Expectation-maximization-type Algorithms for the Analysis of System Lifetime Data

Hon Keung Tony Ng

Department of Statistical Science
Southern Methodist University
Dallas, Texas, U.S.A.

Co-authored with Dr. Yandan Yang (Targetbase, Irving, Texas)
and Professor N. Balakrishnan (McMaster University, Canada)

The 3rd Pacific Rim Statistical Conference for Production Engineering
December 14, 2018

Outline

Outline

Problem of interest

- The lifetimes of the integrated circuits (IC) can be observed.
- We can't open the IC once it is build.
- What is the lifetime distribution of the electronic element?

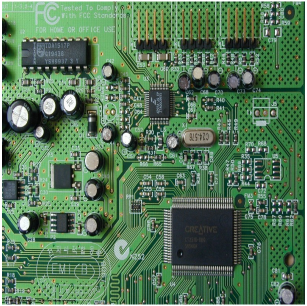


Figure: An integrated circuit

Problem of interest

- System is built from components, we want to study both the system lifetime distribution and the component lifetime distribution
- Observe system lifetime $T_1, \ldots, T_m$
- How to perform statistical inference on component lifetime distribution if only the system lifetimes are observed?
- How do the system lifetimes and component lifetimes link to each other?
- What can we do if we have censored sample?

What is a system?

- System is a collection of components arranged in some fashion in order to achieve desired functions
  - A television
  - An automobile
  - A computer
  - ...
- Relevant: If the status of system depends on all the components
- Monotone: If repairing a failed component will not make the system worse
- Coherent system: If each of its components is relevant and if its structure function is monotone

System Lifetime vs. Component Lifetime

- Factors affect system lifetime
  - Component lifetimes
  - System structure
- Examples:

Figure: 3-Component Series System

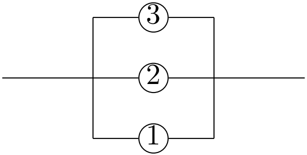


Figure: 3-Component Parallel System

Outline

Order Statistics

- Sample of size n: $X_1, X_2, \ldots, X_n$
- Ordered sample $X_{1:n} \leq X_{2:n} \leq \ldots \leq X_{n:n}$
- If the sample is i.i.d. with p.d.f. $f_X(x)$, c.d.f. $F_X(x)$ and s.f. $\bar{F}_X(x)$
  - The joint density of the order statistics $X_{1:n} \leq X_{2:n} \leq \ldots, X_{n:n}$ is

$$f_{X_{1:n},X_{2:n},\ldots,X_{n:n}}(x_1, x_2, \ldots, x_n) = n! \prod_{i=1}^{n} f_X(x_i), -\infty < x_1 < \cdots < x_n < \infty.$$

  - The p.d.f. of $X_{i:n}$ $(1 \leq i \leq n)$

$$f_{X_{i:n}}(x) = \frac{n!}{(i-1)!(n-i)!}[F_X(x)]^{i-1}[1 - F_X(x)]^{n-i} f_X(x), \quad -\infty < x < \infty.$$

  - The c.d.f. of $X_{i:n}$ $(1 \leq i \leq n)$

$$F_{X_{i:n}}(x) = \sum_{r=i}^{n} \binom{n}{r} \{F_X(x)\}^r \{1 - F_X(x)\}^{n-r}, \quad -\infty < x < \infty.$$

System Signature

- Denote the system lifetime as T, and the n-component lifetimes in a system as $X_1, \ldots, X_n$
- Proposed by Samaniego (1985, *IEEE Transactions on Reliability*)
- Probability vector: $\mathbf{s} = (s_1, s_2, \ldots, s_n)$, for $i = 1, 2, \ldots, n$, $s_i \geq 0$ and $\sum_{i=1}^{n} s_i = 1$
- s_i is the probability that the i-th component failure causes the system failure: $s_i = \Pr(T = X_{i:n})$

Example: 3-component Series-Parallel System

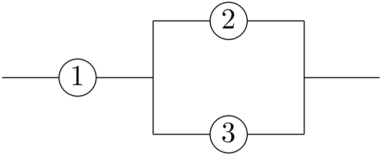


Figure: A 3-Component Series-Parallel System

Arrangements	System Lifetime T
$X_1 \leq X_2 \leq X_3$	$X_{1:3}$
$X_1 \leq X_3 \leq X_2$	$X_{1:3}$
$X_2 \leq X_1 \leq X_3$	$X_{2:3}$
$X_2 \leq X_3 \leq X_1$	$X_{2:3}$
$X_3 \leq X_1 \leq X_2$	$X_{2:3}$
$X_3 \leq X_2 \leq X_1$	$X_{2:3}$

$$s_1 = \Pr(T = X_{1:3}) = 2/6 = 1/3$$
$$s_2 = \Pr(T = X_{2:3}) = 4/6 = 2/3$$
$$s_3 = \Pr(T = X_{3:3}) = 0/6 = 0$$

System Lifetime Distribution Based on System Signature

- Assume the components are i.i.d. in a system, then the distribution of system lifetime T can be expressed as

$$F_T(t) = \Pr(T \leq t) = \sum_{i=1}^{n} \Pr(T = X_{i:n}) \Pr(T \leq t | T = X_{i:n})$$

- The c.d.f. and p.d.f. of system lifetime T, in form of component lifetimes F_X, are respectively

$$F_T(t) = \sum_{i=1}^{n} s_i \sum_{j=i}^{n} \binom{n}{j} [F_X(t)]^j [\bar{F}_X(t)]^{n-j},$$

$$f_T(t) = \sum_{i=1}^{n} s_i \binom{n}{i} i [F_X(t)]^{i-1} f_X(t) [\bar{F}_X(t)]^{n-i}.$$

Ordered System Signature

- Ordered system lifetimes $T_{1:m} < T_{2:m} < \ldots < T_{m:m}$
- Ordered components lifetime for kth ordered system lifetime are $X_{k,1:n} < X_{k,2:n} < \ldots < X_{k,n:n}$.
- Ordered system signature for the kth system lifetime $T_{k:m}$ is $\mathbf{s}^{(k:m)} = (s_1^{(k:m)}, \ldots, s_n^{(k:m)})$ (Balakrishnan and Volterman, 2014, *Journal of Applied Probability*), where

$$s_i^{(k:m)} = \sum_{l=1}^{m} \Pr(T_{k:m} = X_{l,i:n}).$$

- 3-component series-parallel system (with system signature (5/15, 10/15, 0)): sample size $m = 2$, ordered system signature

$T_{1:2}$: $\mathbf{s}^{(1:2)} = (s_1^{(1:2)}, s_2^{(1:2)}, s_3^{(1:2)}) = (7/15, 8/15, 0)$
$T_{2:2}$: $\mathbf{s}^{(2:2)} = (s_1^{(2:2)}, s_2^{(2:2)}, s_3^{(2:2)}) = (3/15, 12/15, 0)$

Minimal Signature

- The minimal signature, proposed by Navarro et al. (2007, *Communications in Statistics-Theory and Methods*), is an expression of the system structure by considering the system lifetime as a generalized mixture of first order statistic
- Denoted by **a**$=(a_1,\ldots,a_n)$, where $a_1,\ldots,a_n$ are some negative and non-negative integers, and $\sum_{i=1}^n a_i = 1$
- p.d.f. and s.f. of system lifetime T

$$
\begin{aligned}
f_T(t|\mathbf{a}) &= \sum_{i=1}^n a_i f_{1:i}(t) = \sum_{i=1}^n a_i i f_X(t)[\bar{F}_X(t)]^{i-1}, \\
\bar{F}_T(t|\mathbf{a}) &= \sum_{i=1}^n a_i \bar{F}_{1:i}(t) = \sum_{i=1}^n a_i[\bar{F}_X(t)]^i.
\end{aligned}
$$

- Example: minimal signature of 3-component series-parallel system **a**$=(0,2,-1)$

Properties

- System signature, ordered system signature and minimal signature are *distribution-free* representations of the system structure
  - Do not rely on the lifetime distribution of the components
  - Only relate to the design of the system
  - Ordered system signature could be approximated by Monte Carlo methods.
- There is a one-to-one mapping between the system signature and minimal signature, i.e., the minimal signature can be expressed in terms of system signature

$$a_l = \sum_{i=n-l+1}^{n} \sum_{j=n-l}^{i-1} s_i \binom{n}{j} \binom{j}{n-l} (-1)^{j+l-n}$$

for $l = 1, \ldots, n$

Example: Water-reservoir control system

- The failure times of a simple software control logic for a water-reservoir control system (Pham and Pham, 1991 and Teng and Pham, 2002)
- Water is supplied via a source pipe controlled by a source valve and removed via a drain pipe controlled by a drain valve.
- There are two level sensors in the control system that maintain the water level between the high and low limits in the water reservoir.
- The water-reservoir control system achieves fault tolerance by using a two-version programming software control logic in which the system fails only when both of its components (software versions) fail.
- The control system can be treated as a two-component parallel system with system signature $s = (0, 1)$.

Example: Phosphor acid filter system

- Frenkel and Khvatskin (2006) considered a prototype for n-component phosphor acid filter system used in Rotem/Deshanim Chemical Processing Facility, Arad, Israel.
- The filter system has n identical turning rollers and a system failure occurs when two adjacent rollers stop working according to the technical specification.
- It is a consecutive 2-out-of-n: F system.
- Consider a consecutive 2-out-of-8: F system with components in the system have equal chance of failure, then the system signature is s $=$ (0, 1/4, 11/28, 2/7, 1/14, 0, 0, 0).

Outline

Maximum Likelihood Estimation (MLE)

For a random sample $\mathbf{X} = (X_1, \ldots, X_n)$ from a specified model with parameter vector $\boldsymbol{\theta}$. The log-likelihood function could be written as

$$\ell(\boldsymbol{\theta}; \mathbf{x}) = \sum_{i=1}^{n} \log f_X(x_i; \boldsymbol{\theta}).$$

The MLE of $\boldsymbol{\theta}$, noted as $\hat{\boldsymbol{\theta}}$, is the value of $\boldsymbol{\theta}$ that maximize the log-likelihood over the parameter space Θ. Let $\mathcal{I}(\boldsymbol{\theta})$ denote the large-sample average amount of information per observation,

$$\mathcal{I}(\boldsymbol{\theta}) = \lim_{n\to\infty} \left\{ \frac{1}{n} E\left[-\frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} \right] \right\}.$$

Expectation-Maximization (EM) Algorithm

- Suppose observed data **U** are generated from the distribution $f(u|\boldsymbol{\theta})$, and missing data are **V**
- Maximize the log-likelihood based on **U**: $\ell(\boldsymbol{\theta}) = \ln f(\mathbf{u}|\boldsymbol{\theta})$
- Log-likelihood based on complete data $(\mathbf{U}, \mathbf{V})$: $\ln f(\mathbf{u}, \mathbf{v}|\boldsymbol{\theta})$
- Consider the expectation of log-likelihood based on complete data over the unobserved **V** given the observed **U**,

$$Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{(h)}) = E_{\mathbf{v}|\mathbf{u}} \left(\ln f(\mathbf{u}, \mathbf{v}|\boldsymbol{\theta})|\mathbf{u}, \boldsymbol{\theta}^{(h)} \right),$$

where $\theta^{(h)}$ is the current parameter.

EM Algorithm

Start with initial value $\boldsymbol{\theta}^{(0)}$. Let $\boldsymbol{\theta}^{(h)}$ denote the estimate of $\boldsymbol{\theta}$ after the h-th iteration, the $(h+1)$-th iteration in EM algorithm is

- E-step: Evaluate

$$
\begin{aligned}
Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{(h)}) &= E_{\mathbf{v}|\mathbf{u}}\left(\ln f(\mathbf{u}, \mathbf{v}|\boldsymbol{\theta})|\mathbf{u}, \boldsymbol{\theta}^{(h)}\right) \\
&= \int_{\mathbf{v}} \ln f(\mathbf{u}, \mathbf{v}|\boldsymbol{\theta}) dF(\mathbf{v}|\mathbf{u}, \boldsymbol{\theta}^{(h)});
\end{aligned}
$$

- M-step: maximizing the $Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{(h)})$ function obtained from the E-step, i.e.,

$$
\boldsymbol{\theta}^{(h+1)} = \arg\max_{\boldsymbol{\theta}} Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{(h)}).
$$

Monte Carlo EM (MCEM) Algorithm

The MCEM algorithm consists the MC E-step and M-step. After the h-th iteration, the $(h+1)$th iteration is:

- Monte Carlo E-step: Draw a sample of the missing data **V**, say, $\mathbf{v}^1, \ldots, \mathbf{v}^M$ from $f_{V|U}(v|u, \boldsymbol{\theta}^{(h)})$, and approximate the $Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(h)})$ function by

$$\bar{Q}(\boldsymbol{\theta}, \boldsymbol{\theta}^{(h)}) = \frac{1}{M} \sum_{m=1}^{M} \ln f_{\boldsymbol{\theta}|X}(\boldsymbol{\theta}|u, \mathbf{v}^m);$$

- M-step: Maximize the $\bar{Q}(\boldsymbol{\theta}, \boldsymbol{\theta}^{(h)})$ and update the estimate $\boldsymbol{\theta}^{(h+1)}$ for next iteration.

Stochastic EM (SEM) Algorithm

Stochastic EM (SEM) algorithm is the case of MCEM when $M = 1$. The procedure of SEM algorithm is

- S-step: Draw a single sample of $\mathbf{V}$, say $\mathbf{v}$, from $f_{V|U}(\mathbf{v}|\mathbf{u}, \boldsymbol{\theta}^{(h)})$, and replace the missing value by the sample $\mathbf{v}$ to obtain the "pseudo" complete likelihood function $\ell(\boldsymbol{\theta}^{(h)}|\mathbf{u}, \mathbf{v})$;
- M-step: Maximize the "pseudo" complete likelihood $\ell(\boldsymbol{\theta}^{(h)}|\mathbf{u}, \mathbf{v})$ with respect to $\boldsymbol{\theta}$, and then update the estimate $\boldsymbol{\theta}^{(h+1)}$.

Outline

Introduction

- Suppose a Type-II censored sample $t_{1:m} < t_{2:m} < \ldots < t_{r:m}$ is observed, then the likelihood function can be expressed as

$$L(\boldsymbol{\theta}; t_{1:m}, t_{2:m}, \ldots, t_{r:m}) \propto \left\{ \prod_{j=1}^{r} f_T(t_{j:m}) \right\} [1 - F_T(t_{r:m})]^{(m-r)},$$

  where

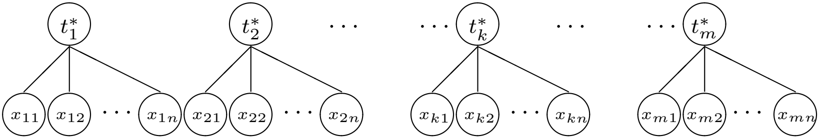
$$F_T(t) = \sum_{i=1}^{n} s_i \sum_{j=i}^{n} \binom{n}{j} [F_X(t)]^j [\bar{F}_X(t)]^{n-j},$$

$$f_T(t) = \sum_{i=1}^{n} s_i \binom{n}{i} i [F_X(t)]^{i-1} f_X(t) [\bar{F}_X(t)]^{n-i}.$$

- Direct maximization can be used, but it may not be numerically stable.

Introduction

- EM algorithm can be hard to implement, due to the conditional expectations involved in E-step.
- EM may not converge to the global maximum.
- Prefer less mathematical complications.
- S-step: replaces the expectation of the missing data step by a single draw from the distribution of the missing data.



Pseudo-Complete Data

- If we observe all the component lifetimes

$$
\begin{array}{ccc}
x_{11}, & \cdots & , x_{1n} \\
x_{21}, & \cdots & , x_{2n} \\
 & \vdots & \\
x_{m1}, & \cdots & , x_{mn}
\end{array}
$$

- Log-likelihood based on all component lifetimes

$$
\ell(\boldsymbol{\theta}) = \sum_{k=1}^{m} \sum_{i=1}^{n} \ln f_X(x_{ki}|\boldsymbol{\theta}).
$$

Observed System Lifetime

- Let I_k be the number of components failed at the time of the system k failure, $k = 1, \ldots, m$.
- If I_k is observed, $t_k = x_{k(I_k:n)}$.
- Denote the component lifetimes in system k as $x_{k1} < x_{k2} < \ldots < x_{k(I_k-1)} < t_k = x_{kI_k} < x_{k(I_k+1)} < \ldots < x_{k,n}$
- Denote the observed system lifetime by U_k, i.e., $U_k = T_k$
- The missing data are $\mathbf{V}_k = (V_{k1}, \mathbf{V}_{k2})$, where $V_{k1} = I_k$ and $\mathbf{V}_{k2} = (X_{k1}, X_{k2}, \ldots, X_{k(I_k-1)}, X_{k(I_k+1)}, \ldots, X_{kn})$

Observed System Lifetime

- Distributions of the missing data I_k and X_{kj}, $k = 1, \ldots, m, \; j = 1, \ldots, n$, given observed data t_k are

$$\Pr(I_k = i) = s_i,$$

$$f(x_{kj}|\boldsymbol{\theta}) = \begin{cases} \dfrac{f_X(x_{kj}|\boldsymbol{\theta})}{F_X(t_k|\boldsymbol{\theta})} & \text{if } j < I_k, \\ \dfrac{f_X(x_{kj}|\boldsymbol{\theta})}{\bar{F}_X(t_k|\boldsymbol{\theta})} & \text{if } j > I_k. \end{cases} \tag{1}$$

- The log-likelihood function of the complete data $(U_k, V_{k1}, \mathbf{V}_{k2})$ for system k is

$$\ell_k(\boldsymbol{\theta}|U_k, V_{k1}, \mathbf{V}_{k2}) = \sum_{j=1}^{I_k-1} \ln f_X(x_{kj}|\boldsymbol{\theta}) + \ln f_X(t_k|\boldsymbol{\theta}) + \sum_{j=I_k+1}^{n} \ln f_X(x_{kj}|\boldsymbol{\theta}). \tag{2}$$

Censored System Lifetime

- Λ_k denotes the number of failed components in system k before time τ_k, $\Lambda_k \in \{0, 1, \ldots, n-1\}$.
- Observed data is $U_k = \tau_k$
- Missing data is $\mathbf{V}_k = (V_{k1}, \mathbf{V}_{k2})$, where $V_{k1} = \Lambda_k$ and $\mathbf{V}_{k2} = (X_{k1}, \ldots, X_{k\Lambda_k}, X_{k(\Lambda_k+1)}, \ldots, X_{kn})$.
- Distributions of the missing data X_{kj}, given observed data τ_k are

$$f(x_{kj}|\boldsymbol{\theta}) = \begin{cases} \frac{f_X(x_{kj}|\boldsymbol{\theta})}{F_X(\tau_k|\boldsymbol{\theta})} & \text{if } j \leq \Lambda_k, \\ \frac{f_X(x_{kj}|\boldsymbol{\theta})}{\bar{F}_X(\tau_k|\boldsymbol{\theta})} & \text{if } j > \Lambda_k. \end{cases} \tag{3}$$

Censored System Lifetime

- Distribution of Λ

$$\begin{aligned} & \Pr(\Lambda = \lambda | T_k > \tau_k, \boldsymbol{\theta}) \\ = \quad & \frac{(1 - \sum_{l=1}^{\lambda} s_l) \binom{n}{\lambda} F_X(\tau_k|\boldsymbol{\theta})^{\lambda} \bar{F}_X(\tau_k|\boldsymbol{\theta})^{n-\lambda}}{\bar{F}_T(\tau_k|\boldsymbol{\theta})}. \end{aligned} \tag{4}$$

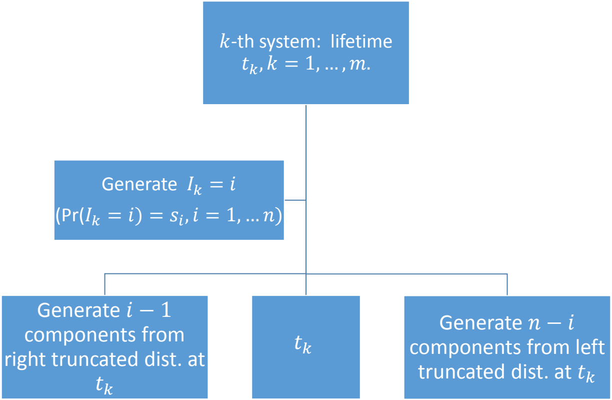
- The log-likelihood function of the complete data $(U_k, V_{k1}, \mathbf{V}_{k2})$ is

$$\ell_k(\boldsymbol{\theta}|U_k, V_{k1}, \mathbf{V}_{k2}) = \sum_{j=1}^{\Lambda_k} \ln f_X(x_{kj}|\boldsymbol{\theta}) + \sum_{j=\Lambda_k+1}^{n} \ln f_X(x_{kj}|\boldsymbol{\theta}).$$

Implementations of the SEM Algorithm

- Complete system lifetime sample: $T_1, \ldots, T_m$
  - SEM based on ordinary system signature (SEM-SS)
  - SEM based on ordered system signature (SEM-OSS)
- Type-II censored system lifetime sample: $T_1, \ldots, T_r,\ r < m$
  - Impute component lifetime directly (SEM-I)
  - Impute the censored system lifetime (SEM-II)

SEM-SS



k-th system: lifetime $t_k, k = 1, \dots, m$.
Generate $I_k = i$ $(\Pr(I_k = i) = s_i, i = 1, \dots n)$
Generate $i - 1$ components from right truncated dist. at t_k
t_k
Generate $n - i$ components from left truncated dist. at t_k

SEM Algorithm Based on System Signature (SEM-SS)

S-step Given current estimate $\boldsymbol{\theta}^{(h)}$, the $(h+1)$-th iteration

1. For k-th system, generate a discrete random variable I_k with probability $\Pr(I_k = i) = s_i$, $i = 1, 2, \ldots, n$, $k = 1, \ldots, m$;
2. Generate $i-1$ random variates from the right truncated distribution in Eq. (1) with $\boldsymbol{\theta} = \boldsymbol{\theta}^{(h)}$, say $\tilde{x}_{k1}, \tilde{x}_{k2}, \ldots, \tilde{x}_{k(i-1)}$;
3. Generate $n-i$ random variates from the left truncated distribution in Eq. (1) with $\boldsymbol{\theta} = \boldsymbol{\theta}^{(h)}$, say $\tilde{x}_{k(i+1)}, \tilde{x}_{k(i+2)}, \ldots, \tilde{x}_{kn}$.
4. The pseudo-complete sample for system k is then

$$\tilde{\mathbf{x}}_k = (\tilde{x}_{k1}, \tilde{x}_{k2}, \ldots, \tilde{x}_{k(i-1)}, \tilde{x}_{ki}, \tilde{x}_{k(i+1)}, \tilde{x}_{k(i+2)}, \ldots, \tilde{x}_{kn})'$$

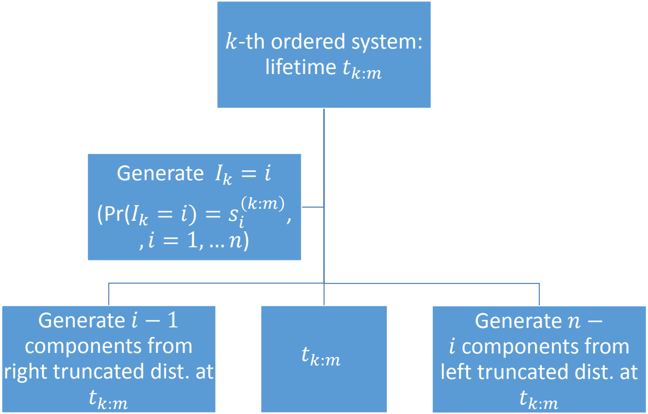
where $\tilde{x}_{ki} = t_k$. Repeat Steps 1 - 3 for $k = 1, \ldots, m$.

M-step Maximize the log-likelihood function,

$$\ln L(\boldsymbol{\theta} | \tilde{\mathbf{x}}_1, \tilde{\mathbf{x}}_2, \ldots, \tilde{\mathbf{x}}_m) = \sum_{k=1}^{m} \sum_{i=1}^{n} \ln f_X(\tilde{x}_{ki} | \boldsymbol{\theta}),$$

with respect to $\boldsymbol{\theta}$ to obtain $\boldsymbol{\theta}^{(h+1)}$ for the next cycle.

SEM-OSS


k-th ordered system: lifetime $t_{k:m}$
Generate $I_k = i$ $(\Pr(I_k = i) = s_i^{(k:m)}, , i = 1, \dots n)$
Generate $i - 1$ components from right truncated dist. at $t_{k:m}$
$t_{k:m}$
Generate $n - i$ components from left truncated dist. at $t_{k:m}$

SEM Algorithm Based on the Ordered System Signature (SEM-OSS)

- The ordered system signature is approximated by

$$\hat{s}_i^{(k:m)} = \frac{\text{no. of times that event } A \text{ occured}}{N},$$

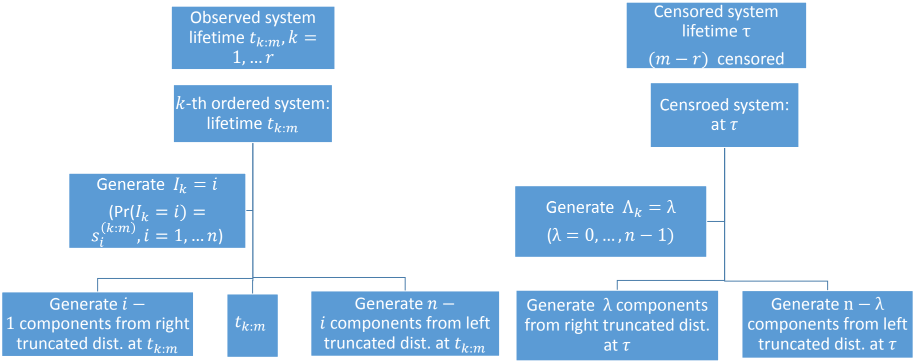
where $A =$
$\{$the k-th ordered system failed due to its i-th component failure$\}$,
for $i = 1, 2, \ldots, n$, $k = 1, 2, \ldots, m$.

S-Step

1.* For the k-th ordered system , generate a discrete random variable I_k with probability $\Pr(I_k = i) = \hat{s}_i^{(k:m)}$, $i = 1, 2, \ldots, n$

Outline

SEM-I

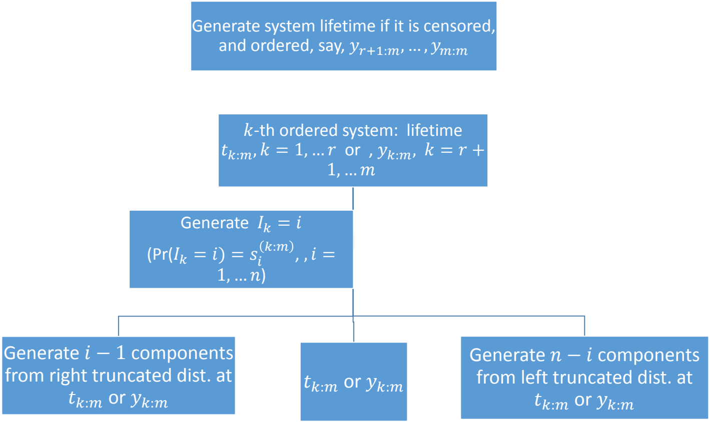

Observed system lifetime $t_{k:m}, k = 1, \dots r$
k-th ordered system: lifetime $t_{k:m}$
Generate $I_k = i$ $(\Pr(I_k = i) = s_i^{(k:m)}, i = 1, \dots n)$
Generate $i -$ 1 components from right truncated dist. at $t_{k:m}$
$t_{k:m}$
Generate $n -$ i components from left truncated dist. at $t_{k:m}$
Censored system lifetime τ
$(m - r)$ censored
Censroed system: at τ
Generate $\Lambda_k = \lambda$ $(\lambda = 0, \dots, n-1)$
Generate λ components from right truncated dist. at τ
Generate $\mathrm{n} - \lambda$ components from left truncated dist. at τ

Impute Component Lifetimes Directly (SEM-I)

S-step

1. For the failed systems, impute component lifetimes based on ordered system signatures;
2. For the censored system with rank j, $j = r+1, r+2, \ldots, m$, the component lifetimes can be imputed as follows:
   - **i.** Generate a random variate Λ, given $\tau = t_{r:m}$, with probability $\Pr(\Lambda = \lambda | T > \tau, \boldsymbol{\theta}^{(h)})$, $\lambda = 0, 1, \ldots, n-1$ in Eq. (4)
   - **ii.** Generate λ random variates from the right truncated distribution in Eq. (3) with $t_k = \tau$ and $\boldsymbol{\theta} = \boldsymbol{\theta}^{(h)}$, say $\tilde{x}_{j1}, \tilde{x}_{j2}, \ldots, \tilde{x}_{j\lambda}$;
   - **iii.** Generate $n - \lambda$ random variates from the left truncated distribution in Eq. (3) with $t_k = \tau$ and $\boldsymbol{\theta} = \boldsymbol{\theta}^{(h)}$, say $\tilde{x}_{j(\lambda+1)}, \tilde{x}_{j(\lambda+2)}, \ldots, \tilde{x}_{jn}$;
   - **iv.** The pseudo-complete sample for j-th ranked censored system is then
   $$\tilde{\mathbf{x}}_j = (\tilde{x}_{j1}, \tilde{x}_{j2}, \ldots, \tilde{x}_{jn});$$
   - **v.** Repeat Steps **i** – **iv** until we have the imputed samples $\tilde{\mathbf{x}}_j, j = r+1, r+2, \ldots, m$.

SEM-II



Generate system lifetime if it is censored, and ordered, say, $y_{r+1:m}, \dots, y_{m:m}$
k-th ordered system: lifetime $t_{k:m}, k = 1, \dots r$ or , $y_{k:m}$, $k = r + 1, \dots m$
Generate $I_k = i$
$(\Pr(I_k = i) = s_i^{(k:m)}, , i = 1, \dots n)$
Generate $i - 1$ components from right truncated dist. at $t_{k:m}$ or $y_{k:m}$
$t_{k:m}$ or $y_{k:m}$
Generate $n - i$ components from left truncated dist. at $t_{k:m}$ or $y_{k:m}$

Impute the Censored System Lifetimes (SEM-II)

- Generate the $(m-r)$ censored system lifetimes from a left-truncated distribution with c.d.f.

$$H_T(y|\boldsymbol{\theta}) = \frac{F_T(y|\boldsymbol{\theta}) - F_T(\tau|\boldsymbol{\theta})}{1 - F_T(\tau|\boldsymbol{\theta})}, y > \tau. \tag{5}$$

- Treat the sample $(t_{1:m} < t_{2:m} < \ldots < t_{r:m} < y_{r+1:m} < y_{r+2:m} < \ldots < y_{m:m})$ as a pseudo-complete system lifetime data, and apply SEM-OSS.

Monte Carlo Simulation

- The log-component lifetimes are generated from SEV distribution with $\mu = 0$, and $\sigma = 1$
- System signature considered fixed as $sys1 = (1/4, 1/4, 1/5, 0)$, $sys2 = (0, 1/2, 1/4, 1/4)$, and $sys3 = (0, 1/3, 2/3)$
- Sample size $m = 10, 20, 30, 50, 100$

Monte Carlo Simulation: SEM Algorithm for Complete Sample

Table: Estimated distance between SEM estimates and direct numerical MLEs, with different signatures

		μ				σ			
		SEM-SS vs. MLE		SEM-OSS vs. MLE		SEM-SS vs. MLE		SEM-OSS vs. MLE	
s	m	Diff.	MSD	Diff.	MSD	Diff.	MSD	Diff.	MSD
$sys1$	10	0.1135	0.0155	0.0009	0.0004	0.2829	0.0954	-0.0015	0.0010
	20	0.1246	0.0168	0.0006	0.0002	0.3342	0.1219	0.0003	0.0006
	30	0.1293	0.0176	0.0003	0.0001	0.3562	0.1344	0.0009	0.0004
	50	0.1330	0.0182	0.0002	0.0001	0.3730	0.1440	0.0006	0.0003
	100	0.1353	0.0186	0.0001	0.0000	0.3854	0.1511	0.0004	0.0001
$sys2$	10	-0.0399	0.0020	-0.0145	0.0004	-0.0142	0.0007	-0.0327	0.0015
	20	0.0851	0.0077	0.0123	0.0006	0.3244	0.1084	0.0047	0.0003
	30	0.0840	0.0074	0.0080	0.0004	0.3447	0.1211	0.0040	0.0002
	50	0.0829	0.0071	0.0050	0.0002	0.3603	0.1313	0.0028	0.0001
	100	0.0820	0.0068	0.0025	0.0001	0.3724	0.1395	0.0016	0.0000
$sys3$	10	0.0308	0.0012	0.0112	0.0002	0.1092	0.0131	-0.0150	0.0006
	20	0.0281	0.0009	0.0063	0.0001	0.1425	0.0211	-0.0061	0.0002
	30	0.0271	0.0008	0.0041	0.0001	0.1547	0.0245	-0.0034	0.0001
	50	0.0261	0.0007	0.0026	0.0000	0.1642	0.0273	-0.0017	0.0000
	100	0.0255	0.0007	0.0012	0.0000	0.1723	0.0299	-0.0007	0.0000

Monte Carlo Simulation: Censored System Lifetime

Table: Estimated distance between SEM and direct numerical MLEs, for censored sample with signature ($\mathbf{s} = (1/4, 1/4, 1/2, 0)$)

			μ				σ			
			SEM-I vs. MLE		SEM-II vs. MLE		SEM-I vs. MLE		SEM-II vs. MLE	
s	m	q	Diff.	MSD	Diff.	MSD	Diff.	MSD	Diff.	MSD
sys1	10	10%	-0.0187	0.0008	-0.0069	0.0005	0.0098	0.0013	0.0147	0.0013
		20%	-0.0312	0.0017	-0.0077	0.0006	-0.0007	0.0011	0.0014	0.0009
		30%	-0.0496	0.0038	-0.0173	0.0011	-0.0135	0.0013	-0.0231	0.0014
		40%	-0.0753	0.0082	-0.0412	0.0033	-0.0307	0.0023	-0.0607	0.0050
		50%	-0.1119	0.0176	-0.0869	0.0168	-0.0540	0.0050	-0.1130	0.0174
	30	20%	-0.0232	0.0008	-0.0015	0.0002	-0.0064	0.0005	-0.0065	0.0004
		30%	-0.0399	0.0020	-0.0145	0.0004	-0.0142	0.0007	-0.0327	0.0015
		40%	-0.0632	0.0048	-0.0467	0.0025	-0.0251	0.0013	-0.0768	0.0065
		50%	-0.0941	0.0104	-0.1134	0.0136	-0.0383	0.0024	-0.1439	0.0220
	50	20%	-0.0214	0.0006	-0.0001	0.0001	-0.0075	0.0004	-0.0080	0.0003
		30%	-0.0379	0.0017	-0.0139	0.0003	-0.0144	0.0005	-0.0347	0.0015

Conclusions

- For the complete data case, the SEM-OSS algorithm results in estimates very close to the direct numerical MLEs.
- For Type-II censored data case, SEM-I algorithm and SEM-II algorithm perform similarly in the sense that both approximate the MLEs very well.
- SEM-II algorithm runs slower than SEM-I algorithm.
- When censoring proportion is small to moderate (say, $q \leq 40\%$), SEM-II algorithm is recommended.
- When censoring proportion is large (say, $q > 40\%$), SEM-I algorithm is recommended.

Unknown System Structure

- We can relax the assumption that the system structure is known in the statistical analysis of coherent systems.
- We can estimate parameter vector $\boldsymbol{\theta}$ and the system structure (system signature) when we have the information that the systems are coherent systems.
- Suppose m independent n-component systems with the same structure are placed on a life-testing experiment and the censoring time for the k-th system is τ_k.
- That is, the experiment is terminated at $T_k^* = \min\{T_k,\ \tau_k\}$ for the k-th system.

Unknown System Structure: Right-Censored Data

- Let d_k be the censoring indicator, i.e., $d_k = 1$ if $T_k < \tau_k$ and $d_k = 0$ otherwise.
- The observation can be expressed as $T_k^* = d_k T_k + (1 - d_k)\tau_k$, $k = 1, 2, \ldots, m$, and we denote the observed values as $t_k^* = d_k t_k + (1 - d_k)\tau_k$, $k = 1, 2, \ldots, m$.
- The log-likelihood based on the observation $(\mathbf{t}^*, \mathbf{d}) = ((t_1^*, d_1), \ldots, (t_m^*, d_m))$ is

$$\ell(\boldsymbol{\theta}, \mathbf{s}) = \sum_{k=1}^{m} d_k \ln\left(\sum_{i=1}^{n} s_i f_{i:n}(t_k|\boldsymbol{\theta})\right) + \sum_{k=1}^{m} (1 - d_k) \ln\left(\sum_{i=1}^{n} s_i \bar{F}_{i:n}(\tau_k|\boldsymbol{\theta})\right). \tag{6}$$

Unknown System Structure: Right-Censored Data

- Directly maximizing Eq.(6) with respect to $\boldsymbol{\theta}$ and **s** at the same time is not an easy task.
- Consider an iterative procedure in which we fix one parameter vector at a time and maximize the log-likelihood function with respect to the other parameter vector.
- Fixed the parameter vector $\boldsymbol{\theta}$, maximize the log-likelihood function with respect to **s** with the constraint $\sum_{i=1}^{n} s_i = 1$:

$$\sum_{k=1}^{m} d_k \ln\left(\sum_{i=1}^{n} s_i f_{i:n}(t_k|\boldsymbol{\theta})\right) + \sum_{k=1}^{m} (1-d_k) \ln\left(\sum_{i=1}^{n} s_i \bar{F}_{i:n}(\tau_k|\boldsymbol{\theta})\right) - \upsilon\left(\sum_{i=1}^{n} s_i - 1\right), \tag{7}$$

where υ is the Lagrange multiplier.

Unknown System Structure: Right-Censored Data

- Take the derivative of Eq.(7) with respect to s_i, and set the derivative to 0, we have

$$\sum_{k=1}^{m} d_k \frac{f_{i:n}(t_k|\boldsymbol{\theta})}{\sum_{l=1}^{n} s_l f_{l:n}(t_k|\boldsymbol{\theta})} + \sum_{k=1}^{m} (1-d_k) \frac{\bar{F}_{i:n}(\tau_k|\boldsymbol{\theta})}{\sum_{l=1}^{n} s_l \bar{F}_{l:n}(\tau_k|\boldsymbol{\theta})} - \upsilon = 0, \tag{8}$$

for $i = 1, \ldots, n$.

- Multiply both sides of Eq.(8) by s_i, and sum over all s_i, for $i = 1, \ldots, n$, we can obtain $\upsilon = m$ and

$$s_i = \frac{1}{m} \left(\sum_{k=1}^{m} d_k \frac{s_i f_{i:n}(t_k|\boldsymbol{\theta})}{\sum_{l=1}^{n} s_l \ f_{l:n}(t_k|\boldsymbol{\theta})} + \sum_{k=1}^{m} (1-d_k) \frac{s_i \bar{F}_{i:n}(\tau_k|\boldsymbol{\theta})}{\sum_{l=1}^{n} s_l \bar{F}_{l:n}(\tau_k|\boldsymbol{\theta})} \right), \tag{9}$$

for $i = 1, \ldots, n$.

Unknown System Structure: Right-Censored Data

- Fixed the system signature **s**, maximize the log-likelihood function with respect to $\boldsymbol{\theta}$:

$$\begin{aligned}\frac{\partial \ell(\boldsymbol{\theta}, \mathbf{s})}{\partial \boldsymbol{\theta}} &= \sum_{k=1}^{m} d_k \left(\sum_{i=1}^{n} \delta_{ki}(\boldsymbol{\theta}, \mathbf{s}) \frac{\partial \ln f_{i:n}(t_k|\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right) \\ &+ \sum_{k=1}^{m} (1 - d_k) \left(\sum_{i=1}^{n} \tilde{\delta}_{ki}(\boldsymbol{\theta}, \mathbf{s}) \frac{\partial \ln \bar{F}_{i:n}(\tau_k|\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right).\end{aligned}$$

where

$$\delta_{ki}(\boldsymbol{\theta}, \mathbf{s}) = \frac{s_i f_{i:n}(t_k|\boldsymbol{\theta})}{\sum_{l=1}^{n} s_l f_{l:n}(t_k|\boldsymbol{\theta})}, \text{ and } \tilde{\delta}_{ki}(\boldsymbol{\theta}, \mathbf{s}) = \frac{s_i \bar{F}_{i:n}(\tau_k|\boldsymbol{\theta})}{\sum_{l=1}^{n} s_l \bar{F}_{l:n}(\tau_k|\boldsymbol{\theta})},$$

for $k = 1, \ldots, m, \ i = 1, \ldots, n$.

Outline

Unknown System Structure: EM-algorithm

E-step: Evaluate the conditional probability of latent variable $\mathbf{I}$ as

$$\begin{aligned}
\Pr(I_k = i | T_k = t_k, \mathbf{s}^{(h)}, \boldsymbol{\theta}^{(h)}) &= \delta_{ki}(\boldsymbol{\theta}^{(h)}, \mathbf{s}^{(h)}) && (10)\\
&= \frac{s_i^{(h)} f_{i:n}(t_k|\boldsymbol{\theta}^{(h)})}{\sum_{l=1}^{n} s_l^{(h)} f_{l:n}(t_k|\boldsymbol{\theta}^{(h)})} \text{ for } d_k = 1, \\
\Pr(I_k = i | T_k > \tau_k, \mathbf{s}^{(h)}, \boldsymbol{\theta}^{(h)}) &= \tilde{\delta}_{ki}(\boldsymbol{\theta}^{(h)}, \mathbf{s}^{(h)}) && (11)\\
&= \frac{s_i^{(h)} \bar{F}_{i:n}(\tau_k|\boldsymbol{\theta}^{(h)})}{\sum_{l=1}^{n} s_l^{(h)} \bar{F}_{l:n}(\tau_k|\boldsymbol{\theta}^{(h)})} \text{ for } d_k = 0,
\end{aligned}$$

$k = 1, 2, \ldots, m$, $i = 1, 2, \ldots, n$.

- Then, we obtain the expectation of the complete-data on system-level log-likelihood as

$$\begin{aligned}
& Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{(h)}, \mathbf{s}^{(h)}) && (12)\\
= & \sum_{k=1}^{m} \sum_{i=1}^{n} d_k \Pr(I_k = i | T_k = t_k, \mathbf{s}^{(h)}, \boldsymbol{\theta}^{(h)}) \ln f_{i:n}(T_k = t_k | \boldsymbol{\theta}) \\
& + \sum_{k=1}^{m} \sum_{i=1}^{n} (1 - d_k) \Pr(I_k = i | T_k > \tau_k, \mathbf{s}^{(h)}, \boldsymbol{\theta}^{(h)}) \ln \bar{F}_{i:n}(\tau_k | \boldsymbol{\theta}).
\end{aligned}$$

Unknown System Structure: EM-algorithm

M-step: Step 1. Maximize $Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{(h)}, \mathbf{s}^{(h)})$ with respect to $\boldsymbol{\theta}$ to obtain $\boldsymbol{\theta}^{(h+1)}$.
Step 2. The updated estimate of $\mathbf{s}$ can be obtained by

$$\begin{aligned} s_i^{(h+1)} &= \frac{1}{m} \sum_{k=1}^{m} d_k \frac{s_i^{(h)} f_{i:n}(t_k|\boldsymbol{\theta}^{(h)})}{\sum_{l=1}^{n} s_l^{(h)} f_{l:n}(t_k|\boldsymbol{\theta}^{(h)})} \\ &+ \frac{1}{m} \sum_{k=1}^{m} (1-d_k) \frac{s_i^{(h)} \bar{F}_{i:n}(\tau_k|\boldsymbol{\theta}^{(h)})}{\sum_{l=1}^{n} s_l^{(h)} \bar{F}_{l:n}(\tau_k|\boldsymbol{\theta}^{(h)})}, \end{aligned}$$

for $i = 1, \ldots, n$.

- The above E-step and M-step are repeated until convergence occurs.

Unknown System Structure: EM-algorithm

Table: Average of the MLEs obtained by EM algorithms with unknown system signature based on 5,000 simulations from Weibull(e^{μ}, $1/\sigma$)

$\boldsymbol{\theta}$	**s**	q	$\hat{\mu}$	$\hat{\sigma}$	estimate of system signature
(0, 1)	(.25, .25, .5, 0)	0%	0.0025	1.1675	(0.3719, 0.0054, 0.6226, 0.0000)
		10%	0.0081	1.1915	(0.3660, 0.0064, 0.6276, 0.0000)
		30%	0.1365	1.3421	(0.3554, 0.0317, 0.6129, 0.0000)
(0, 1)	(.4, .3, .2, .1)	0%	0.0162	1.1644	(0.3656, 0.0385, 0.5924, 0.0034)
		10%	-0.7414	1.1930	(0.2598, 0.0068, 0.1431, 0.5902)
		30%	-0.6633	1.3367	(0.2183, 0.0797, 0.1659, 0.5361)
(2, 3)	(.25, .25, .5, 0)	0%	2.0011	3.1678	(0.3729, 0.0000, 0.6271, 0.0000)
		10%	2.0064	3.1913	(0.3667, 0.0032, 0.6302, 0.0000)
		30%	2.1251	3.3414	(0.3564, 0.0197, 0.6238, 0.0000)
(2, 3)	(.4, .3, .2, .1)	0%	2.0012	3.1676	(0.3721, 0.0041, 0.6238, 0.0000)
		10%	2.2599	3.1931	(0.2617, 0.0010, 0.1497, 0.5876)
		30%	2.6645	3.3371	(0.2179, 0.0800, 0.1649, 0.5371)

Outline

Unknown System Structure: coherent systems

- In the study of the coherent system, there are only finite coherent systems if the number of components is given.
- For example, if a system is a 3-component coherent system, then there are only 5 possible arrangements:

Table: 3-component coherent systems

System	System Lifetime	System Signature
Series	min $(X_1, X_2, X_3){=}X_{1:3}$	(1, 0, 0)
Series-parallel	min $(X_1$, max $(X_2, X_3))$	(1/3, 2/3, 0)
2-out-of-3	max$_{1\leq i<i\leq 3}$(min $(X_i, X_j)){=}X_{2:3}$	(0, 1, 0)
Parallel-series	max $(X_1$, min $(X_2, X_3))$	(0, 2/3, 1/3)
Parallel	max $(X_1, X_2, X_3){=}X_{3:3}$	(0,0,1)

Monte Carlo Simulation: 3-component coherent system

Table: The proportion of correct detecting system signature when the sample size is 10 or 100

$m = 10$			Selected		
Simulated	Series	Series-Parallel	2-out-of-3	Parallel-Series	Parallel
Series	0.295	0.292	0.076	0.033	0.304
Series-Parallel	0.285	0.313	0.075	0.037	0.290
2-out-of-3	0.271	0.195	0.081	0.044	0.409
Parallel-Series	0.239	0.171	0.082	0.045	0.462
Parallel	0.122	0.227	0.123	0.044	0.484
$m = 100$			Selected		
Simulated	Series	Series-Parallel	2-out-of-3	Parallel-Series	Parallel
Series	0.392	0.349	0.193	0.027	0.039
Series-Parallel	0.119	0.525	0.131	0.030	0.195
2-out-of-3	0.211	0.079	0.329	0.135	0.247
Parallel-Series	0.093	0.030	0.258	0.188	0.431
Parallel	0.057	0.021	0.227	0.163	0.532

Unknown System Structure: coherent systems

- The values in tables represent the percentage that the data sampled from the row category have the largest likelihood value if assuming it is sampled from the column category.
- The preliminary results are promising because the EM algorithm can correctly detecting which system the data is generated from when the sample size is large.
- There are rooms for improvement in the performance of the EM algorithm in identifying the correct system for small sample size situations.

Future Research Directions

- The SEM algorithm could be generalized to handle the case when the system signature is unknown.
- The nonparametric and semiparametric models can be applied to the analysis of system lifetime distribution by applying the EM and SEM algorithms
- Resampling approaches can be used to adjust for the detecting based on likelihood method when system signature is unknown.
- Extend the SEM algorithm to INID case or non-IID case (see, for example, Navarro, Ng and Balakrishnan, 2011, Naval Research Logistics).

Major References

Yang, Y., Ng, H. K. T. and Balakrishnan, N. (2016). A stochastic expectation-maximization algorithm for the analysis of system lifetime data with known signature, *Computational Statistics*, 31, 609641.

Yang, Y., Ng, H. K. T. and Balakrishnan, N.. Expectation-Maximization Algorithm for System-based Lifetime Data with Unknown System Structure, to appear in *AStA Advances in Statistical Analysis*.

References

DAVID, H.A, NAGARAJA, H.N. (2003), Order Statistics, 2nd Edition,John Wiley. &Sons, Hoboken, New Jersey.

BALAKRISHNAN, N., Ng, H. K. T., and NAVARRO, J. (2011b), Linear inference for Type-II censored lifetime data of reliability systems with known signatures, *IEEE Transactions on Reliability*, 60, $426 - 440$.

BALAKRISHNAN, N., AND VOLTERMAN, W. On the signatures of ordered system lifetimes. *Journal of Applied Probability* 51, (03 2014), 82-91.

BIRNBAUM, Z., AND SAUNDERS, S. C. A new family of life distributions. *Journal of applied probability* (1969), $319 - 327$.

CELEUX, G., AND DIEBOLT, J. The SEM algorithm: a probabilistic teacher algorithm derived from the EM algorithm for the mixture problem. *Computational statistics quarterly* 2, (1985), $73 - 82$.

DEMPSTER, A.P., LAIRD, N.M., AND RUBIN, D.B. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society Series B*, 39(1977), $1 - 38$.

FEODOR NIELSEN, S., ET AL. The stochastic EM algorithm: estimation and asymptotic results. *Bernoulli* 6, (2000), $457 - 489$.

KOCHAR, S., MUKERJEE, H., AND SAMANIEGO, F. J. The signature of a coherent system and its application to comparisons among systems. *Naval Research Logistics* 46, (1999),$507 - 523$.

NAVARRO, J., NG, H. K. T. and BALAKRISHNAN, N., Parametric Inference for Component Distributions from Lifetimes of Systems with Dependent Components, *Naval Research Logistics* 59 (2012), 487-496.

NAVARRO, J., RUIZ, J. M., AND SANDOVAL, C. J. Properties of coherent systems with dependent components. *Communications in Statistics Theory and Methods* 36, (2007), $175 - 191$.

SAMANIEGO, F. J. On closure of the IFR class under formation of coherent systems. *IEEE Transactions on Reliability* 34, (1985), $69 - 72$.

SAMANIEGO, F. J. *System signatures and their applications in engineering reliability*, vol. 110. Springer, 2007.

WU, C. F. J. On the convergence properties of the EM algorithm. *The Annals of Statistics*, 11, (1983), $95 - 103$.

Thank you!!!