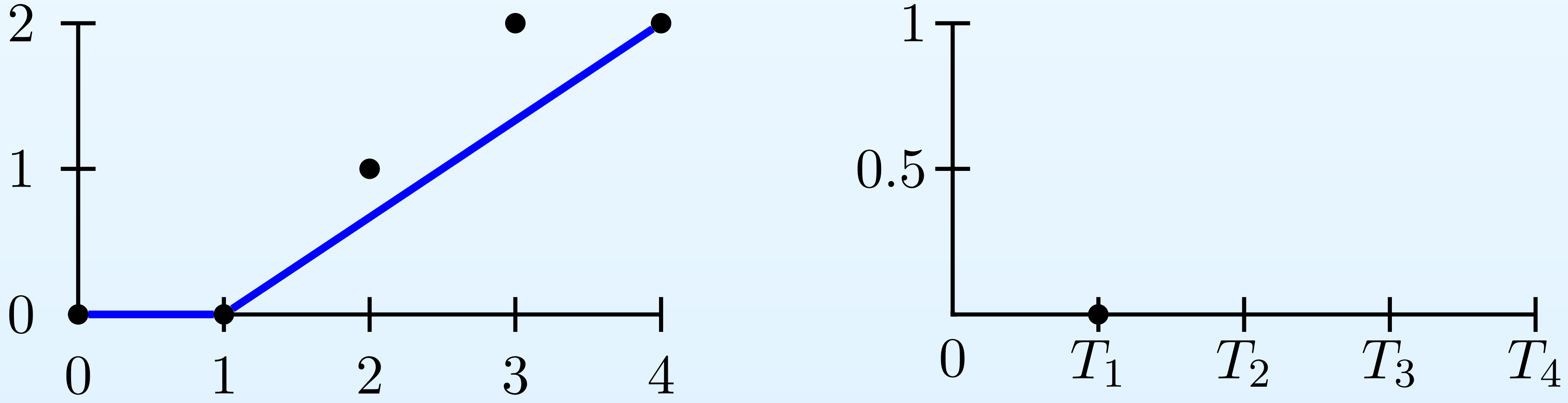


Convex minorant characterizations (1)

- Assume for notational simplicity that the T_i are ordered and have no ties: $T_1 < T_2 < \cdots < T_n$
- The naive estimator $\tilde{F}_{nk}$ can be characterized as the left derivative of the greatest convex minorant of the points

$$\left(i, \sum_{j=1}^{i} \Delta_{kj}\right), \quad i = 0, \ldots, n$$

Example: $n = 4$ and $(\Delta_{k1}, \Delta_{k2}, \Delta_{k3}, \Delta_{k4}) = (0, 1, 1, 0)$

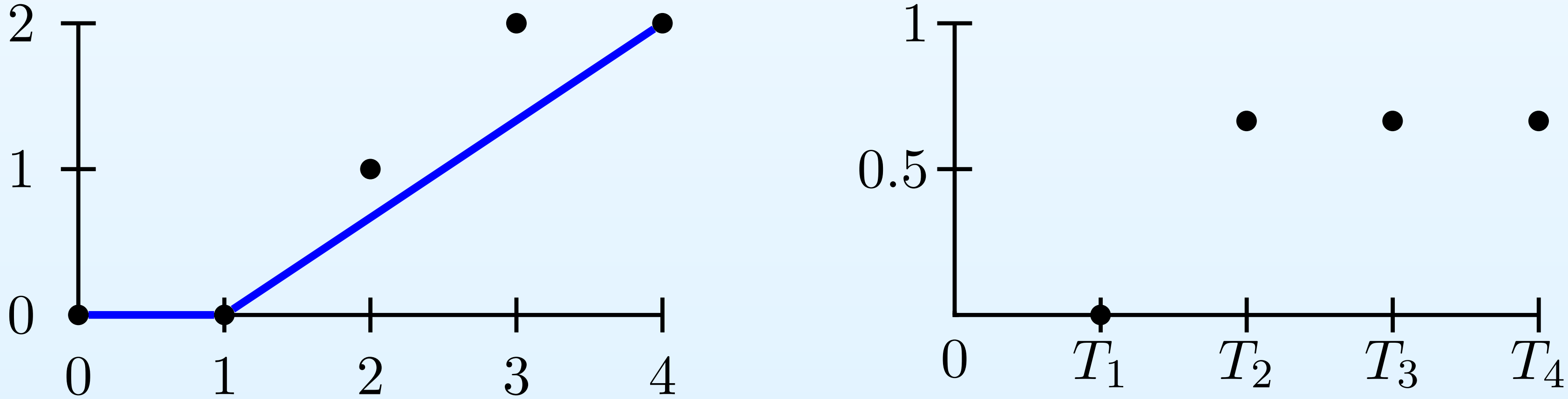


Convex minorant characterizations (1)

- Assume for notational simplicity that the T_i are ordered and have no ties: $T_1 < T_2 < \cdots < T_n$
- The naive estimator $\tilde{F}_{nk}$ can be characterized as the left derivative of the greatest convex minorant of the points

$$\left(i, \sum_{j=1}^{i} \Delta_{kj}\right), \quad i = 0, \ldots, n$$

Example: $n = 4$ and $(\Delta_{k1}, \Delta_{k2}, \Delta_{k3}, \Delta_{k4}) = (0, 1, 1, 0)$



Convex minorant characterizations (1)

- Assume for notational simplicity that the T_i are ordered and have no ties: $T_1 < T_2 < \cdots < T_n$
- The naive estimator $\tilde{F}_{nk}$ can be characterized as the left derivative of the greatest convex minorant of the points

$$\left(i, \sum_{j=1}^{i} \Delta_{kj}\right), \quad i = 0, \ldots, n$$

Convex minorant characterizations (1)

- Assume for notational simplicity that the T_i are ordered and have no ties: $T_1 < T_2 < \cdots < T_n$
- The naive estimator $\tilde{F}_{nk}$ can be characterized as the left derivative of the greatest convex minorant of the points

$$\left(i, \sum_{j=1}^{i} \Delta_{kj} \right), \quad i = 0, \ldots, n$$

- The MLE $\hat{F}_{nk}$ can be characterized as the left derivative of the greatest convex minorant of the points

$$\left(i + \sum_{j=1}^{i} \frac{\hat{F}_{n+}(T_j) - \Delta_{+j}}{1 - \hat{F}_{n+}(T_j)} + \lambda_n 1\{i = n\}, \sum_{j=1}^{i} \Delta_{kj} \right), \quad i = 0, \ldots, n$$

Convex minorant characterizations (1)

- Assume for notational simplicity that the T_i are ordered and have no ties: $T_1 < T_2 < \cdots < T_n$
- The naive estimator $\tilde{F}_{nk}$ can be characterized as the left derivative of the greatest convex minorant of the points

$$\left(i, \sum_{j=1}^{i} \Delta_{kj}\right), \quad i = 0, \dots, n$$

- The MLE $\hat{F}_{nk}$ can be characterized as the left derivative of the greatest convex minorant of the points

$$\left(i + \sum_{j=1}^{i} \frac{\hat{F}_{n+}(T_j) - \Delta_{+j}}{1 - \hat{F}_{n+}(T_j)} + \lambda_n 1\{i = n\}, \sum_{j=1}^{i} \Delta_{kj}\right), \quad i = 0, \dots, n$$

Convex minorant characterizations (2)

$$\left(i + \sum_{j=1}^{i} \frac{\hat{F}_{n+}(T_j) - \Delta_{+j}}{1 - \hat{F}_{n+}(T_j)} + \lambda_n 1\{i = n\}, \sum_{j=1}^{i} \Delta_{kj} \right), \quad i = 0, \ldots, n$$

Convex minorant characterizations (2)

$$\left(i + \sum_{j=1}^{i} \frac{\hat{F}_{n+}(T_j) - \Delta_{+j}}{1 - \hat{F}_{n+}(T_j)} + \lambda_n 1\{i = n\}, \sum_{j=1}^{i} \Delta_{kj} \right), \quad i = 0, \ldots, n$$

- Self induced system

Convex minorant characterizations (2)

$$\left(i + \sum_{j=1}^{i} \frac{\hat{F}_{n+}(T_j) - \Delta_{+j}}{1 - \hat{F}_{n+}(T_j)} + \lambda_n 1\{i = n\}, \sum_{j=1}^{i} \Delta_{kj} \right), \quad i = 0, \ldots, n$$

- Self induced system
- It shows the relation between $\hat{F}_{n+}$ and the components $\hat{F}_{nk}$

Convex minorant characterizations (2)

$$\left(i + \sum_{j=1}^{i} \frac{\hat{F}_{n+}(T_j) - \Delta_{+j}}{1 - \hat{F}_{n+}(T_j)} + \lambda_n 1\{i = n\}, \sum_{j=1}^{i} \Delta_{kj} \right), \quad i = 0, \ldots, n$$

- Self induced system
- It shows the relation between $\hat{F}_{n+}$ and the components $\hat{F}_{nk}$
- It gives insight in the relation between the MLE and the naive estimator

Convex minorant characterizations (2)

$$\left(i + \sum_{j=1}^{i} \frac{\hat{F}_{n+}(T_j) - \Delta_{+j}}{1 - \hat{F}_{n+}(T_j)} + \lambda_n 1\{i = n\}, \sum_{j=1}^{i} \Delta_{kj}\right), \quad i = 0, \ldots, n$$

- Self induced system
- It shows the relation between $\hat{F}_{n+}$ and the components $\hat{F}_{nk}$
- It gives insight in the relation between the MLE and the naive estimator
- Computation: iterative convex minorant algorithms

Convex minorant characterizations (2)

$$\left(i + \sum_{j=1}^{i} \frac{\hat{F}_{n+}(T_j) - \Delta_{+j}}{1 - \hat{F}_{n+}(T_j)} + \lambda_n 1\{i = n\}, \sum_{j=1}^{i} \Delta_{kj} \right), \quad i = 0, \ldots, n$$

- Self induced system
- It shows the relation between $\hat{F}_{n+}$ and the components $\hat{F}_{nk}$
- It gives insight in the relation between the MLE and the naive estimator
- Computation: iterative convex minorant algorithms
- Family of convex minorant characterizations & link with sequential quadratic programming

Simulation study: $K = 2$

Model:

$$X \sim \text{Exp}(3), \quad P(Y=1) = 1/3, \quad P(Y=2) = 2/3,$$
$$T \sim \text{Exp}(3/2)$$

1,000 simulations for sample sizes

$$n = 25;\ 250;\ 2{,}500;\ 25{,}000$$

Simulation study: $K = 2$

Model:

$$X \sim \text{Exp}(3), \quad P(Y=1) = 1/3, \quad P(Y=2) = 2/3,$$
$$T \sim \text{Exp}(3/2)$$

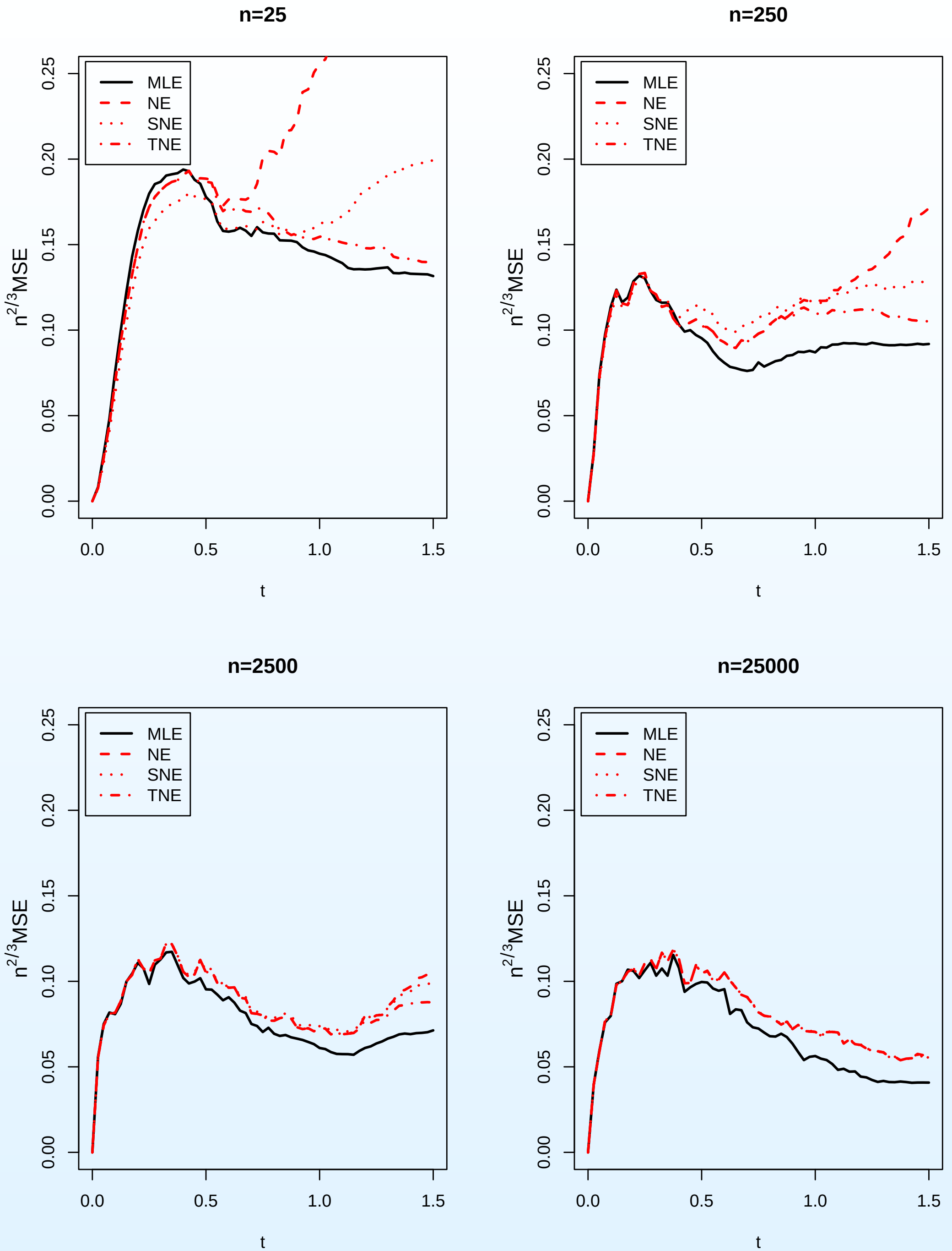
1,000 simulations for sample sizes

$$n = 25;\ 250;\ 2{,}500;\ 25{,}000$$

We compute the MLE, the naive estimator (NE), and two modified versions of the naive estimator:

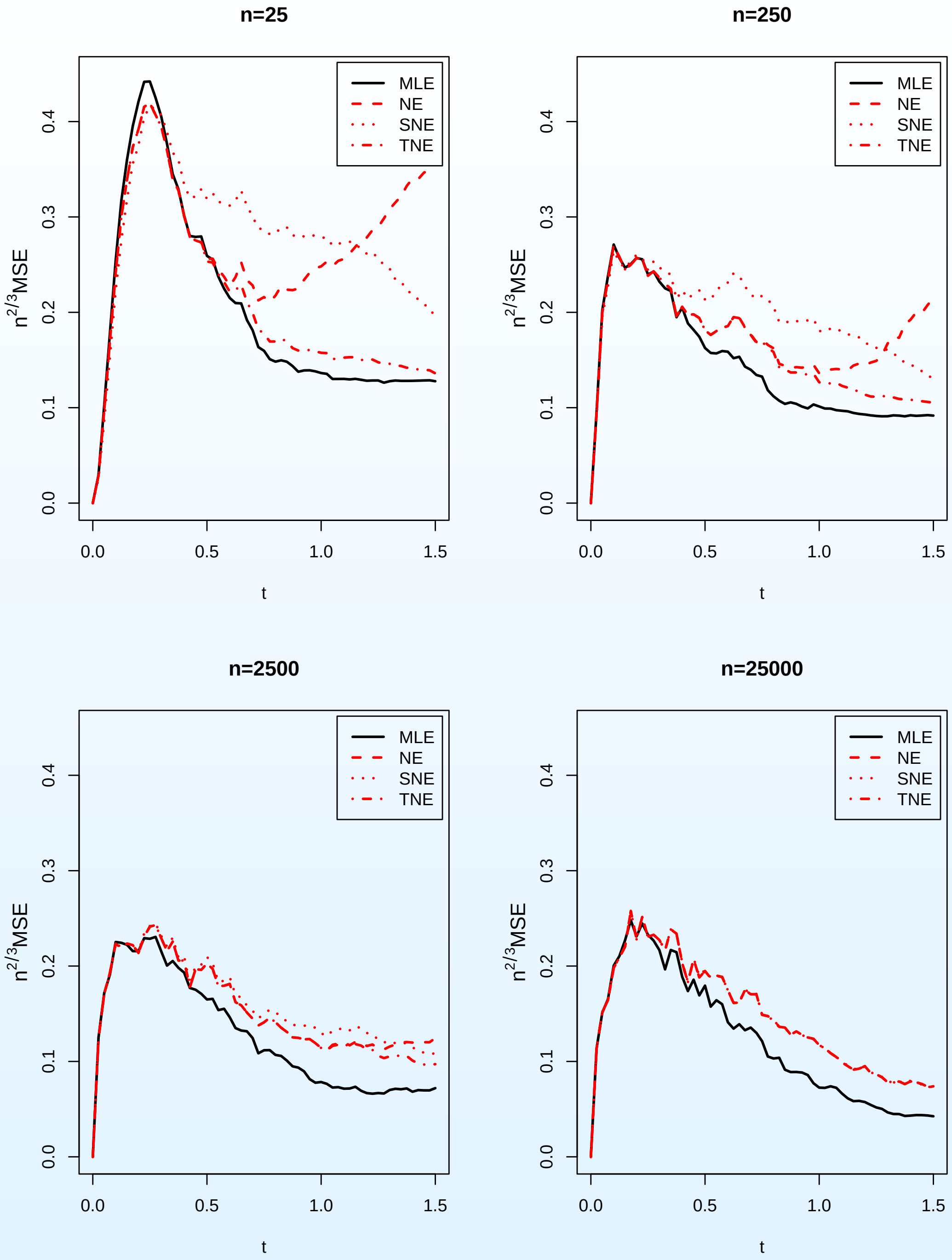
- truncated (TNE)
- scaled (SNE)

$k = 1$



n=25
n=250
n=2500
n=25000
MLE
NE
SNE
TNE
n^{2/3}MSE
t
0.00
0.05
0.10
0.15
0.20
0.25
0.0
0.5
1.0
1.5

$k = 2$



n=25
n=250
n=2500
n=25000
MLE
NE
SNE
TNE
$n^{2/3}$MSE
t
0.0
0.1
0.2
0.3
0.4
0.5
1.0
1.5

Summary

- We developed self-induced systems of convex minorant characterizations for the MLE:
  - give insight in the relation between MLE and naive estimator
  - can be used for computational purposes

marloes@stat.washington.edu

http://www.washington.edu/marloes

marloes@stat.washington.edu

http://www.washington.edu/marloes

Summary

- We developed self-induced systems of convex minorant characterizations for the MLE:
  - give insight in the relation between MLE and naive estimator
  - can be used for computational purposes
- Simulation study
  - The MSE of the MLE tends to be smaller than that of the naive estimator
  - The MLE and the naive estimator do not seem asymptotically equivalent

marloes@stat.washington.edu

http://www.washington.edu/marloes

marloes@stat.washington.edu

http://www.washington.edu/marloes

Summary

- We developed self-induced systems of convex minorant characterizations for the MLE:
  - give insight in the relation between MLE and naive estimator
  - can be used for computational purposes
- Simulation study
  - The MSE of the MLE tends to be smaller than that of the naive estimator
  - The MLE and the naive estimator do not seem asymptotically equivalent

marloes@stat.washington.edu

http://www.washington.edu/marloes

marloes@stat.washington.edu

http://www.washington.edu/marloes