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

Reference
Introduction - What is double-truncation

For instance: Channing House study (Hyde, 1980)

\[ b_0 : \text{birth date} \quad d : \text{date of death} \quad Y : \text{age at death (in months)} = d - b_0 \]

\[ U : \text{age on January 31, 1964 (in months)} \]

\[ V = U + 137 : \text{age on July 1, 1975 (in months)} \]
Introduction - Double-truncation

Data: \{ y_1, y_2, \ldots, y_n \} subject to \( u_i \leq y_i \leq v_i \).

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

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

Double-truncation

Nonparametric method

Parametric method
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-Álvarez (2010)
- Use bootstrap to construct the interval estimation.
Moreira and Keliegom (2013)
- A kernel method to estimate the density function.

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_{\eta}(y) = \exp\{ \eta^T \cdot t(y) - \phi(\eta) \}, \quad y \in \mathbb{Y}.\]

- $\mathbb{Y} \subset \mathbb{R}$ is the support of $Y$

- $t(y) = (y, y^2, \cdots, y^k)^T$

- $\eta = (\eta_1, \eta_2, \cdots, \eta_k)^T \in \Theta \subset \mathbb{R}^k$

- $\phi(\eta) = \log[\int_y \exp\{ \eta^T \cdot t(y) \} dy]$
Model-Special exponential family

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

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

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

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 $y_1, y_2, \ldots, y_n$ suffer from double-truncation?

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

$$f(y_i | y_i \in R_i) = \begin{cases} \frac{f(y_i)}{F_i} & \text{if } y_i \in R_i, \\ 0 & \text{if } y_i \not\in R_i. \end{cases}$$

where

$$F_i = \int_{u_i}^{v_i} f(y) \, dy$$
Methodology-Newton-Raphson algorithm

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

Step 2: Set

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

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

\( \ell(\eta) \) = log-likelihood function of  \( \eta \)
Objective: Solve \( S(\eta) \equiv \frac{\partial}{\partial \eta} \ell(\eta) = 0 \rightarrow \eta = g(\eta) \)

Algorithm:

Step 1: Choose the initial value \( \eta^{(0)} \).
Step 2: The iterative process \( \eta^{(p+1)} = g(\eta^{(p)}) \), \( p = 0, 1, 2, \ldots \).

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 \( S(\eta_1, \eta_2) \equiv \frac{\partial}{\partial \eta} \ell(\eta_1, \eta_2) = 0 \) \( \rightarrow \eta_1 = g(\eta_1, \eta_2) \) and \( \eta_2 = q(\eta_1, \eta_2) \)

**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 \( \eta_2^{(p+1)} = q(\eta_1^{(p)}, \eta_2^{(p)}) \),

\[ p = 0, 1, 2, \ldots \]

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.

Two dimensional case
Methodology - One-parameter SEF \( (k = 1) \)

The density of one-parameter SEF is

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

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

- \( f_\eta(y) = -\eta \exp\{\eta(y - \tau_1)\}, \quad y \in \mathcal{Y} = [\tau_1, \infty) \), \quad \text{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

\[
F_i(\eta) = \left\{ \frac{\exp(\eta v_i) - \exp(\eta u_i)}{\exp(\eta \tau_2)} \right\}^{\delta_i} \left\{ \frac{\exp(\eta \tau_2) - \exp(\eta u_i)}{\exp(\eta \tau_2)} \right\}^{1-\delta_i}
\]

The log-likelihood function is

\[
\ell(\eta) = n \log \eta + \eta \sum_{i=1}^{n} y_i - \sum_{i=1}^{n} \delta_i \left\{ \log \left\{ \frac{\exp(\eta v_i) - \exp(\eta u_i)}{\exp(\eta \tau_2)} \right\} \right\} \\
- \sum_{i=1}^{n} (1-\delta_i) \left\{ \log \left\{ \frac{\exp(\eta \tau_2) - \exp(\eta u_i)}{\exp(\eta \tau_2)} \right\} \right\}.
\]
Methodology- One-parameter SEF \((k = 1)\)

The likelihood function for doubly truncated data when \(y_i \in (u_i, v_i)\) is

\[
L(\eta) = \prod_{i=1}^{n} \frac{\exp(\eta v_i) - \exp(\eta u_i)}{\exp(\eta v_i) - \exp(\eta \tau_2)} \cdot \frac{\exp(\eta \tau_2) - \exp(\eta u_i)}{\exp(\eta \tau_2)}^{\delta_i}
\]

where

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

and

\[
F_i(\eta) = \frac{\left\{ \exp(\eta v_i) - \exp(\eta u_i) \right\}^{\delta_i} \left\{ \exp(\eta \tau_2) - \exp(\eta u_i) \right\}^{1-\delta_i}}{\exp(\eta \tau_2)}
\]

The log-likelihood function is

\[
\ell(\eta) = \sum_{i=1}^{n} \left\{ \log\left( \frac{\exp(\eta v_i) - \exp(\eta u_i)}{\exp(\eta v_i) - \exp(\eta \tau_2)} \right) \cdot \frac{\exp(\eta \tau_2) - \exp(\eta u_i)}{\exp(\eta \tau_2)}^{\delta_i} \right\}
\]

and

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

\[
\frac{\partial^2}{\partial \eta^2} \ell(\eta) = \frac{-n}{\eta^2} - \sum_{i=1}^{n} \delta_i \left[ \frac{v_i^2 \exp(\eta v_i) - u_i^2 \exp(\eta u_i)}{\exp(\eta v_i) - \exp(\eta u_i)} - \left\{ \frac{v_i \exp(\eta v_i) - u_i \exp(\eta u_i)}{\exp(\eta v_i) - \exp(\eta u_i)} \right\}^2 \right] \\
- \sum_{i=1}^{n} (1 - \delta_i) \left[ \frac{\tau_2^2 \exp(\eta \tau_2) - u_i^2 \exp(\eta u_i)}{\exp(\eta \tau_2) - \exp(\eta u_i)} - \left\{ \frac{\tau_2 \exp(\eta \tau_2) - u_i \exp(\eta u_i)}{\exp(\eta \tau_2) - \exp(\eta u_i)} \right\}^2 \right].
\]
Methodology- Two-parameter SEF \((k = 2)\)

The density of two-parameter SEF is

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

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

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

\[
L(\eta) = \prod_{i=1}^{n} \frac{f_{\eta}(y_i)}{F_{i}(\eta)} = \frac{\exp \left\{ \sum_{i=1}^{n} \left( \eta_1 y_i + \eta_2 y_i^2 \right) \right\}}{\prod_{i=1}^{n} \left\{ \Phi \left( \frac{v_i + \frac{\eta_1}{2\eta_2}}{\sqrt{-1}} \right) - \Phi \left( \frac{u_i + \frac{\eta_1}{2\eta_2}}{\sqrt{-1}} \right) \right\}}.
\]

where

\[
F_{i}(\eta) = \int_{u_i}^{v_i} \exp \left\{ \eta_1 y + \eta_2 y^2 + \frac{\eta_i^2}{4\eta_2} - \log \left( \frac{-\pi}{2\eta_2} \right) \right\} \, dy = \Phi \left( \frac{v_i + \frac{\eta_1}{2\eta_2}}{\sqrt{-1}} \right) - \Phi \left( \frac{u_i + \frac{\eta_1}{2\eta_2}}{\sqrt{-1}} \right).
\]
Methodology- Two-parameter SEF ($k = 2$)

Define notations

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

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

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

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

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

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

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

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

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

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

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

\[
+ n \eta_2^2 \left( \frac{\partial^2}{\partial \eta_1^2} \ell(\eta) \right) - \frac{n}{2\eta_2^2}
\]
Methodology - Two-parameter SEF \((k = 2)\)

\[
\frac{\partial^2}{\partial \eta_1 \partial \eta_2} \ell(\eta) = \\
- \frac{n \eta_1}{2 \eta_2^2} + (-2 \eta_2) \sum_{i=1}^{n} \{ H_{i1}(\eta_1, \eta_2) - H_{i2}(\eta_1, \eta_2) \} \\
- \frac{1}{\sqrt{-2 \eta_2}} \sum_{i=1}^{n} H_{i1}(\eta_1, \eta_2) \left( \frac{v_i + \eta_1}{2 \eta_2} \right) \left( \frac{-v_i}{\sqrt{-2 \eta_2}} - (-2 \eta_2)^{-\frac{3}{2}} \eta_1 \right) \\
+ \frac{1}{\sqrt{-2 \eta_2}} \sum_{i=1}^{n} H_{i2}(\eta_1, \eta_2) \left( \frac{u_i + \eta_1}{2 \eta_2} \right) \left( \frac{-u_i}{\sqrt{-2 \eta_2}} - (-2 \eta_2)^{-\frac{3}{2}} \eta_1 \right) \\
- \frac{1}{\sqrt{-2 \eta_2}} \sum_{i=1}^{n} \{ H_{i1}(\eta_1, \eta_2) - H_{i2}(\eta_1, \eta_2) \} \left[ H_{i1}(\eta_1, \eta_2) \left( \frac{-v_i}{\sqrt{-2 \eta_2}} - (-2 \eta_2)^{-\frac{3}{2}} \eta_1 \right) \\
- H_{i2}(\eta_1, \eta_2) \left( \frac{-u_i}{\sqrt{-2 \eta_2}} - (-2 \eta_2)^{-\frac{3}{2}} \eta_1 \right) \right].
\]
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 \((\eta_1, \eta_2)\).
- 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

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

\[-\frac{1}{2\eta_2} = \bar{y}^2 - \frac{\eta_1^2}{4\eta_2^2} \]

\[-\frac{1}{n} \sum_{i=1}^{n} \left\{ H_{i1}( \eta_1, \eta_2 ) \left( \frac{-v_i}{\sqrt{-2\eta_2}} - \eta_1 \frac{\sqrt{-2\eta_2}}{4\eta_2^2} \right) \right\} + \frac{1}{n} \sum_{i=1}^{n} \left\{ H_{i2}( \eta_1, \eta_2 ) \left( \frac{-u_i}{\sqrt{-2\eta_2}} - \eta_1 \frac{\sqrt{-2\eta_2}}{4\eta_2^2} \right) \right\}.\]
Methodology- Two-parameter SEF \((k = 2)\)

Use another setting can rewrite the above equation and by fixed-point iteration to fine the MLE of \(\mu\) and \(\sigma^2\).

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

\[
\sigma^2 = \bar{y}^2 - \mu^2
\]

\[
-\frac{1}{n} \sum_{i=1}^{n} \left\{ H_{i1}(\eta_1, \eta_2) \left( \frac{-v_i}{\sqrt{-2\eta_2}} - \eta_1 \frac{\sqrt{-2\eta_2}}{4\eta_2^2} \right) \right\} + \frac{1}{n} \sum_{i=1}^{n} \left\{ H_{i2}(\eta_1, \eta_2) \left( \frac{-u_i}{\sqrt{-2\eta_2}} - \eta_1 \frac{\sqrt{-2\eta_2}}{4\eta_2^2} \right) \right\}.
\]
Methodology - Two-parameter SEF $^{(k = 2)}$

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

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

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

$$
|\eta_i^{(p+1)} - \eta_i^{(p)}| < 10^{-4}, \quad i = 1, 2
$$
The density of cubic SEF is

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

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

In application, \( y_{(n)} \leq \tau_2 \leq \sup_{i} v_i \).
Methodology - Cubic SEF \((k = 3)\)

The density of cubic SEF is

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

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

In application, \(\inf \inf u_i \leq \tau_1 \leq y_{(1)}\).
Methodology - Cubic SEF $(k = 3)$

The likelihood function for doubly truncated data is

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

where

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

The log-likelihood function is

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

Define notation

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

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

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

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

\[ \frac{\partial}{\partial \eta_3} \ell(\eta) = \sum_{i=1}^{n} \{ y_i^3 - E_i^3(\eta) / E_i^0(\eta) \} \]
The second-order derivative of the log-likelihood function is

\[
\frac{\partial^2}{\partial \eta_i^2} \ell(\eta) = \sum_{i=1}^{n} \left[ - E_i^2(\eta) / E_i^0(\eta) + \{ E_i^1(\eta) / E_i^0(\eta) \}^2 \right]
\]

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

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

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

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

\[
\frac{\partial^2}{\partial \eta_3 \partial \eta_2} \ell(\eta) = \sum_{i=1}^{n} \left[ - E_i^5(\eta) / E_i^0(\eta) + \{ E_i^3(\eta) / E_i^0(\eta) \} \{ E_i^2(\eta) / E_i^0(\eta) \} \right]
\]
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 \( \eta = (\eta_1^{(0)}, \eta_2^{(0)}, \eta_3^{(0)})^T \).

Step 2: Set

\[
\begin{bmatrix}
\eta_1^{(p+1)} \\
\eta_2^{(p+1)} \\
\eta_3^{(p+1)}
\end{bmatrix} =
\begin{bmatrix}
\eta_1^{(p)} \\
\eta_2^{(p)} \\
\eta_3^{(p)}
\end{bmatrix} -
\begin{bmatrix}
\frac{\partial^2}{\partial \eta_1^2} \ell(\eta) & \frac{\partial^2}{\partial \eta_1 \eta_2} \ell(\eta) & \frac{\partial^2}{\partial \eta_1 \eta_3} \ell(\eta) \\
\frac{\partial^2}{\partial \eta_2 \eta_1} \ell(\eta) & \frac{\partial^2}{\partial \eta_2^2} \ell(\eta) & \frac{\partial^2}{\partial \eta_2 \eta_3} \ell(\eta) \\
\frac{\partial^2}{\partial \eta_3 \eta_1} \ell(\eta) & \frac{\partial^2}{\partial \eta_3 \eta_2} \ell(\eta) & \frac{\partial^2}{\partial \eta_3^2} \ell(\eta)
\end{bmatrix}^{-1}
\begin{bmatrix}
\frac{\partial \ell(\eta)}{\partial \eta_1} \\
\frac{\partial \ell(\eta)}{\partial \eta_2} \\
\frac{\partial \ell(\eta)}{\partial \eta_3}
\end{bmatrix}
\]

\([\eta_1^{(p)}, \eta_2^{(p)}, \eta_3^{(p)}] \]

The iteration procedure then continues 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 the target value.
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 $(\eta_1, \eta_2, \eta_3)^T$ with $(\eta_1 + u_1, \eta_2 + u_2, \eta_3)^T$, 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

\[ P(U \leq Y \leq V) = \iiint_{u \leq y \leq v} f_Y(y) \cdot f_V(v) \cdot f_U(u) \, du \, dy \, dv \]

We set the condition that all the simulations conducted under

\[ P(U \leq Y \leq V) \approx 0.5 \]
## Methodology-Simulation

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

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

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

AI= The average number of iterations until convergence
### Methodology-Simulation

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

<table>
<thead>
<tr>
<th>(n)</th>
<th>(\eta_1, \eta_2)</th>
<th>Method</th>
<th>(E(\hat{\eta}_1))</th>
<th>(E(\hat{\eta}_2))</th>
<th>(MSE(\hat{\eta}_1))</th>
<th>(MSE(\hat{\eta}_2))</th>
<th>AI</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>((5, -0.5))</td>
<td>FPI</td>
<td>5.259</td>
<td>-0.526</td>
<td>1.955</td>
<td>0.0191</td>
<td>28.2</td>
</tr>
<tr>
<td></td>
<td>((5, -0.5))</td>
<td>NR</td>
<td>5.259</td>
<td>-0.526</td>
<td>1.955</td>
<td>0.0191</td>
<td>5.0</td>
</tr>
<tr>
<td></td>
<td>((\bar{y} - 1) / (s^2, 2s^2))</td>
<td>FPI</td>
<td>5.259</td>
<td>-0.526</td>
<td>1.955</td>
<td>0.0191</td>
<td>31.3</td>
</tr>
<tr>
<td></td>
<td>((\bar{y} - 1) / (s^2, 2s^2))</td>
<td>NR</td>
<td>5.259</td>
<td>-0.526</td>
<td>1.955</td>
<td>0.0191</td>
<td>6.0</td>
</tr>
<tr>
<td>200</td>
<td>((5, -0.5))</td>
<td>FPI</td>
<td>5.140</td>
<td>-0.514</td>
<td>0.926</td>
<td>0.0091</td>
<td>25.8</td>
</tr>
<tr>
<td></td>
<td>((5, -0.5))</td>
<td>NR</td>
<td>5.140</td>
<td>-0.514</td>
<td>0.926</td>
<td>0.0091</td>
<td>4.8</td>
</tr>
<tr>
<td></td>
<td>((\bar{y} - 1) / (s^2, 2s^2))</td>
<td>FPI</td>
<td>5.140</td>
<td>-0.514</td>
<td>0.926</td>
<td>0.0091</td>
<td>29.8</td>
</tr>
<tr>
<td></td>
<td>((\bar{y} - 1) / (s^2, 2s^2))</td>
<td>NR</td>
<td>5.140</td>
<td>-0.514</td>
<td>0.926</td>
<td>0.0091</td>
<td>6.0</td>
</tr>
<tr>
<td>300</td>
<td>((5, -0.5))</td>
<td>FPI</td>
<td>5.125</td>
<td>-0.513</td>
<td>0.622</td>
<td>0.0061</td>
<td>24.8</td>
</tr>
<tr>
<td></td>
<td>((5, -0.5))</td>
<td>NR</td>
<td>5.124</td>
<td>-0.513</td>
<td>0.622</td>
<td>0.0061</td>
<td>4.7</td>
</tr>
<tr>
<td></td>
<td>((\bar{y} - 1) / (s^2, 2s^2))</td>
<td>FPI</td>
<td>5.125</td>
<td>-0.513</td>
<td>0.622</td>
<td>0.0061</td>
<td>29.2</td>
</tr>
<tr>
<td></td>
<td>((\bar{y} - 1) / (s^2, 2s^2))</td>
<td>NR</td>
<td>5.124</td>
<td>-0.513</td>
<td>0.622</td>
<td>0.0061</td>
<td>6.0</td>
</tr>
</tbody>
</table>

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

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

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

**AI** = Absolute Index
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)?
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
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

$$\frac{1}{n} \sum_{i=1}^{n} (Y_i - EY_i) \xrightarrow{p} 0.$$
Lindeberg-Feller central limit theorem (CLT)

For each $n$ let $D_{n,1}, \ldots, D_{n,n}$ be independent random vectors with finite variance such that, as $n \to \infty$,

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

$$\sum_{i=1}^{n} \text{Cov} \ D_{n,i} \to \Sigma$$

Then the sequence $\sum_{i=1}^{n} (D_{n,i} - ED_{n,i})$ converges in distribution to a multivariate normal distribution with mean $\mathbf{0}$ and covariance matrix $\Sigma$. 
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 $\omega$ of $\Theta$ containing the true parameter point

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

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

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

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

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

where

$$m_{ijkl} = E_{\eta^0}\{M_{jkl}(Y_i)\} < \infty \text{ for all } j,k,l \text{ and } i = 1,2,\ldots,n.$$
Assumption (D) Suppose that there exists a measurable function $W_{jk}$ such that

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

where

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

Also, for some $w_{jk}$,

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

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

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

Also, for any $y$,

$$\max_{y_i \in S_{\eta^n}} A_j^2(Y_i) \leq \sup_y A_j^2(y) < \infty.$$
Theorem 1

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

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

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

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

$$\sqrt{n}(\hat{\eta}_{jn} - \eta_j) \xrightarrow{d} N(0, [\{I(\eta)\}^{-1}]_{jj})$$
Objective: Construct the confidence interval

Using: Asymptotically efficient

Target: 

\[(1 - \alpha)100\% \text{ confidence interval for } \eta_j \text{ is } \]

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

where

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

\[Z_p \text{ is the } p\text{-th upper quantile for } N(0, 1)\]
Methodology-coverage probability

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

<table>
<thead>
<tr>
<th>((\eta_1, \eta_2, \eta_3))</th>
<th>Sample size</th>
<th>(sd(\hat{\eta}_1))</th>
<th>(E{s\hat{e}(\hat{\eta}_1)})</th>
<th>Coverage probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>(5, -0.5, 0.005)</td>
<td>100</td>
<td>7.692</td>
<td>7.106</td>
<td>0.907</td>
</tr>
<tr>
<td></td>
<td>200</td>
<td>5.256</td>
<td>4.936</td>
<td>0.896</td>
</tr>
<tr>
<td></td>
<td>300</td>
<td>4.266</td>
<td>4.033</td>
<td>0.91</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>((\eta_1, \eta_2, \eta_3))</th>
<th>Sample size</th>
<th>(sd(\hat{\eta}_2))</th>
<th>(E{s\hat{e}(\hat{\eta}_2)})</th>
<th>Coverage probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>(5, -0.5, 0.005)</td>
<td>100</td>
<td>1.505</td>
<td>1.370</td>
<td>0.906</td>
</tr>
<tr>
<td></td>
<td>200</td>
<td>1.017</td>
<td>0.952</td>
<td>0.902</td>
</tr>
<tr>
<td></td>
<td>300</td>
<td>0.824</td>
<td>0.777</td>
<td>0.913</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>((\eta_1, \eta_2, \eta_3))</th>
<th>Sample size</th>
<th>(sd(\hat{\eta}_3))</th>
<th>(E{s\hat{e}(\hat{\eta}_3)})</th>
<th>Coverage probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>(5, -0.5, 0.005)</td>
<td>100</td>
<td>0.097</td>
<td>0.087</td>
<td>0.912</td>
</tr>
<tr>
<td></td>
<td>200</td>
<td>0.064</td>
<td>0.060</td>
<td>0.899</td>
</tr>
<tr>
<td></td>
<td>300</td>
<td>0.052</td>
<td>0.049</td>
<td>0.915</td>
</tr>
</tbody>
</table>

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

- Channing House retirement center data (Hyde, 1980)

\[ Y : \text{age at death} \quad \leftarrow \quad \text{Estimation} \]

\[ U : \text{age at 31 Jan, 1964} \]

\[ V = U + 137 : \text{age at 1 Jul, 1975} \]

\[ n : \text{sample size} = 167 \]
Data analysis

- Channing House retirement center data (Hyde, 1980)

\[
Y : \text{age at death} \quad \leftarrow \quad \text{Estimation}
\]

\[
U : \text{age at 31 Jan, 1964}
\]

\[
V = U + 137 : \text{age at 1 Jul, 1975}
\]

\[
n : \text{sample size} = 167
\]

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
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 \( D = \max_y \{ | \hat{S}_{NPMLE}(y) - \hat{S}_\eta(y) | \} \)

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

- \( \hat{S}_\eta(y) = P(Y > y) \) is the survival curve for the parametric estimators.
## Data analysis-results

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

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

Thank you for your listening