Maximum likelihood estimation for double-truncation data under a special exponential family

Advisor: Takeshi Emura

Presenter: Ya-Hsuan Hu

6/23/2014

Graduate Institute of Statistics National Central University

Outline

- Introduction
- Methodology
- Theory
- Data analysis
- Conclusion

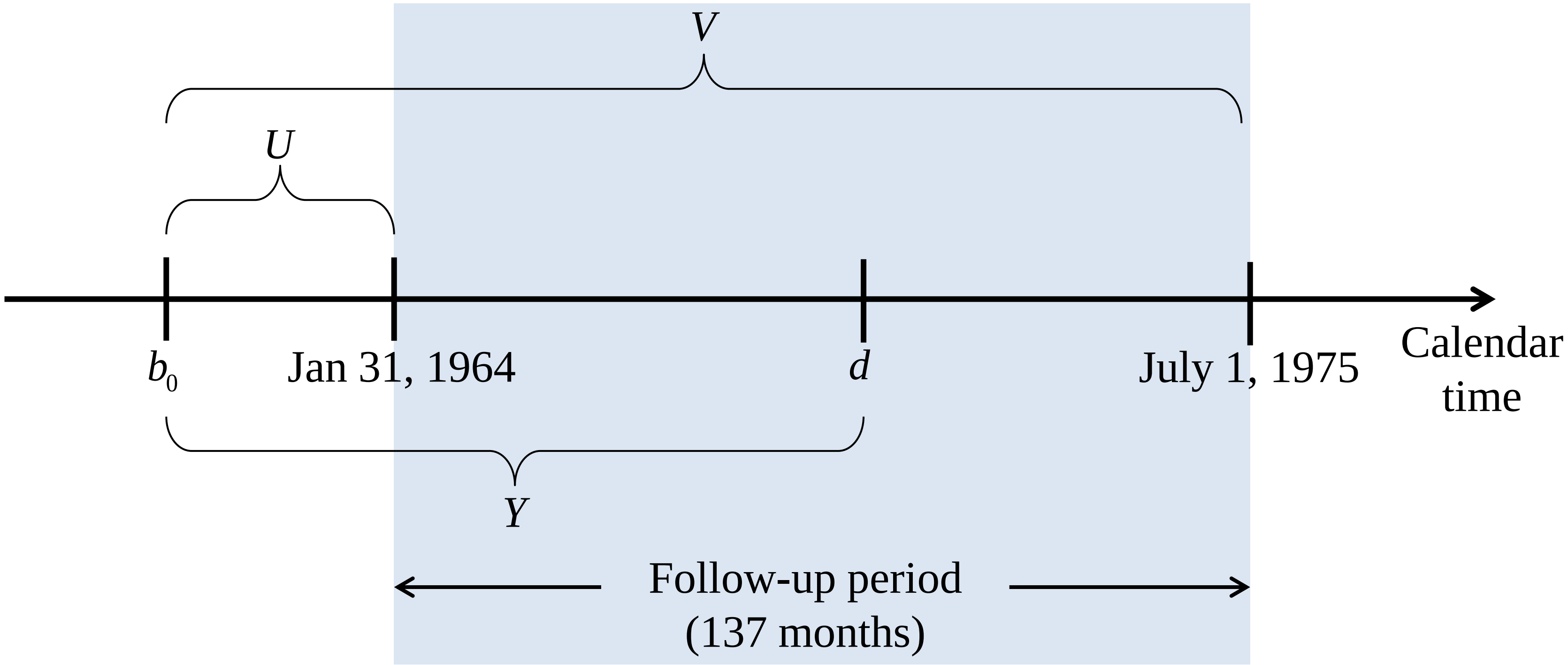
Introduction- What is double-truncation

For instance: Channing House study (Hyde, 1980)

b_0 : birth date d : date of death Y : age at death (in months) $\equiv d - b_0$

U : age on January 31, 1964 (in months)

$V = U + 137$: age on July 1, 1975 (in months)



Introduction- Double-truncation

Data: $\{ y_1, y_2, \cdots, y_n \}$ subject to $u_i \le y_i \le v_i$.

Target: Estimate of $f_Y(y) = \frac{d}{dy} P(Y \le y)$.

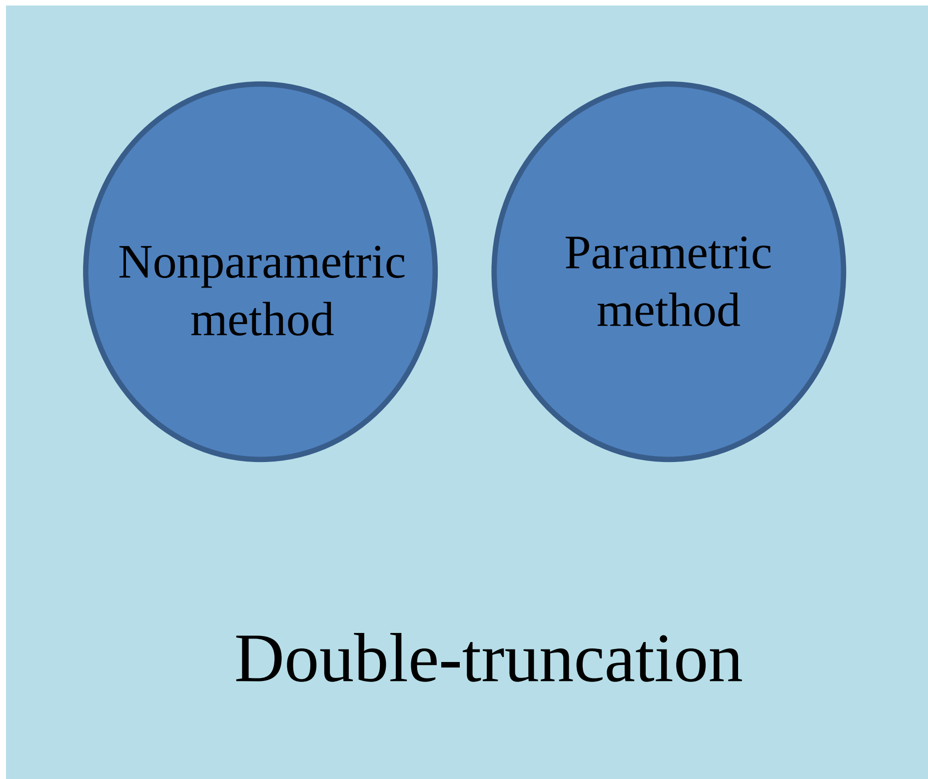
Example:

u_i = age at 31 Jan, 1964

y_i = age at death

$v_i = u_i + 137$ = age at 1 July, 1975

Introduction-Statistical inference



Introduction- nonparametric method

Nonparametric method:

Efron and Petrosian (1999)

- Proposed nonparametric maximum likelihood estimator (NPMLE).

Shen (2010)

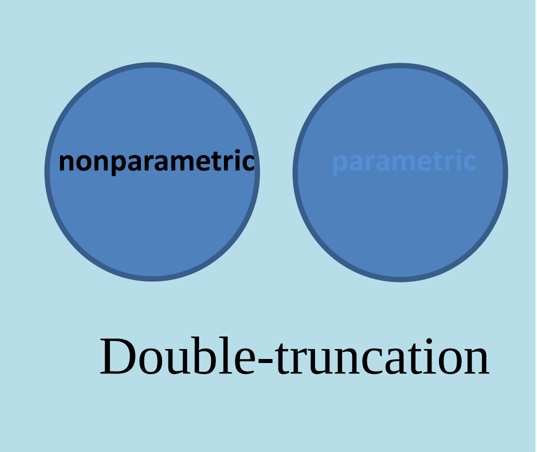
- Derived the uniform consistency and weak convergence of NPMLE

Moreira and Uña-Á lvarez (2010)

- Use bootstrap to construct the interval estimation

Moreira and Keliegom (2013)

- A kernel method to estimate the density function



Introduction- parametric method

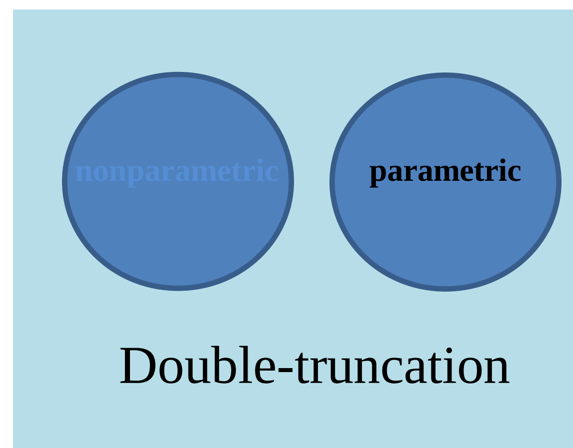
There are scarcer paper on parametric method under double-truncation, only for Efron and Petrosian (1999), Stovring and Wang (2007), Emura and Konno (2012).

Application of normal distribution approach under double-truncation.

Use of parametric approaches under doubly truncated data

Our research:

Efron and Petrosian (1999) proposed special exponential family (SEF) to do the estimation.



Model-Special exponential family

Assume that the lifetime variable Y follows a continuous distribution with a density function

$$f_{\boldsymbol{\eta}}(y) = \exp\{\, \boldsymbol{\eta}^{\mathrm{T}} \cdot \mathbf{t}(y) - \phi(\,\boldsymbol{\eta}\,)\,\}, \quad y \in \mathcal{Y}.$$

- $\mathcal{Y} \subset \Re$ is the support of Y
- $\mathbf{t}(\,y\,) = (\,y, y^2, \cdots, y^k\,)^{\mathrm{T}}$
- $\boldsymbol{\eta} = (\eta_1, \eta_2, \cdots, \eta_k\,)^{\mathrm{T}} \in \Theta \subset \Re^{\mathrm{k}}$
- $\phi(\,\boldsymbol{\eta}\,) = \log[\,\int_{\mathcal{Y}} \exp\{\,\boldsymbol{\eta}^{\mathrm{T}} \cdot \mathbf{t}(\,y\,)\,\}\, dy\,]$

Model-Special exponential family

Assume that the lifetime variable Y follows a continuous distribution with a density function

$$f_{\boldsymbol{\eta}}(y) = \exp\{\, \boldsymbol{\eta}^{\mathrm{T}} \cdot \mathbf{t}(y) - \phi(\boldsymbol{\eta}) \,\}, \quad y \in \mathcal{Y}.$$

- $\mathcal{Y} \subset \Re$ is the support of Y
- $\mathbf{t}(y) = (y, y^2, \cdots, y^k)^{\mathrm{T}}$
- $\boldsymbol{\eta} = (\eta_1, \eta_2, \cdots, \eta_k)^{\mathrm{T}} \in \Theta \subset \Re^{\mathrm{k}}$
- $\phi(\boldsymbol{\eta}) = \log[\int_{\mathcal{Y}} \exp\{\, \boldsymbol{\eta}^{\mathrm{T}} \cdot \mathbf{t}(y) \,\}\, dy \,]$

In our research, we consider $k = 1, 2, 3$.

Methodology

Our purpose:

Model: SEF

Data: doubly truncated data

Objective: Find MLE of parameters

Method:
-Newton-Raphson method (NR) (Efron and Petrosian, 1999)
-Fixed-point iteration method (FPI) (Burden and Faires, 2011)

Methodology

What is the density $f(y)$ change when the samples $y_1, y_2, \cdots, y_n$ suffer from double-truncation?

Assume that the truncation interval $R_i = [u_i, v_i]$, then the truncated density is

$$f(y_i \mid y_i \in R_i) = \begin{cases} f(y_i)/F_i & \text{if } y_i \in R_i, \\ 0 & \text{if } y_i \notin R_i. \end{cases}$$

where

$$F_i = \int_{u_i}^{v_i} f(y)\,dy$$

Methodology-Newton-Raphson algorithm

Step 1: Choose the initial value $\boldsymbol{\eta}^{(0)} = (\eta_1^{(0)}, \eta_2^{(0)}, \cdots, \eta_k^{(0)})$.

Step 2: Set

$$\boldsymbol{\eta}^{(p+1)} = \boldsymbol{\eta}^{(p)} - \left[\frac{\partial^2}{\partial \boldsymbol{\eta}^2} \ell(\boldsymbol{\eta}^{(p)}) \right]^{-1} \frac{\partial}{\partial \boldsymbol{\eta}} \ell(\boldsymbol{\eta}^{(p)}) ,\ p = 0, 1, 2, \cdots$$

If $|\eta_i^{(p+1)} - \eta_i^{(p)}| < 10^{-4}, i = 1, 2, \cdots, k$ stop the algorithm, and set $\hat{\boldsymbol{\eta}} = \boldsymbol{\eta}^{(p+1)}$.

$\ell(\boldsymbol{\eta})$ = log-likelihood function of $\boldsymbol{\eta}$

Methodology- Fixed-point iteration

Objective: Solve $S(\eta) \equiv \frac{\partial}{\partial \eta} \ell(\eta) = 0 \longrightarrow \eta = g(\eta)$

One dimensional case

Algorithm:

Step 1: Choose the initial value $\eta^{(0)}$.

Step 2: The iterative process $\eta^{(p+1)} = g(\eta^{(p)}), \; p = 0, 1, 2, \cdots$.

If $|\eta^{(p+1)} - \eta^{(p)}| < 10^{-4}$ stop the algorithm, and set $\hat{\eta} = \eta^{(p+1)}$.

- Chen (2009) also use the similar method for finding MLE

  (Weighted Breslow-type and maximum likelihood estimation in semiparametric transformation models)

Methodology- Fixed-point iteration

Objective: Solve $S(\eta_1,\eta_2)\equiv\frac{\partial}{\partial\boldsymbol{\eta}}\ell(\eta_1,\eta_2)=0 \longrightarrow \eta_1=g(\eta_1,\eta_2)$ and $\eta_2=q(\eta_1,\eta_2)$

Two dimensional case

Algorithm:

Step 1: Choose the initial value $\eta_1^{(0)}$ and $\eta_2^{(0)}$.

Step 2: The iterative process $\eta_1^{(p+1)}=g(\eta_1^{(p)},\eta_2^{(p)})$ and $\eta_2^{(p+1)}=q(\eta_1^{(p)},\eta_2^{(p)})$, $p=0,1,2,\cdots$.

If $|\eta_i^{(p+1)}-\eta_i^{(p)}|<10^{-4}, i=1,2$ stop the algorithm, then $\eta_1^{(p+1)},\eta_2^{(p+1)}$ are the solutions.

Methodology- One-parameter SEF $(k=1)$

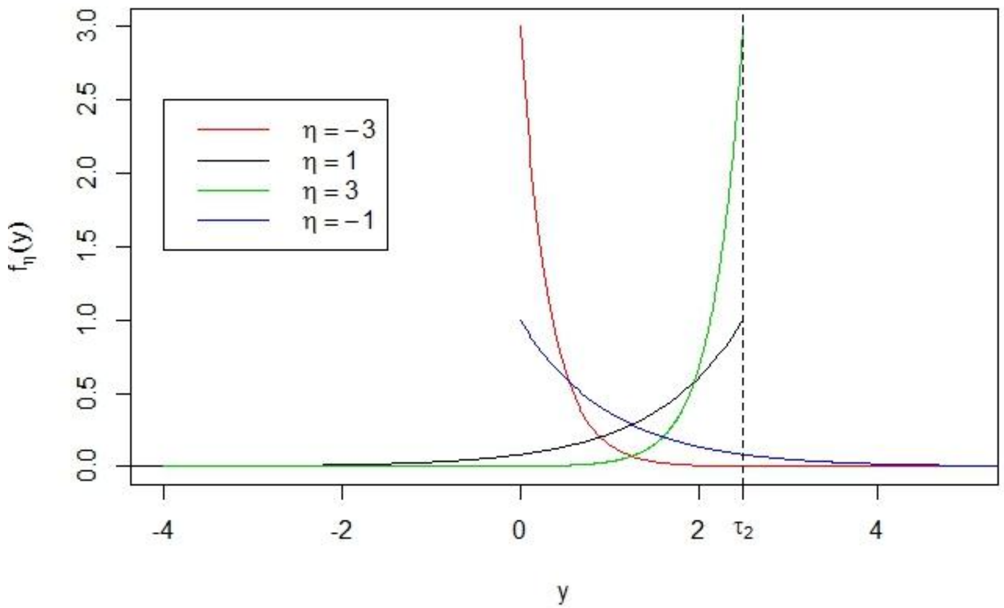
The density of one-parameter SEF is

- $f_\eta(y) = \eta \exp\{\eta(y-\tau_2)\}, \quad y \in \mathcal{Y} = (-\infty, \tau_2]$ → In application, $\tau_2 = y_{(n)}$

with parameter space $\Theta = \{\eta : \eta > 0\}$.

- $f_\eta(y) = -\eta \exp\{\eta(y-\tau_1)\}, \quad y \in \mathcal{Y} = [\tau_1, \infty)$, → In application, $\tau_1 = y_{(1)}$

with parameter space $\Theta = \{\eta : \eta < 0\}$



Methodology- One-parameter SEF $(k=1)$

The likelihood function for doubly truncated data when $\eta > 0$ is

$$L(\eta) = \prod_{i=1}^{n} \frac{f_\eta(y_i)}{F_i(\eta)} = \prod_{i=1}^{n} \left\{ \frac{\eta \exp(\eta y_i)}{\exp(\eta v_i) - \exp(\eta u_i)} \right\}^{\delta_i} \times \left\{ \frac{\eta \exp(\eta y_i)}{\exp(\eta \tau_2) - \exp(\eta u_i)} \right\}^{1-\delta_i}$$

where

$$\delta_i = \begin{cases} 1 & \text{if } v_i < \tau_2, \\ 0 & \text{if } v_i \geq \tau_2. \end{cases}$$

and

$$F_i(\eta) = \frac{\{\exp(\eta v_i) - \exp(\eta u_i)\}^{\delta_i} \{\exp(\eta \tau_2) - \exp(\eta u_i)\}^{1-\delta_i}}{\exp(\eta \tau_2)}$$

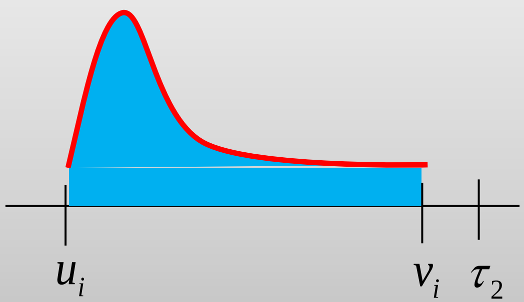
The log-likelihood function is

$$\ell(\eta) = n \log \eta + \eta \sum_{i=1}^{n} y_i - \sum_{i=1}^{n} \delta_i [\log\{\exp(\eta v_i) - \exp(\eta u_i)\}]$$
$$- \sum_{i=1}^{n} (1-\delta_i)[\log\{\exp(\eta \tau_2) - \exp(\eta u_i)\}].$$

Methodology- One-parameter SEF $(k=1)$

The likelihood function for doubly trun... 0 is

$L(\eta) = \ldots$



$$F_i = \int_{u_i}^{v_i} f(y)\,dy$$

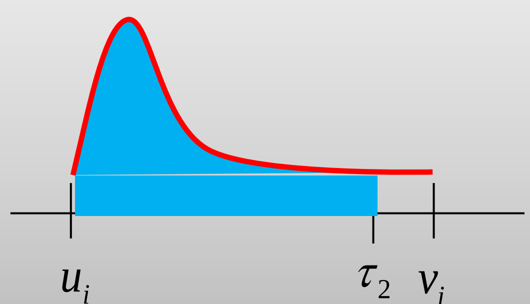
where

$$\delta_i = \begin{cases} 1 & \text{if } v_i < \tau_2, \\ 0 & \text{if } v_i \ge \tau_2. \end{cases}$$

and

$$F_i(\eta) = \frac{\{\exp(\eta v_i) - \exp(\eta u_i)\}^{\delta_i}\{\exp(\eta\tau_2) - \exp(\eta u_i)\}^{1-\delta_i}}{\exp(\eta\tau_2)}$$

The log-likel...



$$F_i = \int_{u_i}^{\tau_2} f(y)\,dy$$

$\ell \ldots \{\exp(\eta v \ldots$

$\sum_{i=1} \ldots {}_2) - \exp(\eta \ldots$

Methodology- One-parameter SEF $(k=1)$

The first-order derivative of the log-likelihood function is

$$\frac{\partial}{\partial \eta}\ell(\eta)=\frac{n}{\eta}+\sum_{i=1}^{n}y_i-\sum_{i=1}^{n}\delta_i\left\{\frac{v_i\exp(\eta v_i)-u_i\exp(\eta u_i)}{\exp(\eta v_i)-\exp(\eta u_i)}\right\}$$
$$-\sum_{i=1}^{n}(1-\delta_i)\left\{\frac{\tau_2\exp(\eta\tau_2)-u_i\exp(\eta u_i)}{\exp(\eta\tau_2)-\exp(\eta u_i)}\right\}.$$

The second-order derivative of the log-likelihood function is

$$\frac{\partial^2}{\partial \eta^2}\ell(\eta)=\frac{-n}{\eta^2}-\sum_{i=1}^{n}\delta_i\left[\frac{v_i^2\exp(\eta v_i)-u_i^2\exp(\eta u_i)}{\exp(\eta v_i)-\exp(\eta u_i)}-\left\{\frac{v_i\exp(\eta v_i)-u_i\exp(\eta u_i)}{\exp(\eta v_i)-\exp(\eta u_i)}\right\}^2\right]$$
$$-\sum_{i=1}^{n}(1-\delta_i)\left[\frac{\tau^2\exp(\eta\tau_2)-u_i^2\exp(\eta u_i)}{\exp(\eta\tau_2)-\exp(\eta u_i)}-\left\{\frac{\tau_2\exp(\eta\tau_2)-u_i\exp(\eta u_i)}{\exp(\eta\tau_2)-\exp(\eta u_i)}\right\}^2\right].$$

Methodology- Two-parameter SEF $(k=2)$

The density of two-parameter SEF is

$$f_{\boldsymbol{\eta}}(y) = \exp\left\{ \eta_1 y + \eta_2 y^2 + \frac{\eta_1^2}{4\eta_2} - \log\left(\sqrt{\frac{-\pi}{\eta_2}} \right) \right\}, \quad y \in \mathcal{Y} = (-\infty, \infty)$$

with parameter space $\Theta = \{ (\eta_1, \eta_2) : \eta_1 \in \Re, \eta_2 < 0 \}$.

Reparameterization

Setting $\mu = -\eta_1 / 2\eta_2$ and $\sigma^2 = -1/2\eta_2$, this produces a normal distribution. (Castillo,1994)

Methodology- Two-parameter SEF $(k=2)$

The likelihood function for doubly truncated data is

$$L(\boldsymbol{\eta})=\prod_{i=1}^{n}\frac{f_{\boldsymbol{\eta}}(y_i)}{F_i(\boldsymbol{\eta})}=\frac{\exp\{\sum_{i=1}^{n}(\eta_1 y_i+\eta_2 y_i^2)\}}{\exp(-\frac{n\eta_1^2}{4\eta_2})(\sqrt{\frac{-\pi}{\eta_2}})^n\prod_{i=1}^{n}\left\{\Phi\left(\frac{v_i+\frac{\eta_1}{2\eta_2}}{\sqrt{\frac{-1}{2\eta_2}}}\right)-\Phi\left(\frac{u_i+\frac{\eta_1}{2\eta_2}}{\sqrt{\frac{-1}{2\eta_2}}}\right)\right\}}.$$

where

$$F_i(\boldsymbol{\eta})=\int_{u_i}^{v_i}\exp\left\{\eta_1 y+\eta_2 y^2+\frac{\eta_1^2}{4\eta_2}-\log(\sqrt{\frac{-\pi}{\eta_2}})\right\}dy=\Phi\left(\frac{v_i+\frac{\eta_1}{2\eta_2}}{\sqrt{\frac{-1}{2\eta_2}}}\right)-\Phi\left(\frac{u_i+\frac{\eta_1}{2\eta_2}}{\sqrt{\frac{-1}{2\eta_2}}}\right)$$

Methodology- Two-parameter SEF $(k=2)$

Define notations

$$H_{i1}(\eta_1,\eta_2)=\frac{\phi\left(\frac{v_i+\frac{\eta_1}{2\eta_2}}{\sqrt{\frac{-1}{2\eta_2}}}\right)}{\Phi\left(\frac{v_i+\frac{\eta_1}{2\eta_2}}{\sqrt{\frac{-1}{2\eta_2}}}\right)-\Phi\left(\frac{u_i+\frac{\eta_1}{2\eta_2}}{\sqrt{\frac{-1}{2\eta_2}}}\right)},\quad H_{i2}(\eta_1,\eta_2)=\frac{\phi\left(\frac{u_i+\frac{\eta_1}{2\eta_2}}{\sqrt{\frac{-1}{2\eta_2}}}\right)}{\Phi\left(\frac{v_i+\frac{\eta_1}{2\eta_2}}{\sqrt{\frac{-1}{2\eta_2}}}\right)-\Phi\left(\frac{u_i+\frac{\eta_1}{2\eta_2}}{\sqrt{\frac{-1}{2\eta_2}}}\right)}$$

These are the hazard function of the normal distribution with doubly truncated samples (Sankaran and Sunoj, 2004)

Methodology- Two-parameter SEF $(k=2)$

The first-order derivative of the log-likelihood function is

$$\frac{\partial}{\partial \eta_1}\ell(\boldsymbol{\eta}) = \sum_{i=1}^{n} y_i + \frac{n\eta_1}{2\eta_2} + \frac{1}{\sqrt{-2\eta_2}}\left\{\sum_{i=1}^{n} H_{i1}(\eta_1,\eta_2) - \sum_{i=1}^{n} H_{i2}(\eta_1,\eta_2)\right\}$$

$$\frac{\partial}{\partial \eta_2}\ell(\boldsymbol{\eta}) = \sum_{i=1}^{n} y_i^2 - \frac{n\eta_1^2}{4\eta_2^2} + \frac{n}{2\eta_2} - \sum_{i=1}^{n}\left\{ H_{i1}(\eta_1,\eta_2)\cdot\left(\frac{-v_i}{\sqrt{-2\eta_2}} - \eta_1\frac{\sqrt{-2\eta_2}}{4\eta_2^2}\right)\right\}$$
$$+\sum_{i=1}^{n}\left\{ H_{i2}(\eta_1,\eta_2)\cdot\left(\frac{-u_i}{\sqrt{-2\eta_2}} - \eta_1\frac{\sqrt{-2\eta_2}}{4\eta_2^2}\right)\right\}$$

Methodology- Two-parameter SEF $(k=2)$

The second-order derivative of the log-likelihood function is

$$\frac{\partial^2}{\partial \eta_1^2}\ell(\boldsymbol{\eta}) = \frac{n}{2\eta_2} - \frac{1}{2\eta_2}\sum_{i=1}^{n}\left\{ H_{i1}(\eta_1,\eta_2)\left(\frac{v_i + \frac{\eta_1}{2\eta_2}}{\sqrt{\frac{-1}{2\eta_2}}}\right) - H_{i2}(\eta_1,\eta_2)\left(\frac{u_i + \frac{\eta_1}{2\eta_2}}{\sqrt{\frac{-1}{2\eta_2}}}\right)\right\}$$

$$\frac{\partial^2}{\partial \eta_2^2}\ell(\boldsymbol{\eta}) = -\frac{1}{2\eta_2}\sum_{i=1}^{n}\left\{ H_{i1}(\eta_1,\eta_2) - H_{i2}(\eta_1,\eta_2)\right\}^2,$$

$$\sum_{i=1}^{n}\left\{ H_{i1}(\eta_1,\eta_2)\left(\frac{-v_i}{\sqrt{-2\eta_2}} - \eta_1\frac{\sqrt{-2\eta_2}}{4\eta_2^2}\right) - H_{i2}(\eta_1,\eta_2)\left(\frac{-u_i}{\sqrt{-2\eta_2}} - \eta_1\frac{\sqrt{-2\eta_2}}{4\eta_2^2}\right)\right\}^2$$

$$+\sum_{i=1}^{n} H_{i1}(\eta_1,\eta_2)\left\{\left(\frac{v_i + \frac{\eta_1}{2\eta_2}}{\sqrt{\frac{-1}{2\eta_2}}}\right)\left(\frac{-v_i}{\sqrt{-2\eta_2}} - \eta_1\frac{\sqrt{-2\eta_2}}{4\eta_2^2}\right)^2 - \left(\frac{-v_i}{\sqrt{(-2\eta_2)^3}} - \frac{3\eta_1}{\sqrt{(-2\eta_2)^5}}\right)\right\}$$

$$-\sum_{i=1}^{n} H_{i2}(\eta_1,\eta_2)\left\{\left(\frac{u_i + \frac{\eta_1}{2\eta_2}}{\sqrt{\frac{-1}{2\eta_2}}}\right)\left(\frac{-u_i}{\sqrt{-2\eta_2}} - \eta_1\frac{\sqrt{-2\eta_2}}{4\eta_2^2}\right)^2 - \left(\frac{-u_i}{\sqrt{(-2\eta_2)^3}} - \frac{3\eta_1}{\sqrt{(-2\eta_2)^5}}\right)\right\}$$

$$+\frac{n\eta_1^2}{2\eta_2^3} - \frac{n}{2\eta_2^2},$$

Methodology- Two-parameter SEF $(k=2)$

$$\frac{\partial^2}{\partial\eta_1\partial\eta_2}\ell(\boldsymbol{\eta})=$$

$$-\frac{n\eta_1}{2\eta_2^2}+(-2\eta_2)^{\frac{-3}{2}}\sum_{i=1}^{n}\{H_{i1}(\eta_1,\eta_2)-H_{i2}(\eta_1,\eta_2)\}$$

$$-\frac{1}{\sqrt{-2\eta_2}}\sum_{i=1}^{n}\left[H_{i1}(\eta_1,\eta_2)\left(\frac{v_i+\frac{\eta_1}{2\eta_2}}{\sqrt{\frac{-1}{2\eta_2}}}\right)\left\{\frac{-v_i}{\sqrt{-2\eta_2}}-(-2\eta_2)^{\frac{-3}{2}}\eta_1\right\}\right]$$

$$+\frac{1}{\sqrt{-2\eta_2}}\sum_{i=1}^{n}\left[H_{i2}(\eta_1,\eta_2)\left(\frac{u_i+\frac{\eta_1}{2\eta_2}}{\sqrt{\frac{-1}{2\eta_2}}}\right)\left\{\frac{-u_i}{\sqrt{-2\eta_2}}-(-2\eta_2)^{\frac{-3}{2}}\eta_1\right\}\right]$$

$$-\frac{1}{\sqrt{-2\eta_2}}\sum_{i=1}^{n}\left[\{H_{i1}(\eta_1,\eta_2)-H_{i2}(\eta_1,\eta_2)\}\left[\begin{matrix}H_{i1}(\eta_1,\eta_2)\left\{\frac{-v_i}{\sqrt{-2\eta_2}}-(-2\eta_2)^{\frac{-3}{2}}\eta_1\right\}\\-H_{i2}(\eta_1,\eta_2)\left\{\frac{-u_i}{\sqrt{-2\eta_2}}-(-2\eta_2)^{\frac{-3}{2}}\eta_1\right\}\end{matrix}\right]\right].$$

Methodology- Two-parameter SEF $(k = 2)$

We encounter the problem in FPI method !

- We encounter that the cases that the algorithm diverges when we directly estimate (η_1, η_2).
- Use reparameterization of two-parameter SEF proposed by Castillo (1994), that is $\mu = -\eta_1 / 2\eta_2$ and $\sigma^2 = -1/2\eta_2$.

Methodology- Two-parameter SEF $(k=2)$

Rewrite the first-order derivative of log-likelihood function as

$$-\frac{\eta_1}{2\eta_2} = \bar{y} + \frac{1}{n\sqrt{-2\eta_2}} \sum_{i=1}^{n} \{ H_{i1}(\eta_1, \eta_2) - H_{i2}(\eta_1, \eta_2) \}$$

$$-\frac{1}{2\eta_2} = \overline{y^2} - \frac{\eta_1^2}{4\eta_2^2} - \frac{1}{n}\sum_{i=1}^{n}\left\{ H_{i1}(\eta_1, \eta_2)\left(\frac{-v_i}{\sqrt{-2\eta_2}} - \eta_1 \frac{\sqrt{-2\eta_2}}{4\eta_2^2} \right) \right\} + \frac{1}{n}\sum_{i=1}^{n}\left\{ H_{i2}(\eta_1, \eta_2)\left(\frac{-u_i}{\sqrt{-2\eta_2}} - \eta_1 \frac{\sqrt{-2\eta_2}}{4\eta_2^2} \right) \right\}.$$

Methodology- Two-parameter SEF $(k=2)$

Use another setting can rewrite the above equation and by fixed-point iteration to fine the MLE of μ and σ^2 .

$$\mu = \bar{y} + \frac{1}{\sqrt{-2\eta_2}} \frac{1}{n} \sum_{i=1}^{n} \{ H_{i1}(\eta_1, \eta_2) - H_{i2}(\eta_1, \eta_2) \}$$

$$\sigma^2 = \overline{y^2} - \mu^2 - \frac{1}{n} \sum_{i=1}^{n} \left\{ H_{i1}(\eta_1, \eta_2) \left(\frac{-v_i}{\sqrt{-2\eta_2}} - \eta_1 \frac{\sqrt{-2\eta_2}}{4\eta_2^2} \right) \right\} + \frac{1}{n} \sum_{i=1}^{n} \left\{ H_{i2}(\eta_1, \eta_2) \left(\frac{-u_i}{\sqrt{-2\eta_2}} - \eta_1 \frac{\sqrt{-2\eta_2}}{4\eta_2^2} \right) \right\}.$$

Methodology- Two-parameter SEF $(k = 2)$

By the invariance of MLEs, the MLE of $(\eta_1, \eta_2)^T$ is

$$\begin{bmatrix} \eta_1^{(p+1)} \\ \eta_2^{(p+1)} \end{bmatrix} = \begin{bmatrix} \mu^{(p+1)} / \sigma^{2(p+1)} \\ -1/2\sigma^{2(p+1)} \end{bmatrix}, \; p = 0, 1, 2, \cdots$$

The iteration continuous until convergence, i.e., until

$$|\eta_i^{(p+1)} - \eta_i^{(p)}| < 10^{-4}, \quad i = 1, 2$$

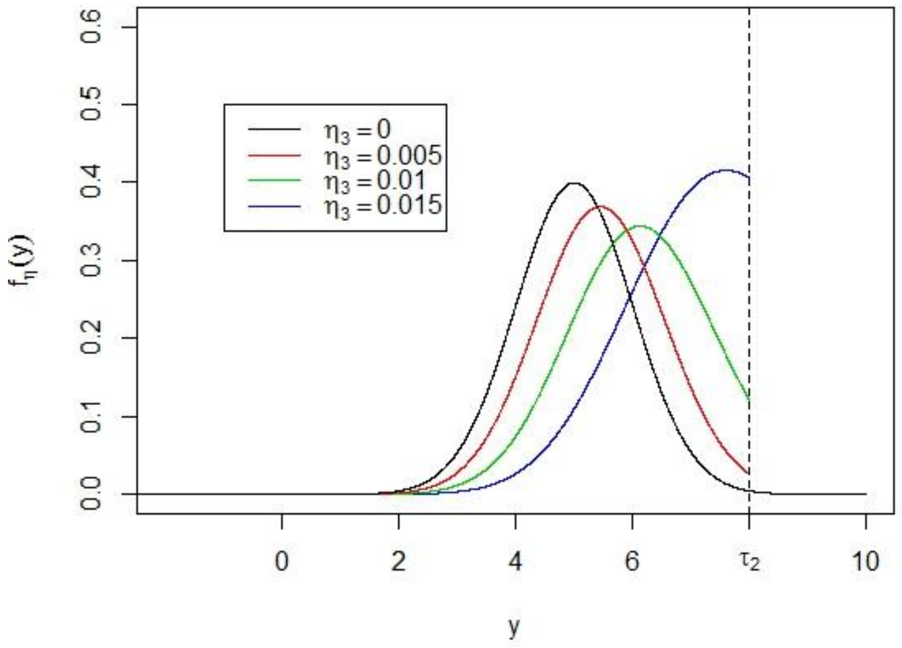
Methodology- Cubic SEF $(k=3)$

The density of cubic SEF is

In application, $y_{(n)} \le \tau_2 \le \sup_i v_i$

$$f_{\boldsymbol{\eta}}(y) = \exp[\,\eta_1 y + \eta_2 y^2 + \eta_3 y^3 - \phi(\boldsymbol{\eta})\,], \quad y \in \mathcal{Y} = (-\infty, \tau_2\,],$$

with parameter space $\Theta = \{\, (\eta_1, \eta_2, \eta_3) : \eta_1 \in \Re, \eta_2 \in \Re, \ \eta_3 > 0 \,\}$.

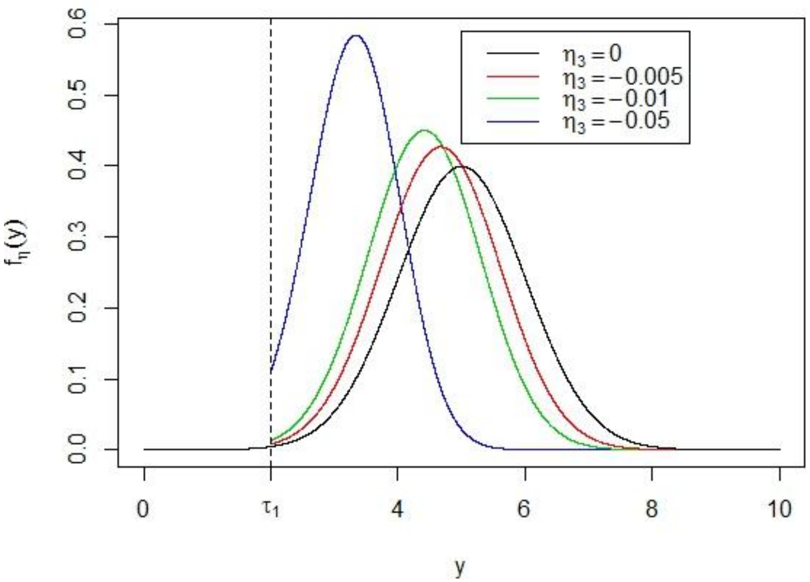


Methodology- Cubic SEF $(k=3)$

The density of cubic SEF is → In application, $\inf_i u_i \le \tau_1 \le y_{(1)}$

$$f_{\boldsymbol{\eta}}(y) = \exp[\,\eta_1 y + \eta_2 y^2 + \eta_3 y^3 - \phi(\boldsymbol{\eta})\,], \quad y \in \mathcal{Y} = [\,\tau_1, \infty\,),$$

with parameter space $\Theta = \{\,(\eta_1, \eta_2, \eta_3) : \eta_1 \in \Re, \eta_2 \in \Re, \ \eta_3 < 0\,\}.$



Methodology- Cubic SEF $(k=3)$

The likelihood function for doubly truncated data is

$$L(\boldsymbol{\eta}) = \prod_{i=1}^{n} \frac{f_{\boldsymbol{\eta}}(y_i)}{F_i(\boldsymbol{\eta})} = \frac{\exp\{\sum_{i=1}^{n}(\eta_1 y_i + \eta_2 y_i^2 + \eta_3 y_i^3)\}}{\prod_{i=1}^{n}\int_{u_i}^{v_i} \exp(\eta_1 y + \eta_2 y^2 + \eta_3 y^3)\, dy_i}$$

where

$$F_i(\boldsymbol{\eta}) = \int_{u_i}^{v_i} f_{\boldsymbol{\eta}}(y)\, dy = \frac{\int_{u_i}^{v_i} \exp(\eta_1 y + \eta_2 y^2 + \eta_3 y^3)\, dy}{\int_{y} \exp(\eta_1 y + \eta_2 y^2 + \eta_3 y^3)\, dy}$$

The log-likelihood function is

$$\ell(\boldsymbol{\eta}) = \log L(\boldsymbol{\eta}) = \sum_{i=1}^{n}(\eta_1 y_i + \eta_2 y_i^2 + \eta_3 y_i^3) - \sum_{i=1}^{n} \log\{\int_{u_i}^{v_i} \exp(\eta_1 y + \eta_2 y^2 + \eta_3 y^3)\, dy\}$$

Methodology- Cubic SEF $(k=3)$

Define notation

$$E_i^k(\boldsymbol{\eta}) = \int_{u_i}^{v_i} y^k \exp(\eta_1 y + \eta_2 y^2 + \eta_3 y^3)\, dy, \quad k = 0, 1, \cdots, 6$$

The first-order derivative of the log-likelihood function is

$$\frac{\partial}{\partial \eta_1} \ell(\boldsymbol{\eta}) = \sum_{i=1}^{n} \{ y_i - E_i^1(\boldsymbol{\eta}) / E_i^0(\boldsymbol{\eta}) \}$$

$$\frac{\partial}{\partial \eta_2} \ell(\boldsymbol{\eta}) = \sum_{i=1}^{n} \{ y_i^2 - E_i^2(\boldsymbol{\eta}) / E_i^0(\boldsymbol{\eta}) \}$$

$$\frac{\partial}{\partial \eta_3} \ell(\boldsymbol{\eta}) = \sum_{i=1}^{n} \{ y_i^3 - E_i^3(\boldsymbol{\eta}) / E_i^0(\boldsymbol{\eta}) \}$$

Methodology- Cubic SEF $(k=3)$

The second-order derivative of the log-likelihood function is

$$\frac{\partial^2}{\partial \eta_1^2}\ell(\boldsymbol{\eta})=\sum_{i=1}^{n}[-E_i^2(\boldsymbol{\eta})/E_i^0(\boldsymbol{\eta})+\{E_i^1(\boldsymbol{\eta})/E_i^0(\boldsymbol{\eta})\}^2]$$

$$\frac{\partial^2}{\partial \eta_2^2}\ell(\boldsymbol{\eta})=\sum_{i=1}^{n}[-E_i^4(\boldsymbol{\eta})/E_i^0(\boldsymbol{\eta})+\{E_i^2(\boldsymbol{\eta})/E_i^0(\boldsymbol{\eta})\}^2]$$

$$\frac{\partial^2}{\partial \eta_3^2}\ell(\boldsymbol{\eta})=\sum_{i=1}^{n}[-E_i^6(\boldsymbol{\eta})/E_i^0(\boldsymbol{\eta})+\{E_i^3(\boldsymbol{\eta})/E_i^0(\boldsymbol{\eta})\}^2]$$

$$\frac{\partial^2}{\partial \eta_2 \partial \eta_1}\ell(\boldsymbol{\eta})=\sum_{i=1}^{n}[-E_i^3(\boldsymbol{\eta})/E_i^0(\boldsymbol{\eta})+\{E_i^2(\boldsymbol{\eta})/E_i^0(\boldsymbol{\eta})\}\{E_i^1(\boldsymbol{\eta})/E_i^0(\boldsymbol{\eta})\}]$$

$$\frac{\partial^2}{\partial \eta_3 \partial \eta_1}\ell(\boldsymbol{\eta})=\sum_{i=1}^{n}[-E_i^4(\boldsymbol{\eta})/E_i^0(\boldsymbol{\eta})+\{E_i^3(\boldsymbol{\eta})/E_i^0(\boldsymbol{\eta})\}\{E_i^1(\boldsymbol{\eta})/E_i^0(\boldsymbol{\eta})\}]$$

$$\frac{\partial^2}{\partial \eta_3 \partial \eta_2}\ell(\boldsymbol{\eta})=\sum_{i=1}^{n}[-E_i^5(\boldsymbol{\eta})/E_i^0(\boldsymbol{\eta})+\{E_i^3(\boldsymbol{\eta})/E_i^0(\boldsymbol{\eta})\}\{E_i^2(\boldsymbol{\eta})/E_i^0(\boldsymbol{\eta})\}]$$

Methodology- Cubic SEF $(k=3)$

We encounter the problem in NR method !

We encounter that the cases that the algorithm diverges
→ We need to adjust the initial value, and we proposed the Randomized Newton-Raphson algorithm (RNR).

Methodology- Cubic SEF $(k=3)$

Randomized Newton-Raphson Algorithm:

Step 1: Choose the initial value $\boldsymbol{\eta}=(\eta_1^{(0)},\eta_2^{(0)},\eta_3^{(0)})^{\mathrm{T}}$.

Step 2: Set

$$\begin{bmatrix}\eta_1^{(p+1)}\\ \eta_2^{(p+1)}\\ \eta_3^{(p+1)}\end{bmatrix}=\begin{bmatrix}\eta_1^{(p)}\\ \eta_2^{(p)}\\ \eta_3^{(p)}\end{bmatrix}-\left[\begin{matrix}\frac{\partial^2}{\partial\eta_1^2}\ell(\boldsymbol{\eta}) & \frac{\partial^2}{\partial\eta_1\partial\eta_2}\ell(\boldsymbol{\eta}) & \frac{\partial^2}{\partial\eta_1\partial\eta_3}\ell(\boldsymbol{\eta})\\ \frac{\partial^2}{\partial\eta_2\partial\eta_1}\ell(\boldsymbol{\eta}) & \frac{\partial^2}{\partial\eta_2^2}\ell(\boldsymbol{\eta}) & \frac{\partial^2}{\partial\eta_2\partial\eta_3}\ell(\boldsymbol{\eta})\\ \frac{\partial^2}{\partial\eta_3\partial\eta_1}\ell(\boldsymbol{\eta}) & \frac{\partial^2}{\partial\eta_3\partial\eta_2}\ell(\boldsymbol{\eta}) & \frac{\partial^2}{\partial\eta_3^2}\ell(\boldsymbol{\eta})\end{matrix}\right]^{-1}\Bigg|_{(\eta_1^{(p)},\eta_2^{(p)},\eta_3^{(p)})}\begin{bmatrix}\frac{\partial}{\partial\eta_1}\ell(\boldsymbol{\eta})\\ \frac{\partial}{\partial\eta_2}\ell(\boldsymbol{\eta})\\ \frac{\partial}{\partial\eta_3}\ell(\boldsymbol{\eta})\end{bmatrix}\Bigg|_{(\eta_1^{(p)},\eta_2^{(p)},\eta_3^{(p)})}$$

The iteration procedure then continuous until convergence, i.e., until $|\eta_i^{(p+1)}-\eta_i^{(p)}|<10^{-4}$ for $i=1,2,3$. Then $(\eta_1^{(p+1)},\eta_2^{(p+1)},\eta_3^{(p+1)})$ is target value.

Methodology- Cubic SEF $(k = 3)$

Step 3: If $|\eta_1^{(p+1)} - \eta_1^{(p)}| > 20$ or $|\eta_2^{(p+1)} - \eta_2^{(p)}| > 10$ or $|\eta_3^{(p+1)} - \eta_3^{(p)}| > 1$, replace $(\eta_1, \eta_2, \eta_3)^{\mathrm{T}}$ with $(\eta_1 + u_1, \eta_2 + u_2, \eta_3)^{\mathrm{T}}$, where $u_1 \sim U(-6, 6)$ and $u_2 \sim U(-0.5, 0.5)$, and return to Step 1.

Methodology-Simulation

- Inclusion probability

$$P(U \le Y \le V) = \iiint\limits_{u \le y \le v} f_Y(y) \cdot f_V(v) \cdot f_U(u)\,du\,dy\,dv$$

We set the condition that all the simulations conducted under

$$P(U \le Y \le V) \approx 0.5$$

Methodology-Simulation

One-parameter SEF with $\eta = 3$, 500 repetitions

	True	Initial value	Method	$E(\hat{\eta})$	$MSE(\hat{\eta})$	AI
$n=100$	$\eta=3$	$\eta^{(0)}=3$	FPI	3.0804	0.1880	12.61
			NR	3.0803	0.1881	4.39
		$\eta^{(0)}=\frac{1}{y_{(n)}-\bar{y}}$	FPI	3.0804	0.1881	12.22
			NR	3.0803	0.1881	4.29
$n=200$	$\eta=3$	$\eta^{(0)}=3$	FPI	3.0442	0.0917	12.10
			NR	3.0442	0.0917	4.20
		$\eta^{(0)}=\frac{1}{y_{(n)}-\bar{y}}$	FPI	3.0443	0.0917	11.84
			NR	3.0442	0.0917	4.1
$n=300$	$\eta=3$	$\eta^{(0)}=3$	FPI	3.0364	0.0607	11.76
			NR	3.0364	0.0608	4.09
		$\eta^{(0)}=\frac{1}{y_{(n)}-\bar{y}}$	FPI	3.0365	0.0608	11.64
			NR	3.0364	0.0608	4.01

$MSE(\hat{\eta}) = E(\hat{\eta}-\eta)^2$

AI= The average number of iterations until convergence

Methodology-Simulation

Two-parameter SEF with $(\eta_1, \eta_2) = (5, -0.5)$, 500 repetitions

	True (η_1, η_2)	Initial (η_1^0, η_2^0)	Method	$E(\hat{\eta}_1)$	$E(\hat{\eta}_2)$	$MSE(\hat{\eta}_1)$	$MSE(\hat{\eta}_2)$	AI
$n = 100$	(5,–0.5)	(5,–0.5)	FPI	5.259	-0.526	1.955	0.0191	28.2
			NR	5.259	-0.526	1.955	0.0191	5.0
		$(\frac{\bar{y}}{s^2}, \frac{-1}{2s^2})$	FPI	5.259	-0.526	1.955	0.0191	31.3
			NR	5.259	-0.526	1.955	0.0191	6.0
$n = 200$	(5,–0.5)	(5,–0.5)	FPI	5.140	-0.514	0.926	0.0091	25.8
			NR	5.140	-0.514	0.926	0.0091	4.8
		$(\frac{\bar{y}}{s^2}, \frac{-1}{2s^2})$	FPI	5.140	-0.514	0.926	0.0091	29.8
			NR	5.140	-0.514	0.926	0.0091	6.0
$n = 300$	(5,–0.5)	(5,–0.5)	FPI	5.125	-0.513	0.622	0.0061	24.8
			NR	5.124	-0.513	0.622	0.0061	4.7
		$(\frac{\bar{y}}{s^2}, \frac{-1}{2s^2})$	FPI	5.125	-0.513	0.622	0.0061	29.2
			NR	5.124	-0.513	0.622	0.0061	6.0

$$MSE(\hat{\eta}_1) = E(\hat{\eta}_1 - \eta_1)^2$$
$$MSE(\hat{\eta}_2) = E(\hat{\eta}_2 - \eta_2)^2$$

AI= The average number of iterations until convergence

Methodology-Simulation

Cubic SEF with $(\eta_1, \eta_2, \eta_3) = (5, -0.5, 0.005)$

Define notations:

$$MSE(\hat{\eta}_1) = E(\hat{\eta}_1 - \eta_1)^2$$

$$MSE(\hat{\eta}_2) = E(\hat{\eta}_2 - \eta_2)^2$$

$$MSE(\hat{\eta}_3) = E(\hat{\eta}_3 - \eta_3)^2$$

RNR= Randomized Newton-Raphson algorithm

AI= The average number of iterations until convergence

Methodology-Simulation

	Initial value $(\eta_1^{(0)}, \eta_2^{(0)}, \eta_3^{(0)})$	Method	$E(\hat{\eta}_1)$	$E(\hat{\eta}_2)$	$E(\hat{\eta}_3)$	$MSE(\hat{\eta}_1)$	$MSE(\hat{\eta}_2)$	$MSE(\hat{\eta}_3)$	AI
100	(5,–0.5,0.005)	RNR	5.277	-0.507	0.0027	44.49	1.63	0.0065	6.1
		Optim	5.285	-0.509	0.0028	44.44	1.62	0.0065	175.3
	$(\frac{\bar{y}}{s^2}, \frac{-1}{2s^2}, 0)$	RNR	5.277	-0.507	0.0027	44.49	1.63	0.0065	7.5
		Optim				Un-convergent			
	(–3,–0.5,0.005)	RNR	5.277	-0.507	0.0027	44.49	1.63	0.0065	7.3
		Optim	3.497	-0.162	0.0191	49.61	1.88	0.0077	212.4
200	(5,–0.5,0.005)	RNR	5.011	-0.478	0.0022	22.18	0.82	0.0033	5.8
		Optim	5.010	-0.478	0.0021	22.17	0.82	0.0033	168.1
	$(\frac{\bar{y}}{s^2}, \frac{-1}{2s^2}, 0)$	RNR	5.011	-0.478	0.0022	22.18	0.82	0.0033	7.5
		Optim				Un-convergent			
	(–3,–0.5,0.005)	RNR	5.011	-0.478	0.0022	22.18	0.82	0.0033	7.3
		Optim	3.225	-0.129	0.0200	33.60	1.29	0.0053	220.9
300	(5,–0.5,0.005)	RNR	4.964	-0.477	0.0026	14.30	0.53	0.0021	5.5
		Optim	4.964	-0.477	0.0026	14.30	0.53	0.0021	163.5
	$(\frac{\bar{y}}{s^2}, \frac{-1}{2s^2}, 0)$	RNR	4.964	-0.477	0.0026	14.30	0.53	0.0021	7.5
		Optim				Un-convergent			
	(–3,–0.5,0.005)	RNR	4.964	-0.477	0.0026	14.30	0.53	0.0021	7.3
		Optim	2.768	-0.047	0.0247	28.98	1.12	0.0046	219.2

Theory

Is the MLE has the property such as consistency, efficiency, normality when the samples are independent but not identically distributed (i.n.i.d)?

Theory

Is the MLE has the property such as consistency, efficiency, normality when the samples are independent but not identically distributed (i.n.i.d)?

Yes

Theory- Theorem

Weak law of large numbers of non-identical sequence (WLLN)

Let $Y_1, Y_2, \cdots$ be independent random variable with $E|Y_i| < \infty$. If there is a constant $p \in [1, 2]$ such that

$$\lim_{n\to\infty} \frac{1}{n^p} \sum_{i=1}^{n} E|Y_i|^p = 0,$$

then

$$\frac{1}{n} \sum_{i=1}^{n} (Y_i - EY_i) \xrightarrow{p} 0.$$

Theory- Theorem

Lindeberg-Feller central limit theorem (CLT)

For each n let $\mathbf{D}_{n,1},\cdots,\mathbf{D}_{n,n}$ be independent random vectors with finite variance such that, as $n \to \infty$,

$$\sum_{i=1}^{n} E\,\|\mathbf{D}_{n,i}\|^2\, 1\{\|\mathbf{D}_{n,i}\| > \varepsilon\} \to 0, \qquad \text{every} \quad \varepsilon > 0,$$

$$\sum_{i=1}^{n} Cov\,\mathbf{D}_{n,i} \to \Sigma$$

Then the sequence $\sum_{i=1}^{n}(\mathbf{D}_{n,i} - E\mathbf{D}_{n,i})$ converges in distribution to a multivariate normal distribution with mean $\mathbf{0}$ and covariance matrix Σ.

Theory

Consistency proof

- WLLN for i.n.i.d

Asymptotic normality proof

- Lindeberg-Feller CLT (non i.i.d)

Theory- Assumption

Assumption (A) There exists an open subset ω of Θ containing the true parameter point

$$\boldsymbol{\eta}^0 = (\eta_1^0, \eta_2^0, \eta_3^0).$$

Assumption (B) There exist a 3×3 positive definite matrix $I(\boldsymbol{\eta}) = \{I_{jk}(\boldsymbol{\eta})\}_{j,k=1,2,3}$ such that,

$$I_{ijk}(\boldsymbol{\eta}) = E_{\boldsymbol{\eta}}\{\frac{\partial}{\partial \eta_j}\log f_i(Y_i \mid \boldsymbol{\eta})\frac{\partial}{\partial \eta_k}\log f_i(Y_i \mid \boldsymbol{\eta})\}$$

as $n \to \infty$,

$$\sum_{i=1}^{n} I_{ijk}(\boldsymbol{\eta})/n \to I_{jk}(\boldsymbol{\eta}) \text{ for } j,k = 1,2,3 \text{ and all } \boldsymbol{\eta} \in \omega.$$

Assumption (C) Suppose that there exists a measurable function M_{jkl} such that

$$\left| \frac{\partial^3}{\partial \eta_j \partial \eta_k \partial \eta_l} \log f_i(y_i \mid \boldsymbol{\eta}) \right| \le M_{jkl}(y_i) \text{ for all } i = 1,2,\cdots,n \text{ and } \boldsymbol{\eta} \in \omega,$$

where

$$m_{ijkl} = E_{\boldsymbol{\eta}^0}\{M_{jkl}(Y_i)\} < \infty \text{ for all } j,k,l \text{ and } i = 1,2,\cdots,n.$$

Also, for some m_{jkl}^2, $\sum_{i=1}^{n} m_{ijkl}{}^2/n \to m_{jkl}^2$ and $\sum_{i=1}^{n} m_{ijkl}/n \to m_{jkl}$, as $n \to \infty$.

Theory- Assumption

Assumption (D) Suppose that there exists a measurable function W_{jk} such that

$$\left| \frac{\partial^2}{\partial \eta_j \partial \eta_k} \log f_i(y_i \mid \boldsymbol{\eta}) \right| \le W_{jk}(y_i) \quad \text{for all} \quad i = 1, 2, \cdots, n \quad \text{and} \quad \boldsymbol{\eta} \in \omega,$$

where

$$w_{ijk} = E_{\eta^0}\{W_{jk}(Y_i)\} < \infty \quad \text{for all} \quad j, k \quad \text{and} \quad i = 1, 2, \cdots, n.$$

Also, for some w_{jk}, $\sum_{i=1}^{n} w_{ijk} / n \to w_{jk}$, as $n \to \infty$.

Assumption (E) Suppose that there exists a measurable function A_j such that

$$\left| \frac{\partial}{\partial \eta_j} \log f_i(y_i \mid \boldsymbol{\eta}) \right| \le A_j(y_i) \quad \text{for all} \quad i = 1, 2, \cdots, n \quad \text{and} \quad \boldsymbol{\eta} \in \omega.$$

Also, for any y, $\max_{1 \le i \le n} A_j^2(Y_i) \le \sup_y A_j^2(y) < \infty$.

Theory- Main Theorem

Theorem 1

If Assumptions (A)-(E) hold, then

(a) $\hat{\eta}_{jn}$ is consistent for estimating η_j, that is $\lim_{n\to\infty} P_{\boldsymbol{\eta}}(|\hat{\eta}_{jn}-\eta_j|\le\varepsilon)=1$ for $j=1,2,3$ and $\varepsilon>0$.

(b) $\sqrt{n}(\hat{\boldsymbol{\eta}}_n-\boldsymbol{\eta})\xrightarrow{d} N_3(\mathbf{0}, I(\boldsymbol{\eta})^{-1})$.

(c) $\hat{\eta}_{jn}$ is asymptotically efficient, that is

$$\sqrt{n}(\hat{\eta}_{jn}-\eta_j)\xrightarrow{d} N(0,[\{I(\boldsymbol{\eta})\}^{-1}]_{jj})$$

Theory- confidence interval

Objective: Construct the confidence interval

Using: Asymptotically efficient

Target: $(1-\alpha)100\%$ confidence interval for η_j is

$$[\, \hat{\eta}_j - Z_{\alpha/2} \cdot s\hat{e}\,(\hat{\eta}_j), \hat{\eta}_j + Z_{\alpha/2} \cdot s\hat{e}\,(\hat{\eta}_j)\,],$$

Z_p is the p-th upper quantile for $N(0,1)$

where

$$s\hat{e}\,(\hat{\eta}_j) = \sqrt{\frac{\{[\,\hat{I}(\hat{\boldsymbol{\eta}})\,]^{-1}\}_{jj}}{n}} = \sqrt{\left[\left\{-\frac{\partial^2}{\partial \boldsymbol{\eta}^2}\ell_n(\hat{\boldsymbol{\eta}})\right\}^{-1}\right]_{jj}}$$

Methodology-coverage probability

Cubic SEF with $(\eta_1, \eta_2, \eta_3) = (5, -0.5, 0.005)$

(η_1, η_2, η_3)	Sample size	$sd(\hat{\eta}_1)$	$E\{s\hat{e}(\hat{\eta}_1)\}$	Coverage probability	
				$1-\alpha=0.90$	$1-\alpha=0.95$
$(5, -0.5, 0.005)$	100	7.692	7.106	0.907	0.940
	200	5.256	4.936	0.896	0.949
	300	4.266	4.033	0.91	0.951

(η_1, η_2, η_3)	Sample size	$sd(\hat{\eta}_2)$	$E\{s\hat{e}(\hat{\eta}_2)\}$	Coverage probability	
				$1-\alpha=0.90$	$1-\alpha=0.95$
$(5, -0.5, 0.005)$	100	1.505	1.370	0.906	0.944
	200	1.017	0.952	0.902	0.945
	300	0.824	0.777	0.913	0.948

(η_1, η_2, η_3)	Sample size	$sd(\hat{\eta}_3)$	$E\{s\hat{e}(\hat{\eta}_3)\}$	Coverage probability	
				$1-\alpha=0.90$	$1-\alpha=0.95$
$(5, -0.5, 0.005)$	100	0.097	0.087	0.912	0.949
	200	0.064	0.060	0.899	0.951
	300	0.052	0.049	0.915	0.947

$sd(\hat{\eta}_j)$ =sample standard deviation

Data analysis

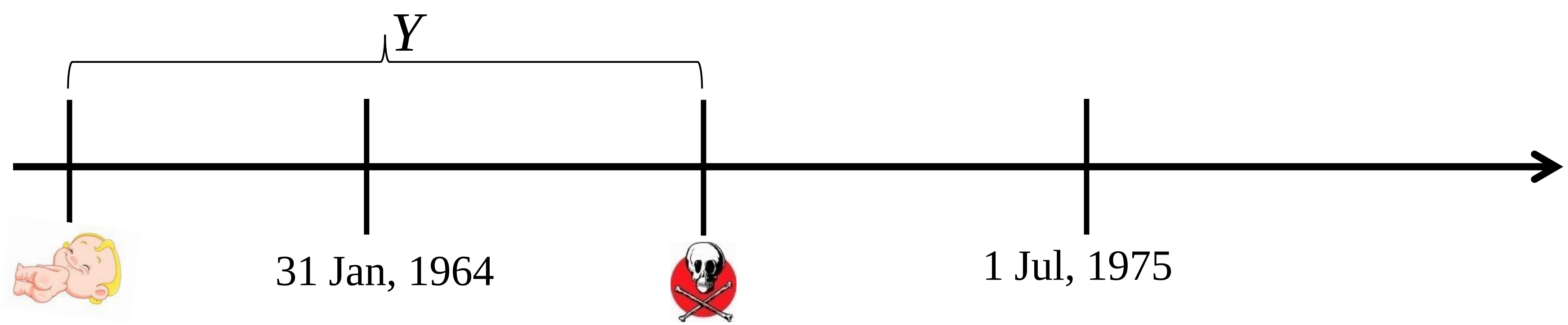
- Channing House retirement center data (Hyde, 1980)

Y : age at death ← **Estimation**

U : age at 31 Jan, 1964

$V = U + 137$: age at 1 Jul, 1975

n : sample size = 167



Data analysis

- Channing House retirement cent

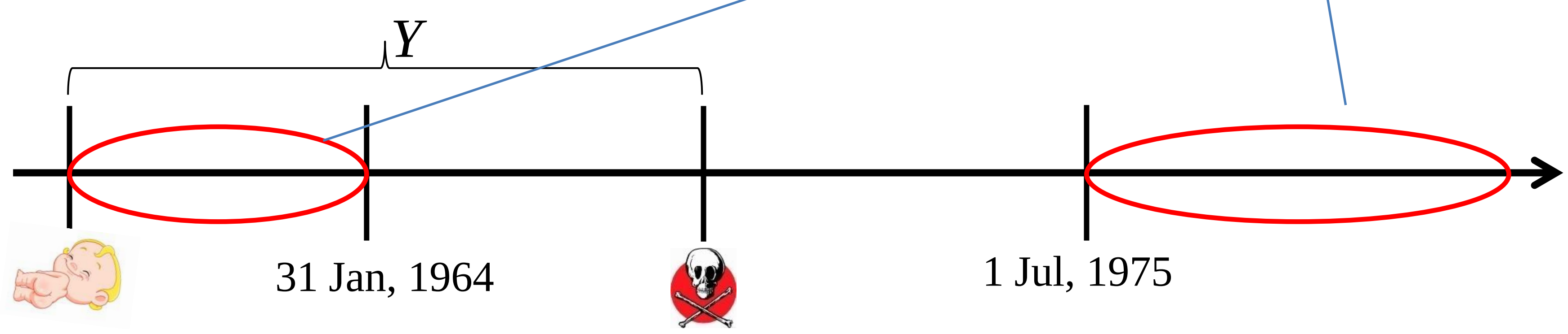
Y : age at death ← **Estimat**

U : age at 31 Jan, 1964

$V = U + 137$: age at 1 Jul, 1975

n : sample size = 167

No information if the person is retirement and died in this period.



Data analysis

Model

Model (a): one-parameter SEF $(\eta_1 > 0)$
Model (b): one-parameter SEF $(\eta_1 < 0)$
Model (c): two-parameter SEF
Model (d): cubic SEF $(\eta_3 < 0)$

Model selection

-Akaike information criterion (AIC) (Akaike, 1973)
-Kolmogorov-Smirnov statistic

Data analysis-AIC

Akaike information criterion (AIC) :

$$AIC = -2\log L + 2k$$

- k is the number of unknown parameters in the model
- L maximized value of likelihood function

Data analysis- Kolmogorov-Smirnov statistic

Define $D = \max_{y} \{ | \hat{S}_{NPMLE}(y) - S_{\hat{\boldsymbol{\eta}}}(y) | \}$

- $\hat{S}_{NPMLE} = \hat{P}(Y > y) = \sum_{y_i > y} \hat{f}_i$, where $\hat{\mathbf{f}} = (\hat{f}_1, \hat{f}_2, \cdots, \hat{f}_n)$ is obtained by self-consistency algorithm. (Efron and Petrosian, 1999)

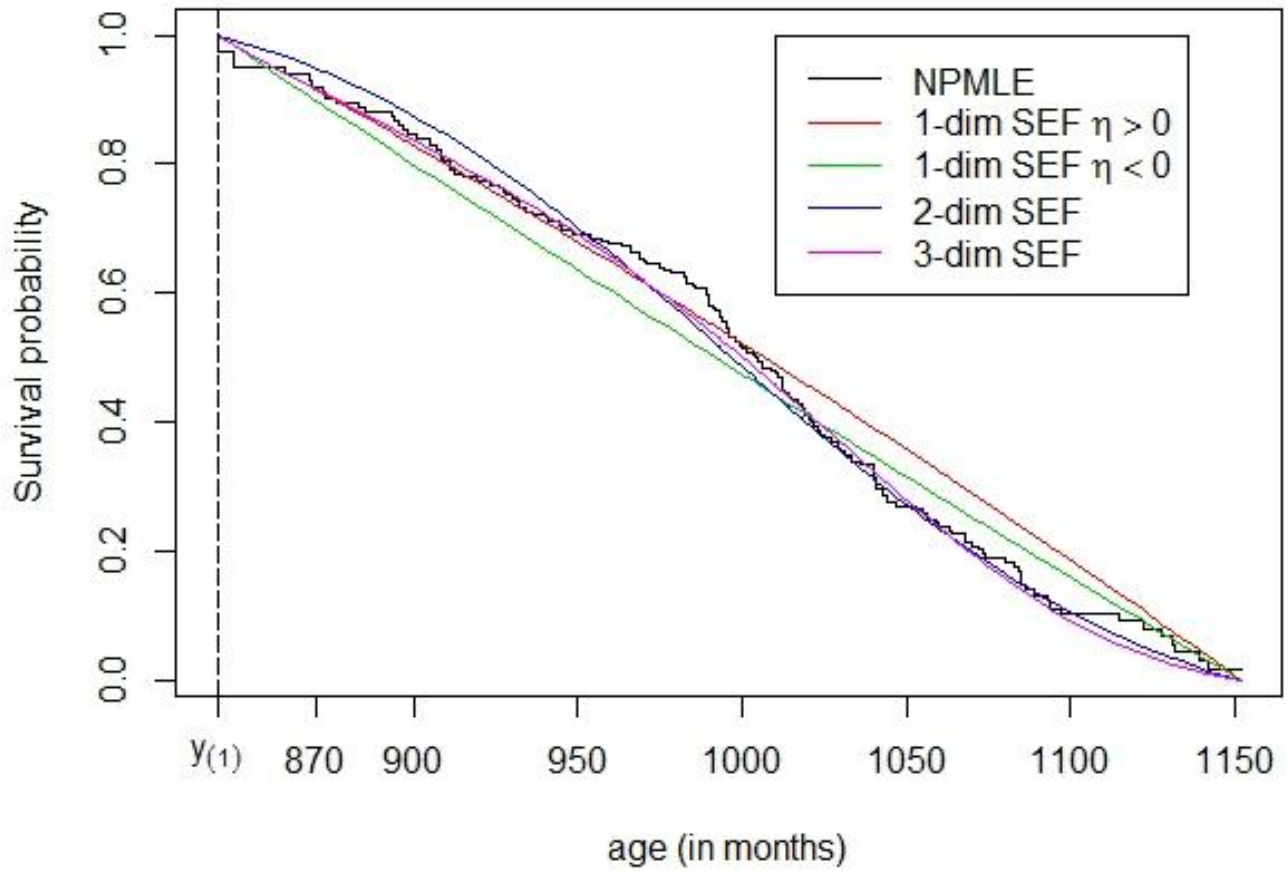
- $S_{\hat{\boldsymbol{\eta}}}(y) = P(Y > y)$ is the survival curve for the parametric estimators.

Data analysis-results

Model	$\hat{\eta}_1$	$\hat{\eta}_2$	$\hat{\eta}_3$	$\log L$	AIC	K-S statistic
(a) 1 par. SEF ($\eta_1 > 0$)	0.0009	0	0	-817.8	1637.6	0.102
(b) 1 par. SEF ($\eta_1 < 0$)	-0.0003	0	0	-819.7	1641.5	0.100
(c) 2 par. SEF	0.0946	-4.74×10^{-5}	0	-817.2	1638.4	0.072
(d) Cubic SEF ($\eta_3 < 0$)	-0.8972	9.44×10^{-4}	-3.29×10^{-7}	-814.5	1635.0	0.061

Smallest

Data analysis-survival function



Prefer the model of cubic SEF with $\eta_3 < 0$

Conclusion

- Newton-Raphson method converges more quickly than fixed-point iteration.
- R optim is sensitive to the initial value. We proposed to use the Randomized Newton-Raphson algorithm.
- For i.n.i.d case, the MLE has the property of consistency and asymptotic efficiency.
- In real data analysis, the cubic SEF gives the best fit.

Reference

- Akaike H (1973) Information theory and an extension of the maximum likelihood principle, Petrov BN and Csaki F, *Proc. 2nd International Symposium on Information Theory*, Akademiai Kiado, Budapest, pp.267-281.
- Burden RL, Faires JD (2011) Numerical Analysis. *Cengage Learning*, Boston.
- Chen YH (2009) Weighted Breslow-type and maximum likelihood estimation in semiparametric transformation models. *Biometrika* 96: 235-251
- Castillo JD (1994) The singly truncated normal distribution: A non-steep exponential family. *Annals of the Institute of Statistical Mathematics*, 46: 57-66.
- Efron B, Petrosian R (1999) Nonparametric methods for doubly truncated data. *Journal of the American Statistical Association* 94: 824-834.
- Emura T, Konno Y (2012) Multivariate normal distribution approaches for dependently truncated data. *Statistical Papers* 53:133-149.
- Miller RG, Efron B, Brown BW, Moses LE (1980) Survival analysis with incomplete observations, Hyde J, *Biostatistics Casebook*, Wiley, New York, pp. 31-46.
- Moreira C, Uña-Á lvarez J (2010) Bootstrapping the NPMLE for doubly truncated data. *Journal of Nonparametric Statistics* 22: 567-583.
- Moreira C, Uña-Á lvarez J (2012) Kernel density estimation with doubly-truncated data. *Electronic Journal of Statistics* 6: 501-521.
- Shen PS (2010) Nonparametric analysis of doubly truncated data. *Annals of the Institute of Statistical Mathematics* 62: 835-853.
- Stovring H, Wang MC (2007) A new approach of nonparametric estimation of incidence and lifetime risk based on birth rates and incidence events. *BMC Medical Research Methodology* 7: 53.
- Sankaran PG, Sunoj SM (2004) Identification of models using failure rate and mean residual life of doubly truncated random variables. *Statistical Papers* 45: 97-109.

Thank you for your listening