Asymptotic inference for MLEs under the special exponential family with double-truncation

Takeshi Emura (江村剛志)
Graduate Institute of Statistics
National Central University

Joint work with
Ya-Hsuan Hu, Yoshihiko Konno

11/5/2015
• Presentation is based on our papers:


Childhood cancer data
(Moreira and de Uña-Álvarez 2010)

\[ Y^* : \text{Age at cancer (in days)} \]
\[ U^* : \text{Age at recruitment start (in days)} \]
\[ V^* = U^* + 1825 : \text{Age at recruitment end (in months)} \]
• Double-truncation: 
\[(U^*, Y^*, V^*) \text{ with } (U^*, V^*) \perp Y^*\]

* If \(U^* \leq Y^* \leq V^*\) \(\Rightarrow\) observed; otherwise, nothing is available!

• Observation: \(y_1, y_2, \ldots, y_n\)

subject to \(u_i \leq y_i \leq v_i, i = 1, 2, \ldots, n\)

• Target of Estimation: \(S(y) = P(Y^* > y)\)

\[f(y) = \frac{d}{dy} P(Y^* \leq y)\]
Nonparametric approaches to double-truncation

Efron and Petrosian (1999, JASA)
- Nonparametric maximum likelihood estimator (NPMLE).
Shen (2010, AISM)
- Uniform consistency and weak convergence of NPMLE.
Moreira and Uña-Álvarez (2010, J of Nonpar)
- Bootstrap confidence interval.
Moreira and Keliegom (2013, CSDA)
- A kernel density estimation.
Shen (2012 J. Nonpar), Austin et al. (2014 LIDA)
- Independence test \((U^*, V^*) \perp Y^*)
Emura et al. (2015 LIDA)
- Explicit formula of asymptotic variance of NPMLE
Parametric approaches to double-truncation

Efron and Petrosian (1999, JASA)
- Maximum likelihood estimator (MLE) under the series exponential family (SEF).

Hu and Emura (2015, Computational Statistics)
- Newton Raphson algorithm to obtain the MLE under the SEF

Motivation:
• Asymptotic theory of the MLE under double-truncation has not been studied
• Asymptotic theory does not follows from the usual central limit theorems due to double-truncation

Why? ➔ Data are not i.i.d
Special Exponential family (SEF) introduced by Efron and Petrosian (1999, JASA)

Lifetime $Y^*$ follows a continuous distribution with a density

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

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

I focus on $k = 3$
Special Exponential family \((k = 3)\)

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

Parameter space \(\Theta^+ = \{ (\eta_1, \eta_2, \eta_3) : \eta_1 \in \mathbb{R}, \eta_2 \in \mathbb{R}, \eta_3 > 0 \} \)

\[
y_{(1)} = \max(y_{i}) = \text{maximum observed lifetime}
\]
Special Exponential family \( (k = 3) \)

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

Parameter space \( \Theta^- = \{ (\eta_1, \eta_2, \eta_3) : \eta_1 \in \mathbb{R}, \eta_2 \in \mathbb{R}, \eta_3 < 0 \} \)

Lower bound for the support

\[
y_{(1)} = \min(y_i) = \text{minimum observed lifetime}
\]
Special exponential family (SEF): Summary

\[ f_\eta(y) = \exp\left[ \eta_1 y + \eta_2 y^2 + \cdots + \eta_k y^k - \phi(\eta) \right] \]

<table>
<thead>
<tr>
<th>Distribution</th>
<th>Parameter(s) Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Location-exponential tail distribution</td>
<td>1 parameter SEF ((\eta_1 &gt; 0))</td>
</tr>
<tr>
<td>Location-Scale exponential distribution</td>
<td>1 parameter SEF ((\eta_1 &lt; 0))</td>
</tr>
<tr>
<td>Normal distribution</td>
<td>2 parameter SEF ((\eta_2 &lt; 0))</td>
</tr>
<tr>
<td>U-shaped distribution</td>
<td>2 parameter SEF ((\eta_2 &gt; 0))</td>
</tr>
<tr>
<td>Positively skewed normal distribution</td>
<td>3 parameter (cubic) SEF ((\eta_3 &lt; 0))</td>
</tr>
<tr>
<td>Negatively skewed normal distribution</td>
<td>3 parameter (cubic) SEF ((\eta_3 &gt; 0))</td>
</tr>
</tbody>
</table>

Resemble to **flexible skew normal class**


, and many other skew normal distributions
Likelihood under double-truncation

**Truncation interval:** \( R_i = [u_i, v_i] \)

No likelihood to be observed outside the interval!
(Trunbull, 1976 JRSSB; Efron and Petrosian, 1999 JASA)

**Truncated density**
\[
 f_i(y | \eta) \equiv \frac{f_{\eta}(y)}{F_i(\eta)} \mathbf{1}\{y \in R_i\}
\]

(conditional density given \( y \in R_i \))

where \( F_i = \int_{u_i}^{v_i} f(y) \, dy \)
Likelihood under double-truncation

• Log-likelihood:
\[
\ell_n(\eta) = \log \left\{ \prod_{i=1}^{n} f_i(y_i | \eta) \right\} = \sum_{i=1}^{n} \{ \log f_\eta(y_i) - \log F_i(\eta) \}
\]

• Under the cubic SEF (k=3): \( \eta_3 < 0 \); Positive skew
\[
\ell_n(\eta) = \sum_{i=1}^{n} \left( \eta_1 y_i + \eta_2 y_i^2 + \eta_3 y_i^3 \right)
\]
\[
- \sum_{i=1}^{n} \delta_i \log \left\{ \int_{u_i}^{v_i} \exp \left( \eta_1 y + \eta_2 y^2 + \eta_3 y^3 \right) dy \right\}
\]
\[
- \sum_{i=1}^{n} (1 - \delta_i) \log \left\{ \int_{\tau_i}^{v_i} \exp \left( \eta_1 y + \eta_2 y^2 + \eta_3 y^3 \right) dy \right\}.
\]
Maximum likelihood estimation - Cubic SER $(k = 3)$

Newton-Raphson Algorithm

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

**Step 2**: Repeat

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

until convergence, $|\eta_i^{(p+1)} - \eta_i^{(p)}| < \varepsilon_i \quad \forall i = 1, 2, 3$

Then, $\hat{\eta}_n = (\eta_1^{(p+1)}, \eta_2^{(p+1)}, \eta_3^{(p+1)})$ is the MLE

Stopping criterion: $\varepsilon_1 = 0.0001$, $\varepsilon_2 = 0.0001$, $\varepsilon_3 = 0.0000001$
Randomized Newton-Raphson Algorithm
(Hu and Emura, 2015 Computational Statistics):

**Step 4:** If the Newton Raphson diverges, i.e.,

\[ |\eta_1^{(p+1)} - \eta_1^{(p)}| > D_1 \text{ or } |\eta_2^{(p+1)} - \eta_2^{(p)}| > D_2 \text{ or } |\eta_3^{(p+1)} - \eta_3^{(p)}| > D_3 \]

(“Diameters” \(D_1 = 20\), \(D_1 = 10\), \(D_3 = 1\))

**Return to Step 1:** by replacing \((\eta_1^{(0)}, \eta_2^{(0)}, \eta_3^{(0)})^T\) with

\((\eta_1^{(0)} + u_1, \eta_2^{(0)} + u_2, \eta_3^{(0)})^T\)

where \(u_1 \sim U(-d_1, d_1)\) and \(u_2 \sim U(-d_2, d_2)\)

We chose \(d_1 = 6\) and \(d_2 = 0.5\).
Maximum likelihood estimation- Cubic SER \((k = 3)\)

\[
\text{MLE } \hat{\eta}_n : \quad \frac{\partial}{\partial \eta_j} \ell_n (\eta) = 0, \quad j = 1, 2, 3
\]

- Estimator of density

\[
f_{\hat{\eta}_n} (t) = \exp \{ \hat{\eta}_{n1} t + \hat{\eta}_{n2} t^2 + \hat{\eta}_{n3} t^3 - \phi(\hat{\eta}_n) \}, \quad t \in \mathbb{Y}.
\]

- Estimator of survival function

\[
S_{\hat{\eta}_n} (y) = \int_y^\infty \exp \{ \hat{\eta}_{n1} t + \hat{\eta}_{n2} t^2 + \hat{\eta}_{n3} t^3 - \phi(\hat{\eta}_n) \} dt, \quad y \geq \tau_1
\]

where \(\phi(\eta) = \log \left\{ \int_{\tau_1}^\infty \exp (\eta_1 y + \eta_2 y^2 + \eta_3 y^3) \, dy \right\}\)

We always need SE and confidence intervals for \(f(t)\) and \(S(y)\)
Asymptotic analysis of the MLE

MLE \( \hat{\eta}_n : \frac{\partial}{\partial \eta_j} \ell_n(\eta) = 0, \quad j = 1, 2, \ldots, k \)

- Existence of the MLE
- Consistency, Asymptotic normality
- \( \text{SE and confidence interval} \)
  have not been studied

Our goal:
Asymptotic analysis based on independent but not identically distributed (i.n.i.d) data.
Likelihood under double-truncation

- **Likelihood:**
  \[
  f_i(y | \eta) = \frac{f_\eta(y)}{F_i(\eta)} \mathbb{1}\{u_i \leq y \leq v_i\}
  \]

  \[\begin{array}{ccc}
  u_i & y_i & v_i \\
  \end{array}\]

  Intervals are heterogeneous between samples

  \[u_i \leq y_i \leq v_i, \ i = 1, 2, \ldots, n\]

- **Score function:** not the sum of i.i.d. terms

  \[
  \frac{1}{\sqrt{n}} \frac{\partial}{\partial \eta} \ell_n(\eta) = \frac{1}{\sqrt{n}} \frac{\partial}{\partial \eta} \log \left\{ \prod_{i=1}^{n} f_i(y_i | \eta) \right\}
  \]

  \[
  = \sum_{i=1}^{n} \frac{1}{\sqrt{n}} \frac{\partial}{\partial \eta} \{ \log f_\eta(y_i) - \log F_i(\eta) \} \xrightarrow{d} N(0, \Sigma)
  \]
Asymptotic theories of the MLE under i.n.i.d.

• Bradley and Gart (1962 Biometrika):
  - Seminal work for the MLE under i.n.i.d.
  - Proofs are not rigorous (the literature in 1940’s during which the probability theory was not established)

  - Regularity condition fairly technical and less intuitive
  - Example of lifetime model with fixed censored points

• Philippou and Roussas (1975 AISM)
  - Consistency of MLE is assumed at the beginning.

Conclusion:
• Above existing theorems are not reliable and helpful to our case.
• Yet, the idea of Bradley and Gart (1962) is appealing
Our strategy of establishing asymptotics

• **Regularity conditions:**
  Follow the styles of Bradley and Gart (1962 Biometrika): (but not follow their proofs)

• **Tools:** Textbooks of mathematical statistics:
  - Strong Law of Large Number for i.n.i.d. (Shao 2003)
  - Lindeberg-Feller Central Limit Theorem (van der Vaart, 1998)

• **Proofs:** Modify the proof of Lehmann and Casella (1998) for i.i.d. case
  - Handling multi-parameter cases
  - Rigorous + Clear
  - Simultaneously establish: Existence + Consistency + Asymptotic normality
**Strong Law of Large Number (SLLN) for i.n.i.d.**

**Lemma 2**  Let $Y_1, Y_2, \ldots$ be independent random variables with $E[|Y_i|] < \infty$ for $i = 1, 2, \ldots$. If there is a constant $p \in [1, 2]$ s.t. $\lim_{n \to \infty} \frac{1}{n^p} \sum_{i=1}^{n} E[|Y_i|^p] = 0$, then

$$\frac{1}{n} \sum_{i=1}^{n} (Y_i - E[Y_i]) \overset{p}{\to} 0.$$

---

**Lindeberg-Feller Central Limit Theorem (CLT)**

**Lemma 3**  Let $D_{n,1}, \ldots, D_{n,n}$ be independent $k$-dimensional random vectors s.t.

$$\sum_{i=1}^{n} E[\|D_{n,i} - E[D_{n,i}]\|^2 1\{\|D_{n,i} - E[D_{n,i}]\| > \varepsilon\}] \to 0, \quad n \to \infty \quad \text{(Lindeberg Condition)}$$

for every $\varepsilon > 0$, and $\sum_{i=1}^{n} \text{Cov}(D_{n,i}) \to \Sigma$,

Then, $\sum_{i=1}^{n} (D_{n,i} - ED_{n,i}) \overset{d}{\to} N_k(0, \Sigma)$ as $n \to \infty$. 
Characterization of the MLE

Define \( \hat{\eta}_n = (\hat{\eta}_{1n}, \hat{\eta}_{2n}, \ldots, \hat{\eta}_{kn})^T \) to be a solution to the score equations

\[
\frac{\partial}{\partial \eta_j} \ell_n(\eta) = 0, \quad j = 1, 2, \ldots, k, \quad (*)
\]

**Assumption (A)** The parameter space \( \Theta \) is open and contains

the true \( \eta^0 = (\eta^0_1, \eta^0_2, \ldots, \eta^0_k)^T \). Parameter space \( \Theta \) is natural,

i.e., \( \int_y \exp\{\eta^T \cdot t(y)\} \, dy < \infty \), \( \eta \in \Theta \).

**Lemma 1 (Characterization of MLE):** Under Assumption (A),

if the solution \( \hat{\eta}_n = (\hat{\eta}_{1n}, \hat{\eta}_{2n}, \ldots, \hat{\eta}_{kn})^T \) exists, then it is the MLE,

i.e., \( \ell_n(\hat{\eta}_n) \geq \ell_n(\eta) \) for any \( \eta \in \Theta \).
Assumption (B)  There exist a  $k \times k$  positive definite matrix

$$I(\eta) = \{ I_{js}(\eta) \}_{j,s=1,2,\ldots,k} \text{ s.t.}$$

$$\sum_{i=1}^{n} I_{i,js}(\eta) / n \rightarrow I_{js}(\eta), \quad j, s \in \{1, 2, \ldots, k\}, \quad \eta \in \Theta, \quad \text{as} \quad n \rightarrow \infty.$$ 

where the Fisher information of the  $i$ th sample is

$$I_{i,js}(\eta) = E_{\eta} \left\{ \frac{\partial}{\partial \eta_j} \log f_i(Y_i | \eta) \cdot \frac{\partial}{\partial \eta_s} \log f_i(Y_i | \eta) \right\}, \quad i = 1, 2, \ldots, n, \quad j, s = 1, 2, \ldots, k$$
Boundeness conditions:

**Assumption (C)** For \( j, s, l \in \{ 1, 2, \ldots, k \} \), there exist \( M_{jsl}(\cdot) \) s.t.

\[
\left| \frac{\partial^3}{\partial \eta_j \partial \eta_s \partial \eta_l} \log f_i( y \mid \eta ) \right| \leq M_{jsl}(y), \quad m_{i,jsl} = E_{\eta^0} \{ M_{jsl}(Y_i) \} < \infty \quad \text{and} \quad m^2_{i,jsl} = E_{\eta^0} \{ M_{jsl}(Y_i)^2 \} < \infty.
\]

**Assumption (D)** For \( j, s \in \{ 1, 2, \ldots, k \} \), there exist \( W_{js}(\cdot) \) s.t.

\[
\left| \frac{\partial^2}{\partial \eta_j \partial \eta_s} \log f_i( y \mid \eta ) \right| \leq W_{js}(y), \quad w_{i,js} = E_{\eta^0} \{ W_{js}(Y_i) \} < \infty \quad \text{and} \quad w^2_{i,js} = E_{\eta^0} \{ W_{js}(Y_i)^2 \} < \infty.
\]

**Assumption (E)** For \( j \in \{ 1, 2, \ldots, k \} \), there exist \( A_j(\cdot) \) s.t.

\[
\left| \frac{\partial}{\partial \eta_j} \log f_i( y \mid \eta ) \right| \leq A_j(y), \quad \text{with} \quad \sup_y A^2_j(y) < \infty.
\]

• Assumption (E): Similar to Bradley and Gart (1962 Biometrika) to regulate i.n.i.d. samples
Main result: Asymptotic theory

**Theorem 1:** If Assumptions (A)-(E) hold, then

(a) Existence and consistency: There exists a solution \( \hat{\eta}_n \)

with probability tending to one \( \text{s.t. } \hat{\eta}_n \xrightarrow{P} \eta^0 \text{ as } n \to \infty. \)

(b) Asymptotic normality: \( \sqrt{n}( \hat{\eta}_n - \eta^0 ) \xrightarrow{d} N_k(0, I(\eta^0)^{-1}) \text{ as } n \to \infty. \)

• Valid approximation to the Fisher information

\[
I_{js}(\eta^0) \approx \frac{1}{n} \sum_{i=1}^{n} I_{i,js}(\eta^0) = \frac{1}{n} \sum_{i=1}^{n} E_{\eta^0} \left\{ \frac{\partial}{\partial \eta_j} \log f_i(Y_i|\eta^0) \cdot \frac{\partial}{\partial \eta_s} \log f_i(Y_i|\eta^0) \right\} \\
= \frac{1}{n} \sum_{i=1}^{n} E_{\eta^0} \left\{ -\frac{\partial^2}{\partial \eta_j \partial \eta_s} \log f_i(Y_i|\eta^0) \right\} \approx \frac{1}{n} \frac{\partial^2}{\partial \eta_j \partial \eta_s} \ell_n(\eta) \bigg|_{\eta=\hat{\eta}_n} \equiv \hat{I}_{js}(\hat{\eta}_n).
\]
Asymptotic inference for density

- **Standard Error**

\[
SE \{ f_{\hat{\eta}_n}(y) \} = \sqrt{\left\{ \frac{\partial}{\partial \eta} f_\eta(y) \right\}^{\top} \left\{ -\frac{\partial^2}{\partial \eta^2} \ell_n(\eta) \right\}^{-1} \left. \frac{\partial}{\partial \eta} f_\eta(y) \right|_{\eta=\hat{\eta}_n} },
\]

Where

\[
\frac{\partial}{\partial \eta} f_\eta(y) = \begin{bmatrix}
y - e^1(\eta)/e^0(\eta) \\
\vdots \\
y^k - e^k(\eta)/e^0(\eta)
\end{bmatrix} \cdot f_\eta(y),
\]

and where \( e^j(\eta) = \int_y y^j \exp \left\{ \eta^{\top} \cdot t(y) \right\} dy \), \( j \in \{ 0, 1, 2, \ldots, k \} \).

- \( (1 - \alpha) \) 100% confidence interval for \( f_\eta(y) \) is,

\[
[ f_{\hat{\eta}_n}(y) - Z_{\alpha/2} \cdot SE\{ f_{\hat{\eta}_n}(y) \}, \quad f_{\hat{\eta}_n}(y) + Z_{\alpha/2} \cdot SE\{ f_{\hat{\eta}_n}(y) \} ].
\]
Asymptotic inference for survival function

- **Standard Error**

\[
SE \{ S_{\hat{\eta}_n}(y) \} = \sqrt{\left\{ \frac{\partial}{\partial \eta} S_\eta(y) \right\}^T \left\{ -\frac{\partial^2}{\partial \eta^2} \ell_n(\eta) \right\}^{-1} \frac{\partial}{\partial \eta} S_\eta(y) \bigg|_{\eta=\hat{\eta}_n}},
\]

Where \( \frac{\partial}{\partial \eta} S_\eta(y) = \int_{y \leq t < y} \left[ \begin{array}{c} t - e^1(\eta) / e^0(\eta) \\ \vdots \\ t^k - e^k(\eta) / e^0(\eta) \end{array} \right] \cdot f_\eta(t) \, dt \).

- **(1 - \alpha) 100\% confidence interval for** \( S_\eta(y) \) **is**

\[
\left[ S_{\hat{\eta}_n}(y) - Z_{\alpha/2} \cdot SE \{ S_{\hat{\eta}_n}(y) \}, \quad S_{\hat{\eta}_n}(y) + Z_{\alpha/2} \cdot SE \{ S_{\hat{\eta}_n}(y) \} \right].
\]
Interpretation of Assumption (G)

• Boundedness of truncation interval

• **Sufficient follow-up**: Interval cannot be too short (should be longer than \( v_0 - u_0 \))
Simulation results

the cubic SEF based on 1000 repetitions (under the inclusion probability $P(U^* \leq Y^* \leq V^*) \approx 0.50$).

<table>
<thead>
<tr>
<th>$(\eta_1, \eta_2, \eta_3)$</th>
<th>n</th>
<th>$E(\hat{\eta}_1)$</th>
<th>$SD(\hat{\eta}_1)$</th>
<th>$E[ SE(\hat{\eta}_1) ]$</th>
<th>95%Cov</th>
</tr>
</thead>
<tbody>
<tr>
<td>$(5, -0.5, 0.005)$</td>
<td>100</td>
<td>5.856</td>
<td>7.282</td>
<td>7.436</td>
<td>0.936</td>
</tr>
<tr>
<td></td>
<td>200</td>
<td>5.484</td>
<td>5.146</td>
<td>5.165</td>
<td>0.944</td>
</tr>
<tr>
<td></td>
<td>300</td>
<td>5.378</td>
<td>4.125</td>
<td>4.207</td>
<td>0.951</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$(\eta_1, \eta_2, \eta_3)$</th>
<th>n</th>
<th>$E(\hat{\eta}_2)$</th>
<th>$SD(\hat{\eta}_2)$</th>
<th>$E[ SE(\hat{\eta}_2) ]$</th>
<th>95%Cov</th>
</tr>
</thead>
<tbody>
<tr>
<td>$(5, -0.5, 0.005)$</td>
<td>100</td>
<td>-0.622</td>
<td>1.397</td>
<td>1.437</td>
<td>0.940</td>
</tr>
<tr>
<td></td>
<td>200</td>
<td>-0.573</td>
<td>0.995</td>
<td>0.998</td>
<td>0.945</td>
</tr>
<tr>
<td></td>
<td>300</td>
<td>-0.561</td>
<td>0.797</td>
<td>0.813</td>
<td>0.946</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$(\eta_1, \eta_2, \eta_3)$</th>
<th>n</th>
<th>$E(\hat{\eta}_3)$</th>
<th>$SD(\hat{\eta}_3)$</th>
<th>$E[ SE(\hat{\eta}_3) ]$</th>
<th>95%Cov</th>
</tr>
</thead>
<tbody>
<tr>
<td>$(5, -0.5, 0.005)$</td>
<td>100</td>
<td>0.0101</td>
<td>0.089</td>
<td>0.091</td>
<td>0.944</td>
</tr>
<tr>
<td></td>
<td>200</td>
<td>0.0084</td>
<td>0.063</td>
<td>0.063</td>
<td>0.950</td>
</tr>
<tr>
<td></td>
<td>300</td>
<td>0.0081</td>
<td>0.051</td>
<td>0.052</td>
<td>0.949</td>
</tr>
</tbody>
</table>
Simulation results

The cubic SEF based on 1000 repetitions (under the inclusion probability $P(U^* \leq Y^* \leq V^*) \approx 0.50$).

<table>
<thead>
<tr>
<th>$(\eta_1, \eta_2, \eta_3)$</th>
<th>$n$</th>
<th>$E{ S_{\eta}(t) }$</th>
<th>$SD{ S_{\eta}(t) }$</th>
<th>$E[SE { S_{\eta}(t) }]$</th>
<th>95%Cov</th>
</tr>
</thead>
<tbody>
<tr>
<td>$(5, -0.5, 0.005)$</td>
<td>100</td>
<td>0.499</td>
<td>0.071</td>
<td>0.070</td>
<td>0.944</td>
</tr>
<tr>
<td></td>
<td>200</td>
<td>0.500</td>
<td>0.049</td>
<td>0.049</td>
<td>0.939</td>
</tr>
<tr>
<td></td>
<td>300</td>
<td>0.502</td>
<td>0.038</td>
<td>0.039</td>
<td>0.947</td>
</tr>
<tr>
<td>$S_{\eta}(y) = 0.5$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$(5, -0.5, -0.005)$</td>
<td>100</td>
<td>0.504</td>
<td>0.062</td>
<td>0.065</td>
<td>0.941</td>
</tr>
<tr>
<td></td>
<td>200</td>
<td>0.503</td>
<td>0.044</td>
<td>0.045</td>
<td>0.957</td>
</tr>
<tr>
<td></td>
<td>300</td>
<td>0.502</td>
<td>0.036</td>
<td>0.037</td>
<td>0.949</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$(\eta_1, \eta_2, \eta_3)$</th>
<th>$n$</th>
<th>$E{ f_{\eta}(t) }$</th>
<th>$SD{ f_{\eta}(t) }$</th>
<th>$E[ SE{ f_{\eta}(t) } ]$</th>
<th>95%Cov</th>
</tr>
</thead>
<tbody>
<tr>
<td>$(5, -0.5, 0.005)$</td>
<td>100</td>
<td>0.367</td>
<td>0.054</td>
<td>0.057</td>
<td>0.969</td>
</tr>
<tr>
<td>$f_{\eta}(y) = 0.369$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>200</td>
<td>0.367</td>
<td>0.036</td>
<td>0.039</td>
<td>0.967</td>
</tr>
<tr>
<td></td>
<td>300</td>
<td>0.367</td>
<td>0.030</td>
<td>0.031</td>
<td>0.961</td>
</tr>
<tr>
<td>$(5, -0.5, -0.005)$</td>
<td>100</td>
<td>0.430</td>
<td>0.057</td>
<td>0.060</td>
<td>0.961</td>
</tr>
<tr>
<td>$f_{\eta}(y) = 0.427$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>200</td>
<td>0.428</td>
<td>0.040</td>
<td>0.041</td>
<td>0.947</td>
</tr>
<tr>
<td></td>
<td>300</td>
<td>0.427</td>
<td>0.032</td>
<td>0.033</td>
<td>0.958</td>
</tr>
</tbody>
</table>
Data analysis of childhood cancer data
(Moreira and de Uña-Álvarez 2010)

\[ Y^* : \text{Age at cancer (in days)} \leftarrow \text{Estimation} \]

\[ U^* : \text{Age at recruitment start (in days)} \]

\[ V^* = U^* + 1825 : \text{Age at recruitment end (in months)} \]

\[ Y^* = 1825 + \begin{array}{c} \text{Follow-up} \\
\end{array} \]

Recruitment Start: 1999/1/1
Recruitment End: 2003/12/31
5 year (1825 days) follow-up
First, model selection

1) **Kolmogorov-Smirnov distance**

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

- \( \hat{S}_{NPMLE} = \hat{P}(Y > y) = \sum_{y_i > y} \hat{f}_i \) : **Model-free** survival function
  where \( \hat{f} = (\hat{f}_1, \hat{f}_2, \ldots, \hat{f}_n) \) is the NPMLE
  (Efron and Petrosian, 1999)

- \( \hat{S}_\eta(y) = P(Y > y) \) : **Model-based** survival function

2) **AIC (Akaike Information Criterion)**

\[
AIC = -2 \log L + 2k
\]

- \( k \) : the number of unknown parameters
- \( \log L \) : maximized value of likelihood function
KS-distance between MLE and NPMLE

Best Model = Smallest KL distance = Cubic SEF with $\eta_3 < 0$
The maximum likelihood inference for the childhood cancer data.

<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>$8.74 \times 10^{-5}$</td>
<td>0</td>
<td>0</td>
<td>-3013.6</td>
<td>6029.2</td>
<td>0.206</td>
</tr>
<tr>
<td>(b) 1 par. SEF ($\eta_1 &lt; 0$)</td>
<td>$-3.85 \times 10^{-4}$</td>
<td>0</td>
<td>0</td>
<td>-2999.6</td>
<td>6001.1</td>
<td>0.121</td>
</tr>
<tr>
<td>(c) 2 par. SEF</td>
<td>$7.71 \times 10^{-4}$</td>
<td>$-1.87 \times 10^{-7}$</td>
<td>0</td>
<td>-3027.6</td>
<td>6059.2</td>
<td>0.132</td>
</tr>
<tr>
<td>(d) Cubic SEF ($\eta_3 &lt; 0$)</td>
<td>$-7.90 \times 10^{-4}$</td>
<td>$3.38 \times 10^{-7}$</td>
<td>$-4.87 \times 10^{-11}$</td>
<td>-2991.6</td>
<td>5989.2</td>
<td>0.084</td>
</tr>
</tbody>
</table>

- Model (a) = The one-parameter SEF ($\eta_1 > 0$)
- Model (b) = The one-parameter SEF ($\eta_1 < 0$)
- Model (c) = The two-parameter SEF
- Model (d) = The cubic SEF ($\eta_3 < 0$)
- $\log L$ = The maximized log-likelihood
- AIC = Akaike information criterion, defined as $AIC = -2 \log L + 2k$
- K-S statistic = The Kolmogorov-Smirnov distance between the MLE and the NPMLE

**Best model**
Data analysis under the cubic SEF (best model)

\[
S_{\hat{\eta}}(t) = \int f_{\hat{\eta}}(y)dy = \int \exp[\hat{\eta}_1 y + \hat{\eta}_2 y^2 + \hat{\eta}_3 y^3 - \phi(\hat{\eta})]dy
\]

\[y_{(1)} = \min(y_i) = \text{minimum observed lifetime}\]
Asymptotic inference under the cubic SEF

\[ f_{\hat{\eta}}(y) = \exp\left[ \hat{\eta}_1 y + \hat{\eta}_2 y^2 + \hat{\eta}_3 y^3 - \phi(\hat{\eta}) \right] \]

---

High risk of developing cancer in early ages

: Same finding found in Emura et al. (2015 LIDA)
Q1: Assumption (G) hold for real data example?

Answer

Follow-up length: Fixed at 5 years $d_0 = 1825$ (days).

But Assumption (G) requires $d_0 > 7300$ (days)

So Assumption (G) does not hold.

But it is easy to be checked by user.

Other target quantities under double-truncation

- Predictive survival $S(t + w | t) = S(Y > t + w | Y > t)$
  (Klein & Moeschberger, 2003, with left-truncation only)

- Mean / median residual life $m(u) = E(Y - u | Y > u)$
  (Chi et al. 2014 Com.Stat-Simulations, with left-truncation only)

$m(u, v) = E(Y - v | u > Y > v)$

with double-truncation (Sankaran & Sunoj, 2004 Stat Papers)
Thank you for your listening