AISM (2008) 60:151–171
DOI 10.1007/s10463-006-0070-8

N. Balakrishnan · Erhard Cramer

Progressive censoring from heterogeneous distributions with applications to robustness

Received: 1 August 2005 / Revised: 26 April 2006 / Published online: 19 July 2006

Abstract Progressively censored order statistics from heterogeneous distributions are introduced and their properties are investigated. After deriving the joint density function, some properties are established. In particular, the case of proportional hazards leads to an interesting connection to the model of generalized order statistics. Finally, the special case of exponential distribution is considered and some known results are generalized to this heterogeneous case, and their implications to robustness are highlighted.

Keywords Progressive censoring · Order statistics from non-identically distributed random sample · Generalized order statistics · Spacings · Single outlier · Exponential distribution · Permanents · Proportional hazards

1 Introduction

Order statistics from independent and identically distributed (iid) random variables $X_1, \ldots, X_n$ have been discussed extensively in the literature; see, for example, David and Nagaraja (2003) and Arnold et al. (1992). Although the assumptions of independence and identical distribution are restrictive, the model is useful in many applications. Moreover, it serves as an important tool in probability and statistics. However, these assumptions are important only in the probabilistic analysis of order statistics $X_{1:n} \leq \cdots \leq X_{n:n}$ of real-valued random variables $X_1, \ldots, X_n$, which is easy to handle in the iid case.

N. Balakrishnan
Department of Mathematics and Statistics, McMaster University, 1280 Main Street West, Hamilton, ON, Canada L8S 4K1

E. Cramer (✉)
Institute of Statistics, RWTH Aachen University, 52056 Aachen, Germany
E-mail: Erhard.cramer@rwth-aachen.de

In order to get more flexible models of order statistics, some relaxations of these assumptions have been considered in the literature; see, for example, Harter and Balakrishnan (1998). In this paper, we are interested in the case when $X_1, \ldots, X_n$ are independent but non-identically distributed (inid). Properties of order statistics in this general set-up have been discussed in the literature first in the early 1970s by Sen (1970) and Pledger and Proschan (1971). Most of the available results are based on explicit expressions for density functions of order statistics. For example, the joint density function of all order statistics in this case is given by

$$f_{X_{1:n},\ldots,X_{n:n}}(t_1, \ldots, t_n) = \sum_{\pi \in S_n} \prod_{j=1}^{n} f_{\pi(j)}(t_j), \quad t_1 \leq \cdots \leq t_n, \tag{1}$$

where

$$S_n = \left\{(i_1, \ldots, i_n) \in \{1, 2, \ldots, n\}^n \mid i_j \neq i_k, j \neq k, 1 \leq j, k \leq n\right\}$$

denotes the set of all permutations π of $(1, 2, \ldots, n)$. Although the distribution theory, such as marginal distributions, is well-known, Eq. 1 reveals that the resulting expressions may often be complicated if not intractable. Due to these difficulties, the available literature under this general setting is somewhat scarce. Surveys and related references are provided by Arnold and Balakrishnan (1989) and Harter and Balakrishnan (1998). An useful tool to handle Eq. 1 and the marginal distributions is provided by the theory of permanents. Its connection to marginal density functions of order statistics from inid random variables was first noted by Vaughan and Venables (1972). After their observation, the "permanent approach" has been exploited by some authors including Balakrishnan (1988, 1989), Bapat and Beg (1989a, b), and Bapat (1990). Other papers on properties of order statistics from inid variables make use of the explicit expressions of the density functions; see, for example, Kochar and Kirmani (1995), Kochar and Korwar (1996), and Kochar and Rojo (1996). An important application of order statistics from inid variables is in the modelling of outliers and robustness studies. In particular, the case of a single outlier, viz., $F_1 = \cdots = F_{n-1} = F$ and $F_n = G \neq F$, has received great attention in the literature; see, for example, Arnold and Balakrishnan (1989), Barnett and Lewis (1993), and Balakrishnan (1994).

In this paper, we aim to generalize the model of progressive censoring to the case of inid variables. In the model of progressively Type-II censored order statistics, some of the underlying random variables $X_1, \ldots, X_n$ are censored during the observation. In particular, this means that in a life-testing experiment with n independent units, a pre-fixed number R_1 of surviving units are randomly censored from the sample after the first failure time, $\min\{X_1, \ldots, X_n\}$. Then, at the first failure time of the remaining $n - R_1 - 1$ units, R_2 units are censored, and so on. Finally, at the time of the m-th failure, all the remaining $R_m = n - m - R_1 - \cdots - R_{m-1}$ units are censored. For a detailed description of this progressive censoring scheme and related developments, one may refer to the book by Balakrishnan and Aggarwala (2000).

Note that while carrying out this life-test, it is assumed that the units being tested have iid life-times $X_1, \ldots, X_n$ with distribution function $F(\cdot)$. In this paper, we develop results when the underlying life-times have heterogeneous distributions, viz., $X_j \sim F_j,\ 1 \leq j \leq n$. In Sect. 2, we first derive the joint density function of all m progressively censored order statistics. In Sect. 3, we present some examples by considering some special cases. In Sect. 4, we give a permanent expression for the joint density function of the m progressively censored order statistics as well as the joint density function of the first p progressively censored order statistics. This generalizes the result of Vaughan and Venables (1972) for the usual order statistics. In the case of proportional hazards, we establish that the density functions can be represented as a mixture of densities of generalized order statistics. This yields immediately some properties of progressively censored order statistics based on heterogeneous distributions. For example, we show that the spacings have the decreasing failure rate (DFR) property (if the underlying distribution function F is DFR) which is a generalization of a result of Gupta and Kirmani (1988) for the usual order statistics; see also Kamps (1995). In Sect. 5, we focus on the special case of exponential distributions and establish that the first spacing is independent of all other spacings. Moreover, the resulting one-dimensional marginal distributions turn out to be a mixture of exponential distributions. This extends the results of Gross et al. (1986) and Kochar and Korwar (1996) for the usual order statistics from inid exponential random variables. Finally, in Sect. 6, we illustrate some implications of these results in robustness studies.

Due to the framework of life-time experiments, the results presented subsequently are mainly formulated in terms of absolutely continuous life distributions. However, most results are valid for general absolutely continuous distributions, too.

2 Joint density of progressively censored order statistics

Let $F_1, \ldots, F_n$ be absolutely continuous distribution functions with densities $f_1, \ldots, f_n$, respectively. We consider a sample of independent random variables $X_1, \ldots, X_n$, where $X_r \sim F_r,\ 1 \leq r \leq n$. Moreover, let $\boldsymbol{R} = (R_1, \ldots, R_m) \in \mathbb{N}_0^m$ be the progressive censoring scheme with $n = m + \sum_{i=1}^m R_i$. For brevity, we denote by $\gamma_1 = \sum_{i=1}^m (R_i + 1) = n$ the sample size and by $\gamma_j = \sum_{i=j}^m (R_i + 1)$ the number of units remaining in the experiment after the $(j-1)$-th failure for $2 \leq j \leq m$. Let $X_{1:m:n}^{\boldsymbol{R}}, \ldots, X_{m:m:n}^{\boldsymbol{R}}$ denote the progressively Type-II censored order statistics observed from such a progressively censored life-test.

Theorem 1 *For $n \in \mathbb{N}$, let S_n be the set of all permutations π of $(1, 2, \ldots, n)$. For brevity, let $\rho_r = R_1 + \cdots + R_r,\ 1 \leq r \leq m$, with $\rho_0 = 0$ and $\rho_m = n - m$. Then, the joint density of $X_{1:m:n}^{\boldsymbol{R}}, \ldots, X_{m:m:n}^{\boldsymbol{R}}$ is given by*

$$f_{X_{1:m:n}^{\boldsymbol{R}}, \ldots, X_{m:m:n}^{\boldsymbol{R}}}(t_1, \ldots, t_m) = \frac{1}{(n-1)!} \left(\prod_{j=2}^m \gamma_j \right) \sum_{\pi \in S_n} \prod_{j=1}^m f_{\pi(j)}(t_j) \times \left\{ \prod_{r=m+\rho_{j-1}+1}^{m+\rho_j} \bar{F}_{\pi(r)}(t_j) \right\}, \quad t_1 \leq \cdots \leq t_m, \tag{2}$$

where $\pi(i)$ is the i-th component of the permutation vector $\pi \in S_n$, $1 \le i \le n$.

Proof In order to prove this result, we first consider the joint distribution function

$$P\left(X^{\boldsymbol{R}}_{1:m:n} \le x_1, \dots, X^{\boldsymbol{R}}_{m:m:n} \le x_m\right) \quad \text{for} \qquad x_1 \le \dots \le x_m. \tag{3}$$

Otherwise, the joint distribution function reduces to a marginal distribution function, which means that at least one of the variables x_j drops out. Since the expression in (3) will be differentiated later on, the density will be zero at such a point. Hence, without loss of any generality, we can assume the order $x_1 \le \dots \le x_m$ in (3).

The definition of progressively Type-II censored order statistics is based on the following construction. First, the minimum $\min\{X_1, \dots, X_n\}$ is observed. Then, R_1 of the $n-1$ surviving units are randomly censored. Afterwards, the minimum of the remaining $n - R_1 - 1$ surviving units is observed, at which time R_2 of the $n-2-R_1$ surviving units are randomly censored, and so on. For now, let us assume that we know exactly which units fail and which are censored. We consider the following specific outcome of the progressive censoring procedure:

$$\begin{array}{lll}
X_1 \to \text{failure}, & X_{m+1}, \dots, X_{m+\rho_1} & \to \text{censored} \\
X_2 \to \text{failure}, & X_{m+\rho_1+1}, \dots, X_{m+\rho_2} & \to \text{censored} \\
\quad\vdots & & \quad\vdots \\
X_m \to \text{failure}, & X_{m+\rho_{m-1}+1}, \dots, X_{m+\rho_m} & \to \text{censored.}
\end{array}$$

This means that, before the r-th failure, the units with indices $r, \dots, m$ and $m + \rho_{r-1} + 1, \dots, n$ are still in the experiment. Moreover, the previous set-up fixes that the r-th failure is assigned to the unit number r, $1 \le r \le m$. In order to simplify the notation, we introduce the random vectors $Z_r = \left(X_{m+\rho_{r-1}+1}, \dots, X_{m+\rho_r}\right)$, $1 \le r \le m$. The components of Z_r represent the life-times of those units which are censored immediately after the r-th failure. Using this notation, we can write

$$\min\left\{X_{r+1}, \dots, X_m, X_{m+\rho_{r-1}+1}, \dots, X_n\right\} = \min\{X_{r+1}, \dots, X_m, Z_r, \dots, Z_m\}.$$

Then,

$$\begin{aligned} &P\left(X_r \le \min\{X_{r+1}, \dots, X_m, Z_r, \dots, Z_m\}, X_r \le x_r, 1 \le r \le m\right), \\ &\quad x_1 \le \dots \le x_m, \end{aligned} \tag{4}$$

denotes the probability that the random variables X_r represent the failure times (which are supposed to be less than x_r) and that the units corresponding to the components of Z_r are censored (after the r-th failure), $1 \le r \le m$.

The probability in (4) can be calculated as follows:

$$
\begin{aligned}
&P\left(X_r \le \min\left\{X_{r+1}, \ldots, X_m, Z_r, \ldots, Z_m\right\}, X_r \le x_r, 1 \le r \le m\right) \\
&\quad = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_m} P\left(t_r \le \min\left\{t_{r+1}, \ldots, t_m, Z_r, \ldots, Z_m\right\}, 1 \le r \le m\right) \\
&\qquad \times \prod_{j=1}^{m} f_j(t_j)\, dt_m \ldots dt_1 \\
&\quad = \int_{-\infty}^{x_1} \int_{t_1}^{x_2} \cdots \int_{t_{m-1}}^{x_m} P\left(t_r \le \min\left\{Z_r, \ldots, Z_m\right\}, 1 \le r \le m\right) \\
&\quad \times \prod_{j=1}^{m} f_j(t_j)\, \mathrm{d}t_m \cdots \mathrm{d}t_1.
\end{aligned} \tag{5}
$$

Now, let us consider the probability term in the integrand in (5). Using the definition of the minimum and $t_1 \le \cdots \le t_m$, we obtain the following expression:

$$
\begin{aligned}
&P\left(t_r \le \min\left\{Z_r, \ldots, Z_m\right\},\ 1 \le r \le m\right) \\
&\quad = P\left(t_1 \le \min\{Z_1\}, t_r \le \min\{Z_2\}, 1 \le r \le 2, \ldots, t_r \le \min\{Z_m\}, 1 \le r \le m\right) \\
&\quad = P\left(t_1 \le \min\{Z_1\}\right)\ P\left(t_2 \le \min\{Z_2\}\right) \ldots P\left(t_m \le \min\{Z_m\}\right) \\
&\quad = \prod_{j=1}^{m} \prod_{r=m+\rho_{j-1}+1}^{m+\rho_j} \bar{F}_r(t_j), \quad t_1 \le \cdots \le t_m.
\end{aligned}
$$

Hence, we obtain

$$
\begin{aligned}
&P\left(X_r \le \min\left\{X_{r+1}, \ldots, X_m, Z_r, \ldots, Z_m\right\}, X_r \le x_r, 1 \le r \le m\right) \\
&\quad = \int_{-\infty}^{x_1} \int_{t_1}^{x_2} \cdots \int_{t_{m-1}}^{x_m} \prod_{j=1}^{m} f_j(t_j) \left\{ \prod_{r=m+\rho_{j-1}+1}^{m+\rho_j} \bar{F}_r(t_j) \right\} \mathrm{d}t_m \cdots \mathrm{d}t_1.
\end{aligned}
$$

Differentiation of (5) with respect to $x_1, \ldots, x_m$ yields the function

$$
h_I(x_1, \ldots, x_m) = \prod_{j=1}^{m} f_j(x_j) \left\{ \prod_{r=m+\rho_{j-1}+1}^{m+\rho_j} \bar{F}_r(x_j) \right\}, \quad x_1 \le \cdots \le x_m,
$$

where $I = (1, \ldots, n)$. Choosing a permutation $\pi(I) = (\pi(1), \ldots, \pi(n))$ of $(1, \ldots, n)$, this leads to the expression

$$
h_{\pi(I)}(x_1, \ldots, x_m) = \prod_{j=1}^{m} f_{\pi(j)}(x_j) \left\{ \prod_{r=m+\rho_{j-1}+1}^{m+\rho_j} \bar{F}_{\pi(r)}(x_j) \right\}, \quad x_1 \le \cdots \le x_m.
$$

In the next step, we have to take into account that those units that are censored at the r-th failure are censored at random. The procedure works as follows. First, we specify a number i_1 out of $n = \gamma_1$ units and assign this number to the first failure. Then, from the remaining numbers, we choose randomly R_1 values $i_2, \ldots, i_{R_1+1}$ out of $\gamma_1 - 1$ (with ordering!) and remove the associated units from the experiment. The corresponding probability is $\frac{(\gamma_1-R_1-1)!}{(\gamma_1-1)!} = \frac{\gamma_2!}{(\gamma_1-1)!}$. Then, we choose a new failure time $X_{i_{R_1+2}}$ out of γ_2 possible random variables, and so on. Continuing this process, we obtain a permutation $(i_1, \ldots, i_n)$ of $(1, \ldots, n)$, i.e, $(i_1, \ldots, i_n) = \pi(I)$.

The probability to choose a specific permutation π leading to the previous outcome is then given by

$$\frac{\gamma_2!}{(\gamma_1-1)!} \times \frac{\gamma_3!}{(\gamma_2-1)!} \times \cdots \times \frac{\gamma_m!}{(\gamma_{m-1}-1)!} \times \frac{1}{(\gamma_m-1)!}$$
$$= \frac{1}{(\gamma_1-1)!} \prod_{j=2}^{m} \gamma_j = \frac{1}{(n-1)!} \prod_{j=2}^{m} \gamma_j .$$

Denoting by the event A_π the assignment of the permutation π to the unit indices, we obtain

$$\begin{aligned} &P\left(X_{1:m:n}^{\boldsymbol{R}} \le x_1, \ldots, X_{m:m:n}^{\boldsymbol{R}} \le x_m\right) \\ &= \sum_{\pi \in S_n} P\left(X_{1:m:n}^{\boldsymbol{R}} \le x_1, \ldots, X_{m:m:n} \le x_m \mid A_\pi\right) P(A_\pi) \\ &= \frac{1}{(n-1)!} \prod_{j=2}^{m} \gamma_j \sum_{\pi \in S_n} P\big(X_{\pi(r)} \\ &\le \min\left\{X_{\pi(r+1)}, \ldots, X_{\pi(m)}, Z_r^\pi, \ldots, Z_m^\pi\right\}, \\ &X_{\pi(r)} \le x_r, 1 \le r \le m\big), \end{aligned}$$

where $Z_j^\pi = \left(X_{\pi(m+\rho_{j-1}+1)}, \ldots, X_{\pi(m+\rho_j)}\right),\ 1 \le j \le m$. By construction, A_π is independent of the random variables $X_1, \ldots, X_n$ so that the condition can be omitted after specifying the associated outcome of the progressive censoring procedure. Any term in the above sum is of the form defined in (4). In order to apply the preceding results, the specific permutation of the indices has to be taken into account only. Therefore, differentiation of the preceding expression yields

$$\frac{1}{(n-1)!} \prod_{j=2}^{m} \gamma_j \sum_{\pi \in S_n} h_{\pi(I)}(x_1, \ldots, x_m),$$

which is the joint density of $X_{1:m:n}^{\boldsymbol{R}}, \ldots, X_{m:m:n}^{\boldsymbol{R}}$ presented in Eq. 2. $\square$

Remark 1 The proof of representation (2) mimics the construction process of progressively censored order statistics. It is rigorous and, thus, quite technical and complicated. The following argument due to an anonymous referee gives a short intuitive explanation. Given $\pi \in S_n$, the probability density function $f_{X_{1:m:n}^{\boldsymbol{R}}, \ldots, X_{m:m:n}^{\boldsymbol{R}}}(t_1, \ldots, t_m)$ has three factors:

- The product of the constants represents the experimenter randomly removing the specific ordered set of censored items. The probability $\prod_{j=1}^{m} \gamma_j/n!$ to remove a certain ordered set results from the arguments presented in the proof.
- The product of the probability density functions represents the likelihood for the observed failure times,
- The associated product of the survival probabilities represents the probability of the chosen items being alive when they are removed.

Remark 2 It should be noted that the number of terms in the sum $\sum_{\pi \in S_n}$ can be reduced. This can be seen from the following argument. Let $j \in \{1, \dots, n\}$ and π be a permutation with

$$\pi(r) \in B_j = \{m + \rho_{j-1} + 1, \dots, m + \rho_j\} \quad \text{and} \quad \pi(r) = r, r \notin B_j,$$

i.e., π has fix points in the complement of B_j. Then, $\prod_{r=m+\rho_{j-1}+1}^{m+\rho_j} \bar{F}_{\pi(r)}(t_j) = \prod_{r=m+\rho_{j-1}+1}^{m+\rho_j} \bar{F}_r(t_j)$ is independent of π and thus $h_{\pi(I)} = h_I$ obviously holds. Since $R_j!$ permutations have this property, $R_j!$ terms of the sum $\sum_{\pi \in S_n}$ are identical.

3 Examples

Example 1 In the case of the usual order statistics $X_{1:n}, \dots, X_{n:n}$ from non-identical distributions, we have $m = n$ as well as $\gamma_j = n - j + 1$ and $R_j = 0$ for $1 \le j \le n$, so that the normalizing constant becomes 1. Hence, the well-known representation in (1) results; see, for example, Arnold and Balakrishnan (1989, p. 135).

Example 2 Considering the progressively Type-II censored order statistics from identical distributions, we arrive at the representation

$$f_{X^{\boldsymbol{R}}_{1:m:n}, \dots, X^{\boldsymbol{R}}_{m:m:n}}(t_1, \dots, t_m) = \left(\prod_{j=1}^{m} \gamma_j \right) \prod_{j=1}^{m} f(t_j) \left\{ \bar{F}(t_j) \right\}^{R_j}, \quad t_1 \le \dots \le t_m,$$

since $\gamma_1 = n$ and $\sum_{\pi \in S_n} 1 = n!$, which is the same as presented, for example, in Balakrishnan and Aggarwala (2000, p. 8).

Note that Theorem 1 also gives a formal derivation of the joint density of progressively Type-II censored order statistics from identical distributions introduced in Cohen (1963).

Example 3 The case when $F_1 = \dots = F_{n-1} = F$ and $F_n = G$ is of special interest in the modelling of a single outlier; see Barnett and Lewis (1993). With f and g as the corresponding densities, the joint density function of $X^{\boldsymbol{R}}_{1:m:n}, \dots, X^{\boldsymbol{R}}_{m:m:n}$ in (2) simplifies considerably. For $t_1 \le \dots \le t_m$, we obtain

$$
\begin{aligned}
& f_{X^{\boldsymbol{R}}_{1:m:n},\ldots,X^{\boldsymbol{R}}_{m:m:n}}(t_1,\ldots,t_m) \\
&\quad = \frac{1}{(n-1)!}\left(\prod_{j=2}^{m}\gamma_j\right)\sum_{\pi\in S_n}\prod_{j=1}^{m} f_{\pi(j)}(t_j)\left\{\prod_{r=m+\rho_{j-1}+1}^{m+\rho_j}\bar{F}_{\pi(r)}(t_r)\right\} \\
&\quad = \frac{1}{(n-1)!}\prod_{j=2}^{m}\gamma_j\left[\sum_{\nu=1}^{m}\left(\sum_{\pi\in S_n,\pi(\nu)=n}\prod_{j=1}^{m} f_{\pi(j)}(t_j)\left\{\prod_{r=m+\rho_{j-1}+1}^{m+\rho_j}\bar{F}_{\pi(r)}(t_r)\right\}\right)\right. \\
&\qquad \left. + \sum_{\nu=m+1}^{n}\left(\sum_{\pi\in S_n,\pi(\nu)=n}\prod_{j=1}^{m} f_{\pi(j)}(t_j)\left\{\prod_{r=m+\rho_{j-1}+1}^{m+\rho_j}\bar{F}_{\pi(r)}(t_j)\right\}\right)\right] \\
&\quad = \frac{1}{(n-1)!}\prod_{j=2}^{m}\gamma_j\left\{\sum_{\nu=1}^{m}\left(\sum_{\pi\in S_n,\pi(\nu)=n} g(t_\nu)\left\{\bar{F}(t_\nu)\right\}^{R_\nu}\prod_{\substack{j=1\\ j\neq\nu}}^{m} f(t_j)\left\{\bar{F}(t_j)\right\}^{R_j}\right)\right. \\
&\qquad + \sum_{i=1}^{m}\sum_{\nu=m+\rho_{i-1}+1}^{m+\rho_i}\left(\sum_{\pi\in S_n,\pi(\nu)=n} f(t_i)\left\{\bar{F}(t_i)\right\}^{R_i-1}\bar{G}(t_i)\right. \\
&\qquad\qquad \left.\left. \times\prod_{\substack{j=1\\ j\neq i}}^{m} f(t_j)\left\{\bar{F}(t_j)\right\}^{R_j}\right)\right\} \\
&\quad = \left(\prod_{j=2}^{m}\gamma_j\right)\sum_{\nu=1}^{m}\left\{g(t_\nu)\bar{F}(t_\nu)+R_\nu f(t_\nu)\bar{G}(t_\nu)\right\}\left\{\bar{F}(t_\nu)\right\}^{R_\nu-1} \\
&\qquad \times\prod_{\substack{j=1\\ j\neq\nu}}^{m} f(t_j)\left\{\bar{F}(t_j)\right\}^{R_j}.
\end{aligned}
$$

In the preceding calculations, we have used the facts that

$$\sum_{\pi\in S_n,\pi(\nu)=n} 1 = \sum_{\pi\in S_{n-1}} 1 = (n-1)!,\quad n\in\mathbb{N}.$$

In the special case when $m = n$ and $R_j = 0$ for $j = 1,\ldots,n$, the above readily reduces to the well-known formula for distributions of order statistics from a single-outlier model presented, for example, in Kale and Sinha (1971) and Joshi (1972).

4 Permanent representation and proportional hazards case

Vaughan and Venables (1972) displayed that densities of usual order statistics from non-identical distributions can be written in terms of a permanent

$$\overset{+}{|}A\overset{+}{|}=\sum_{\pi\in S_n}\prod_{i=1}^{n}a_{i\pi(i)},$$

where $A=(a_{ij})$ is an appropriately defined $n\times n$ matrix (more information on permanents can be found in Minc, 1978). Using the same notation as in Sect. 2, a similar representation holds for the joint density of progressively Type-II censored order statistics from non-identical distributions as well:

$$f_{X^{\boldsymbol{R}}_{1:m:n},\dots,X^{\boldsymbol{R}}_{m:m:n}}(t_1,\dots,t_m)=\frac{1}{(n-1)!}\prod_{j=2}^{m}\gamma_j\quad\times\overset{+\qquad\qquad\qquad +}{\left|\begin{matrix} f_1(t_1) & \cdots & f_n(t_1)\\ \bar{F}_1(t_1) & \cdots & \bar{F}_n(t_1)\\ \vdots & & \vdots\\ \bar{F}_1(t_1) & \cdots & \bar{F}_n(t_1)\\ \vdots & & \vdots\\ f_1(t_m) & \cdots & f_n(t_m)\\ \bar{F}_1(t_m) & \cdots & \bar{F}_n(t_m)\\ \vdots & & \vdots\\ \bar{F}_1(t_m) & \cdots & \bar{F}_n(t_m)\end{matrix}\right|}\begin{matrix} \\ \left.\begin{matrix}\\ \\ \\ \\ \end{matrix}\right\}R_1\text{ rows}\\ \\ \\ \left.\begin{matrix}\\ \\ \\ \\ \end{matrix}\right\}R_m\text{ rows}\end{matrix}\quad. \tag{6}$$

A similar expression turns out to be valid for the marginal joint density of the first p progressively Type-II censored order statistics. From the proof of Theorem 1, it is directly seen that the expression yields

$$\begin{aligned} f_{X^{\boldsymbol{R}}_{1:m:n},\dots,X^{\boldsymbol{R}}_{p:m:n}}(t_1,\dots,t_p) &= \frac{1}{(n-1)!}\left(\prod_{j=2}^{p}\gamma_j\right)\sum_{\pi\in S_n}\prod_{j=1}^{p-1}f_{\pi(j)}(t_j)\\ &\quad\times\left\{\prod_{r=p+\rho_{j-1}+1}^{p+\rho_j}\bar{F}_{\pi(r)}(t_j)\right\}\\ &\quad\times f_{\pi(p)}(t_p)\left\{\prod_{r=p+\rho_{p-1}+1}^{n}\bar{F}_{\pi(r)}(t_p)\right\},\\ &\quad t_1\le\cdots\le t_p, \end{aligned}\tag{7}$$

where $1\le p\le m$. Moreover, a similar expression in terms of a permanent as in (6) can be established in this case:

$$f_{X^{\boldsymbol{R}}_{1:m:n},\ldots,X^{\boldsymbol{R}}_{p:m:n}}(t_1,\ldots,t_p) = \frac{1}{(n-1)!}\prod_{j=2}^{p}\gamma_j \times$$

$$\begin{array}{|ccc|l} \multicolumn{3}{c}{+ \qquad\qquad\qquad +} & \\ f_1(t_1) & \cdots & f_n(t_1) & \multirow{4}{*}{$\left.\right\}R_1 \text{ rows}$} \\ \bar{F}_1(t_1) & \cdots & \bar{F}_n(t_1) & \\ \vdots & & \vdots & \\ \bar{F}_1(t_1) & \cdots & \bar{F}_n(t_1) & \\ \vdots & & \vdots & \\ f_1(t_{p-1}) & \cdots & f_n(t_{p-1}) & \\ \bar{F}_1(t_{p-1}) & \cdots & \bar{F}_n(t_{p-1}) & \quad R_{p-1} \text{ rows} \\ \vdots & & \vdots & \\ \bar{F}_1(t_{p-1}) & \cdots & \bar{F}_n(t_{p-1}) & \\ f_1(t_p) & \cdots & f_n(t_p) & \\ \bar{F}_1(t_p) & \cdots & \bar{F}_n(t_p) & \quad R_p^* \text{ rows} \\ \vdots & & \vdots & \\ \bar{F}_1(t_p) & \cdots & \bar{F}_n(t_p) & \end{array}, \tag{8}$$

where $R_p^* = n - p - R_1 - \cdots - R_{p-1}$ and $t_1 \le \cdots \le t_p$.

Eventhough the joint density of the first p progressively Type-II censored order statistics from heterogeneous distributions can be expressed as a permanent as in (8), it is an open problem whether such an expression holds for arbitrary marginal density functions. It has to be mentioned that, in general, the derivation of explicit expressions for the marginal distributions does not seem to be possible. However, considering distribution functions $F_1, \ldots, F_n$ generated by

$$F_r(t) = 1 - \{1 - F(t)\}^{\lambda_r} = 1 - \left\{\bar{F}(t)\right\}^{\lambda_r}, \quad t \in \mathbb{R},\ 1 \le r \le n,$$

with an absolutely continuous distribution function F, density function f, and parameters $\lambda_1, \ldots, \lambda_n > 0$, explicit expressions for marginal densities and distribution functions can be derived. In this case of proportional hazards, the corresponding density functions are given by

$$f_r(t) = \lambda_r f(t)\left\{\bar{F}(t)\right\}^{\lambda_r - 1}, \quad t \in \mathbb{R}, \qquad 1 \le r \le n.$$

Applying this particular structure into Eq. 2, we obtain

$$\begin{aligned} &f_{X^{\boldsymbol{R}}_{1:m:n},\ldots,X^{\boldsymbol{R}}_{m:m:n}}(t_1,\ldots,t_m) \\ &= \frac{1}{(n-1)!}\left(\prod_{j=2}^{m}\gamma_j\right)\sum_{\pi\in S_n}\prod_{j=1}^{m}\lambda_{\pi(j)}f(t_j) \\ &\qquad \times \left\{\bar{F}(t_j)\right\}^{\lambda_{\pi(j)}-1}\left\{\prod_{r=m+\rho_{j-1}+1}^{m+\rho_j}\left\{\bar{F}(t_j)\right\}^{\lambda_{\pi(r)}}\right\} \end{aligned}$$

$$= \frac{1}{(n-1)!} \left(\prod_{j=2}^{m} \gamma_j \right) \sum_{\pi \in S_n} \prod_{j=1}^{m} \lambda_{\pi(j)} f(t_j) \left\{ \bar{F}(t_j) \right\}^{m_{j,\pi}}, \quad t_1 \leq \cdots \leq t_m, \qquad (9)$$

where

$$m_{j,\pi} = \lambda_{\pi(j)} + \sum_{r=m+\rho_{j-1}+1}^{m+\rho_j} \lambda_{\pi(r)} - 1 > -1, \quad 1 \leq j \leq m, \quad \pi \in S_n.$$

In order to exploit the representation in (9), we introduce the sums $\gamma_{j,\pi} = \sum_{i=j}^{m}(m_{i,\pi}+1)$, $1 \leq j \leq m$, which are obviously decreasingly ordered, viz., $\gamma_{1,\pi} > \cdots > \gamma_{m,\pi} > 0$. Then,

$$f_{*,\pi}(t_1, \ldots, t_m) = \left(\prod_{j=1}^{m} \gamma_{j,\pi} \right) \prod_{j=1}^{m} f(t_j) \left\{ \bar{F}(t_j) \right\}^{m_{j,\pi}}, \qquad t_1 \leq \cdots \leq t_m,$$

denotes the joint density function of the generalized order statistics $X_{*,\pi}^{(1)}, \ldots, X_{*,\pi}^{(m)}$ based on the distribution function F and parameters $\gamma_{1,\pi}, \ldots, \gamma_{m,\pi}$; see, Kamps (1995). Hence, we can express the joint density in (9) as a mixture of the densities $f_{*,\pi}$, $\pi \in S_n$, as follows:

$$f_{X_{1:m:n}^{\boldsymbol{R}}, \ldots, X_{m:m:n}^{\boldsymbol{R}}}(t_1, \ldots, t_m) = \frac{1}{(n-1)!} \prod_{j=2}^{m} \gamma_j \sum_{\pi \in S_n} \left\{ \prod_{j=1}^{m} \frac{\lambda_{\pi(j)}}{\gamma_{j,\pi}} \right\} \times f_{*,\pi}(t_1, \ldots, t_m), \quad t_1 \leq \cdots \leq t_m. \qquad (10)$$

This representation is useful to derive marginal densities and moments of the progressively Type-II censored order statistics $X_{1:m:n}^{\boldsymbol{R}}, \ldots, X_{m:m:n}^{\boldsymbol{R}}$ in this special case of proportional hazards. For example, the marginal density of $X_{*,\pi}^{(p)}$, $1 \leq p \leq m$, is given by (see Kamps and Cramer, 2001)

$$f_{X_{*,\pi}^{(p)}}(t) = \left(\prod_{j=1}^{p} \gamma_{j,\pi} \right) f(t) \sum_{j=1}^{p} \frac{\left\{ \bar{F}(t) \right\}^{\gamma_{j,\pi}-1}}{\prod_{\substack{\nu=1 \\ \nu \neq j}}^{p} (\gamma_{\nu,\pi} - \gamma_{j,\pi})}, \quad t \in \mathbb{R}.$$

Hence, integration of (10) with respect to all the variables except t_p yields an expression for the marginal density of $X_{p:m:n}^{\boldsymbol{R}}$ as

$$f_{X_{p:m:n}^{\boldsymbol{R}}}(t) = \frac{1}{(n-1)!} \prod_{j=2}^{m} \gamma_j \sum_{\pi \in S_n} \left\{ \prod_{j=1}^{m} \frac{\lambda_{\pi(j)}}{\gamma_{j,\pi}} \right\} f_{X_{*,\pi}^{(p)}}(t), \quad t \in \mathbb{R}. \qquad (11)$$

Similarly, an expression for the distribution function of $X_{p:m:n}^{\boldsymbol{R}}$ can be established.

Equation 10 can also be readily applied to obtain the expectation of $h\left(X_{1:m:n}^{\boldsymbol{R}} \ldots, X_{m:m:n}^{\boldsymbol{R}}\right)$, provided it exists, where $h : \mathbb{R}^m \to \mathbb{R}$ is a real-valued function, as follows:

$$E\left[h\left(X_{1:m:n}^{\boldsymbol{R}},\ldots,X_{m:m:n}^{\boldsymbol{R}}\right)\right]=\frac{1}{(n-1)!}\prod_{j=2}^{m}\gamma_j\sum_{\pi\in S_n}\left\{\prod_{j=1}^{m}\frac{\lambda_{\pi(j)}}{\gamma_{j,\pi}}\right\}\times E\left[h\left(X_{*,\pi}^{(1)},\ldots,X_{*,\pi}^{(m)}\right)\right]. \quad (12)$$

From (10), upon integrating with respect to $t_1,\ldots,t_m$, we obtain

$$\sum_{\pi\in S_n}p_\pi=1, \qquad \text{where } p_\pi=\frac{1}{(n-1)!}\prod_{j=2}^{m}\gamma_j\left\{\prod_{j=1}^{m}\frac{\lambda_{\pi(j)}}{\gamma_{j,\pi}}\right\}, \qquad \pi\in S_n.$$

Hence, $(p_\pi)_{\pi\in S_n}$ defines a probability distribution on the set S_n. Suppose now Ξ is a random variable defined on S_n with distribution $(p_\pi)_{\pi\in S_n}$, we can use a result of Cramer and Kamps (2003) to obtain a stochastic representation for $X_{1:m:n}^{\boldsymbol{R}},\ldots,X_{m:m:n}^{\boldsymbol{R}}$ as follows.

Let $U_1,\ldots,U_m$ be independent Uniform(0,1) random variables which are independent of Ξ. Then, we have the stochastic representation

$$(X_{1:m:n}^{\boldsymbol{R}},\ldots,X_{m:m:n}^{\boldsymbol{R}})\sim\left(F^{-1}\left(1-U_1^{1/\gamma_{1,\Xi}}\right),\ldots,F^{-1}\left(1-\prod_{j=1}^{m}U_j^{1/\gamma_{j,\Xi}}\right)\right),$$

where $F^{-1}(\cdot)$ denotes the quantile function of F. This representation can be easily applied to obtain explicit expressions for moments. Also, utilizing the results of Cramer et al. (2002), upper bounds for moments can be obtained by the Moriguti method. For example, with μ and σ denoting the mean and standard deviation of the distribution function F, respectively, it can be shown that

$$E\left(\frac{X_{*,\pi}^{(p)}-\mu}{\sigma}\right)\le B_\pi,$$

where B_π is independent of F. A representation of the bound could be deduced from formulae presented in Balakrishnan et al. (2001), since the parameters $\gamma_{1,\pi},\ldots,\gamma_{m,\pi}$ are pairwise different. Using the mixture representation in (11), a bound for $E\left(X_{p:m:n}^{\boldsymbol{R}}\right)$ then results. It needs to be mentioned here, however, that this bound is not sharp.

Another application of the connection to generalized order statistics is as follows. A distribution function F with density function f is said to have the DFR-property if the hazard rate $f/(1-F)$ is a decreasing function. Kamps (1995, p. 177) presented an extension of a result of Gupta and Kirmani (1988) proving that the DFR-property of the underlying distribution function F holds for spacings of generalized order statistics. Since the DFR-property is preserved under mixing (see Barlow and Proschan, 1975, p. 103), we readily obtain the following result.

Theorem 2 *Let F have the DFR-property. Then, the spacings $X_{r:m:n}^{\boldsymbol{R}}-X_{r-1:m:n}^{\boldsymbol{R}}$, $2\le r\le m$, also have DFR-distributions.*

5 The exponential case

In the case of exponential distributions with

$$F_r(t) = 1 - \mathrm{e}^{-\lambda_r t}, \quad t \geq 0, \qquad 1 \leq r \leq n,$$

which clearly belongs to the proportional hazards case, Eq. 9 yields

$$\begin{aligned} f_{X^{\boldsymbol{R}}_{1:m:n},\ldots,X^{\boldsymbol{R}}_{m:m:n}}(t_1,\ldots,t_m) = \frac{1}{(n-1)!}\prod_{j=2}^{m}\gamma_j \sum_{\pi\in S_n}\prod_{j=1}^{m}\lambda_{\pi(j)} \\ \times \exp\left\{-\sum_{j=1}^{m}(m_{j,\pi}+1)t_j\right\}, \quad t_1 \leq \cdots \leq t_m. \end{aligned} \tag{13}$$

From this joint density function, we obtain the following theorem.

Theorem 3 *Let $X^{\boldsymbol{R}}_{1:m:n},\ldots,X^{\boldsymbol{R}}_{m:m:n}$ be the progressively Type-II censored order statistics arising from exponential distributions with scale parameters $\lambda_1,\ldots,\lambda_n > 0$. Then,*

$$X^{\boldsymbol{R}}_{1:m:n} \quad \text{and} \quad \left(X^{\boldsymbol{R}}_{2:m:n} - X^{\boldsymbol{R}}_{1:m:n},\ldots,X^{\boldsymbol{R}}_{m:m:n} - X^{\boldsymbol{R}}_{1:m:n}\right)$$

are stochastically independent. Moreover, $X^{\boldsymbol{R}}_{1:m:n}$ is exponentially distributed with scale parameter $\sum_{j=1}^{n}\lambda_j$. Consequently, the normalized spacing $D_1 = \gamma_1 X^{\boldsymbol{R}}_{1:m:n}$ is independent of the normalized spacings $D_2,\ldots,D_m$, where

$$D_r = \gamma_r\left(X^{\boldsymbol{R}}_{r:m:n} - X^{\boldsymbol{R}}_{r-1:m:n}\right), \quad r = 2,\ldots,m.$$

Proof From Eq. 13 and an use of the transformation formula, we have

$$\begin{aligned} & f_{X^{\boldsymbol{R}}_{1:m:n},X^{\boldsymbol{R}}_{2:m:n}-X^{\boldsymbol{R}}_{1:m:n},\ldots,X^{\boldsymbol{R}}_{m:m:n}-X^{\boldsymbol{R}}_{1:m:n}}(t_1,t_2,\ldots,t_m) \\ & \quad = f_{X^{\boldsymbol{R}}_{1:m:n},X^{\boldsymbol{R}}_{2:m:n},\ldots,X^{\boldsymbol{R}}_{m:m:n}}(t_1,t_2+t_1,\ldots,t_m+t_1), \end{aligned}$$

which readily yields, for $t_1,\ldots,t_m \geq 0$,

$$\begin{aligned} & f_{X^{\boldsymbol{R}}_{1:m:n},X^{\boldsymbol{R}}_{2:m:n}-X^{\boldsymbol{R}}_{1:m:n},\ldots,X^{\boldsymbol{R}}_{m:m:n}-X^{\boldsymbol{R}}_{1:m:n}}(t_1,t_2,\ldots,t_m) \\ & = \frac{1}{(n-1)!}\prod_{j=2}^{m}\gamma_j\sum_{\pi\in S_n}\prod_{j=1}^{m}\lambda_{\pi(j)}\exp\left\{-t_1\sum_{j=1}^{m}(m_{j,\pi}+1) - \sum_{j=2}^{m}(m_{j,\pi}+1)t_j\right\} \\ & = \exp\left\{-t_1\sum_{j=1}^{n}\lambda_j\right\}\frac{1}{(n-1)!}\prod_{j=2}^{m}\gamma_j\sum_{\pi\in S_n}\prod_{j=1}^{m}\lambda_{\pi(j)}\exp\left\{-\sum_{j=2}^{m}(m_{j,\pi}+1)t_j\right\}. \end{aligned}$$

Since the term following the summation $\sum_{\pi\in S_n}$, on the RHS of the above equation, is independent of t_1, the factorization

$$\begin{aligned} & f_{X^{\boldsymbol{R}}_{1:m:n}, X^{\boldsymbol{R}}_{2:m:n}-X^{\boldsymbol{R}}_{1:m:n},\ldots,X^{\boldsymbol{R}}_{m:m:n}-X^{\boldsymbol{R}}_{1:m:n}}(t_1,t_2,\ldots,t_m) \\ &\quad = f_{X^{\boldsymbol{R}}_{1:m:n}}(t_1)\; f_{X^{\boldsymbol{R}}_{2:m:n}-X^{\boldsymbol{R}}_{1:m:n},\ldots,X_{m:m:n}-X^{\boldsymbol{R}}_{1:m:n}}(t_2,\ldots,t_m) \end{aligned}$$

results, which establishes the main assertion of the theorem. □

The result in Theorem 3 can be used to obtain an expression for the correlation between the smallest and other progressively Type-II censored order statistics. Using the fact that $\operatorname{Cov}\left(X^{\boldsymbol{R}}_{1:m:n}, X^{\boldsymbol{R}}_{p:m:n}-X^{\boldsymbol{R}}_{1:m:n}\right)=0$ for $p=2,\ldots,m$, we get $\operatorname{Cov}\left(X^{\boldsymbol{R}}_{1:m:n}, X^{\boldsymbol{R}}_{p:m:n}\right)=\operatorname{Var}\left(X^{\boldsymbol{R}}_{1:m:n}\right)$, and consequently

$$\operatorname{Corr}\left(X^{\boldsymbol{R}}_{1:m:n}, X^{\boldsymbol{R}}_{p:m:n}\right)=\sqrt{\frac{\operatorname{Var}\left(X^{\boldsymbol{R}}_{1:m:n}\right)}{\operatorname{Var}\left(X^{\boldsymbol{R}}_{p:m:n}\right)}},\quad 2\le p\le m.$$

This generalizes a result of Gross et al. (1986) for the usual order statistics wherein the smallest and largest order statistics are considered; see also Joshi (1988).

Moreover, from Eq. 12, we can obtain the single and product moments of the progressively Type-II censored order statistics. For example, the expectation of the p-th generalized order statistic $X^{(p)}_{*,\pi}$ based on an exponential distribution and parameters $\gamma_{1,\pi},\ldots,\gamma_{m,\pi}$ is given by

$$E\left(X^{(p)}_{*,\pi}\right)=\sum_{i=1}^{p}\frac{1}{\gamma_{i,\pi}},\quad 1\le p\le m;$$

using this expression in (12), we immediately obtain

$$E\left(X^{\boldsymbol{R}}_{p:m:n}\right)=\frac{1}{(n-1)!}\left(\prod_{j=2}^{m}\gamma_j\right)\sum_{\pi\in S_n}\left(\prod_{j=1}^{m}\frac{\lambda_{\pi(j)}}{\gamma_{j,\pi}}\right)\sum_{i=1}^{p}\frac{1}{\gamma_{i,\pi}},\quad 1\le p\le m.$$

Similar expressions for the variances and covariances of progressively Type-II censored order statistics can be obtained from Eq. 12. In addition, expressions for the characteristic function and the moment generating function, similar to those for the usual order statistics given by Bapat and Beg (1989b), can also be derived.

Kochar and Korwar (1996) proved that, in the case of usual order statistics, the density of a single spacing is a mixture of independent exponentially distributed random variables. A similar result holds under the general progressive censoring scenario as well. To this end, let $\pi\in S_n$ be a fixed permutation. Then, as mentioned above, $f_{*,\pi}$ is the joint density of generalized order statistics based on the standard exponential distribution and parameters $\gamma_{1,\pi},\ldots,\gamma_{m,\pi}$. It is well-known that the spacings (not necessarily normalized) $D^{\pi}_1,\ldots,D^{\pi}_m$ of generalized order statistics $X^{(1)}_{*,\pi},\ldots,X^{(m)}_{*,\pi}$ form a sequence of independent exponential random variables; see Kamps (1995, p. 81). Hence, we can write the corresponding joint density as

$$f_{D^{\pi}_1,\ldots,D^{\pi}_m}(d_1,\ldots,d_m)=\prod_{j=1}^{m}\left(\gamma_{j,\pi}e^{-\gamma_{j,\pi}d_j}\right),\quad d_1,\ldots,d_m\ge 0.$$

Taking into account the normalization factors $\gamma_1, \ldots, \gamma_m$, we obtain from Eq. 13 that

$$f_{D_1,\ldots,D_m}(d_1,\ldots,d_m) = \frac{1}{n!}\sum_{\pi\in S_n}\left(\prod_{j=1}^{n}\frac{\lambda_{\pi(j)}}{\gamma_{j,\pi}}\right)\prod_{j=1}^{m}\left(\gamma_{j,\pi}\mathrm{e}^{-\gamma_{j,\pi}d_j/\gamma_j}\right), \\ d_1,\ldots,d_m \geq 0. \tag{14}$$

Integrating out all the variables other than d_p in (14), we obtain the marginal density function of D_p as

$$f_{D_p}(d_p) = \frac{1}{n!}\left(\prod_{\substack{j=1\\ j\neq p}}^{m}\gamma_j\right)\sum_{\pi\in S_n}\left(\prod_{j=1}^{m}\frac{\lambda_{\pi(j)}}{\gamma_{j,\pi}}\right)\left(\gamma_{p,\pi}\mathrm{e}^{-\gamma_{p,\pi}d_p/\gamma_p}\right), \quad d_p \geq 0,$$

revealing clearly that the density f_{D_p} is indeed a mixture of exponential density functions. Moreover, expressions for joint density functions $f_{D_{i_1},\ldots,D_{i_p}}$, $1 \leq i_1 < \cdots < i_p \leq m$, similar to those for the usual order statistics given by Kochar and Korwar (1996), can be derived in an analogous manner. This mixture property leads to the following result which extends Theorem 2.2 of Kochar and Korwar (1996). The proof follows directly from the preceding mixture representation since an exponential distribution has a log-convex density and that this property is preserved under mixing; see, for example, Marshall and Olkin (1979, p. 452). In addition, this implies directly the DFR-property of the spacings (see Theorem 2).

Theorem 4 *The spacings of the progressively Type-II censored order statistics from heterogeneous exponential distributions have log-convex densities. Further, the normalized spacings $D_1, \ldots, D_m$ have the DFR-property.*

6 Progressive censoring scheme for robust estimation for exponential distribution

In this section, we will focus on a single-outlier model from an exponential distribution and discuss the progressive censoring scheme that will facilitate the robust estimation of the exponential scale parameter. To fix the ideas, let us assume that $X_i \sim \mathrm{Exp}(\lambda)$, $i \in \{1,\ldots,n\}\backslash\{\ell\}$ and $X_\ell \sim \mathrm{Exp}(\mu)$ with $\mu \leq \lambda$; that is, X_ℓ corresponds to the outlier in the sample. Let A_j denote the event that the outlier (viz., X_ℓ) is the j-th failure time (viz., $X^{\boldsymbol{R}}_{j:m:n}$) when the progressive censoring scheme employed in the life-testing experiment is $\boldsymbol{R} = (R_1, \ldots, R_m)$.

First, from Eq. 4, we find for $j = 1, \ldots, m$,

$$P(A_j) = P\left(X_\ell = X^{\boldsymbol{R}}_{j:m:n}\right) \\ = \frac{1}{(n-1)!}\prod_{i=2}^{m}\gamma_i \sum_{\pi\in S_n,\pi(j)=\ell} P\left(X_{\pi(r)} \leq \min\left\{X_{\pi(r+1)},\ldots,X_{\pi(m)},\right.\right. \\ \left.\left. Z_r^{\pi},\ldots,Z_m^{\pi}\right\}, 1 \leq r \leq m\right)$$

$$= \frac{1}{(n-1)!} \prod_{i=2}^{m} \gamma_i \sum_{\pi \in S_n, \pi(j)=\ell} \int_0^\infty \int_{t_1}^\infty \cdots \int_{t_{m-1}}^\infty h_{\pi(I)}(t_1, \ldots, t_m)\, dt_m \cdots dt_1, \tag{15}$$

where

$$\begin{aligned} h_{\pi(I)}(t_1, \ldots, t_m) &= \prod_{i=1}^{m} f_{\pi(I)}(t_i) \left\{ \prod_{r=m+\rho_{i-1}+1}^{m+\rho_i} \bar{F}_{\pi(r)}(t_i) \right\} \\ &= \lambda^{m-1}\mu \, \exp\left\{ -\lambda \sum_{i=1}^{m} (R_i + 1) t_i - (\mu - \lambda) t_j \right\}, \\ &\qquad 0 < t_1 \le \cdots \le t_m. \end{aligned} \tag{16}$$

Since the expression in Eq. 16 does not depend on the permutation π, we obtain from Eq. 15 that

$$P(A_j) = \prod_{i=2}^{m} \gamma_i \int_0^\infty \int_{t_1}^\infty \cdots \int_{t_{m-1}}^\infty \lambda^{m-1}\mu \; g(t_1, \ldots, t_m)\, dt_m \ldots dt_1, \tag{17}$$

where

$$g(t_1, \ldots, t_m) = \exp\left\{ -\lambda \sum_{i=1}^{m} (R_i + 1) t_i - (\mu - \lambda) t_j \right\}. \tag{18}$$

Upon introducing the notation $\eta_i = \lambda(R_i + 1)$, $i \neq j$, and $\eta_j = \lambda(R_j + 1) + \mu - \lambda$, the function $g(t_1, \ldots, t_m)$ in (18) can be expressed as

$$\begin{aligned} g(t_1, \ldots, t_m) &= \exp\left\{ -\sum_{i=1}^{m} \eta_i \; t_i \right\} \\ &= \left(\prod_{i=1}^{m} \tilde{\gamma}_i \right)^{-1} \prod_{i=1}^{m} \tilde{\gamma}_i \prod_{i=1}^{m} \mathrm{e}^{-t_i} \left(\mathrm{e}^{-t_i} \right)^{\eta_i - 1} \\ &= \left(\prod_{i=1}^{m} \tilde{\gamma}_i \right)^{-1} u(t_1, \ldots, t_m), \quad t_1 \le \cdots \le t_m, \end{aligned}$$

where $\tilde{\gamma}_i = \sum_{\nu=i}^{m} \eta_\nu$, $i = 1, \ldots, m$, and u is the joint density of the generalized order statistics from the standard exponential distribution and parameters $\tilde{\gamma}_1, \ldots, \tilde{\gamma}_m$. Using these facts and carrying out the integration in (17), we obtain

$$P(A_j) = \frac{\mu \lambda^{m-1} \prod_{i=2}^{m} \gamma_i}{\prod_{i=1}^{m} \tilde{\gamma}_i}, \tag{19}$$

where

$$\tilde{\gamma}_i = \begin{cases} \lambda \gamma_i, & i > j \\ \lambda(\gamma_i - 1) + \mu, & i \le j. \end{cases}$$

Upon substituting the above expression of $\tilde{\gamma}_i$ and that $\gamma_1 = n$ in Eq. 19, we obtain an explicit expression for $P(A_j)$ as

$$P(A_j) = \frac{\mu\lambda^{m-1}\prod_{i=2}^{m}\gamma_i}{\prod_{i=1}^{j}\{\mu + \lambda(\gamma_i - 1)\}\prod_{i=j+1}^{m}(\lambda\gamma_i)}$$
$$= \frac{(\mu/\lambda)}{n - 1 + (\mu/\lambda)}\prod_{i=2}^{j}\frac{\gamma_i}{\gamma_i - 1 + (\mu/\lambda)}. \tag{20}$$

Remark 3 If we set $\mu = \lambda$ in Eq. 20, we find $P(A_j) = \frac{1}{n}$, which is to be expected since this corresponds to the iid case.

Remark 4 If we consider the conventional Type-II right censoring, we have $\gamma_i = n - i + 1$ and in this case Eq. 20 reduces to

$$P(A_j) = \frac{(\mu/\lambda)}{n - 1 + (\mu/\lambda)}\prod_{i=2}^{j}\frac{n - i + 1}{n - i + (\mu/\lambda)} \tag{21}$$

which shows that $P(A_j)$'s are increasing w.r.t. j, as expected.

Remark 5 With regard to robust censoring scheme, we conclude that the probability $P(A_m)$ (viz., the probability that the largest observation $X_{m:m:n}^{\boldsymbol{R}}$ is the outlier) is *minimal* if the progressive censoring scheme is given by $R_1 = \cdots = R_{m-1} = 0, R_m = n - m$. It is *maximal* for the progressive censoring scheme $R_1 = n - m, R_2 = \cdots = R_m = 0$. This simply means that the most robust progressive censoring scheme is the conventional Type-II right censoring scheme while the least robust progressive censoring scheme is when all the censoring occurs right after the first failure. In order to observe these results, we first note that the function $k(x) = \frac{1}{1-(\alpha/x)}$, for $\alpha > 0$, is decreasing in $(0, \infty)$. Now, since $\mu \le \lambda$ by assumption, upon writing

$$P(A_m) = \frac{(\mu/\lambda)}{n - 1 + (\mu/\lambda)}\prod_{i=2}^{m}\frac{1}{1 - \frac{1-(\mu/\lambda)}{\gamma_i}},$$

we readily observe that each term in the product is minimal if γ_i is maximal and vice versa.

Remark 6 If $P(A_j)$ is to be minimized or maximized, the same argument as the one above applies to the first j progressive censoring numbers $R_1, \ldots, R_j$.

Remark 7 If we wish to minimize the probability that the outlier is one of the observations (viz., $\sum_{j=1}^{m} P(A_j)$), then we will obtain the same optimal progressive censoring schemes as mentioned above in Remark 5.

In particular, we have

$$\sum_{j=1}^{m} P(A_j) = \frac{(\mu/\lambda)}{n - 1 + (\mu/\lambda)}\left(1 + \sum_{j=2}^{m}\prod_{i=2}^{j}\frac{\gamma_i}{\gamma_i - 1 + (\mu/\lambda)}\right).$$

Writing $\alpha_i = \frac{\gamma_i}{\gamma_i - 1 + (\mu/\lambda)}$, we obtain the representation

$$\sum_{j=1}^{m} P(A_j) = \frac{(\mu/\lambda)}{n - 1 + (\mu/\lambda)} \left(1 + \alpha_1(1 + \alpha_2(1 + \cdots))\right).$$

This proves directly that $\sum_{j=1}^{m} P(A_j)$ is minimal (maximal) if each α_i is minimal (maximal). Hence, this criterion which is more natural for judging robustness leads to the same extremal schemes as $P(A_m)$. Thus, it is natural to consider the following robustness measures. Given n and m, a censoring scheme $\tilde{R} = (R_1, \ldots, R_m)$ is said to be more robust than a censoring scheme $\tilde{S} = (S_1, \ldots, S_m)$ if

Criterion $\mathcal{A}$: $P_{\tilde{R}}(A_m) \le P_{\tilde{S}}(A_m)$

Criterion $\mathcal{B}$: $\sum_{j=1}^{m} P_{\tilde{R}}(A_j) \le \sum_{j=1}^{m} P_{\tilde{S}}(A_j)$.

Although both criteria yield the same extremal schemes, they could lead to different decisions. For instance, consider $m = 3$, $n = 10$, $\mu = \lambda/2$, and $(R_1, R_2, R_3) = (2, 1, 4)$, $(S_1, S_2, S_3) = (0, 4, 3)$. Then, we obtain the following probabilities (rounded to five decimals)

Criterion/scheme	(2, 1, 4)	(0, 4, 3)
$\mathcal{A}$	0.06298	0.06369
$\mathcal{B}$	0.17229	0.17205

Hence, the scheme (2, 1, 4) is more robust than (0, 4, 3) according to $\mathcal{A}$ whereas $\mathcal{B}$ yields the other decision. The complete results for $\mathcal{A}$ and $\mathcal{B}$ are presented in Table 1.

Remark 8 Since

$$\frac{P(A_{j+1})}{P(A_j)} = \frac{\gamma_{j+1}}{\gamma_{j+1} - 1 + (\mu/\lambda)} \ge 1 \Longleftrightarrow \mu \le \lambda, \qquad j \in \{1, \ldots, m\},$$

we observe that $P(A_j)$ is increasing in j (see also Remark 4); hence, it is more likely that a larger observation is generated by $\mathrm{Exp}(\mu)$, as expected.

Remark 9 Since the probability to observe an outlier is minimal for the scheme $(R_1, \ldots, R_m) = (0, \ldots, 0, n - m)$, it is favorable to use right censoring for robust estimation of the parameter λ. Hence, if progressive censoring is carried out by design, this scheme should be used in order to minimize the probability that the outlier is contained in the sample. If one could choose between some censoring schemes, criteria $\mathcal{A}$ and $\mathcal{B}$ may be taken into consideration. However, the distributions of the presented linear estimators are complicated functions of the censoring scheme so that the impact of the censoring scheme is not obvious. Thus, our proposal is a first attempt to tackle this problem: it is favorable to use that linear estimator which excludes the outlier with highest probability. A detailed account to this problem will be subject of future research.

Table 1 Results for robustness criteria $\mathcal{A}$ and $\mathcal{B}$ for $n = 10$, $m = 3$, and $\mu = \lambda/2$

Censoring scheme	Criterion $\mathcal{A}$	Criterion $\mathcal{B}$
(0, 0, 7)	0.05944	0.16780
(0, 1, 6)	0.06001	0.16837
(0, 2, 5)	0.06079	0.16915
(0, 3, 4)	0.06192	0.17028
(0, 4, 3)	0.06369	0.17205
(0, 5, 2)	0.06687	0.17523
(0, 6, 1)	0.07430	0.18266
(0, 7, 0)	0.11146	0.21981
(1, 0, 6)	0.06046	0.16923
(1, 1, 5)	0.06124	0.17002
(1, 2, 4)	0.06238	0.17115
(1, 3, 3)	0.06416	0.17293
(1, 4, 2)	0.06737	0.17614
(1, 5, 1)	0.07485	0.18363
(1, 6, 0)	0.11228	0.22105
(2, 0, 5)	0.06183	0.17114
(2, 1, 4)	0.06298	0.17229
(2, 2, 3)	0.06478	0.17409
(2, 3, 2)	0.06802	0.17733
(2, 4, 1)	0.07557	0.18489
(2, 5, 0)	0.11336	0.22267
(3, 0, 4)	0.06380	0.17384
(3, 1, 3)	0.06562	0.17567
(3, 2, 2)	0.06890	0.17895
(3, 3, 1)	0.07656	0.18660
(3, 4, 0)	0.11483	0.22488
(4, 0, 3)	0.06683	0.17794
(4, 1, 2)	0.07018	0.18129
(4, 2, 1)	0.07797	0.18908
(4, 3, 0)	0.11696	0.22807
(5, 0, 2)	0.07218	0.18496
(5, 1, 1)	0.08020	0.19298
(5, 2, 0)	0.12030	0.23308
(6, 0, 1)	0.08421	0.20000
(6, 1, 0)	0.12632	0.24211
(7, 0, 0)	0.14035	0.26316

The robustness aspect supplements some results obtained in experimental design for progressive censoring. For instance, Balakrishnan and Aggarwala (2000) showed that in the exponential case any censoring scheme leads to the same variance of the BLUE for $1/\lambda$. Thus, the choice of the scheme is immaterial in terms of precision of the estimate. Using the preceding robustness considerations, we can pick one scheme leading to both least variance and robustness. In case of other distributions than exponential, it is not clear in which way this additional criterion influences the resulting estimator. The impact of both criteria on the optimal choice of censoring schemes will be considered in future research. More information on experimental design in progressive censoring is provided by Burkschat et al. (2006).

Acknowledgment This research was initiated while Dr. Erhard Cramer visited McMaster University, Canada, and he is grateful to the Deutsche Forschungsgemeinschaft (DFG) for supporting this visit.

References

Arnold, B. C., & Balakrishnan, N. (1989). *Relations, bounds and approximations for order statistics*. Lecture Notes in Statistics (Vol. 53). Berlin Heidelberg New York: Springer

Arnold, B. C., & Balakrishnan, N., & Nagaraja, H.N. (1992). *A first course in order statistics*. New York: Wiley

Balakrishnan, N. (1988). Recurrence relations for order statistics from n independent and non-identically distributed random variables. *Annals of the Institute of Statistical Mathematics*, *40*, 273–277

Balakrishnan, N. (1989). Recurrence relations among moments of order statistics from two related sets of independent and non-identically distributed random variables. *Annals of the Institute of Statistical Mathematics*, *41*, 323–329

Balakrishnan, N. (1994). Order statistics from non-identical exponential random variables and some applications (with discussion). *Computational Statistics & Data Analysis*, *18*, 203–253

Balakrishnan, N., Aggarwala, R. (2000). *Progressive censoring: theory, methods, and applications*. Boston: Birkhäuser

Balakrishnan, N., Cramer, E., Kamps, U. (2001). Bounds for means and variances of progressive Type II censored order statistics. *Statistics & Probability Letters*, *54*, 301–315

Bapat, R. (1990). Permanents in probability and statistics. *Linear Algebra and its Applications*, *127*, 3–25

Bapat, R., & Beg, M. I. (1989a). Identities and recurrence relations for order statistics corresponding to nonidentically distributed variables. *Communications in Statistics – Theory and Methods*, *18*, 1993–2004

Bapat, R., & Beg, M. I. (1989b). Order statistics for nonidentically distributed variables and permanents. *Sankhyā, Series A*, *51*, 79–93

Barlow, R. E., & Proschan, F. (1975). *Statistical theory of reliability and life testing: probability models*. New York: Holt-Rinehart and Winston

Barnett, V., & Lewis, T. (1993). *Outliers in statistical data* (3rd ed.). Chichester: Wiley

Burkschat, M., Cramer, E., & Kamps, U. (2006). On optimal schemes in progressive censoring. *Statistics & Probability Letters*, *76*, 1032–1036

Cohen, A. C. (1963). Progressively censored samples in life testing. *Technometrics*, *5*, 327–339

Cramer, E., & Kamps, U. (2003). Marginal distributions of sequential and generalized order statistics. *Metrika*, *58*, 293-310

Cramer, E., Kamps, U., & Rychlik, T. (2002). Evaluations of expected generalized order statistics in various scale units. *Applied Mathematics*, *29*, 285–295

David, H. A., Nagaraja, H. N. (2003). *Order statistics* (3rd ed.). Hoboken: Wiley

Gross, A. J., Hunt, H. H., & Odeh, R. E. (1986). The correlation coefficient between the smallest and largest observations when $(n-1)$ of the n observations are i.i.d. exponentially distributed. *Communications in Statistics – Theory and Methods*, *15*, 1113–1123

Gupta, R. C., Kirmani, S. N. U. A. (1988). Closure and monotonicity properties of nonhomogeneous Poisson processes and record values. *Probability in Engineering and Information Sciences*, *2*, 475–484

Harter, H. L., & Balakrishnan, N. (1998). Order statistics: a historical perspective. In: N. Balakrishnan and C.R. Rao (Eds.) Handbook of Statistics (Vol. 16) (pp. 25–64) Amsterdam: Elsevier

Joshi, P. C. (1972). Efficient estimation of the mean of an exponential distribution when an outlier is present. *Technometrics*, *14*, 137–143

Joshi, P. C. (1988). Estimation and testing under exchangeable exponential model with a single outlier. *Communications in Statistics – Theory and Methods*, *7*, 2315–2326

Kale, B. K., & Sinha, S. K. (1971). Estimation of expected life in the presence of an outlier observation. *Technometrics*, *13*, 755–759

Kamps, U. (1995). *A concept of generalized order statistics*. Stuttgart: Teubner

Kamps, U., & Cramer, E. (2001). On distributions of generalized order statistics. *Statistics*, *35*, 269–280

Kochar, S. C., Kirmani, S. N. U. A. (1995). Some results on normalized spacings from restricted families of distributions. *Journal of Statistical Planning and Inference*, *46*, 47–57

Kochar, S. C., Korwar, R. (1996). Stochastic orders for spacings of heterogeneous exponential random variables. *Journal of Multivariate Analysis*, *57*, 69–83

Kochar, S. C., Rojo, J. (1996). Some new results on stochastic comparisons of spacings from heterogeneous exponential distributions. *Journal of Multivariate Analysis*, *59*, 272–281

Marshall, A. W., & Olkin, I. (1979). *Inequalities: theory of majorization and its applications*. New York: Academic

Minc, H. (1978). *Permanents*. Reading: Addison-Wesley

Pledger, G., Proschan, F. (1971). Comparisons of order statistics and of spacings from heterogeneous distributions. In J. S. Rustagi (Ed.) *Optimization methods in statistics* , (pp. 89–113) New York: Academic

Sen, P. K. (1970). A note on order statistics for heterogeneous distributions. *Annals of Mathematical Statistics*, *41*, 2137–2139

Vaughan, R. J., Venables, W. N. (1972). Permanent expressions for order statistic densities. *Journal of the Royal Statistical Society, Series B*, *34*, 308–310