Metrika (2015) 78:205–218
DOI 10.1007/s00184-014-0496-6

Linearity of regression for overlapping order statistics

Adam Dołęgowski · Jacek Wesołowski

Received: 14 October 2013 / Published online: 10 May 2014

Abstract We consider a problem of characterization of continuous distributions for which linearity of regression of overlapping order statistics, $\mathbb{E}(X_{i:m}|X_{j:n}) = aX_{j:n} + b, m \le n$, holds. Due to a new representation of conditional expectation $\mathbb{E}(X_{i:m}|X_{j:n})$ in terms of conditional expectations $\mathbb{E}(X_{l:n}|X_{j:n}), l = i, \dots, n-m+i$, we are able to use the already known approach based on the Rao-Shanbhag version of the Cauchy integrated functional equation. However this is possible only if $j \le i$ or $j \ge n-m+i$. In the remaining cases the problem essentially is still open.

Keywords Order statistics · Overlapping samples · Linearity of regression · Characterization of probability distributions · Gamma distribution · Power distribution · Pareto distribution

1 Introduction

Consider a sequence $(X_k)_{k\ge1}$ of independent identically distributed continuous random variables. For an arbitrary $n \ge 1$ denote order statistics for the sample of size n by $X_{1:n} \le X_{2:n} \le \dots \le X_{n:n}$. In this paper we are interested in linearity of regression of overlapping order statistics, that is, we consider the condition

$$\mathbb{E}(X_{i:m}|X_{j:n}) = aX_{j:n} + b, \qquad 1 \le i \le m, \quad 1 \le j \le n, \tag{1}$$

A. Dołęgowski · J. Wesołowski (✉)
Wydział Matematyki i Nauk Informacyjnych, Politechnika Warszawska,
ul. Koszykowa 75, 00-662 Warsaw, Poland
e-mail: j.wesolowski@mini.pw.edu.pl; wesolo@mini.pw.edu.pl

A. Dołęgowski
e-mail: dolegowskia@student.mini.pw.edu.pl

where a, b are some real constants, and we want to describe the family of parent distribution for which (1) holds.

The problem has a long history. It goes back to Fisz (1958) who considered the case $m = n = i = 2,\ j = 1,\ a = 1$ and characterized the exponential distribution. This setting was extended in Rogers (1963) with characterization of the exponential distribution by (1) with $m = n,\ i = j+1,\ a = 1$. The case of adjacent order statistics was completed in Ferguson (1967) who considered the case $m = n,\ i = j+1$ with no restriction on a and characterized three families of distributions: exponential for $a = 1$, Pareto for $a > 1$ and power for $0 < a < 1$. Similar result was obtained in the PhD thesis of Pudeg (1991) and independently in Ahsanullah and Wesołowski (1997) for (1) with $m = n$ and $i = j+2$. Other trials in the non-adjacent case where given in Dembińska and Wesołowski ((1997)) and López-Blázquez and Moreno-Rebollo (1997). Finally the problem for $m = n$ was completely solved in Dembińska and Wesołowski (1998), denoted in the sequel by DW, where the same triplet of exponential, Pareto and power distributions or their symmetric (about zero) versions were characterized by (1) with arbitrary $j < i$ or $j > i$, respectively. Various recent extensions and complements of this result can be found e.g. in Ahsanullah and Hamedani (2012), Ahsanullah et al. (2012), Beg et al. (2013), Bieniek and Szynal (2003), Cramer et al. (2004), Ferguson (2002) or Gupta and Ahsanullah (2004).

All the previously mentioned papers were concerned with the case of one sample, i.e. $m = n$. We were able to trace in the literature only two papers dealing with the case $m \neq n$. In Ahsanullah and Nevzerov (1999) the authors claim that (1) with $i = j = 1$ and $n > m$ characterizes the triplet of exponential, Pareto and power distributions as above. In Wesołowski and Gupta (2001) only a very special case $i = m = 1$ was considered—see Sect. 5 below for more details.

In the present paper we will give the characterization of both the triplet families (exponential, Pareto, power or their symmetric versions) in the case $m \le n$ and $j \le i$ or $j \ge n + m - i$. Note that it does not cover the case considered in Wesołowski and Gupta (2001) but it covers the result announced in Ahsanullah and Nevzerov (1999). It appears that in the case considered, to prove the characterization one can apply Rao-Shanbhag version of integrated Cauchy functional equation (see Rao and Shanbhag 1994), similarly as in DW. This is done in Sect. 4. However, to reduce the problem to one to which this method can be applied we need to prove a representation of the conditional expectation $\mathbb{E}(X_{i:m}|X_{j:n})$ through conditional expectations from a single sample of size n. This is done, even in a more general setting, that is with no restrictions on relations between i and j, in Sect. 2. In Sect. 3 we observe that suitable form of linearity of regression (1) for $m \le n$ holds for both considered triplets of distributions. In Sect. 5 we make some comments regarding the case $i < j < n - m + i$ which still remains unsolved.

2 A representation of conditional expectation for overlapping order statistics

In this section we are interested in the conditional moment $\mathbb{E}(X_{i:m}|X_{j:n})$ for different values of $i, j \in \mathbb{N}$, $m < n \in \mathbb{N}$. We will express it as a convex combination of conditional moments of the form $\mathbb{E}(X_{l:n}|X_{j:n})$, $l = i, i+1, \ldots, n-m+i$.

Theorem 1 *Let $X_1, \ldots, X_n$ be a sequence of continuous, independent, identically distributed and integrable random variables. Then for any $m < n \in \mathbb{N}$, $1 \le i \le m$, $1 \le j \le n$*

$$\mathbb{E}(X_{i:m}|X_{j:n}) = \sum_{l=i}^{n-m+i} \frac{\binom{l-1}{i-1}\binom{n-l}{m-i}}{\binom{n}{m}} \mathbb{E}(X_{l:n}|X_{j:n}). \tag{2}$$

Proof Let us denote the set of all subsets of size m of $\{1, \ldots, n\}$ by $\mathbb{C}_m^n$. Of course, $\# \mathbb{C}_m^n = \binom{n}{m}$. We can number the elements of $\mathbb{C}_m^n$ arbitrarily and define $C(k)$ as the k-th element of $\mathbb{C}_m^n$, where $1 \le k \le \binom{n}{m}$. Denote by $X_{i:m}^{(k)}$ the i-th order statistic from $(X_i,\ i \in C(k))$. Due to the fact that the joint distribution of $(X_1, \ldots, X_n)$ is invariant under permutations, we can write:

$$\mathbb{E}(X_{i:m}|X_{j:n}) = \mathbb{E}(X_{i:m}^{(k)}|X_{j:n}), \qquad k = 1, \ldots, \binom{n}{m}.$$

Consequently, denoting $S_i = X_{i:m}^{(1)} + X_{i:m}^{(2)} + \cdots + X_{i:m}^{\left(\binom{n}{m}\right)}$,we have

$$\mathbb{E}(X_{i:m}|X_{j:n}) = \frac{\mathbb{E}(S_i|X_{j:n})}{\binom{n}{m}}. \tag{3}$$

Let us consider the event $A = \{X_1 < X_2 < \cdots < X_n\}$ and an arbitrary $l \in \{1, \ldots, n\}$. Obviously, on the event A we have $X_l = X_{l:n}$. Note that if $l \in \{1, \ldots, i-1\} \cup \{n-m+i+1, \ldots, n\}$ then on A the variable X_l cannot appear in the sum S_i. Otherwise, on A the variable X_l appears in the sum S_i as many times as there are m-elementary combinations of elements of $\{1, \ldots, n\}$ which consist of: l, exactly $(i-1)$ numbers smaller than l and exactly $(m-i)$ numbers greater than l. That is,

$$S_i\, I_A = \sum_{l=i}^{n-m+i} \frac{\binom{l-1}{i-1}\binom{n-l}{m-i}}{\binom{n}{m}} X_{l:n}\, I_A.$$

By (3) we get

$$\mathbb{E}(X_{i:m}\, I_A|X_{j:n}) = \sum_{l=i}^{n-m+i} \frac{\binom{l-1}{i-1}\binom{n-l}{m-i}}{\binom{n}{m}} \mathbb{E}(X_{l:n}\, I_A|X_{j:n}). \tag{4}$$

Let $\mathfrak{S}_n$ denote the set of permutations of $\{1, \ldots, n\}$. We may repeat the same reasoning for any event $A_\sigma = (X_{\sigma(1)} < \cdots < X_{\sigma(n)})$, where $\sigma \in \mathfrak{S}_n$. Consequently, (4) holds with A changed into A_σ for any $\sigma \in \mathfrak{S}_n$. Since the sets $A_\sigma, \sigma \in \mathfrak{S}_n$, are disjoint, we get

$$\begin{aligned}\mathbb{E}(X_{i:m}|X_{j:n}) &= \sum_{\sigma\in\mathfrak{S}_n} \mathbb{E}(X_{i:m}\, I_{A_\sigma}|X_{j:n}) \\ &= \sum_{l=i}^{n-m+i} \frac{\binom{l-1}{i-1}\binom{n-l}{m-i}}{\binom{n}{m}} \mathbb{E}(X_{l:n} \sum_{\sigma\in\mathfrak{S}_n} I_{A_\sigma}|X_{j:n}). \end{aligned} \tag{5}$$

Now (2) follows due to the identity $\sum_{\sigma\in\mathfrak{S}_n} I_{A_\sigma} = 1$ holding $\mathbb{P}$-a.s. □

Remark 1 Note that the coefficients which appear at the right hand side of (2) have a clear probabilistic interpretation. Namely, for any $1 \le i \le m \le n$

$$\mathbb{P}(X_{i:m} = X_{l:n}) = \begin{cases} \frac{\binom{l-1}{i-1}\binom{n-l}{m-i}}{\binom{n}{m}} & for\ l \in \{i, \ldots, n-m+i\}, \\ 0 & for\ l \in \{1, \ldots, i-1\} \cup \{n-m+i+1, \ldots, n\}. \end{cases} \tag{6}$$

Thus $\sum_{l=i}^{n-m+i} \mathbb{P}(X_{l:n} = X_{i:m}) = 1$.

To see that Remark 1 holds true, note that the event $\{X_{i:m} = X_{l:n}\}$ consists only of special permutations of $X_1, \ldots, X_n$: The variables $X_1, \ldots, X_m$ have to appear only at: position l, $i-1$ positions chosen from $\{1, \ldots, l-1\}$ (on $\binom{l-1}{i-1}$ ways) and $m-i$ positions chosen from $\{l+1, \ldots, n\}$ (on $\binom{n-l}{m-i}$ ways). Now it remains to permute the variables $X_1, \ldots, X_m$ at already fixed m positions (on $m!$ ways) and to permute the variables $X_{m+1}, \ldots, X_n$ at the remaining $n-m$ positions (on $(n-m)!$ ways). Therefore, there are

$$\binom{l-1}{i-1}\binom{n-l}{m-i}\, m!\,(n-m)!$$

permutations of $X_1, \ldots, X_n$ for which $X_{i:m} = X_{l:n}$. Since every permutation of $X_1, \ldots, X_n$ is equally likely, we arrive at (6).

3 Linearity of regression for exponential, Pareto and power distributions

By PAR$(\theta; \mu; \delta)$ we denote the Pareto distribution with the density

$$f(x) = \frac{\theta(\mu+\delta)^{\theta-1}}{(x+\delta)^{\theta+1}}\, I_{(\mu,\infty)}(x),$$

where $\theta > 0$, μ, δ are some real constants such that $\mu + \delta > 0$.

By EXP(λ; γ) we denote the exponential distribution with the density

$$f(x) = \lambda \exp(-\lambda(x-\gamma))\, I_{(\gamma,\infty)}(x),$$

where $\lambda > 0$, γ are some real constants.

By POW(θ; μ; ν) we denote the power distribution with the density

$$f(x) = \frac{\theta(\nu - x)^{\theta-1}}{(\nu-\mu)^{\theta}}\, I_{(\mu,\nu)}(x),$$

where $\theta > 0$, $-\infty < \mu < \nu < \infty$ are some real constants.

It is well known, see e.g. DW, that for each of the above distributions for $l > j$

$$\mathbb{E}(X_{l:n}|X_{j:n}) = \alpha\, X_{j:n} + \beta, \tag{7}$$

where α and β are some constants depending on the distribution and on l, j, n—the formulas for these constants are given on pp. 217–218 of DW. These formulas together with the representation, (2) imply for $j < i$ that

$$\mathbb{E}(X_{i:m}|X_{j:n}) = a\, X_{j:n} + b, \tag{8}$$

where a and b are suitable constants, which in each of special cases are listed below.

- For the exponential distribution EXP(λ; γ)

$$a = 1, \qquad b = \frac{(n-j)!}{\binom{n}{m}}\lambda \sum_{l=i}^{n-m+i} \frac{\binom{l-1}{i-1}\binom{n-l}{m-i}}{(n-l)!} \sum_{s=0}^{l-j-1} \frac{(-1)^s}{s!(l-j-1-s)!(n-l+s+1)^2}. \tag{9}$$

- For the Pareto distribution PAR(θ; μ; δ)

$$a = \frac{\theta(n-j)!}{\binom{n}{m}} \sum_{l=i}^{n-m+i} \frac{\binom{l-1}{i-1}\binom{n-l}{m-i}}{(n-l)!} \sum_{s=0}^{l-j-1} \frac{(-1)^s}{s!(l-j-1-s)![\theta(n-l+1+s)-1]},$$

$$b = \frac{\delta(n-j)!}{\binom{n}{m}} \sum_{l=i}^{n-m+i} \frac{\binom{l-1}{i-1}\binom{n-l}{m-i}}{(n-l)!} \sum_{s-0}^{l-j-1} \frac{(-1)^s}{s!(l-j-1-s)!(n-l+s+1)[\theta(n-l+s+1)-1]}. \tag{10}$$

- For the power distribution POW$(\theta;\mu;\nu)$

$$a=\frac{\theta(n-j)!}{\binom{n}{m}}\sum_{l=i}^{n-m+i}\frac{\binom{l-1}{i-1}\binom{n-l}{m-i}}{(n-l)!}\sum_{s=0}^{l-j-1}\frac{(-1)^s}{s!(l-j-1-s)![\theta(n-l+1+s)+1]},$$

$$b=\frac{\nu(n-j)!}{\binom{n}{m}}\sum_{l=i}^{n-m+i}\frac{\binom{l-1}{i-1}\binom{n-l}{m-i}}{(n-l)!}\sum_{s=0}^{l-j-1}\frac{(-1)^s}{s!(l-j-1-s)!(n-l+1+s)[\theta(n-l+1+s)+1]}. \tag{11}$$

For any distribution μ of a random variable X, denote by μ_- the distribution of $-X$. Since for $Y_i=-X_i$, $i=1,\ldots,n$, we have $Y_{i:n}=-X_{n-i+1:n}$ it follows that (7) holds for $l<j$ if the distribution of X_i's is one of the triplet PAR$_-$, EXP$_-$ or POW$_-$. Consequently, (8) holds for this triplet in the case $j\in\{n-m+i,\ldots,n\}$.

4 Characterization in the case $j\le i$ or $j\ge n-m+i$

These three distributions of type μ or related of type μ_- appear to be the only possible distributions for X_i's for which (8) holds with $j\le i$ or, respectively, with $j\ge n-m+i$.

Before we give the proof of our main result we recall a result on possible solutions of the integrated Cauchy functional equation. Following the method from DW we will use this result in the proof of the characterization. Let λ denote the Lebesgue measure on $\mathbb{R}_+$.

Theorem 2 (Rao and Shanbhag (1994)) *Consider the integral equation:*

$$\int_{\mathbb{R}_+} H(x+y)\mu(dy)=H(x)+c,$$

where μ is a non-arithmetic σ-finite measure on $\mathbb{R}_+$ and $H:\mathbb{R}_+\to\mathbb{R}_+$ is a Borel measurable, either non-decreasing or non-increasing λ-a.e. function that is locally λ-integrable and is not identically equal zero λ-a.e. Then there exists $\eta\in\mathbb{R}$ such that

$$\int_{\mathbb{R}_+}\exp(\eta x)\mu(dx)=1$$

and H has the form

$$H(x)=\begin{cases}\gamma+\alpha(1-\exp(\eta x)) & \lambda-a.e., \text{ if } \eta\neq 0,\\ \gamma+\beta x & \lambda-a.e. \quad\quad \text{ if } \eta=0,\end{cases}$$

where α,β,γ are some constants. If $c=0$, then $\gamma=-\alpha$ and $\beta=0$.

Now we are ready to state and then to prove our main result which is a characterization of both the triplets of distributions described in Sect. 3 by linearity of regression of order statistics from overlapping samples.

Theorem 3 *Let $X_1, \ldots, X_n$ be independent random variables with a common continuous distribution μ. Assume that $\mathbb{E}(|X_1|) < \infty$. If for some $i, m, n \in \mathbb{N}$ such that $1 \le i \le m < n \in \mathbb{N}$ linearity of regression (8) holds for some*

- *$j \in \{1, \ldots, i\}$ then only one of the following cases is possible:*
  (1) $a = 1$ and $\mu = \mathrm{EXP}$,
  (2) $a < 1$ and $\mu = \mathrm{POW}$,
  (3) $a > 1$ and $\mu = \mathrm{PAR}$.
- *$j \in \{n - m + i + 1, \ldots, n\}$ then only one of the following cases is possible:*
  (1) $a = 1$ and $\mu = \mathrm{EXP}_-$,
  (2) $a < 1$ and $\mu = \mathrm{POW}_-$,
  (3) $a > 1$ and $\mu = \mathrm{PAR}_-$.

Proof Let us note that if X has a continuous distribution function F then in the case $j < l$ the conditional distribution of $X_{l:n}$ given $X_{j:n}$ has the form

$$dF_{X_{l:n}|X_{j:n}=x}(y) = \frac{(n-j)!}{(l-j-1)!(n-l)!}\left[\frac{F(y)-F(x)}{1-F(x)}\right]^{l-j-1}\left[\frac{1-F(y)}{1-F(x)}\right]^{n-l}\frac{dF(y)}{1-F(x)}, \quad (12)$$

$l_F \le x \le y \le r_F$, where $l_F = \inf\{x \in \mathbb{R} : F(x) > 0\}$ and $r_F = \sup\{x \in \mathbb{R} : F(x) < 1\}$. Alternatively, for continuous F the conditional distribution $X_{l:n}|X_{j:n} = x$ is the same as the distribution of $Y_{l-j:n-j}$ for the Y_i, $i = 1, \ldots, n-j$, which are iid and their common distribution function is $F_Y(y) = \frac{F(y)-F(x)}{1-F(x)}$, $y \ge x$ and $F_Y(y) = 0$, otherwise. This fact seems to be well known for continuous parent distribution (in particular, it was used in DW). Since in basic monographs by Arnold et al. (1992), David and Nagaraja (2003) it is stated only in the absolutely continuous case, while in Nevzerov (2001) it is formulated for continuous distributions but proved only in the absolutely continuous case, for the sake of completeness we sketch its proof here. We note that from the well known general formula for the distribution function of $X_{k:n}$ (see, e.g. (2.2.15) in Arnold et al. (1992), in the continuous case, since then $F(X_i)$ has the uniform distribution on (0, 1), one gets

$$dF_{k:n}(x) = \frac{n!}{(k-1)!(n-k)!}(1-F(x))^{n-k}F^{k-1}(x)\,dF(x)$$

for any $k = 1, \ldots, n$. Therefore, to prove the formula (12) it suffices to check (which is an elementary computation) that with $dF_{X_{l:n}|X_{j:n}=x}(y)$ defined by (12) the following identity holds

$$dF_{l:n}(y) = \int_{-\infty}^{y} dF_{X_{l:n}|X_{j:n}=x}(y)\,dF_{j:n}(x)$$

for any $y \in \mathbb{R}$.

Let us first consider the case when $j < i$. From (2) and (12) we have:

$$\mathbb{E}(X_{i:m}|X_{j:n}=x) = \sum_{l=i}^{n-m+i} A_l B_l \int_x^{\infty} y \left(\frac{\overline{F}(x)-\overline{F}(y)}{\overline{F}(x)}\right)^{l-j-1} \left(\frac{\overline{F}(y)}{\overline{F}(x)}\right)^{n-l} \times d\left(-\frac{\overline{F}(y)}{\overline{F}(x)}\right) = ax+b, \tag{13}$$

where $A_l = \frac{\binom{l-1}{i-1}\binom{n-l}{m-i}}{\binom{n}{m}}$ and $B_l = \frac{(n-j)!}{(l-j-1)!(n-l)!}$, $x \in (l_F,\ r_F)$

Observe that there does not exist an interval (c, d), $l_F < c < d < r_F$, on which F is constant, because the right side of (13) is either strictly increasing or strictly decreasing. Both sides of this equation are continuous, so they could not be equal in the next point of increase of F. Therefore (l_F, r_F) is the support of distribution given by F and F is strictly increasing on this interval. Both sides of the second equation in (13) are continuous with respect to x, so it holds for any $x \in (l_F, r_F)$. After substituting $t = \overline{F}(y)/\overline{F}(x)$, we insert $y = \overline{F}^{-1}(t\overline{F}(x))$ ($\overline{F}^{-1}$ exists, because $\overline{F}$ is strictly decreasing on (l_F, r_F)) into (13) and thus

$$\sum_{l=i}^{n-m+i} A_l B_l \int_0^1 \overline{F}^{-1}(t\overline{F}(x))t^{n-l}(1-t)^{l-j-1}\,dt = ax+b. \tag{14}$$

Note that the left hand side is strictly increasing in x and thus a has to be positive. Substituting again $\overline{F}(x) = w$ in (14), which implies $x = \overline{F}^{-1}(w)$, we get:

$$\sum_{l=i}^{n-m+i} A_l B_l \int_0^1 \overline{F}^{-1}(tw)t^{n-l}(1-t)^{l-j-1}\,dt = a\overline{F}^{-1}(w)+b, \qquad w \in (0,1).$$

Divide both sides of the above equation by a and substitute again $t = e^{-u}$ and $w = e^{-v}$ for $v > 0$ to arrive at

$$\sum_{l=i}^{n-m+i} \frac{A_l B_l}{a} \int_0^{\infty} \overline{F}^{-1}(e^{-(u+v)})(1-e^{-u})^{l-j-1}e^{-(n-l)u}e^{-u}\,du = \overline{F}^{-1}(e^{-v}) + \frac{b}{a}.$$

After changing sum of integrals into integral of sums:

$$\int_0^{\infty} \overline{F}^{-1}(e^{-(u+v)})\left(\sum_{l=i}^{n-m+i} \frac{A_l B_l}{a}(1-e^{-u})^{l-j-1}e^{-(n-l)u}\right)e^{-u}\,du = \overline{F}^{-1}(e^{-v}) + \frac{b}{a}.$$

Let us now define $H(v) = \overline{F}^{-1}(e^{-v})$. Consequently,

$$\int_{\mathbb{R}_+} H(u+v)\mu(du) = H(v) + \frac{b}{a}, \qquad v > 0,$$

where μ is a finite measure on $\mathbb{R}_+$, which is absolutely continuous with respect to the Lebesgue measure and has the form

$$\mu(du) = \left(\sum_{l=i}^{n-m+i} \frac{A_l B_l}{a} (1-e^{-u})^{l-j-1} e^{-(n-l)u} e^{-u} \right) du.$$

Note that H is strictly increasing on $[0, \infty)$ as composition of two strictly decreasing functions. The assumptions of the Rao-Shanbhag theorem are satisfied, so H has the form

$$H(v) = \begin{cases} \gamma + \alpha(1-\exp(\eta v)), & \text{if } \eta \neq 0, \\ \gamma + \beta v, & \text{if } \eta = 0, \end{cases}$$

$v > 0$, where $\alpha, \beta, \gamma, \delta, \eta$ are some constants and

$$\int_{\mathbb{R}_+} \exp(\eta x)\mu(dx) = 1. \tag{15}$$

To find relations between η and a we rewrite (15) as

$$1 = \int_0^\infty e^{\eta x} \left(\sum_{l=i}^{n-m+i} \frac{A_l B_l}{a} (1-e^{-x})^{l-j-1} e^{-(n-l)x} \right) e^{-x}\, dx.$$

After substituting $t = e^{-x}$

$$1 = \int_0^1 \left(\sum_{l=i}^{n-m+i} \frac{A_l B_l}{a} (1-t)^{l-j-1} t^{n-l-\eta} \right) dt.$$

Performing the integration at the right hand side above (note that necessarily $\eta < m - i + 1$, otherwise the integrals are infinite) we get

$$1 = \sum_{l=i}^{n-m+i} \frac{\binom{l-1}{i-1}\binom{n-l}{m-i}}{\binom{n}{m}} \frac{(n-j)!}{a(l-j-1)!(n-l)!} \frac{\Gamma(n-l-\eta+1)\Gamma(l-j)}{\Gamma(n-j-\eta+1)}$$
$$= \sum_{l=i}^{n-m+i} \frac{\binom{l-1}{i-1}\binom{n-l}{m-i}}{\binom{n}{m}} \frac{(n-j)!}{a(n-l)!} \frac{\Gamma(n-l-\eta+1)}{\Gamma(n-j-\eta+1)}.$$

Finally, we get

$$a = \sum_{l=i}^{n-m+i} \frac{\binom{l-1}{i-1}\binom{n-l}{m-i}}{\binom{n}{m}} h_l(\eta), \tag{16}$$

where

$$h_l(\eta) = \frac{(n-j)}{(n-j-\eta)} \frac{(n-j-1)}{(n-j-\eta-1)} \cdots \frac{(n-l+1)}{(n-l-\eta+1)}.$$

Since the function h_l is strictly increasing on $(-\infty, m-i+1)$ it follows from (16) that for a given coefficient a there exists a unique η satisfying (15). Moreover,

- if $\eta = 0$ then $a = 0$,
- if $0 < \eta < m-i+1$ then $a > 1$,
- if $\eta < 0$ then $a < 1$.

Let us now consider the case when $j = i$. From (12) we get

$$\mathbb{E}(X_{i:m}|X_{i:n} = x) = A_i x + \sum_{l=i+1}^{n-m+i} A_l B_l \int_x^\infty y \left(\frac{\overline{F}(x) - \overline{F}(y)}{\overline{F}(x)}\right)^{l-i-1} \left(\frac{\overline{F}(y)}{\overline{F}(x)}\right)^{n-l}$$
$$\times d\left(-\frac{\overline{F}(y)}{\overline{F}(x)}\right),$$

thus instead of (14) we get

$$\sum_{l=i+1}^{n-m+i} A_l B_l \int_x^\infty y \left(\frac{\overline{F}(x) - \overline{F}(y)}{\overline{F}(x)}\right)^{l-i-1} \left(\frac{\overline{F}(y)}{\overline{F}(x)}\right)^{n-l} d\left(-\frac{\overline{F}(y)}{\overline{F}(x)}\right) = (a - A_i)x + b.$$

Similarly, as in the case above we make substitutions and use the Rao-Shabhag theorem to arrive at the solution H. The only difference is the equation for a which now reads

$$a - A_i = \sum_{l=i+1}^{n-m+i} \frac{\binom{l-1}{i-1}\binom{n-l}{m-i}}{\binom{n}{m}} h_l(\eta).$$

This equation gives us the same condition for parameter a as for the case $j < i$.

Before computing the parameters of distributions we arrived at, we will explain why solution of the case $j \le i$ gives also the solution in the case $j \ge n - m + i$. Define $Y_k = -X_k$, $k = 1, \dots, n$ and consider order statistics of the random vector $(Y_1, \dots, Y_n)$. Since $Y_{k:n} = -X_{n-k+1:n}$, so we can write for $j \ge n - m + i$:

$$-aY_{n-j+1:n} + b = aX_{j:n} + b = \mathbb{E}(X_{i:m}|X_{j:n}) = -\mathbb{E}(Y_{m-i+1:m}|Y_{n-j+1:n}).$$

Consequently,

$$\mathbb{E}(Y_{i':m}|Y_{j':n}) = a'Y_{j':n} + b',$$

where $j' = n - j + 1 \le i' = m - i + 1$, $a' = a$ and $b' = -b$.

We will find distribution functions only in the case $j < i$ (For $i = j$ the derivation is almost exactly the same and is skipped. In the case $j \ge n - m + 1$ one has again to refer to the representation $Y_k = -X_k$ and use the results of the case $j \le i$). For $\eta \ne 0$ from the definition of H we get

$$\overline{F}^{-1}(e^{-v}) = \gamma + \alpha(1 - e^{\eta v}).$$

Hence for $z > \gamma$

$$\overline{F}(z) = \left(\frac{1}{1 - \frac{z-\gamma}{\alpha}}\right)^{1/\eta}. \tag{17}$$

Consider now three cases:

(1) $a < 1$ and $\eta < 0$ then (17) for $z \in (\mu, \nu)$ can be written as

$$\overline{F}(z) = \left(\frac{\alpha + \gamma - z}{\alpha}\right)^{-1/\eta} = \left(\frac{\alpha + \gamma - z}{\alpha + \gamma - \gamma}\right)^{-1/\eta} = \left(\frac{\nu - z}{\nu - \mu}\right)^{\theta},$$

where $\nu = \alpha + \gamma$, $\mu = \gamma$, $\theta = -\frac{1}{\eta} > 0$. Notice that α has to be positive. Hence X_1 has POW$(\theta; \mu; \nu)$ distribution and

(a) $\theta = -\frac{1}{\eta}$, where η satisfies (16),

(b) ν may be calculated from (11) with $\theta = -\frac{1}{\eta}$,

(c) μ is a real number such that $\mu < \nu$.

(2) $a > 1$ and $\eta > 0$ then (17) for $z > \mu$ can be written as

$$\overline{F}(z) = \left(\frac{-\alpha}{z - \alpha - \gamma}\right)^{1/\eta} = \left(\frac{\gamma + (-\alpha - \gamma)}{z + (-\alpha - \gamma)}\right)^{1/\eta} = \left(\frac{\mu + \delta}{z + \delta}\right)^{\theta},$$

where and $\delta = -\alpha - \gamma$, $\mu = \gamma$, $\theta = \frac{1}{\eta} > 0$.
Thus X_1 has PAR$(\theta; \mu; \delta)$ distribution and
(a) $\theta = \frac{1}{\eta}$, where η satisfies (16),
(b) δ may be calculated from (10) with $\theta = \frac{1}{\eta}$,
(c) μ is a real number.

(3) $a = 1$ and $\eta = 0$ then by the definition of H we get

$$\overline{F}^{-1}(e^{-v}) = \gamma + \beta v$$

and, consequently,

$$\overline{F}(z) = e^{-(z-\gamma)/\beta} = e^{-\lambda(z-\gamma)}$$

for $z > \gamma$, where $\lambda = \frac{1}{\beta} > 0$
Hence X_1 has EXP$(\lambda; \gamma)$ distribution and
(a) λ may be calculated from the formula for b in (9),
(b) γ is a real number. □

5 The case $i < j < n - m + i$ remains unsolved

As it was already said in the introduction if $i < j < n - m + i$ then only the case $m = i = 1$ was considered in Wesołowski and Gupta (2001) (see also Nagaraja and Nevzerov 1997, and Gupta and Kirmani 2008). More precisely, only the family of distributions for which $\mathbb{E}(X_1|X_{k+1:2k+1}) = aX_{k+1:2k}$ was described. Unexpectedly, this family is completely different than the triplets of distributions described above, e.g. it contains Student distribution with two degrees of freedom.

In the case $j \in \{i+1, \dots, n-m+i-1\}$ it follows from Theorem 1 that

$$\begin{aligned}
&\mathbb{E}(X_{i:m}|X_{j:n} = x) \\
&= \frac{\binom{j-1}{i-1}\binom{n-j}{m-i}}{\binom{n}{m}} x + \sum_{l=i}^{j-1} \frac{\binom{l-1}{i-1}\binom{n-l}{m-i}}{\binom{n}{m}} \frac{(j-1)!}{(l-1)!(j-l-1)!} \\
&\times \int_{-\infty}^{x} \left(\frac{F(y)}{F(x)}\right)^{l-1} \left(\frac{F(x)-F(y)}{F(x)}\right)^{j-l-1} \frac{f(y)}{F(x)} dy + \sum_{l=j+1}^{n-m+i} \frac{\binom{l-1}{i-1}\binom{n-l}{m-i}}{\binom{n}{m}} \\
&\times \int_{x}^{\infty} \frac{(n-j)!}{(l-j-1)!(n-l)!} \left(\frac{F(y)-F(x)}{1-F(x)}\right)^{l-j-1} \left(\frac{1-F(y)}{1-F(x)}\right)^{n-l} \frac{f(y)}{1-F(x)} dy\,.
\end{aligned}$$

Linearity of regression, as in (1) would imply that the right hand side above equals $ax+b$. Such an equation seems to be much harder to solve than the one solved in Sect. 4 above. In particular, it is not visible how to reduce it, through some substitutions, to the Rao-Shanbhag equation.

For $i = 1, j = 2, m = 2, n = 4$ under linearity of regression assumption we obtain the equation

$$\mathbb{E}(X_{1:2}|X_{2:4} = x) = \frac{1}{3}x + \frac{1}{2}\int_{-\infty}^{x} \frac{f(y)}{F(x)}dy + \frac{1}{3}\int_{x}^{\infty} \frac{1 - F(y)}{1 - F(x)}\,\frac{f(y)}{1 - F(x)}dy = ax + b.$$

Similarly, for $i = 2, j = 3, m = 2, n = 4$ we have

$$\mathbb{E}(X_{2:2}|X_{3:4} = x) = \frac{1}{3}x + \frac{1}{6}\int_{-\infty}^{x} \frac{F(y)}{F(x)}\,\frac{f(y)}{F(x)}dy + \frac{2}{3}\int_{x}^{\infty} \frac{f(y)}{1 - F(x)}dy = ax + b.$$

These two last equations seem to be the simplest unsolved cases.

Nevertheless, it can be easily verified that if a sample is taken from a uniform distribution then both the above linearity of regression conditions hold true.

References

Ahsanullah M, Hamedani GG (2012) Characterizations of certain univariate distributions based on the conditional distribution of generalized order statistics. Pakistan J Stat 28(2):253–258

Ahsanullah M, Hamedani GG, Wesołowski J (2012) Linearity of regressions inside top-k-lists and related characterizations. Studia Sci Math Hungar 49(4):436–445

Ahsanullah M, Nevzerov VB (1999) Spacings of order statistics from extended sample. In: Ahsanullah M, Yildrim F (eds) Applied statistical science IV. Nova Sci. Publ, Commack, pp 251–257

Ahsanullah M, Wesołowski J (1997) On characterizing distributions via linearity of regression for order statistics. Aust J Stat 39(1):69–78

Arnold BC, Balakrishnan N, Nagaraja HN (1992) A first course in order statistics. Wiley, New York

Beg MI, Ahsanullah M, Gupta RC (2013) Characterizations via regressions for generalized order statistics. Stat Methodol 12:31–41

Bieniek M, Szynal D (2003) Characterizations of distributions via linearity of regression of generalized order statistics. Metrika 58:259–272

Cramer E, Kamps U, Keseling C (2004) Characterizations via linear regressions of order statistics: a unifying approach. Commun Stat Theory Methods 33:2885–2911

David HA, Nagaraja HN (2003) Order statistics. Wiley, Hoboken

Dembińska A, Wesołowski J (1997) On characterizing the exponential distribution by linearity of regression for non-adjacent order statistics. Demonstratio Mathematica 30:945–952

Dembińska A, Wesołowski J (1998) Linearity of regression for non-adjacent order statistics. Metrika 48:215–222

Ferguson TS (1967) On characterizing distributions by properties of order statistics. Sankhya A 29:265–278

Ferguson TS (2002) On a Rao-Shanbhag characterization of exponential/geometric distribution. Sankya A 64:246–255

Fisz M (1958) Characterizations of some probability distributions. Skand Aktuarietidskr 41:65–70

Gupta RC, Ahsanullah M (2004) Some characterization results based on the conditional expectation of a function of non-adjacent order statistics (record values). Ann Inst Stat Math 56:721–732

Gupta RC, Kirmani SNUA (2008) Characterizations based on convex conditional mean function. J Stat Plan Inference 138:964–970

López-Blázquez F, Moreno-Rebollo JL (1997) A characterization of distributions based on linearity of regression for order statistics and record values. Sankhya A 59:311–323

Nagaraja HN, Nevzerov VB (1997) On characterizations based on records and order statistics. J Stat Plan Inference 63:271–284

Nevzerov VB (2001) Records: mathematical theory. AMS, Providence

Pudeg A (1991) Characterization of probability distributions via distributional properties of order statistics and record values. PhD Dissert., Aachen Univ. Tech., Aachen (in German)

Rao CR, Shanbhag DN (1994) Choquet-Deny type of functional equations with applications to stochastic models. Wiley, New York

Rogers GS (1963) An alternative proof of the characterization of the density Ax^{β}. Am Math Mon 70:857–858

Wesołowski J, Gupta AK (2001) Linearity of convex mean residual life. J Stat Plan Inference 99:183–191