Ordered Data

Jorge Navarro[1]
Universidad de Murcia, Murcia, Spain.

[1]Supported by Ministerio de Ciencia e Innovación of Spain in the project PID2022-137396NB-I00, funded by MICIU/AEI/10.13039/501100011033 and by 'ERDF A way of making Europe'.

References

The talk is based on the following references:

- Navarro J, Arevalillo JM (2023). On connections between skewed, weighted and distorted distributions: applications to model extreme value distributions. Test 32, 1307–1335.
- Navarro J., Cramer E., Balakrishnan N. (2025). Inference under conditionally ordered heterogeneous exponential data. Submitted.
- Lagos G., Navarro J., Olivero H. (2025). Repair policies decomposition for monotone systems with a Lévy frailty Marshall-Olkin Process. In preparation.

Outline

Theoretical results
- Bivariate case
- Multivariate case
- Generalized Order Statistics

Inferencial results
- Case 1: First data (series system)
- Case 2: Second data (parallel system)
- Properties for systems under MO-model.

References

Theoretical results

Notation

- X and Y independent random variables (lifetimes).

Notation

- X and Y independent random variables (lifetimes).
- $F_X(t) = \Pr(X \leq t)$ and $F_Y(t) = \Pr(Y \leq t)$ cumulative distribution functions (CDF).

Notation

- X and Y independent random variables (lifetimes).
- $F_X(t) = \Pr(X \leq t)$ and $F_Y(t) = \Pr(Y \leq t)$ cumulative distribution functions (CDF).
- $\bar{F}_X(t) = 1 - F_X(t) = \Pr(X > t)$ and $\bar{F}_Y(t) = \Pr(Y > t)$ survival (or reliability) functions (SF).

Notation

- X and Y independent random variables (lifetimes).
- $F_X(t) = \Pr(X \leq t)$ and $F_Y(t) = \Pr(Y \leq t)$ cumulative distribution functions (CDF).
- $\bar{F}_X(t) = 1 - F_X(t) = \Pr(X > t)$ and $\bar{F}_Y(t) = \Pr(Y > t)$ survival (or reliability) functions (SF).
- We assume that X and Y have absolutely continuous distributions with probability density functions (PDF) $f_X = F'_X$ and $f_Y = F'_Y$.

Ordered data-Bivariate case

- If they are independent, the joint PDF of (X, Y) is

$$f(x,y) = f_X(x)f_Y(y).$$

For results with dependency see Navarro et al. (2022a).

Ordered data-Bivariate case

- If they are independent, the joint PDF of (X,Y) is
$$f(x,y) = f_X(x)f_Y(y).$$
For results with dependency see Navarro et al. (2022a).
- If we assume that $X<Y$, the joint PDF of $(X,Y|X<Y)$ is
$$g(x,y) = c\ f_X(x)f_Y(y),\ x \le y,$$
where $c = 1/\Pr(X<Y)$.

Ordered data-Bivariate case

- If they are independent, the joint PDF of (X, Y) is

$$f(x, y) = f_X(x) f_Y(y).$$

For results with dependency see Navarro et al. (2022a).

- If we assume that $X < Y$, the joint PDF of $(X, Y|X < Y)$ is

$$g(x, y) = c\ f_X(x) f_Y(y),\ x \leq y,$$

where $c = 1/\Pr(X < Y)$.

- The first marginal PDF of $(X, Y|X < Y)$ is

$$g_1(x) = \int_{-\infty}^{\infty} g(x, y) dy = \int_{x}^{\infty} c\ f_X(x) f_Y(y) dy = c\ f_X(x) \bar{F}_Y(x).$$

Ordered data-Bivariate case

- Analogously, the second marginal PDF of $(X, Y|X < Y)$ is

$$g_2(y) = \int_{-\infty}^{\infty} g(x, y)dx = \int_{-\infty}^{y} c\, f_X(x)f_Y(y)dx = c\, f_Y(y)F_X(y).$$

Ordered data-Bivariate case

- Analogously, the second marginal PDF of $(X, Y|X < Y)$ is

$$g_2(y) = \int_{-\infty}^{\infty} g(x, y)dx = \int_{-\infty}^{y} c\, f_X(x)f_Y(y)dx = c\, f_Y(y)F_X(y).$$

- Therefore, the marginal r.v. of $(X, Y|X < Y)$ are not independent.

Ordered data-Bivariate case

- Analogously, the second marginal PDF of $(X, Y|X < Y)$ is

$$g_2(y) = \int_{-\infty}^{\infty} g(x, y)dx = \int_{-\infty}^{y} c\, f_X(x)f_Y(y)dx = c\, f_Y(y)F_X(y).$$

- Therefore, the marginal r.v. of $(X, Y|X < Y)$ are not independent.
- These marginal PDF are skew versions of f_X and f_Y.

Ordered data-Bivariate case

- Analogously, the second marginal PDF of $(X,Y|X<Y)$ is

$$g_2(y)=\int_{-\infty}^{\infty}g(x,y)dx=\int_{-\infty}^{y}c\,f_X(x)f_Y(y)dx=c\,f_Y(y)F_X(y).$$

- Therefore, the marginal r.v. of $(X,Y|X<Y)$ are not independent.
- These marginal PDF are skew versions of f_X and f_Y.
- The PDF $g_2=c\,f_YF_X$ is the **right skewed distribution** associated to f_Y and F_X, as defined in Navarro and Arevalillo (Test, 2023).

Ordered data-Bivariate case

- Analogously, the second marginal PDF of $(X, Y|X < Y)$ is

$$g_2(y) = \int_{-\infty}^{\infty} g(x, y)dx = \int_{-\infty}^{y} c\, f_X(x)f_Y(y)dx = c\, f_Y(y)F_X(y).$$

- Therefore, the marginal r.v. of $(X, Y|X < Y)$ are not independent.
- These marginal PDF are skew versions of f_X and f_Y.
- The PDF $g_2 = c\, f_Y F_X$ is the **right skewed distribution** associated to f_Y and F_X, as defined in Navarro and Arevalillo (Test, 2023).
- The PDF $g_1 = c\, f_X \bar{F}_Y$ is the **left skewed distribution** associated to f_X and $\bar{F}_Y$, as defined in Navarro and Arevalillo (Test, 2023).

Ordered data-Particular cases: IID

- If they are identically distributed $f_X = f_Y = f$, then $c = 2$,

$$g_1(x) = 2\ f(x)\bar{F}(x).$$

and

$$g_2(y) = 2\ f(y)F(y).$$

Ordered data-Particular cases: IID

- If they are identically distributed $f_X = f_Y = f$, then $c = 2$,

$$g_1(x) = 2\, f(x)\bar{F}(x).$$

and

$$g_2(y) = 2\, f(y)F(y).$$

- Therefore, $\bar{G}_1(x) = \bar{F}^2(x)$, which is the survival function of $\min(X, Y)$ and $G_2(x) = F^2(x)$, which is the CDF of $\max(X, Y)$.

Ordered data-Particular cases: IID

- If they are identically distributed $f_X = f_Y = f$, then $c = 2$,

$$g_1(x) = 2\ f(x)\bar{F}(x).$$

and

$$g_2(y) = 2\ f(y)F(y).$$

- Therefore, $\bar{G}_1(x) = \bar{F}^2(x)$, which is the survival function of $\min(X, Y)$ and $G_2(x) = F^2(x)$, which is the CDF of $\max(X, Y)$.
- Note that both are distorted distributions, that is, $G_i = q_i(F)$.

Ordered data-Particular cases: IID

- If they are identically distributed $f_X = f_Y = f$, then $c = 2$,

$$g_1(x) = 2\ f(x)\bar{F}(x).$$

and

$$g_2(y) = 2\ f(y)F(y).$$

- Therefore, $\bar{G}_1(x) = \bar{F}^2(x)$, which is the survival function of $\min(X, Y)$ and $G_2(x) = F^2(x)$, which is the CDF of $\max(X, Y)$.
- Note that both are distorted distributions, that is, $G_i = q_i(F)$.
- Similar representations are obtained for series and parallel systems with dependent components in Navarro and Arevalillo (Test, 2023) and for general coherent systems in

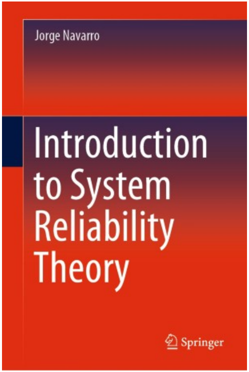


Figure: Publicity of my book on System Reliability Theory.

Ordered data-Particular cases: Skew normal

- If $f_X = \phi$ is the standard normal PDF and

$$\bar{F}_Y(x) = 1 - \Phi(\alpha x) = \Phi(-\alpha x),$$

where Φ is the standard normal CDF and $\alpha > 0$, then $c = 2$ and

$$g_1(x) = 2\ \phi(x)\Phi(-\alpha x)$$

is a skew normal distribution, see Azzalini and Capitanio (2014).

Ordered data-Particular cases: Skew normal

- If $f_X = \phi$ is the standard normal PDF and

$$\bar{F}_Y(x) = 1 - \Phi(\alpha x) = \Phi(-\alpha x),$$

where Φ is the standard normal CDF and $\alpha > 0$, then $c = 2$ and

$$g_1(x) = 2\ \phi(x)\Phi(-\alpha x)$$

is a skew normal distribution, see Azzalini and Capitanio (2014).

- In this case, G_1 is the PDF of $\min(X, Y)$ where (X, Y) has a bivariate normal distribution with standarized marginals, correlation ρ and $\alpha = \sqrt{(1-\rho)/(1+\rho)}$, see Theorem 3.2 in Loperfido, Navarro, Ruiz, and Sandoval (2007).

Ordered data-Particular cases: Skew normal

- If $f_Y = \phi$ is the standard normal PDF and

$$F_X(y) = \Phi(\alpha y),$$

where Φ is the standard normal CDF and $\alpha > 0$, then $c = 2$ and

$$g_2(y) = 2\ \phi(y)\Phi(\alpha y)$$

is a skew normal distribution, see Azzalini and Capitanio (2014).

Ordered data-Particular cases: Skew normal

- If $f_Y = \phi$ is the standard normal PDF and

$$F_X(y) = \Phi(\alpha y),$$

where Φ is the standard normal CDF and $\alpha > 0$, then $c = 2$ and

$$g_2(y) = 2\ \phi(y)\Phi(\alpha y)$$

is a skew normal distribution, see Azzalini and Capitanio (2014).

- In this case, G_2 is the PDF of $\max(X, Y)$ where (X, Y) has a bivariate normal distribution with standarized marginals, correlation ρ and $\alpha = \sqrt{(1-\rho)/(1+\rho)}$, see Roberts (1966).

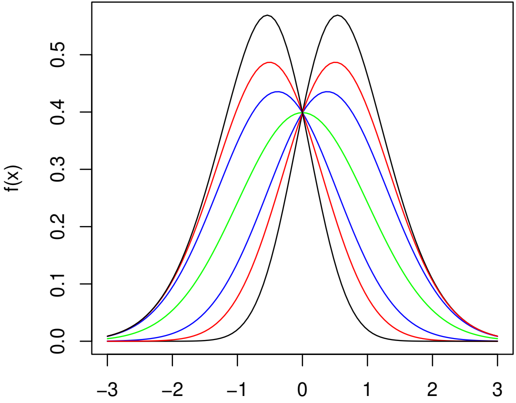


Figure: PDF for $\max(X, Y)$ and $\min(X, Y)$ for $\rho = -0.5$ (black), 0 (red) and 0.5 (blue). The green line represents the standard normal PDF $\rho = 1, \alpha = 0$.

Ordered data-General result under dependency

Proposition (Navarro and Arevalillo, 2023)

Let (X, Y) be an EXC random vector with absolutely continuous copula C and common marginal CDF F and PDF f. Then the PDF of $\max(X, Y)$ *can be written as*

$$f_{2:2}(x) = c_R f(x) G(x)$$

where $G(x) = \partial_1 C(F(x), F(x))$ *and* $c_R = 2$.

Ordered data-General result under dependency

- A similar result holds for $\min(X, Y)$.

Ordered data-General result under dependency

- A similar result holds for $\min(X, Y)$.
- G is not always a CDF.

Ordered data-General result under dependency

- A similar result holds for $\min(X, Y)$.
- G is not always a CDF.
- These results can be extended to the multivariate case, see Navarro and Arevalillo (2023) and Loperfido, Navarro, Ruiz, and Sandoval (2007).

Ordered data-Particular cases: PHR

- If X and Y are independent, $\bar{F}_X = \bar{F}^{\alpha_1}$ and $\bar{F}_Y = \bar{F}^{\alpha_2}$ for $\alpha_1, \alpha_2 > 0$, i.e., they satisfy the Proportional Hazard Rate (PHR) model, then the PDF of $(X|X < Y)$ is

$$g_1(x) = c\ f_X(x)\bar{F}_Y(x) = (\alpha_1 + \alpha_2)\ f(x)\bar{F}^{\alpha_1+\alpha_2-1}(x)$$

with $c = (\alpha_1 + \alpha_2)/\alpha_1$.

Ordered data-Particular cases: PHR

- If X and Y are independent, $\bar{F}_X = \bar{F}^{\alpha_1}$ and $\bar{F}_Y = \bar{F}^{\alpha_2}$ for $\alpha_1, \alpha_2 > 0$, i.e., they satisfy the Proportional Hazard Rate (PHR) model, then the PDF of $(X|X < Y)$ is

$$g_1(x) = c\ f_X(x)\bar{F}_Y(x) = (\alpha_1 + \alpha_2)\ f(x)\bar{F}^{\alpha_1+\alpha_2-1}(x)$$

  with $c = (\alpha_1 + \alpha_2)/\alpha_1$.
- It is both a left skew distribution from f and a distortion of $\bar{F}$ since

$$\bar{G}_1(x) = \bar{F}^{\alpha_1+\alpha_2}(x).$$

Ordered data-Particular cases: PHR

- If X and Y are independent, $\bar{F}_X = \bar{F}^{\alpha_1}$ and $\bar{F}_Y = \bar{F}^{\alpha_2}$ for $\alpha_1, \alpha_2 > 0$, i.e., they satisfy the Proportional Hazard Rate (PHR) model, then the PDF of $(X|X < Y)$ is

$$g_1(x) = c\ f_X(x)\bar{F}_Y(x) = (\alpha_1 + \alpha_2)\ f(x)\bar{F}^{\alpha_1+\alpha_2-1}(x)$$

  with $c = (\alpha_1 + \alpha_2)/\alpha_1$.
- It is both a left skew distribution from f and a distortion of $\bar{F}$ since

$$\bar{G}_1(x) = \bar{F}^{\alpha_1+\alpha_2}(x).$$

- Note that $(X|X < Y)$ belongs to the same PHR model.

Ordered data-Particular cases: PHR

- If X and Y are independent, $\bar{F}_X = \bar{F}^{\alpha_1}$ and $\bar{F}_Y = \bar{F}^{\alpha_2}$ for $\alpha_1, \alpha_2 > 0$, i.e., they satisfy the Proportional Hazard Rate (PHR) model, then the PDF of $(X|X < Y)$ is

$$g_1(x) = c\ f_X(x)\bar{F}_Y(x) = (\alpha_1 + \alpha_2)\ f(x)\bar{F}^{\alpha_1+\alpha_2-1}(x)$$

  with $c = (\alpha_1 + \alpha_2)/\alpha_1$.
- It is both a left skew distribution from f and a distortion of $\bar{F}$ since

$$\bar{G}_1(x) = \bar{F}^{\alpha_1+\alpha_2}(x).$$

- Note that $(X|X < Y)$ belongs to the same PHR model.
- Even more $(X|X < Y)$, $(Y|X > Y)$ and $\min(X, Y)$ have the same distribution!!

Ordered data-Particular cases: PHR

- If X and Y are independent, $\bar{F}_X = \bar{F}^{\alpha_1}$ and $\bar{F}_Y = \bar{F}^{\alpha_2}$ for $\alpha_1, \alpha_2 > 0$, then the PDF of $(Y|X < Y)$ is

$$g_2(y) = c\ f_Y(y)F_X(y) = c\ \alpha_2 f(y)\bar{F}^{\alpha_2-1}(y)(1-\bar{F}^{\alpha_1}(y))$$

where $c = (\alpha_1 + \alpha_2)/\alpha_1$, that is,

$$g_2(y) = c\ \alpha_2 f(y)\bar{F}^{\alpha_2-1}(y) - c\ \alpha_2 f(y)\bar{F}^{\alpha_1+\alpha_2-1}(y)$$

and

$$\bar{G}_2(y) = \frac{\alpha_1+\alpha_2}{\alpha_1}\bar{F}^{\alpha_2}(y) - \frac{\alpha_2}{\alpha_1}\bar{F}^{\alpha_1+\alpha_2}(y).$$

Ordered data-Particular cases: PHR

- If X and Y are independent, $\bar{F}_X = \bar{F}^{\alpha_1}$ and $\bar{F}_Y = \bar{F}^{\alpha_2}$ for $\alpha_1, \alpha_2 > 0$, then the PDF of $(Y|X < Y)$ is

$$g_2(y) = c\ f_Y(y)F_X(y) = c\ \alpha_2 f(y)\bar{F}^{\alpha_2-1}(y)(1-\bar{F}^{\alpha_1}(y))$$

where $c = (\alpha_1 + \alpha_2)/\alpha_1$, that is,

$$g_2(y) = c\ \alpha_2 f(y)\bar{F}^{\alpha_2-1}(y) - c\ \alpha_2 f(y)\bar{F}^{\alpha_1+\alpha_2-1}(y)$$

and

$$\bar{G}_2(y) = \frac{\alpha_1+\alpha_2}{\alpha_1}\bar{F}^{\alpha_2}(y) - \frac{\alpha_2}{\alpha_1}\bar{F}^{\alpha_1+\alpha_2}(y).$$

- It is negative mixture of members of the PHR model.

Ordered data-Particular cases: PHR

- If X and Y are independent, $\bar{F}_X = \bar{F}^{\alpha_1}$ and $\bar{F}_Y = \bar{F}^{\alpha_2}$ for $\alpha_1, \alpha_2 > 0$, then the PDF of $(Y|X < Y)$ is

$$g_2(y) = c\ f_Y(y)F_X(y) = c\ \alpha_2 f(y)\bar{F}^{\alpha_2 - 1}(y)(1 - \bar{F}^{\alpha_1}(y))$$

where $c = (\alpha_1 + \alpha_2)/\alpha_1$, that is,

$$g_2(y) = c\ \alpha_2 f(y)\bar{F}^{\alpha_2 - 1}(y) - c\ \alpha_2 f(y)\bar{F}^{\alpha_1 + \alpha_2 - 1}(y)$$

and

$$\bar{G}_2(y) = \frac{\alpha_1 + \alpha_2}{\alpha_1}\bar{F}^{\alpha_2}(y) - \frac{\alpha_2}{\alpha_1}\bar{F}^{\alpha_1 + \alpha_2}(y).$$

- It is negative mixture of members of the PHR model.
- Note that $(Y|X < Y)$, $(X|X > Y)$ and $\max(X, Y)$ have different distributions!!

Ordered data-Particular cases: exponential model

- If X and Y are independent, $\bar{F}_X(t) = e^{-\alpha_1 t}$ and $\bar{F}_Y(t) = e^{-\alpha_2 t}$ for $t \geq 0$ and $\alpha_1, \alpha_2 > 0$, then $c = (\alpha_1 + \alpha_2)/\alpha_1$, the PDF of $(X|X < Y)$ is

$$g_1(t) = c\ f_X(t)\bar{F}_Y(t) = (\alpha_1 + \alpha_2)\ e^{-(\alpha_1+\alpha_2)t},\ \ t \geq 0.$$

Ordered data-Particular cases: exponential model

- If X and Y are independent, $\bar{F}_X(t) = e^{-\alpha_1 t}$ and $\bar{F}_Y(t) = e^{-\alpha_2 t}$ for $t \geq 0$ and $\alpha_1, \alpha_2 > 0$, then $c = (\alpha_1 + \alpha_2)/\alpha_1$, the PDF of $(X|X < Y)$ is

$$g_1(t) = c\ f_X(t)\bar{F}_Y(t) = (\alpha_1 + \alpha_2)\ e^{-(\alpha_1+\alpha_2)t},\ \ t \geq 0.$$

- Analogously, the survival function of $(Y|X < Y)$ is

$$\bar{G}_2(t) = \frac{\alpha_1 + \alpha_2}{\alpha_1} e^{-\alpha_2 t} - \frac{\alpha_2}{\alpha_1} e^{-(\alpha_1+\alpha_2)t},\ \ t \geq 0.$$

Ordered data-Particular cases: exponential model

- If X and Y are independent, $\bar{F}_X(t) = e^{-\alpha_1 t}$ and $\bar{F}_Y(t) = e^{-\alpha_2 t}$ for $t \geq 0$ and $\alpha_1, \alpha_2 > 0$, then $c = (\alpha_1 + \alpha_2)/\alpha_1$, the PDF of $(X|X < Y)$ is

$$g_1(t) = c\ f_X(t)\bar{F}_Y(t) = (\alpha_1 + \alpha_2)\ e^{-(\alpha_1+\alpha_2)t},\ \ t \geq 0.$$

- Analogously, the survival function of $(Y|X < Y)$ is

$$\bar{G}_2(t) = \frac{\alpha_1 + \alpha_2}{\alpha_1} e^{-\alpha_2 t} - \frac{\alpha_2}{\alpha_1} e^{-(\alpha_1+\alpha_2)t},\ \ t \geq 0.$$

- The first one is an exponential distribution and the second a negative mixture of exponential distributions.

Ordered data-Particular cases: exponential model

- If X and Y are independent, $\bar{F}_X(t) = e^{-\alpha_1 t}$ and $\bar{F}_Y(t) = e^{-\alpha_2 t}$ for $t \geq 0$ and $\alpha_1, \alpha_2 > 0$, then $c = (\alpha_1 + \alpha_2)/\alpha_1$, the PDF of $(X|X < Y)$ is

$$g_1(t) = c\ f_X(t)\bar{F}_Y(t) = (\alpha_1 + \alpha_2)\ e^{-(\alpha_1+\alpha_2)t},\ \ t \geq 0.$$

- Analogously, the survival function of $(Y|X < Y)$ is

$$\bar{G}_2(t) = \frac{\alpha_1 + \alpha_2}{\alpha_1} e^{-\alpha_2 t} - \frac{\alpha_2}{\alpha_1} e^{-(\alpha_1+\alpha_2)t},\ \ t \geq 0.$$

- The first one is an exponential distribution and the second a negative mixture of exponential distributions.
- It is a obtained as a convolution of two independent exponential distributions $(X|X < Y)$ and $(Y - X|X < Y)$ with different parameters.

Multivariate case

Proposition (Navarro, Cramer and Balakrishnan (2025))

Let $X_1, \dots, X_n$ be independent random variables with reliability functions $\overline{F}_i = \overline{F}^{\alpha_i}$ for $\alpha_i > 0$ and $i = 1, \dots, n$. Then, the reliability function $\bar{G}_k$ of $(X_k | X_1 < \cdots < X_n)$ is given by the reliability function of the k-th generalized order statistic based on parameters $\gamma_\ell = \sum_{j=\ell}^{n} \alpha_j$, $1 \le \ell \le n$, that is,

$$\bar{G}_k(t) = \sum_{i=1}^{k} \left(\prod_{\substack{j=1 \\ j \neq i}}^{k} \frac{\gamma_j}{\gamma_j - \gamma_i} \right) \bar{F}^{\gamma_i}(t), \quad t \ge 0. \tag{1.1}$$

If $X_1, \dots, X_n$ are exponentially distributed then, conditionally on $X_1 < \cdots < X_n$, the spacings $X_{k:n} - X_{k-1:n}$ are independent exponentially distributed with scale parameters γ_k for $k = 1, \dots, n$.

Multivariate case

- For $k = 1$, the reliability function of $(X_1|X_1 < \cdots < X_n)$ is

$$\bar{G}_1(t) = \bar{F}^{\alpha_1+\cdots+\alpha_n}(t), \quad t \geq 0 \tag{1.2}$$

and we have

$$\Pr(X_1 < \cdots < X_n) = \frac{\alpha_{n-1}\ldots\alpha_1}{(\alpha_{n-1}+\alpha_n)\ldots(\alpha_1+\cdots+\alpha_n)}. \tag{1.3}$$

Multivariate case

- For $k = 1$, the reliability function of $(X_1|X_1 < \cdots < X_n)$ is

$$\bar{G}_1(t) = \bar{F}^{\alpha_1+\cdots+\alpha_n}(t), \quad t \geq 0 \tag{1.2}$$

and we have

$$\Pr(X_1 < \cdots < X_n) = \frac{\alpha_{n-1}\ldots\alpha_1}{(\alpha_{n-1}+\alpha_n)\ldots(\alpha_1+\cdots+\alpha_n)}. \tag{1.3}$$

- The distribution of $(X_1|X_1 < \cdots < X_n)$ also belongs to the same PHR model with parameter $\alpha_1 + \cdots + \alpha_n$ and coincides with that of $X_{1:n} = \min(X_1, \ldots, X_n)$.

Multivariate case

- For $k = 1$, the reliability function of $(X_1|X_1 < \cdots < X_n)$ is

$$\bar{G}_1(t) = \bar{F}^{\alpha_1+\cdots+\alpha_n}(t), \quad t \geq 0 \tag{1.2}$$

and we have

$$\Pr(X_1 < \cdots < X_n) = \frac{\alpha_{n-1} \ldots \alpha_1}{(\alpha_{n-1} + \alpha_n) \ldots (\alpha_1 + \cdots + \alpha_n)}. \tag{1.3}$$

- The distribution of $(X_1|X_1 < \cdots < X_n)$ also belongs to the same PHR model with parameter $\alpha_1 + \cdots + \alpha_n$ and coincides with that of $X_{1:n} = \min(X_1, \ldots, X_n)$.
- Hence, the distribution of $(X_i|X_i < X_{\sigma(1)} < \cdots < X_{\sigma(n)})$ is the same for any permutation σ of these indices.

Multivariate case

- For $k = 1$, the reliability function of $(X_1|X_1 < \cdots < X_n)$ is

$$\bar{G}_1(t) = \bar{F}^{\alpha_1+\cdots+\alpha_n}(t), \quad t \geq 0 \tag{1.2}$$

and we have

$$\Pr(X_1 < \cdots < X_n) = \frac{\alpha_{n-1} \ldots \alpha_1}{(\alpha_{n-1} + \alpha_n) \ldots (\alpha_1 + \cdots + \alpha_n)}. \tag{1.3}$$

- The distribution of $(X_1|X_1 < \cdots < X_n)$ also belongs to the same PHR model with parameter $\alpha_1 + \cdots + \alpha_n$ and coincides with that of $X_{1:n} = \min(X_1, \ldots, X_n)$.
- Hence, the distribution of $(X_i|X_i < X_{\sigma(1)} < \cdots < X_{\sigma(n)})$ is the same for any permutation σ of these indices.
- This is a quite surprising property since the distributions of $X_1, \ldots, X_n$ are heterogeneous.

Multivariate case

- In particular, for the exponential model, we obtain

$$
\begin{aligned}
E(X_1|X_1 < \cdots < X_n) &= E(X_i|X_i < X_{\sigma(1)} < \cdots < X_{\sigma(n)}) \\
&= E(X_{1:n}) = \frac{1}{\alpha_1 + \cdots + \alpha_n}.
\end{aligned}
$$

Multivariate case

- In particular, for the exponential model, we obtain

$$\begin{aligned} E(X_1|X_1 < \cdots < X_n) &= E(X_i|X_i < X_{\sigma(1)} < \cdots < X_{\sigma(n)}) \\ &= E(X_{1:n}) = \frac{1}{\alpha_1 + \cdots + \alpha_n}. \end{aligned}$$

- In the case $k = 2$ the reliability function of $(X_2|X_1 < \cdots < X_n)$ is

$$\bar{G}_2(t) = \frac{\alpha_1 + \cdots + \alpha_n}{\alpha_1} \bar{F}^{\alpha_2 + \cdots + \alpha_n}(t) - \frac{\alpha_2 + \cdots + \alpha_n}{\alpha_1} \bar{F}^{\alpha_1 + \cdots + \alpha_n}(t), \tag{1.4}$$

Multivariate case

- In particular, for the exponential model, we obtain

$$E(X_1|X_1 < \cdots < X_n) = E(X_i|X_i < X_{\sigma(1)} < \cdots < X_{\sigma(n)})$$
$$= E(X_{1:n}) = \frac{1}{\alpha_1 + \cdots + \alpha_n}.$$

- In the case $k = 2$ the reliability function of $(X_2|X_1 < \cdots < X_n)$ is

$$\bar{G}_2(t) = \frac{\alpha_1 + \cdots + \alpha_n}{\alpha_1} \bar{F}^{\alpha_2 + \cdots + \alpha_n}(t) - \frac{\alpha_2 + \cdots + \alpha_n}{\alpha_1} \bar{F}^{\alpha_1 + \cdots + \alpha_n}(t), \tag{1.4}$$

- For properties of negative mixtures see Navarro and Sarabia (2024).

Exponential case

- If $X_1, \dots, X_n$ are independent and have exponential distributions of parameters $\alpha_1, \dots, \alpha_n$, then

$$E(X_k | X_1 < \cdots < X_n) = \sum_{i=1}^{k} \frac{1}{\gamma_i} = \frac{1}{\alpha_1 + \cdots + \alpha_n} + \cdots + \frac{1}{\alpha_k + \cdots + \alpha_n},$$

$$Var(X_k | X_1 < \cdots < X_n) = \sum_{i=1}^{k} \frac{1}{\gamma_i^2} = \frac{1}{(\alpha_1 + \cdots + \alpha_n)^2} + \cdots + \frac{1}{(\alpha_k + \cdots + \alpha_n)^2}$$

and

$$E(X_k^2 | X_1 < \cdots < X_n) = \sum_{i=1}^{k} \frac{1}{\gamma_i^2} + \left(\sum_{i=1}^{k} \frac{1}{\gamma_i} \right)^2$$

with $\gamma_i = \sum_{j=i}^{n} \alpha_j$, $1 \le i \le n$.

Generalized Order Statistics (GOS)

Theorem

Let $X_1, \ldots, X_n$ be independent random variables with strictly increasing continuous reliability functions $\overline{F}_i = \overline{F}^{\alpha_i}$ for $\alpha_i > 0$ and $i = 1, \ldots, n$. Then,

$$(X_1, \ldots, X_n | X_1 < \cdots < X_n) \stackrel{d}{=} (Y_1, \ldots, Y_n),$$

where $(Y_1, \ldots, Y_n)$ are the GOS based on F and parameters $\gamma_1, \ldots, \gamma_n$ with $\gamma_\ell = \sum_{j=\ell}^{n} \alpha_j$, $1 \leq \ell \leq n$.

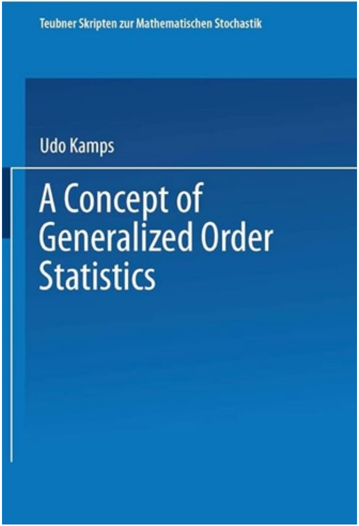


Figure: A tribute to Udo.

Inferencial results

Ordered data

- Suppose that we observe m independent samples of the random variables $X_1, \ldots, X_n$ represented by $X_1^{(j)}, \ldots, X_n^{(j)}$ for $j = 1, \ldots, m$.

Ordered data

- Suppose that we observe m independent samples of the random variables $X_1, \ldots, X_n$ represented by $X_1^{(j)}, \ldots, X_n^{(j)}$ for $j = 1, \ldots, m$.
- The purpose is to estimate the parameters α_i by assuming that $\Pr(X_i^{(j)} \leq t) = \bar{F}^{\alpha_i}(t)$ and that $\bar{F}$ is known.

Ordered data

- Suppose that we observe m independent samples of the random variables $X_1, \ldots, X_n$ represented by $X_1^{(j)}, \ldots, X_n^{(j)}$ for $j = 1, \ldots, m$.
- The purpose is to estimate the parameters α_i by assuming that $\Pr(X_i^{(j)} \leq t) = \bar{F}^{\alpha_i}(t)$ and that $\bar{F}$ is known.
- Furthermore, we assume that not all the data are available, i.e., the samples are subject to censoring.

Ordered data

- Suppose that we observe m independent samples of the random variables $X_1, \ldots, X_n$ represented by $X_1^{(j)}, \ldots, X_n^{(j)}$ for $j = 1, \ldots, m$.
- The purpose is to estimate the parameters α_i by assuming that $\Pr(X_i^{(j)} \leq t) = \bar{F}^{\alpha_i}(t)$ and that $\bar{F}$ is known.
- Furthermore, we assume that not all the data are available, i.e., the samples are subject to censoring.
- For illustration, we consider below some typical cases.

Ordered data

- Suppose that we observe m independent samples of the random variables $X_1, \ldots, X_n$ represented by $X_1^{(j)}, \ldots, X_n^{(j)}$ for $j = 1, \ldots, m$.
- The purpose is to estimate the parameters α_i by assuming that $\Pr(X_i^{(j)} \leq t) = \bar{F}^{\alpha_i}(t)$ and that $\bar{F}$ is known.
- Furthermore, we assume that not all the data are available, i.e., the samples are subject to censoring.
- For illustration, we consider below some typical cases.
- Other cases can be managed in a similar way.

Case 1: First data (series system)

- Suppose that we observe only the minimum in each sample

$$X_{1:n}^{(j)} = \min(X_1^{(j)}, \ldots, X_n^{(j)}), \quad j = 1, \ldots, m.$$

Case 1: First data (series system)

- Suppose that we observe only the minimum in each sample

$$X_{1:n}^{(j)} = \min(X_1^{(j)}, \dots, X_n^{(j)}), \quad j = 1, \dots, m.$$

- Alternatively, we could just observe the first measurement of the first component, that is,

$$X_1^{(j)} | X_1^{(j)} = X_{1:n}^{(j)}, \quad j = 1, \dots, m.$$

Case 1: First data (series system)

- Suppose that we observe only the minimum in each sample

$$X_{1:n}^{(j)} = \min(X_1^{(j)}, \dots, X_n^{(j)}), \quad j = 1, \dots, m.$$

- Alternatively, we could just observe the first measurement of the first component, that is,

$$X_1^{(j)} | X_1^{(j)} = X_{1:n}^{(j)}, \quad j = 1, \dots, m.$$

- We use the random variable $I_i^{(j)}$ indicating if the first failure in sample j is caused by the i-th component.

Case 1: First data (series system)

- Suppose that we observe only the minimum in each sample

$$X_{1:n}^{(j)} = \min(X_1^{(j)}, \dots, X_n^{(j)}), \quad j = 1, \dots, m.$$

- Alternatively, we could just observe the first measurement of the first component, that is,

$$X_1^{(j)} | X_1^{(j)} = X_{1:n}^{(j)}, \quad j = 1, \dots, m.$$

- We use the random variable $I_i^{(j)}$ indicating if the first failure in sample j is caused by the i-th component.
- Thus, $I_i^{(j)} = 1$ if $X_i^{(j)} = X_{1:n}^{(j)}$ and $I_i^{(j)} = 0$ otherwise for $i = 1, \dots, n$ and $j = 1, \dots, m$.

Case 1: First data (series system)

- Suppose that we observe only the minimum in each sample

$$X_{1:n}^{(j)} = \min(X_1^{(j)}, \dots, X_n^{(j)}), \quad j = 1, \dots, m.$$

- Alternatively, we could just observe the first measurement of the first component, that is,

$$X_1^{(j)} | X_1^{(j)} = X_{1:n}^{(j)}, \quad j = 1, \dots, m.$$

- We use the random variable $I_i^{(j)}$ indicating if the first failure in sample j is caused by the i-th component.
- Thus, $I_i^{(j)} = 1$ if $X_i^{(j)} = X_{1:n}^{(j)}$ and $I_i^{(j)} = 0$ otherwise for $i = 1, \dots, n$ and $j = 1, \dots, m$.
- Note that we do not have ties.

Case 1: First data (series system)

- Suppose that we observe only the minimum in each sample

$$X_{1:n}^{(j)} = \min(X_1^{(j)}, \dots, X_n^{(j)}), \quad j = 1, \dots, m.$$

- Alternatively, we could just observe the first measurement of the first component, that is,

$$X_1^{(j)} | X_1^{(j)} = X_{1:n}^{(j)}, \quad j = 1, \dots, m.$$

- We use the random variable $I_i^{(j)}$ indicating if the first failure in sample j is caused by the i-th component.
- Thus, $I_i^{(j)} = 1$ if $X_i^{(j)} = X_{1:n}^{(j)}$ and $I_i^{(j)} = 0$ otherwise for $i = 1, \dots, n$ and $j = 1, \dots, m$.
- Note that we do not have ties.
- Both scenarios can be managed in a similar way.

Case 1: First data (series system)

- Several practical situations fit to these scenarios.

Case 1: First data (series system)

- Several practical situations fit to these scenarios.
- For example, we can consider lifetimes from a series system with n (possibly heterogeneous) components where the other units stop to work when the system fails.

Case 1: First data (series system)

- Several practical situations fit to these scenarios.
- For example, we can consider lifetimes from a series system with n (possibly heterogeneous) components where the other units stop to work when the system fails.
- This scheme can also be applied to survival data of organs (eyes, kidneys, etc.).

Case 1: First data (series system)

- Several practical situations fit to these scenarios.
- For example, we can consider lifetimes from a series system with n (possibly heterogeneous) components where the other units stop to work when the system fails.
- This scheme can also be applied to survival data of organs (eyes, kidneys, etc.).
- It can also be seen as a progressive censoring procedure where the units are arranged in blocks of n elements and after the first failure, the other elements in the block are censored.

Case 1: First data (series system)

- Several practical situations fit to these scenarios.
- For example, we can consider lifetimes from a series system with n (possibly heterogeneous) components where the other units stop to work when the system fails.
- This scheme can also be applied to survival data of organs (eyes, kidneys, etc.).
- It can also be seen as a progressive censoring procedure where the units are arranged in blocks of n elements and after the first failure, the other elements in the block are censored.
- In the literature, this model is called 'first failure censoring'.

Case 1: First data (series system)

- Several practical situations fit to these scenarios.
- For example, we can consider lifetimes from a series system with n (possibly heterogeneous) components where the other units stop to work when the system fails.
- This scheme can also be applied to survival data of organs (eyes, kidneys, etc.).
- It can also be seen as a progressive censoring procedure where the units are arranged in blocks of n elements and after the first failure, the other elements in the block are censored.
- In the literature, this model is called 'first failure censoring'.
- From Balakrishnan and Cramer (2014), p. 529, this scheme can be seen as a standard progressive censoring scheme with a modified censoring plan $\mathcal{S} = n\mathcal{R} + (n-1)(1, \ldots, 1)$.

Case 1: First data (series system)

- In the case $n = 2$, the data of type $(X_1|X_1 < X_2)$ represents data from X_1 with an independent censoring random time X_2.

Case 1: First data (series system)

- In the case $n = 2$, the data of type $(X_1|X_1 < X_2)$ represents data from X_1 with an independent censoring random time X_2.
- Here, we have two options: independent censoring times for each data, that is, $(X_1^{(j)}|X_1^{(j)} < X_2^{(j)})$ or a common censoring time X_2 for all the sample values, that is, $(X_1^{(j)}|X_1^{(j)} < X_2)$.

Case 1: First data (series system)

- In the case $n = 2$, the data of type $(X_1|X_1 < X_2)$ represents data from X_1 with an independent censoring random time X_2.
- Here, we have two options: independent censoring times for each data, that is, $(X_1^{(j)}|X_1^{(j)} < X_2^{(j)})$ or a common censoring time X_2 for all the sample values, that is, $(X_1^{(j)}|X_1^{(j)} < X_2)$.
- In this last case the sample values are dependent since they share a common censoring time X_2.

Case 1: First data (series system)

- In the case $n = 2$, the data of type $(X_1|X_1 < X_2)$ represents data from X_1 with an independent censoring random time X_2.
- Here, we have two options: independent censoring times for each data, that is, $(X_1^{(j)}|X_1^{(j)} < X_2^{(j)})$ or a common censoring time X_2 for all the sample values, that is, $(X_1^{(j)}|X_1^{(j)} < X_2)$.
- In this last case the sample values are dependent since they share a common censoring time X_2.
- In particular, if $X_2 = t_0$ for some $t_0 > 0$ then we have an experiment with truncated data in the fixed period of time $[0, t_0]$.

Case 1: First data (series system)

- In the case $n = 2$, the data of type $(X_1|X_1 < X_2)$ represents data from X_1 with an independent censoring random time X_2.
- Here, we have two options: independent censoring times for each data, that is, $(X_1^{(j)}|X_1^{(j)} < X_2^{(j)})$ or a common censoring time X_2 for all the sample values, that is, $(X_1^{(j)}|X_1^{(j)} < X_2)$.
- In this last case the sample values are dependent since they share a common censoring time X_2.
- In particular, if $X_2 = t_0$ for some $t_0 > 0$ then we have an experiment with truncated data in the fixed period of time $[0, t_0]$.
- This scheme can also represent data from stress-strength models where a data is observed (fails) if and only if the strength $X_1^{(j)}$ is below the stress $X_2^{(j)}$.

Case 1: First data (series system)

- Note that if we only have measurements of the first failure, then the model (parameters) is subject to identifiability issues since $\bar{F}_{1:n} = \bar{F}^{\alpha_1+\cdots+\alpha_n}$.

Case 1: First data (series system)

- Note that if we only have measurements of the first failure, then the model (parameters) is subject to identifiability issues since $\bar{F}_{1:n} = \bar{F}^{\alpha_1+\cdots+\alpha_n}$.
- Hence, estimation of the single parameter α_i is not possible.

Case 1: First data (series system)

- Note that if we only have measurements of the first failure, then the model (parameters) is subject to identifiability issues since $\bar{F}_{1:n} = \bar{F}^{\alpha_1+\cdots+\alpha_n}$.
- Hence, estimation of the single parameter α_i is not possible.
- The same happen with $(X_1|X_1 = X_{1:n})$.

Case 1: First data (series system)

- Note that if we only have measurements of the first failure, then the model (parameters) is subject to identifiability issues since $\bar{F}_{1:n} = \bar{F}^{\alpha_1+\cdots+\alpha_n}$.
- Hence, estimation of the single parameter α_i is not possible.
- The same happen with $(X_1|X_1 = X_{1:n})$.
- Thus, assuming exponentially distributed lifetimes X_i with

$$\bar{X}_{1:n} = \frac{1}{m}\sum_{j=1}^{m} X_{1:n}^{(j)},$$

we get

$$E\left(\bar{X}_{1:n}\right) = \mu_{1:n} = E(X_{1:n}) = \frac{1}{\alpha_1+\cdots+\alpha_n}$$

and $\bar{X}_{1:n}$ is an unbiased estimator of $\mu_{1:n}$.

Case 1: First data (series system)

- Note that if we only have measurements of the first failure, then the model (parameters) is subject to identifiability issues since $\bar{F}_{1:n} = \bar{F}^{\alpha_1+\cdots+\alpha_n}$.
- Hence, estimation of the single parameter α_i is not possible.
- The same happen with $(X_1|X_1 = X_{1:n})$.
- Thus, assuming exponentially distributed lifetimes X_i with

$$\bar{X}_{1:n} = \frac{1}{m}\sum_{j=1}^{m} X_{1:n}^{(j)},$$

we get

$$E\left(\bar{X}_{1:n}\right) = \mu_{1:n} = E(X_{1:n}) = \frac{1}{\alpha_1+\cdots+\alpha_n}$$

and $\bar{X}_{1:n}$ is an unbiased estimator of $\mu_{1:n}$.
- However we cannot identify a single parameters α_i.

Case 1: First data (series system)

- The same problem occurs for

$$\bar{X}_{1|1<} = \frac{1}{m_1} \sum_{j=1}^{m} I_1^{(j)} X_{1:n}^{(j)}, \quad \text{where } m_1 = \sum_{j=1}^{m} I_1^{(j)}.$$

Case 1: First data (series system)

- The same problem occurs for

$$\bar{X}_{1|1<} = \frac{1}{m_1} \sum_{j=1}^{m} I_1^{(j)} X_{1:n}^{(j)}, \quad \text{where } m_1 = \sum_{j=1}^{m} I_1^{(j)}.$$

- Of course, we need to assume that $m_1 > 0$.

Case 1: First data (series system)

- The same problem occurs for

$$\bar{X}_{1|1<} = \frac{1}{m_1} \sum_{j=1}^{m} I_1^{(j)} X_{1:n}^{(j)}, \quad \text{where } m_1 = \sum_{j=1}^{m} I_1^{(j)}.$$

- Of course, we need to assume that $m_1 > 0$.
- In order to calculate the mean of $\bar{X}_{1|1<}$, note that,

$$p_1 = \Pr(I_1^{(1)} = 1) = \Pr(X_1 < \min(X_2, \ldots, X_n)) = \frac{\alpha_1}{\alpha_1 + \cdots + \alpha_n}$$

and the property $m_1 \sim Binom(m, p_1)$ holds, i.e.,

$$\Pr(m_1 = k) = \binom{m}{k} p_1^k (1 - p_1)^{m-k}, \quad k = 0, \ldots, m.$$

Case 1: First data (series system)

- The same problem occurs for

$$\bar{X}_{1|1<} = \frac{1}{m_1} \sum_{j=1}^{m} I_1^{(j)} X_{1:n}^{(j)}, \quad \text{where } m_1 = \sum_{j=1}^{m} I_1^{(j)}.$$

- Of course, we need to assume that $m_1 > 0$.
- In order to calculate the mean of $\bar{X}_{1|1<}$, note that,

$$p_1 = \Pr(I_1^{(1)} = 1) = \Pr(X_1 < \min(X_2, \dots, X_n)) = \frac{\alpha_1}{\alpha_1 + \cdots + \alpha_n}$$

and the property $m_1 \sim Binom(m, p_1)$ holds, i.e.,

$$\Pr(m_1 = k) = \binom{m}{k} p_1^k (1 - p_1)^{m-k}, \quad k = 0, \dots, m.$$

- Then $m_1/m \xrightarrow{a.s.} p_1$.

Case 1: First data (series system)

- We prove that

$$E\left(\bar{X}_{1|1<}|m_1>0\right)=\frac{1}{\alpha_1+\alpha_2+\cdots+\alpha_n}.$$

Case 1: First data (series system)

- We prove that

$$E\left(\bar{X}_{1|1<}|m_1 > 0\right) = \frac{1}{\alpha_1 + \alpha_2 + \cdots + \alpha_n}.$$

- To illustrate the inferential approach, we consider $n = 2$.

Case 1: First data (series system)

- We prove that

$$E\left(\bar{X}_{1|1<}|m_1>0\right)=\frac{1}{\alpha_1+\alpha_2+\cdots+\alpha_n}.$$

- To illustrate the inferential approach, we consider $n=2$.
- Then, we can use the information provided by m_1.

Case 1: First data (series system)

- We prove that

$$E\left(\bar{X}_{1|1<}|m_1>0\right)=\frac{1}{\alpha_1+\alpha_2+\cdots+\alpha_n}.$$

- To illustrate the inferential approach, we consider $n=2$.
- Then, we can use the information provided by m_1.
- Thus, we consider the estimator $\widehat{p}_1=m_1/m$ with

$$E\left(\widehat{p}_1\right)=p_1=\Pr(X_1<X_2)=\frac{\alpha_1}{\alpha_1+\alpha_2}=\alpha_1\mu_{1|1<2}$$

where $\mu_{1|1<2}=E(X_1|X_1<X_2)=1/(\alpha_1+\alpha_2)$.

Case 1: First data (series system)

- We prove that

$$E\left(\bar{X}_{1|1<}|m_1>0\right)=\frac{1}{\alpha_1+\alpha_2+\cdots+\alpha_n}.$$

- To illustrate the inferential approach, we consider $n=2$.
- Then, we can use the information provided by m_1.
- Thus, we consider the estimator $\widehat{p}_1=m_1/m$ with

$$E\left(\widehat{p}_1\right)=p_1=\Pr(X_1<X_2)=\frac{\alpha_1}{\alpha_1+\alpha_2}=\alpha_1\mu_{1|1<2}$$

where $\mu_{1|1<2}=E(X_1|X_1<X_2)=1/(\alpha_1+\alpha_2)$.
- Therefore, α_1 can be estimated by

$$\widehat{\alpha}_1=\frac{\widehat{p}_1}{\bar{X}_{1|1<2}}.$$

Case 1: First data (series system)

- We prove that

$$E\left(\bar{X}_{1|1<}|m_1 > 0\right) = \frac{1}{\alpha_1 + \alpha_2 + \cdots + \alpha_n}.$$

- To illustrate the inferential approach, we consider $n = 2$.
- Then, we can use the information provided by m_1.
- Thus, we consider the estimator $\widehat{p}_1 = m_1/m$ with

$$E\left(\widehat{p}_1\right) = p_1 = \Pr(X_1 < X_2) = \frac{\alpha_1}{\alpha_1 + \alpha_2} = \alpha_1 \mu_{1|1<2}$$

where $\mu_{1|1<2} = E(X_1|X_1 < X_2) = 1/(\alpha_1 + \alpha_2)$.

- Therefore, α_1 can be estimated by

$$\widehat{\alpha}_1 = \frac{\widehat{p}_1}{\bar{X}_{1|1<2}}.$$

- Analogously, α_2 can be estimated with $\widehat{\alpha}_2 = (1 - \widehat{p}_1)/\bar{X}_{1|1<2}$.

Case 1: First data (series system)

- If we prefer to estimate the mean $\mu_1 = E(X_1) = 1/\alpha_1$, we can use

$$\widehat{\mu}_1 = \frac{1}{\widehat{\alpha}_1} = \frac{\bar{X}_{1|1<2}}{\widehat{p}_1} \geq \bar{X}_{1|1<2}$$

provided that $m_1 > 0$.

Case 1: First data (series system)

- If we prefer to estimate the mean $\mu_1 = E(X_1) = 1/\alpha_1$, we can use
$$\widehat{\mu}_1 = \frac{1}{\widehat{\alpha}_1} = \frac{\bar{X}_{1|1<2}}{\widehat{p}_1} \geq \bar{X}_{1|1<2}$$
provided that $m_1 > 0$.
- Note that if we just use the sample mean with the non-censored data $\bar{X}_{1|1<2}$, then we get an under-biased estimation (as expected).

Case 1: First data (series system)

- If we prefer to estimate the mean $\mu_1 = E(X_1) = 1/\alpha_1$, we can use
$$\widehat{\mu}_1 = \frac{1}{\widehat{\alpha}_1} = \frac{\bar{X}_{1|1<2}}{\widehat{p}_1} \geq \bar{X}_{1|1<2}$$
provided that $m_1 > 0$.
- Note that if we just use the sample mean with the non-censored data $\bar{X}_{1|1<2}$, then we get an under-biased estimation (as expected).
- If we have the information about all the first failure times, it is better to replace $\bar{X}_{1|1<2}$ by $\bar{X}_{1:2}$.

Case 1: First data (series system)

- If we prefer to estimate the mean $\mu_1 = E(X_1) = 1/\alpha_1$, we can use
$$\widehat{\mu}_1 = \frac{1}{\widehat{\alpha}_1} = \frac{\bar{X}_{1|1<2}}{\widehat{p}_1} \geq \bar{X}_{1|1<2}$$
provided that $m_1 > 0$.
- Note that if we just use the sample mean with the non-censored data $\bar{X}_{1|1<2}$, then we get an under-biased estimation (as expected).
- If we have the information about all the first failure times, it is better to replace $\bar{X}_{1|1<2}$ by $\bar{X}_{1:2}$.
- In this case, it is easy to get distributional properties for
$$\widehat{\mu}_1 = \frac{1}{\widehat{\alpha}_1} = \frac{\bar{X}_{1:2}}{\widehat{p}_1} = \frac{1}{m_1}\sum_{j=1}^{m} X_{1:2}^{(j)} \geq \bar{X}_{1:2} = \frac{1}{m}\sum_{j=1}^{m} X_{1:2}^{(j)}.$$

Case 1: First data (series system)

- Thus, from the Central Limit Theorem (CLT) we know that

$$\frac{\bar{X}_{1:2} - \mu_{1:2}}{\sigma_{1:2}/\sqrt{m}} \xrightarrow{d} Z,$$

where $\mu_{1:2} = E(X_{1:2}) = 1/(\alpha_1 + \alpha_2)$,
$\sigma^2_{1:2} = \mathrm{Var}(X_{1:2}) = 1/(\alpha_1 + \alpha_2)^2$ and Z has a standard normal distribution.

Case 1: First data (series system)

- Thus, from the Central Limit Theorem (CLT) we know that
$$\frac{\bar{X}_{1:2}-\mu_{1:2}}{\sigma_{1:2}/\sqrt{m}} \xrightarrow{d} Z,$$
where $\mu_{1:2}=E(X_{1:2})=1/(\alpha_1+\alpha_2)$,
$\sigma^2_{1:2}=\mathrm{Var}(X_{1:2})=1/(\alpha_1+\alpha_2)^2$ and Z has a standard normal distribution.
- From the Weak Law of Large Numbers (WLLN), we have that
$$\frac{\widehat{p}_1}{p_1} \xrightarrow{p} 1.$$

Case 1: First data (series system)

- Thus, from the Central Limit Theorem (CLT) we know that

$$\frac{\bar{X}_{1:2}-\mu_{1:2}}{\sigma_{1:2}/\sqrt{m}} \xrightarrow{d} Z,$$

where $\mu_{1:2}=E(X_{1:2})=1/(\alpha_1+\alpha_2)$,
$\sigma^2_{1:2}=\mathrm{Var}(X_{1:2})=1/(\alpha_1+\alpha_2)^2$ and Z has a standard normal distribution.

- From the Weak Law of Large Numbers (WLLN), we have that

$$\frac{\widehat{p}_1}{p_1} \xrightarrow{p} 1.$$

- Hence, from Slutsky's theorem, we get

$$\frac{\sqrt{m}}{\mu_1}\left(\widehat{\mu}_1-\frac{\mu_{1:2}}{\widehat{p}_1}\right)=\frac{\sqrt{m}\left(\bar{X}_{1:2}-\mu_{1:2}\right)/\sigma_{1:2}}{\widehat{p}_1/p_1} \xrightarrow{d} Z, \qquad (2.1)$$

where Z has a standard normal distribution.

Example 1

- To illustrate this case we consider a simulated sample of size $m = 100$ of $n = 2$ independent exponential distributions with $\alpha_1 = 1$ and $\alpha_2 = 2$.

Example 1

- To illustrate this case we consider a simulated sample of size $m = 100$ of $n = 2$ independent exponential distributions with $\alpha_1 = 1$ and $\alpha_2 = 2$.
- We use the statistical program R with seed 333.

Example 1

- To illustrate this case we consider a simulated sample of size $m = 100$ of $n = 2$ independent exponential distributions with $\alpha_1 = 1$ and $\alpha_2 = 2$.
- We use the statistical program R with seed 333.
- Thus we obtain $m_1 = 31$, that is, the series system lifetime $X_{1:2}$ coincides with X_1 in the sample 31 times and with X_2 69 times.

Example 1

- To illustrate this case we consider a simulated sample of size $m = 100$ of $n = 2$ independent exponential distributions with $\alpha_1 = 1$ and $\alpha_2 = 2$.
- We use the statistical program R with seed 333.
- Thus we obtain $m_1 = 31$, that is, the series system lifetime $X_{1:2}$ coincides with X_1 in the sample 31 times and with X_2 69 times.
- The sample mean with all the series system lifetimes is

$$\bar{X}_{1:2} = 0.3402184 \approx \frac{1}{\alpha_1 + \alpha_2} = \frac{1}{3}.$$

Example 1

- To illustrate this case we consider a simulated sample of size $m = 100$ of $n = 2$ independent exponential distributions with $\alpha_1 = 1$ and $\alpha_2 = 2$.
- We use the statistical program R with seed 333.
- Thus we obtain $m_1 = 31$, that is, the series system lifetime $X_{1:2}$ coincides with X_1 in the sample 31 times and with X_2 69 times.
- The sample mean with all the series system lifetimes is

$$\bar{X}_{1:2} = 0.3402184 \approx \frac{1}{\alpha_1 + \alpha_2} = \frac{1}{3}.$$

- As mentioned above, this information is not enough to estimate the unknown parameters α_1 and α_2.

Example 1

- To do so we need to use

$$\widehat{p}_1 = \frac{m_1}{m} = \frac{31}{100} = 0.31 \approx \frac{\alpha_1}{\alpha_1 + \alpha_2} = \frac{1}{3}.$$

Example 1

- To do so we need to use

$$\widehat{p}_1 = \frac{m_1}{m} = \frac{31}{100} = 0.31 \approx \frac{\alpha_1}{\alpha_1 + \alpha_2} = \frac{1}{3}.$$

- Then we can estimate α_1 with

$$\widehat{\alpha}_1 = \frac{\widehat{p}_1}{\bar{X}_{1:2}} = \frac{0.31}{0.3402184} = 0.9111795 \approx \alpha_1 = 1.$$

Example 1

- To do so we need to use

$$\widehat{p}_1 = \frac{m_1}{m} = \frac{31}{100} = 0.31 \approx \frac{\alpha_1}{\alpha_1 + \alpha_2} = \frac{1}{3}.$$

- Then we can estimate α_1 with

$$\widehat{\alpha}_1 = \frac{\widehat{p}_1}{\bar{X}_{1:2}} = \frac{0.31}{0.3402184} = 0.9111795 \approx \alpha_1 = 1.$$

- Analogously, for α_2 we get $\widehat{p}_2 = 1 - \widehat{p}_1$ and

$$\widehat{\alpha}_2 = \frac{\widehat{p}_2}{\bar{X}_{1:2}} = \frac{0.69}{0.3402184} = 2.028109 \approx \alpha_2 = 2.$$

Example 1

- If we prefer to estimate the means we get

$$\widehat{\mu}_1 = \frac{\bar{X}_{1:2}}{\widehat{p}_1} = \frac{0.3402184}{0.31} = 1.097479 \approx \mu_1 = E(X_1) = \frac{1}{\alpha_1} = 1$$

and

$$\widehat{\mu}_2 = \frac{\bar{X}_{1:2}}{\widehat{p}_2} = \frac{0.3402184}{0.69} = 0.4930701 \approx \mu_2 = E(X_2) = \frac{1}{\alpha_2} = 0.5.$$

Example 1

- If we prefer to estimate the means we get

$$\widehat{\mu}_1 = \frac{\bar{X}_{1:2}}{\widehat{p}_1} = \frac{0.3402184}{0.31} = 1.097479 \approx \mu_1 = E(X_1) = \frac{1}{\alpha_1} = 1$$

and

$$\widehat{\mu}_2 = \frac{\bar{X}_{1:2}}{\widehat{p}_2} = \frac{0.3402184}{0.69} = 0.4930701 \approx \mu_2 = E(X_2) = \frac{1}{\alpha_2} = 0.5.$$

- If we just have the 31 censored data obtained from X_1 at the first component failures, then

$$\widehat{\mu}_1 = \frac{1}{\widehat{\alpha}_1} = \frac{\bar{X}_{1|1<2}}{\widehat{p}_1} = \frac{0.4129864}{0.31} = 1.332214 \geq \bar{X}_{1|1<2} = 0.4129864.$$

Example 1

- If we prefer to estimate the means we get

$$\widehat{\mu}_1=\frac{\bar{X}_{1:2}}{\widehat{p}_1}=\frac{0.3402184}{0.31}=1.097479\approx\mu_1=E(X_1)=\frac{1}{\alpha_1}=1$$

and

$$\widehat{\mu}_2=\frac{\bar{X}_{1:2}}{\widehat{p}_2}=\frac{0.3402184}{0.69}=0.4930701\approx\mu_2=E(X_2)=\frac{1}{\alpha_2}=0.5.$$

- If we just have the 31 censored data obtained from X_1 at the first component failures, then

$$\widehat{\mu}_1=\frac{1}{\widehat{\alpha}_1}=\frac{\bar{X}_{1|1<2}}{\widehat{p}_1}=\frac{0.4129864}{0.31}=1.332214\geq\bar{X}_{1|1<2}=0.4129864.$$

- The estimation is not very good since we just have 31 data and it is better to use the information of all the censored times as above getting the estimation 1.097479 for $\mu_1=1$.

Example 1

- Of course, the sample mean with the complete data from X_1 at $X_{1:2}$ underestimate (0.4129864) the mean $\mu_1 = 1$.

Example 1

- Of course, the sample mean with the complete data from X_1 at $X_{1:2}$ underestimate (0.4129864) the mean $\mu_1 = 1$.
- If we just use the data from X_2 when $X_2 = X_{1:2}$ we get

$$\widehat{\mu}_2 = \frac{1}{\widehat{\alpha}_2} = \frac{\bar{X}_{2|2<1}}{\widehat{p}_2} = \frac{0.3075255}{0.69} = 0.4456891 \geq \bar{X}_{2|2<1} = 0.3075255$$

which is a good estimation of $\mu_2 = 0.5$.

Example 1

- Of course, the sample mean with the complete data from X_1 at $X_{1:2}$ underestimate (0.4129864) the mean $\mu_1 = 1$.
- If we just use the data from X_2 when $X_2 = X_{1:2}$ we get

$$\widehat{\mu}_2 = \frac{1}{\widehat{\alpha}_2} = \frac{\bar{X}_{2|2<1}}{\widehat{p}_2} = \frac{0.3075255}{0.69} = 0.4456891 \geq \bar{X}_{2|2<1} = 0.3075255$$

which is a good estimation of $\mu_2 = 0.5$.

- Note that these values can also be used to estimate the means of the other variables obtaining

$$\widehat{\mu}_1 = \frac{\bar{X}_{2|2<1}}{\widehat{p}_1} = \frac{0.3075255}{0.31} = 0.9920176 \approx \mu_1 = 1$$

and

$$\widehat{\mu}_2 = \frac{\bar{X}_{1|1<2}}{\widehat{p}_2} = \frac{0.4129864}{0.69} = 0.598531 \approx \mu_2 = 0.5.$$

Example 1

- The data from X_1 and X_2 are plotted in the following figure.

Example 1

- The data from X_1 and X_2 are plotted in the following figure.
- In the right plot we can see the difference between the series system $L = X_{1:2}$ and the parallel system lifetimes $U = X_{2:2}$.

Example 1

- The data from X_1 and X_2 are plotted in the following figure.
- In the right plot we can see the difference between the series system $L = X_{1:2}$ and the parallel system lifetimes $U = X_{2:2}$.
- Note that we can estimate the parallel system failure time from the first component failure time with

$$E(X_2|X_1 < X_2, X_1 = x) = x + \frac{1}{\alpha_2} \approx m_2(x) = x + \frac{\bar{X}_{1:2}}{\widehat{p}_2} = x + 0.4930701$$

by replacing x with $X_1^{(j)}$ when $X_1^{(j)} < X_2^{(j)}$ or with

$$E(X_1|X_2 < X_1, X_2 = x) = x + \frac{1}{\alpha_1} \approx m_1(x) = x + \frac{\bar{X}_{1:2}}{\widehat{p}_1} = x + 1.097479$$

by replacing x with $X_2^{(j)}$ when $X_2^{(j)} < X_1^{(j)}$.

Example 1

- The data from X_1 and X_2 are plotted in the following figure.
- In the right plot we can see the difference between the series system $L = X_{1:2}$ and the parallel system lifetimes $U = X_{2:2}$.
- Note that we can estimate the parallel system failure time from the first component failure time with

$$E(X_2|X_1 < X_2, X_1 = x) = x + \frac{1}{\alpha_2} \approx m_2(x) = x + \frac{\bar{X}_{1:2}}{\widehat{p}_2} = x + 0.4930701$$

by replacing x with $X_1^{(j)}$ when $X_1^{(j)} < X_2^{(j)}$ or with

$$E(X_1|X_2 < X_1, X_2 = x) = x + \frac{1}{\alpha_1} \approx m_1(x) = x + \frac{\bar{X}_{1:2}}{\widehat{p}_1} = x + 1.097479$$

by replacing x with $X_2^{(j)}$ when $X_2^{(j)} < X_1^{(j)}$.
- These regression lines are plotted in the right plot.

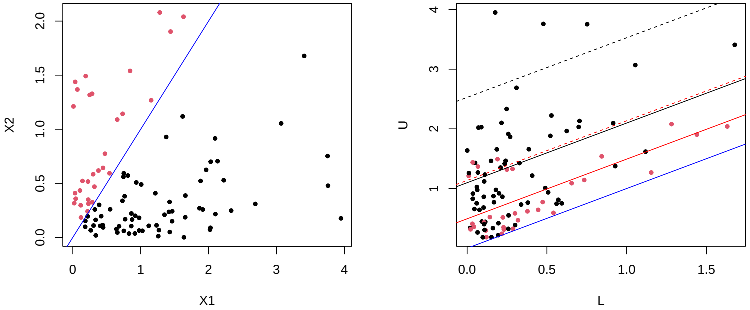


Figure: Censored data (left) for X_1 when $X_1 < X_2$ (red) and for X_2 when $X_2 < X_1$ (black). Data (right) from series $L = X_{1:2}$ and parallel $U = X_{2:2}$ systems. The regression lines are the red and black solid lines.

Example 1

- As the fits are not very good, quantile regression curves might be used to improve the fits since the residual lifetimes have also exponential distributions.

Example 1

- As the fits are not very good, quantile regression curves might be used to improve the fits since the residual lifetimes have also exponential distributions.
- For example, as $(X_2 - x | X_1 < X_2, X_1 = x)$ has an exponential distribution with mean $1/\alpha_2 \approx 0.4930701$, a lower 90% confidence band for X_2, given $X_1 = x < X_2$, is
$$(x, x - 0.4930701 \cdot \log(0.1)) = (x, x + 1.135336).$$

Example 1

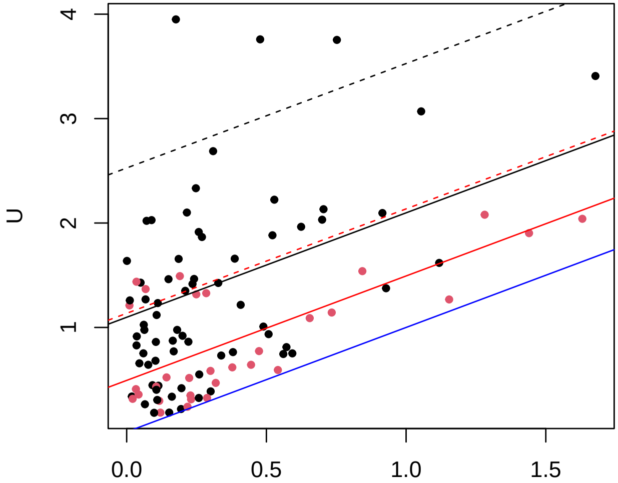
- As the fits are not very good, quantile regression curves might be used to improve the fits since the residual lifetimes have also exponential distributions.
- For example, as $(X_2 - x | X_1 < X_2, X_1 = x)$ has an exponential distribution with mean $1/\alpha_2 \approx 0.4930701$, a lower 90% confidence band for X_2, given $X_1 = x < X_2$, is

$$(x, x - 0.4930701 \cdot \log(0.1)) = (x, x + 1.135336).$$

- The top line is plotted in the right plot (red dashed line) where we see that only four red points out-of-31 points are out of this band (i.e., it contains the 87.09677% of the red points).

Example 1

- As the fits are not very good, quantile regression curves might be used to improve the fits since the residual lifetimes have also exponential distributions.
- For example, as $(X_2 - x|X_1 < X_2, X_1 = x)$ has an exponential distribution with mean $1/\alpha_2 \approx 0.4930701$, a lower 90% confidence band for X_2, given $X_1 = x < X_2$, is
$$(x, x - 0.4930701 \cdot \log(0.1)) = (x, x + 1.135336).$$
- The top line is plotted in the right plot (red dashed line) where we see that only four red points out-of-31 points are out of this band (i.e., it contains the 87.09677% of the red points).
- The 90% lower confidence band for the black points is $(x, x + 2.527039)$ (black dashed line) and it contains 63-out-of-69 black points (i.e., it contains the 95.65217% of the black points).


U
4
3
2
1
0.0
0.5
1.0
1.5

Case 2: Second data with $n = 2$ (parallel system)

- Let us assume now that we have only lifetime data from a parallel system without monitoring the first component failure., that is, we only observe the system lifetimes

$$X_{2:2}^{(j)} = \max(X_1^{(j)}, X_2^{(j)}),\ j = 1, \dots, m.$$

Case 2: Second data with $n=2$ (parallel system)

- Let us assume now that we have only lifetime data from a parallel system without monitoring the first component failure., that is, we only observe the system lifetimes

$$X_{2:2}^{(j)} = \max(X_1^{(j)}, X_2^{(j)}),\ j = 1, \dots, m.$$

- As in case 1, this model is not identifiable, that is, the data $(X_{2:2}^{(j)})_{j=1,\dots,m}$ cannot be used to estimate α_1 and α_2 since

$$\bar{X}_{2:2} = \frac{1}{m}\sum_{j=1}^{m} X_{2:2}^{(j)}$$

satisfies

$$E(\bar{X}_{2:2}) = \mu_{2:2} = \frac{1}{\alpha_1} + \frac{1}{\alpha_2} - \frac{1}{\alpha_1+\alpha_2}. \tag{2.2}$$

Case 2: Second data with $n = 2$ (parallel system)

- Let us assume now that we have only lifetime data from a parallel system without monitoring the first component failure., that is, we only observe the system lifetimes

$$X_{2:2}^{(j)} = \max(X_1^{(j)}, X_2^{(j)}), \ j = 1, \dots, m.$$

- As in case 1, this model is not identifiable, that is, the data $(X_{2:2}^{(j)})_{j=1,\dots,m}$ cannot be used to estimate α_1 and α_2 since

$$\bar{X}_{2:2} = \frac{1}{m} \sum_{j=1}^{m} X_{2:2}^{(j)}$$

satisfies

$$E(\bar{X}_{2:2}) = \mu_{2:2} = \frac{1}{\alpha_1} + \frac{1}{\alpha_2} - \frac{1}{\alpha_1 + \alpha_2}. \tag{2.2}$$

- Again, we need extra information.

Case 2: Second data with $n = 2$ (parallel system)

- For illustration, in the paper we consider three options.

Case 2: Second data with $n = 2$ (parallel system)

- For illustration, in the paper we consider three options.
- In first, we assume that we know which component causes the system failure, that is, we know the indicator random variable $J_i^{(j)}$, where $J_i^{(j)} = 1$ (resp. 0) iff $X_i^{(j)} = X_{2:2}^{(j)}$ (resp. $\neq$) for $i = 1, 2$.

Case 2: Second data with $n = 2$ (parallel system)

- For illustration, in the paper we consider three options.
- In first, we assume that we know which component causes the system failure, that is, we know the indicator random variable $J_i^{(j)}$, where $J_i^{(j)} = 1$ (resp. 0) iff $X_i^{(j)} = X_{2:2}^{(j)}$ (resp. $\neq$) for $i = 1, 2$.
- This kind of data might appear in *autopsy data* from systems that are only available when they fail.

Case 2: Second data with $n = 2$ (parallel system)

- For illustration, in the paper we consider three options.
- In first, we assume that we know which component causes the system failure, that is, we know the indicator random variable $J_i^{(j)}$, where $J_i^{(j)} = 1$ (resp. 0) iff $X_i^{(j)} = X_{2:2}^{(j)}$ (resp. $\neq$) for $i = 1, 2$.
- This kind of data might appear in *autopsy data* from systems that are only available when they fail.
- They could also represent censored data from X_2 that are available just when they exceed X_1.

Case 2: Second data with $n = 2$ (parallel system)

- For illustration, in the paper we consider three options.
- In first, we assume that we know which component causes the system failure, that is, we know the indicator random variable $J_i^{(j)}$, where $J_i^{(j)} = 1$ (resp. 0) iff $X_i^{(j)} = X_{2:2}^{(j)}$ (resp. $\neq$) for $i = 1, 2$.
- This kind of data might appear in *autopsy data* from systems that are only available when they fail.
- They could also represent censored data from X_2 that are available just when they exceed X_1.
- The other options can be seen in the paper.

Case 2: Second data with $n = 2$ (parallel system)

- Obviously, then we also know, which component was the first failure in each experiment, that is, we also know the indicator random variables $I_i^{(j)}$ for $i = 1, 2$ used in case 1.

Case 2: Second data with $n = 2$ (parallel system)

- Obviously, then we also know, which component was the first failure in each experiment, that is, we also know the indicator random variables $I_i^{(j)}$ for $i = 1, 2$ used in case 1.
- Then we could also use m_1 as defined above even if we do not know the lifetimes $X_{1:2}^{(j)}$.

Case 2: Second data with $n = 2$ (parallel system)

- Obviously, then we also know, which component was the first failure in each experiment, that is, we also know the indicator random variables $I_i^{(j)}$ for $i = 1, 2$ used in case 1.
- Then we could also use m_1 as defined above even if we do not know the lifetimes $X_{1:2}^{(j)}$.
- Hence, as in case 1, we can estimate $p_1 = \alpha_1/(\alpha_1 + \alpha_2)$ with $\widehat{p}_1 = m_1/m$. Analogously, we can estimate $p_2 = \alpha_2/(\alpha_1 + \alpha_2)$ with
$$\widehat{p}_2 = 1 - \widehat{p}_1 = \frac{m_2}{m},$$
where $m_2 = m - m_1$.

Case 2: Second data with $n = 2$ (parallel system)

- Here we also know which lifetime data from $X_{2:2}$ belongs to X_2 and then, under exponential distributions, we have

$$\mu_{2|1<2} = E(X_2|X_1 < X_2) = \frac{1}{\alpha_1 + \alpha_2} + \frac{1}{\alpha_2}.$$

Case 2: Second data with $n = 2$ (parallel system)

- Here we also know which lifetime data from $X_{2:2}$ belongs to X_2 and then, under exponential distributions, we have

$$\mu_{2|1<2} = E(X_2|X_1 < X_2) = \frac{1}{\alpha_1 + \alpha_2} + \frac{1}{\alpha_2}.$$

- This quantity can be estimated by the weighted sample mean

$$\bar{X}_{2|1<2} = \frac{1}{m_1} \sum_{j=1}^{m} J_2^{(j)} X_2^{(j)}$$

Case 2: Second data with $n = 2$ (parallel system)

- Here we also know which lifetime data from $X_{2:2}$ belongs to X_2 and then, under exponential distributions, we have

$$\mu_{2|1<2} = E(X_2|X_1 < X_2) = \frac{1}{\alpha_1 + \alpha_2} + \frac{1}{\alpha_2}.$$

- This quantity can be estimated by the weighted sample mean

$$\bar{X}_{2|1<2} = \frac{1}{m_1} \sum_{j=1}^{m} J_2^{(j)} X_2^{(j)}$$

- By using that $\alpha_2 = \frac{1+p_2}{\mu_{2|1<2}}$. we can estimate $\mu_2 = 1/\alpha_2$ by

$$\widehat{\mu}_2 = \frac{\bar{X}_{2|1<2}}{1 + m_2/m} \leq \bar{X}_{2|1<2},$$

where we can use again Slutsky's theorem.

Case 2-Example 1

- Let us consider case 2 with the same simulated data set.

Case 2-Example 1

- Let us consider case 2 with the same simulated data set.
- The sample mean with all the parallel system lifetimes is

$$\bar{X}_{2:2} = 1.172625 \approx \frac{1}{\alpha_1} + \frac{1}{\alpha_2} - \frac{1}{\alpha_1 + \alpha_2} = 1 + \frac{1}{2} - \frac{1}{3} = \frac{7}{6} = 1.166667.$$

Case 2-Example 1

- Let us consider case 2 with the same simulated data set.
- The sample mean with all the parallel system lifetimes is

$$\bar{X}_{2:2} = 1.172625 \approx \frac{1}{\alpha_1}+\frac{1}{\alpha_2}-\frac{1}{\alpha_1+\alpha_2} = 1+\frac{1}{2}-\frac{1}{3} = \frac{7}{6} = 1.166667.$$

- We cannot estimate the parameters α_1 and α_2 from $\bar{X}_{2:2}$.

Case 2-Example 1

- Let us consider case 2 with the same simulated data set.
- The sample mean with all the parallel system lifetimes is

$$\bar{X}_{2:2} = 1.172625 \approx \frac{1}{\alpha_1} + \frac{1}{\alpha_2} - \frac{1}{\alpha_1 + \alpha_2} = 1 + \frac{1}{2} - \frac{1}{3} = \frac{7}{6} = 1.166667.$$

- We cannot estimate the parameters α_1 and α_2 from $\bar{X}_{2:2}$.
- To this purpose we need to separate the data from X_1 and X_2.

Case 2-Example 1

- Let us consider case 2 with the same simulated data set.
- The sample mean with all the parallel system lifetimes is

$$\bar{X}_{2:2} = 1.172625 \approx \frac{1}{\alpha_1} + \frac{1}{\alpha_2} - \frac{1}{\alpha_1 + \alpha_2} = 1 + \frac{1}{2} - \frac{1}{3} = \frac{7}{6} = 1.166667.$$

- We cannot estimate the parameters α_1 and α_2 from $\bar{X}_{2:2}$.
- To this purpose we need to separate the data from X_1 and X_2.
- We also need m_1 and m_2.

Case 2-Example 1

- Let us consider case 2 with the same simulated data set.
- The sample mean with all the parallel system lifetimes is

$$\bar{X}_{2:2} = 1.172625 \approx \frac{1}{\alpha_1} + \frac{1}{\alpha_2} - \frac{1}{\alpha_1 + \alpha_2} = 1 + \frac{1}{2} - \frac{1}{3} = \frac{7}{6} = 1.166667.$$

- We cannot estimate the parameters α_1 and α_2 from $\bar{X}_{2:2}$.
- To this purpose we need to separate the data from X_1 and X_2.
- We also need m_1 and m_2.
- Thus, we obtain the following estimations

$$\widehat{\mu}_1 = \frac{\bar{X}_{1|2<1}}{1 + m_1/m} = \frac{1.305879}{1 + 31/100} = 0.9968541 \approx \mu_1 = 1$$

and

$$\widehat{\mu}_2 = \frac{\bar{X}_{2|1<2}}{1 + m_2/m} = \frac{0.8760269}{1 + 69/100} = 0.5183591 \approx \mu_2 = 0.5.$$

Case 2-Example 1

- Now the estimations from X_1 are a little bit better than that from X_2 since we have more data (69) from X_1 than from X_2 (31) at $X_{2:2}$.

Case 2-Example 1

- Now the estimations from X_1 are a little bit better than that from X_2 since we have more data (69) from X_1 than from X_2 (31) at $X_{2:2}$.
- With the data from X_1 we can estimate μ_2 with

$$\widehat{\mu}_2^* = \frac{\bar{X}_{1|2<1}}{m_2/m_1 + m_2/m} = \frac{1.305879}{69/31 + 69/100} = 0.447862 \approx \mu_2 = 0.5$$

which does not improve the preceding estimation obtained with the data from X_2.

Case 2-Example 1

- Now the estimations from X_1 are a little bit better than that from X_2 since we have more data (69) from X_1 than from X_2 (31) at $X_{2:2}$.
- With the data from X_1 we can estimate μ_2 with

$$\widehat{\mu}_2^* = \frac{\bar{X}_{1|2<1}}{m_2/m_1 + m_2/m} = \frac{1.305879}{69/31 + 69/100} = 0.447862 \approx \mu_2 = 0.5$$

which does not improve the preceding estimation obtained with the data from X_2.

- Conversely, we can estimate μ_1 with the data from X_2 with

$$\widehat{\mu}_1^* = \frac{\bar{X}_{2|1<2}}{m_1/m_2 + m_1/m} = \frac{0.8760269}{(31/69 + 31/100)} = 1.153767 \approx \mu_1 = 1.$$

Comments

- Here we can also predict the first component failures from $(X_2|X_1 < X_2)$ or $(X_1|X_2 < X_1)$.

Comments

- Here we can also predict the first component failures from $(X_2|X_1 < X_2)$ or $(X_1|X_2 < X_1)$.
- In the paper we also consider the case of a 2-out-of-3 system with lifetime $X_{2:3}$.

Comments

- Here we can also predict the first component failures from $(X_2|X_1 < X_2)$ or $(X_1|X_2 < X_1)$.
- In the paper we also consider the case of a 2-out-of-3 system with lifetime $X_{2:3}$.
- The estimations are more complex in this case.

Comments

- Here we can also predict the first component failures from $(X_2|X_1 < X_2)$ or $(X_1|X_2 < X_1)$.
- In the paper we also consider the case of a 2-out-of-3 system with lifetime $X_{2:3}$.
- The estimations are more complex in this case.
- Again, quantile regression techniques can be used to predict $X_{2:3}$ from $X_{1:3}$.

Properties for systems under the MO-model.

- The results are included in the paper:
- Lagos G., Navarro J., Olivero H. (2025). Repair policies decomposition for monotone systems with a Lévy frailty Marshall-Olkin Process. In preparation.

M-O model

- A random vector $(T_1, \ldots, T_n)$ have an exponential Marshall-Olkin (MO) distribution if

$$T_i = \min_{V \subseteq \{1,\ldots,n\} \,:\, i \in V} X_V, \qquad i = 1, \ldots, n, \tag{2.3}$$

where, for all $V \subseteq \{1, \ldots, n\}$, X_V is an exponential random variable with parameter $\lambda_V \geq 0$, and is independent of the other random variables.

M-O model

- A random vector $(T_1, \ldots, T_n)$ have an exponential Marshall-Olkin (MO) distribution if

$$T_i = \min_{V \subseteq \{1,\ldots,n\} : i \in V} X_V, \qquad i = 1, \ldots, n, \tag{2.3}$$

  where, for all $V \subseteq \{1, \ldots, n\}$, X_V is an exponential random variable with parameter $\lambda_V \geq 0$, and is independent of the other random variables.
- The random variables X_V represent the time of arrival of a shock that simultaneously hits all components in the set V.

M-O model

- A random vector $(T_1, \ldots, T_n)$ have an exponential Marshall-Olkin (MO) distribution if

$$T_i = \min_{V \subseteq \{1,\ldots,n\}\,:\,i \in V} X_V, \qquad i = 1, \ldots, n, \tag{2.3}$$

where, for all $V \subseteq \{1, \ldots, n\}$, X_V is an exponential random variable with parameter $\lambda_V \geq 0$, and is independent of the other random variables.

- The random variables X_V represent the time of arrival of a shock that simultaneously hits all components in the set V.
- It is exchangeable when $\lambda_U = \lambda_V$ whenever $|U| = |V|$.

M-O model

- The joint reliability function of $(T_1, \dots, T_n)$ is

$$\Pr(T_1 > t_1, \dots, T_n > t_n) = \exp\left(-\sum_{\emptyset \neq V \subset \{1,\dots,n\}} \lambda_V \max_{i \in V} t_i\right),$$

for $t_1, \dots, t_n \geq 0$.

M-O model

- The joint reliability function of $(T_1, \dots, T_n)$ is

$$\Pr(T_1 > t_1, \dots, T_n > t_n) = \exp\left(-\sum_{\emptyset \neq V \subset \{1,\dots,n\}} \lambda_V \max_{i \in V} t_i\right),$$

  for $t_1, \dots, t_n \geq 0$.
- It is not absolutely continuous (there are ties).

M-O model

- The joint reliability function of $(T_1, \ldots, T_n)$ is

$$\Pr(T_1 > t_1, \ldots, T_n > t_n) = \exp\left(-\sum_{\emptyset \neq V \subset \{1,\ldots,n\}} \lambda_V \max_{i \in V} t_i\right),$$

  for $t_1, \ldots, t_n \geq 0$.
- It is not absolutely continuous (there are ties).
- All the series systems have exponential distributions.

M-O model

- The joint reliability function of $(T_1, \dots, T_n)$ is

$$\Pr(T_1 > t_1, \dots, T_n > t_n) = \exp\left(-\sum_{\emptyset \neq V \subset \{1,\dots,n\}} \lambda_V \max_{i \in V} t_i\right),$$

for $t_1, \dots, t_n \geq 0$.
- It is not absolutely continuous (there are ties).
- All the series systems have exponential distributions.
- All the coherent system distributions are negative mixtures of exponential distributions, Navarro (2022).

Cost of simple repair policies

- At time $t = 0$ the system starts with all its components working.

Cost of simple repair policies

- At time $t = 0$ the system starts with all its components working.
- The repair of j failed components costs c_j.

Cost of simple repair policies

- At time $t = 0$ the system starts with all its components working.
- The repair of j failed components costs c_j.
- The repair of a failed system costs c_{sys} plus the cost of repairing all failed components.

Cost of simple repair policies

- At time $t = 0$ the system starts with all its components working.
- The repair of j failed components costs c_j.
- The repair of a failed system costs c_{sys} plus the cost of repairing all failed components.
- Any failure of a component or the system is detected instantaneously.

Cost of simple repair policies

- At time $t = 0$ the system starts with all its components working.
- The repair of j failed components costs c_j.
- The repair of a failed system costs c_{sys} plus the cost of repairing all failed components.
- Any failure of a component or the system is detected instantaneously.
- Both types of repairs, i.e., of the system and of components, are performed instantaneously.

Cost of simple repair policies

- We will consider the following simple repair policies:

Cost of simple repair policies

- We will consider the following simple repair policies:
- For any $r \in \{1, \dots, n\}$, we define the *r-out-of-n:R repair policy* as the one where all failed components, and the system itself if failed, is repaired in either of the following cases:
  - when r or more components fail;
  - when the system fails.

Cost of simple repair policies

- We will consider the following simple repair policies:
- For any $r \in \{1, \dots, n\}$, we define the *r-out-of-n:R repair policy* as the one where all failed components, and the system itself if failed, is repaired in either of the following cases:
  - when r or more components fail;
  - when the system fails.
- In a system operating under an r-out-of-n:R repair policy, denote by T_r^{sys} and T_r^{rep} the first time of system failure and repair, respectively.

Cost of simple repair policies

- We will consider the following simple repair policies:
- For any $r \in \{1, \ldots, n\}$, we define the *r-out-of-n:R repair policy* as the one where all failed components, and the system itself if failed, is repaired in either of the following cases:
  - when r or more components fail;
  - when the system fails.
- In a system operating under an r-out-of-n:R repair policy, denote by T_r^{sys} and T_r^{rep} the first time of system failure and repair, respectively.
- As the MO-model has the lack of memory property, after a repair the system has the same distribution as a new system.

Example 2

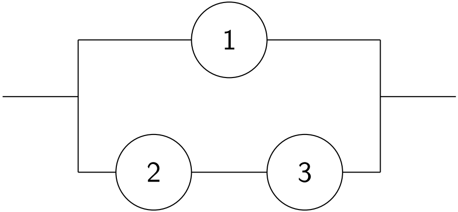


Figure: A coherent system with exchangeable components lifetimes (T_1, T_2, T_3) satisfying the exponential MO-model with $\lambda_1 := \lambda_{\{1\}} = 13/12$, $\lambda_2 := \lambda_{\{1,2\}} = 1/12$, and $\lambda_3 := \lambda_{\{1,2,3\}} = 1/4$.

Example 2: Non-repair case.

- The lifetime T_1 satisfies

$$T_1 = \min(X_{\{1\}}, X_{\{1,2\}}, X_{\{1,3\}}, X_{\{1,2,3\}}).$$

Example 2: Non-repair case.

- The lifetime T_1 satisfies

$$T_1 = \min(X_{\{1\}}, X_{\{1,2\}}, X_{\{1,3\}}, X_{\{1,2,3\}}).$$

- Hence

$$T_1 \sim Exp(\lambda = \lambda_1 + 2\lambda_2 + \lambda_3 = 3/2).$$

Example 2: Non-repair case.

- The lifetime T_1 satisfies
$$T_1 = \min(X_{\{1\}}, X_{\{1,2\}}, X_{\{1,3\}}, X_{\{1,2,3\}}).$$
- Hence
$$T_1 \sim Exp(\lambda = \lambda_1 + 2\lambda_2 + \lambda_3 = 3/2).$$
- If N is the number of broken components when the system fails, we define the **signature of the system in the process** as $w = (w_1, \ldots, w_n)$ where $w_i = \Pr(N = i)$ for $i = 1, \ldots, n$.

Example 2: Non-repair case.

- The lifetime T_1 satisfies
$$T_1 = \min(X_{\{1\}}, X_{\{1,2\}}, X_{\{1,3\}}, X_{\{1,2,3\}}).$$
- Hence
$$T_1 \sim Exp(\lambda = \lambda_1 + 2\lambda_2 + \lambda_3 = 3/2).$$
- If N is the number of broken components when the system fails, we define the **signature of the system in the process** as $w = (w_1, \dots, w_n)$ where $w_i = \Pr(N = i)$ for $i = 1, \dots, n$.
- By using the initial results we get
$$w = (0, 11/20, 9/20).$$

Example 2: Non-repair case.

- The lifetime T_1 satisfies
$$T_1 = \min(X_{\{1\}}, X_{\{1,2\}}, X_{\{1,3\}}, X_{\{1,2,3\}}).$$
- Hence
$$T_1 \sim Exp(\lambda = \lambda_1 + 2\lambda_2 + \lambda_3 = 3/2).$$
- If N is the number of broken components when the system fails, we define the **signature of the system in the process** as $w = (w_1, \ldots, w_n)$ where $w_i = \Pr(N = i)$ for $i = 1, \ldots, n$.
- By using the initial results we get
$$w = (0, 11/20, 9/20).$$
- It does not coincide with the classical (structural) signature (see Navarro (2022)),
$$s = (0, 2/3, 1/3).$$

Case	Lifetimes	N	Probab.	Repair $r=1$	Repair $r=2$
1	$T_1 = T_2 = T_3$	3	$1/15$	No	No
2	$T_1 < T_2 = T_3$	3	$13/360$	**Yes**	No
3	$T_2 < T_1 = T_3$	3	$13/360$	**Yes**	No
4	$T_3 < T_1 = T_2$	3	$13/360$	**Yes**	No
5	$T_1 = T_2 < T_3$	2	$1/45$	No	No
6	$T_1 = T_3 < T_2$	2	$1/45$	No	No
7	$T_2 = T_3 < T_1$	3	$1/45$	**Yes**	**Yes**
8	$T_1 < T_2 < T_3$	2	$91/720$	**Yes**	No
9	$T_1 < T_3 < T_2$	2	$91/720$	**Yes**	No
10	$T_2 < T_1 < T_3$	2	$91/720$	**Yes**	No
11	$T_2 < T_3 < T_1$	3	$91/720$	**Yes**	**Yes**
12	$T_3 < T_1 < T_2$	2	$91/720$	**Yes**	No
13	$T_3 < T_2 < T_1$	3	$91/720$	**Yes**	**Yes**

Table: Options and probabilities needed to compute the signature w of T.

Example 2: Non-repair case.

- The probability of case 1 is

$$p_1 = \Pr(T_1 = T_2 = T_3) = \Pr(X_{\{1,2,3\}} < Z)$$

where $X_{\{1,2,3\}} \sim Exp(\alpha_1 = \lambda_3 = 1/4)$ and

$Z = \min(X_{\{1\}}, X_{\{3\}}, X_{\{3\}}, X_{\{1,2\}}, X_{\{1,3\}}, X_{\{2,3\}}) \sim Exp(\alpha_2 = 7/2).$

Example 2: Non-repair case.

- The probability of case 1 is

$$p_1 = \Pr(T_1 = T_2 = T_3) = \Pr(X_{\{1,2,3\}} < Z)$$

where $X_{\{1,2,3\}} \sim Exp(\alpha_1 = \lambda_3 = 1/4)$ and

$Z = \min(X_{\{1\}}, X_{\{3\}}, X_{\{3\}}, X_{\{1,2\}}, X_{\{1,3\}}, X_{\{2,3\}}) \sim Exp(\alpha_2 = 7/2).$

- As $X_{\{1,2,3\}}$ and Z are independent, then

$$p_1 = \frac{\alpha_1}{\alpha_1 + \alpha_2} = \frac{1/4}{1/4 + 7/2} = \frac{1}{15}.$$

Example 2: Non-repair case.

- The probability of case 1 is

$$p_1 = \Pr(T_1 = T_2 = T_3) = \Pr(X_{\{1,2,3\}} < Z)$$

where $X_{\{1,2,3\}} \sim Exp(\alpha_1 = \lambda_3 = 1/4)$ and

$Z = \min(X_{\{1\}}, X_{\{3\}}, X_{\{3\}}, X_{\{1,2\}}, X_{\{1,3\}}, X_{\{2,3\}}) \sim Exp(\alpha_2 = 7/2)$.

- As $X_{\{1,2,3\}}$ and Z are independent, then

$$p_1 = \frac{\alpha_1}{\alpha_1 + \alpha_2} = \frac{1/4}{1/4 + 7/2} = \frac{1}{15}.$$

- The other probabilities are computed in a similar way (by using the initial results) and thus we can obtain w_2 and w_3.

Example 2: Non-repair case.

- Hence, the expected cost in a cycle without repairs is

$$c(T) = c_{sys} + c_2 w_2 + c_3 w_3 = c_{sys} + c_2 \frac{11}{20} + c_3 \frac{9}{20}.$$

Example 2: Non-repair case.

- Hence, the expected cost in a cycle without repairs is

$$c(T) = c_{sys} + c_2 w_2 + c_3 w_3 = c_{sys} + c_2 \frac{11}{20} + c_3 \frac{9}{20}.$$

- If we choose the values $c_{sys} = 1$ and $c_i = i$, $i = 1, 2, 3$, we get

$$c(T) = 1 + 2w_2 + 3w_3 = \frac{69}{20} = 3.45.$$

Example 2: Non-repair case.

- Hence, the expected cost in a cycle without repairs is

$$c(T) = c_{sys} + c_2 w_2 + c_3 w_3 = c_{sys} + c_2 \frac{11}{20} + c_3 \frac{9}{20}.$$

- If we choose the values $c_{sys} = 1$ and $c_i = i$, $i = 1, 2, 3$, we get

$$c(T) = 1 + 2w_2 + 3w_3 = \frac{69}{20} = 3.45.$$

- The Mean Time To Failure is

$$E(T) = \frac{31}{40} = 0.775.$$

Example 2: Non-repair case.

- Hence, the expected cost in a cycle without repairs is

$$c(T) = c_{sys} + c_2 w_2 + c_3 w_3 = c_{sys} + c_2 \frac{11}{20} + c_3 \frac{9}{20}.$$

- If we choose the values $c_{sys} = 1$ and $c_i = i$, $i = 1, 2, 3$, we get

$$c(T) = 1 + 2w_2 + 3w_3 = \frac{69}{20} = 3.45.$$

- The Mean Time To Failure is

$$E(T) = \frac{31}{40} = 0.775.$$

- To measure the quality of the system we can consider the Mean Cost of a Cycle (MCC) defined as

$$MCC(T) = \frac{c(T)}{E(T)} = \frac{69/20}{31/40} = \frac{138}{31} = 4.451613.$$

Case	Lifetimes	N	Probab.	Repair $r=1$	Repair $r=2$
1	$T_1 = T_2 = T_3$	3	$1/15$	No	No
2	$T_1 < T_2 = T_3$	3	$13/360$	**Yes**	No
3	$T_2 < T_1 = T_3$	3	$13/360$	**Yes**	No
4	$T_3 < T_1 = T_2$	3	$13/360$	**Yes**	No
5	$T_1 = T_2 < T_3$	2	$1/45$	No	No
6	$T_1 = T_3 < T_2$	2	$1/45$	No	No
7	$T_2 = T_3 < T_1$	3	$1/45$	**Yes**	**Yes**
8	$T_1 < T_2 < T_3$	2	$91/720$	**Yes**	No
9	$T_1 < T_3 < T_2$	2	$91/720$	**Yes**	No
10	$T_2 < T_1 < T_3$	2	$91/720$	**Yes**	No
11	$T_2 < T_3 < T_1$	3	$91/720$	**Yes**	**Yes**
12	$T_3 < T_1 < T_2$	2	$91/720$	**Yes**	No
13	$T_3 < T_2 < T_1$	3	$91/720$	**Yes**	**Yes**

Table: Options needed to compute the probability of repair.

Example 2: Repair case with $r = 1$.

- For $r = 1$, the probability of repair is

$$p_R = 3\frac{13}{360} + \frac{1}{45} + 6\frac{91}{720} = \frac{8}{9} = 0.8888889.$$

Example 2: Repair case with $r = 1$.

- For $r = 1$, the probability of repair is

$$p_R = 3\frac{13}{360} + \frac{1}{45} + 6\frac{91}{720} = \frac{8}{9} = 0.8888889.$$

- The expected cost in a cycle with repair is

$$c(T) = c_2 w_2^R + c_3 w_3^R$$

where the conditional signature with repair is

$$\mathrm{w}^R = \left(\frac{39}{40}, \frac{1}{40}, 0\right).$$

Example 2: Repair case with $r = 1$.

- For $r = 1$, the probability of repair is

$$p_R = 3\frac{13}{360} + \frac{1}{45} + 6\frac{91}{720} = \frac{8}{9} = 0.8888889.$$

- The expected cost in a cycle with repair is

$$c(T) = c_2 w_2^R + c_3 w_3^R$$

where the conditional signature with repair is

$$\mathrm{w}^R = \left(\frac{39}{40}, \frac{1}{40}, 0\right).$$

- The probability of "no repair" is $p_{\bar{R}} = 1/9$ and its conditional signature is

$$\mathrm{w}^{\bar{R}} = \left(0, \frac{2}{5}, \frac{3}{5}\right).$$

Example 2: Repair case with $r = 1$.

- Therefore, for $r = 1$, we get the expected cost

$$c(T_{(1)}) = c_{sys} + \frac{p_R}{1-p_R}\sum_{i=r}^{n-1} w_i^R c_i + \sum_{i=1}^{n} w_i^{\bar{R}} c_i = \frac{59}{5} = 11.8$$

and the MTTF

$$E(T_{(1)}) = \frac{p_R}{1-p_R} E(T_{(1)}|R) + E(T_{(1)}|\bar{R}) = \frac{12}{5} = 2.4.$$

Example 2: Repair case with $r = 1$.

- Therefore, for $r = 1$, we get the expected cost

$$c(T_{(1)}) = c_{sys} + \frac{p_R}{1 - p_R} \sum_{i=r}^{n-1} w_i^R c_i + \sum_{i=1}^{n} w_i^{\bar{R}} c_i = \frac{59}{5} = 11.8$$

  and the MTTF

$$E(T_{(1)}) = \frac{p_R}{1 - p_R} E(T_{(1)} | R) + E(T_{(1)} | \bar{R}) = \frac{12}{5} = 2.4.$$

- Hence, the Mean Cost of a Cycle for $r = 1$ is

$$MCC(T_{(1)}) = \frac{c(T_{(1)})}{E(T_{(1)})} = \frac{59/5}{12/5} = \frac{59}{12} = 4.916667.$$

Example 2: Repair case with $r = 2$.

- Analogously, for $r = 2$, we get

$$MCC(T_{(2)}) = \frac{c(T_{(2)})}{E(T_{(2)})} = \frac{4}{0.816092} = 4.901408.$$

Example 2: Repair case with $r = 2$.

- Analogously, for $r = 2$, we get

$$MCC(T_{(2)}) = \frac{c(T_{(2)})}{E(T_{(2)})} = \frac{4}{0.816092} = 4.901408.$$

- Hence

$$MCC(T) = 4.451613 < MCC(T_{(2)}) = 4.901408 < MCC(T_{(1)}) = 4.916667.$$

Example 2: Repair case with $r = 2$.

- Analogously, for $r = 2$, we get

$$MCC(T_{(2)}) = \frac{c(T_{(2)})}{E(T_{(2)})} = \frac{4}{0.816092} = 4.901408.$$

- Hence

$$MCC(T) = 4.451613 < MCC(T_{(2)}) = 4.901408 < MCC(T_{(1)}) = 4.916667.$$

- Therefore, in this case, the best option is the one without repairs.

References

Azzalini A, Capitanio A (2014). The skew-normal and related families. Cambridge University Press, Cambridge.

Balakrishnan N., Cramer E. (2014). The Art of Progressive Censoring. Applications to Reliability and Quality. Birkhäuser, Boston.

Lagos G., Navarro J., Olivero H. (2025). Repair policies decomposition for monotone systems with a Lévy frailty Marshall-Olkin Process. In preparation.

Loperfido N, Navarro J, Ruiz JM, Sandoval CJ (2007). Some relationships between skew-normal distributionsand order statistics from exchangeable normal random vectors. Commun Stat Theory Methods 36, 1719–1733.

Navarro J. (2022). Introduction to System Reliability Theory. Springer.

Navarro J, Arevalillo JM (2023). On connections between skewed, weighted and distorted distributions: applications to model extreme value distributions. Test 32, 1307–1335.

Navarro J., Calì C., Longobardi M., Durante F. (2022a). Distortion representations of multivariate distributions. Statistical Methods and Applications 31, 925–954.

Navarro J., Cramer E., Balakrishnan N. (2025). Inference under conditionally ordered heterogeneous exponential data. Submitted.

Navarro J., Sarabia J.M. (2024). A note on the limiting behaviour of hazard rate functions of generalized mixtures. Journal of Computational and Applied Mathematics 435, 114653.

Roberts C (1966). A correlation model useful in the study of twins. J Am Stat Assoc 61:1184–1190

Final slide

- The slides and more references are in my web page

  https://webs.um.es/jorgenav/miwiki/doku.php

Final slide

- The slides and more references are in my web page

  https://webs.um.es/jorgenav/miwiki/doku.php

- That's all. Thank you for your attention!!

Final slide

- The slides and more references are in my web page

  https://webs.um.es/jorgenav/miwiki/doku.php

- That's all. Thank you for your attention!!
- Questions?