Available online at www.sciencedirect.com

ScienceDirect

Journal of Statistical Planning and Inference 138 (2008) 1358–1367

journal of statistical planning and inference

www.elsevier.com/locate/jspi

Characterizations based on record values and order statistics

Jyh-Cherng Su[a,*,1], Nan-Cheng Su[b,2], Wen-Jang Huang[c,2]

[a]Department of Management Science, R.O.C. Military Academy, Fengshan, Kaohsiung 830, Taiwan, ROC
[b]Department of Applied Mathematics, National Sun Yat-sen University, Kaohsiung 804, Taiwan, ROC
[c]Department of Applied Mathematics, National University of Kaohsiung, Kaohsiung 811, Taiwan, ROC

Received 26 November 2005; received in revised form 10 April 2007; accepted 22 June 2007
Available online 3 July 2007

Abstract

Consider a sequence of independent and identically distributed random variables $\{X_i, i \geqslant 1\}$ with a common absolutely continuous distribution function F. Let $X_{1:n} \leqslant X_{2:n} \leqslant \cdots \leqslant X_{n:n}$ be the order statistics of $\{X_1, X_2, \ldots, X_n\}$ and $\{Y_l, l \geqslant 1\}$ be the sequence of record values generated by $\{X_i, i \geqslant 1\}$. In this work, the conditional distribution of Y_l given $X_{n:n}$ is established. Some characterizations of F based on record values and $X_{n:n}$ are then given.

MSC: Primary 62E10; Secondary 60E05

Keywords: Characterization; Conditional distribution; Conditional expectation; Order statistics; Record values

1. Introduction

Throughout this work, let $\{X_i, i \geqslant 1\}$ be a sequence of independent and identically distributed random variables with a common absolutely continuous distribution function F, a probability density function (p.d.f.) f, a hazard function r and a cumulative hazard function R. Assume that F has support (a, b) and $f(x) > 0$, $a < x < b$, where $-\infty \leqslant a < b \leqslant \infty$. Note that $r(x) = f(x)/(1 - F(x))$, $R(x) = \int_a^x r(u)\,\mathrm{d}u = -\log(1 - F(x))$, $a < x < b$, and $0 < r(x)$, $R(x) < \infty$, $a < x < b$. Let $X_{1:n} \leqslant X_{2:n} \leqslant \cdots \leqslant X_{n:n}$ be the order statistics based on $\{X_1, X_2, \ldots, X_n\}$. In addition, the sequences of record times $U(l)$ and record values Y_l are defined as follows: $U(1) = 1$, $U(l+1) = \min\{i > U(l) : X_i > X_{U(l)}\}$ and $Y_l = X_{U(l)}$, $l \geqslant 1$. Furthermore, let $N_n = \sup\{l | Y_l \leqslant X_{n:n}, l \geqslant 1\}$ be the number of records among the first n observations. The properties and characterizations related to order statistics and record values have been widely studied and some excellent reviews can be found in books such as Arnold et al. (1992, 1998), Rao and Shanbhag (1994), Nevzorov (2001) and David and Nagaraja (2003).

The properties between record values and order statistics have also been extensively investigated. Huang and Su (1999) viewed both record values and order statistics as point processes with order statistics properties and gave some characterization results. This explains why record values and order statistics are closely related. Hence there are many parallel characterizations for record values and order statistics. Related studies have been reported by Deheuvels (1984),

* Corresponding author. Tel.: +886 7 7403060; fax: +886 7 7104697.
E-mail addresses: sujc@cc.cma.edu.tw (J.-C. Su), sunc@math.nsysu.edu.tw (N.-C. Su), huangwj@nuk.edu.tw (W.-J. Huang).
[1] Support for this research was provided in part by the National Science Council of the Republic of China, Grant No. NSC 93-2118-M-145-001.
[2] Support for this research was provided in part by the National Science Council of the Republic of China, Grant No. NSC 94-2118-M-390-003.

0378-3758/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.jspi.2007.06.021

Gupta (1984), Nagaraja (1988), Huang and Su (1999) and Huang et al. (2007), etc. These results can be applied to characterize the uniform distribution using the sequence of order statistics, and the exponential distribution using the sequence of record values, respectively.

On the other hand, Nagaraja and Nevzorov (1997) investigated the conditional distribution of Y_2 given $X_{2:2}$, and obtained a characterization of the exponential distribution based on the independence of $Y_2 - X_{2:2}$ and $X_{2:2}$. In fact, from (5.2) to (5.4) of Nagaraja and Nevzorov (1997), it is easily shown that the distribution function F can be uniquely determinated by the conditional expectation $g(x) = E(Y_2 - X_{2:2}|X_{2:2} = x)$, $a < x < b$. Balakrishnan and Stepanov (2004) recently extended the work of Nagaraja and Nevzorov (1997). They established the conditional distribution of Y_2 given $X_{n:n}$, where $n \geqslant 1$. Some characterizations related to the conditional distribution of Y_2 given $X_{n:n}$ were also studied by Balakrishnan and Stepanov (2004).

In this work, we first provide a generalization of Balakrishnan and Stepanov (2004). More precisely, the conditional distribution of the record value Y_l given $X_{n:n}$, where l and n are positive integers, is established. Characterizations of distributions based on record values and $X_{n:n}$ are then studied. Our results extend some characterization theorems in the literature.

2. The conditional distribution of record values given $X_{n:n}$

First, we give the following two lemmas which can be used to prove Theorem 1, and can be found in books by Arnold et al. (1998) or Nevzorov (2001).

Lemma 1. *For any positive integer n, N_n and $X_{n:n}$ are independent, and*

$$f_{N_n}(k) = \frac{1}{n} \sum_{1<t_1<\cdots<t_{k-1}\leqslant n} \left(\prod_{i=1}^{k-1} (t_i - 1) \right)^{-1}, \quad 1 \leqslant k \leqslant n, \tag{1}$$

where $\sum_{1<t_1<\cdots<t_{k-1}\leqslant n} (\prod_{i=1}^{k-1}(t_i-1))^{-1}$ is defined to be 1 for $k=1$.

It is known that $f_{N_n}(k)$ is the coefficient of the term s^k in the expansion of $\prod_{j=1}^{n}(s+j-1)/n!$.

Lemma 2. *For positive integers $k \leqslant n$ and $l \geqslant k+1$,*

$$\begin{aligned} f_{Y_l}(y|N_n = k, X_{n:n} = x) &= f_{Y_l}(y|Y_k = x) \\ &= \frac{1}{(l-k-1)!}(R(y) - R(x))^{l-k-1} \frac{f(y)}{1-F(x)}, \quad a < x < y < b. \end{aligned}$$

In the following, we give the conditional distribution of Y_l given $X_{n:n}$. Note that the conditional distributions for cases $l = 1$ and 2 have been given by Feller (1970) and Balakrishnan and Stepanov (2004), respectively. Hereafter, let $p_{n,k} = f_{N_n}(k)$, $1 \leqslant k \leqslant n$. Obviously $\sum_{k=1}^{n} p_{n,k} = 1$. For convenience, in the following, let $p_{n,0} = 0$, $n \geqslant 1$, and $p_{0,0} = 1$. Note that for a random vector where some of the variables are discrete and the others are continuous, the joint p.d.f., conditional p.d.f. and marginal p.d.f. are defined as usual. That is, the probabilities are calculated by summing for discrete variables, and integrating for continuous variables.

Theorem 1. *Let l, n be positive integers and $a < x$, $y < b$.*

(i) *For $l > n$,*

$$f_{Y_l}(y|X_{n:n} = x) = \begin{cases} \sum_{k=1}^{n} \frac{p_{n,k}}{(l-k-1)!}(R(y) - R(x))^{l-k-1} \frac{f(y)}{1-F(x)}, & y \geqslant x, \\ 0, & y < x. \end{cases} \tag{2}$$

(ii) *For* $l \leqslant n$, $P(Y_l = x|X_{n:n} = x) = p_{n,l}$, *and*

$$f_{Y_l}(y|X_{n:n}=x) = \begin{cases} \sum_{k=1}^{l-1} \frac{p_{n,k}}{(l-k-1)!}(R(y)-R(x))^{l-k-1}\frac{f(y)}{1-F(x)}, & y>x, \\ \sum_{k=l}^{n-1} p_{k-1,l-1}\left(1-\frac{k}{n}\right)\left(\frac{F(y)}{F(x)}\right)^{k-1}\frac{f(y)}{F(x)}, & y<x. \end{cases}$$

Proof. First, we have

$$f_{Y_l}(y|X_{n:n}=x) = \sum_{k=1}^{n} f_{N_n,Y_l}(k, y|X_{n:n}=x). \tag{3}$$

Assume $l > n$, it can be seen that $P(Y_l < x|X_{n:n} = x) = 0$, and for $y \geqslant x$, it turns out

$$f_{Y_l}(y|X_{n:n}=x) = \sum_{k=1}^{n} f_{Y_l}(y|N_n=k, X_{n:n}=x) f_{N_n}(k|X_{n:n}=x).$$

The assertion of part (i) now follows immediately from Lemmas 1 and 2.

Next, assume $l \leqslant n$. It is clear that

$$P(Y_l = x|X_{n:n}=x) = f_{N_n}(l|X_{n:n}=x) = p_{n,l}.$$

Now if $y > x$, it can easily be seen that

$$f_{Y_l}(y|N_n=k, X_{n:n}=x) = 0, \quad k \geqslant l. \tag{4}$$

In view of (3) and (4), we obtain

$$\begin{aligned} f_{Y_l}(y|X_{n:n}=x) &= \sum_{k=1}^{l-1} f_{Y_l}(y|N_n=k, X_{n:n}=x) f_{N_n}(k|X_{n:n}=x) \\ &= \sum_{k=1}^{l-1} \frac{p_{n,k}}{(l-k-1)!}(R(y)-R(x))^{l-k-1}\frac{f(y)}{1-F(x)}, \quad y>x. \end{aligned}$$

On the other hand, if $y < x$, it can easily be seen that for $l = n$, we have $f_{Y_n}(y|X_{n:n} = x) = 0$, and for $l < n$ and $a < y_1 < \cdots < y_{l-1} < y < x < b$, we have

$$\begin{aligned} &f_{Y_1,\ldots,Y_{l-1},Y_l,X_{n:n}}(y_1,\ldots,y_{l-1},y,x) \\ &\quad = \sum_{1<t_1<\cdots<t_{l-1}<n} \left(\prod_{i=1}^{l-1}(F(y_i))^{t_i-t_{i-1}-1} f(y_i)\right) f(y)(n-t_{l-1})(F(x))^{n-t_{l-1}-1} f(x), \end{aligned} \tag{5}$$

where $t_0 = 1$ and again $\sum_{1<t_1<\cdots<t_{l-1}<n}\prod_{i=1}^{l-1}(F(y_i))^{t_i-t_{i-1}-1}f(y_i)$ is defined as 1 for $l = 1$. Taking the integrations with respect to $y_1, \ldots, y_{l-1}$, yields

$$\begin{aligned} f_{Y_l,X_{n:n}}(y,x) &= \sum_{1<t_1<\cdots<t_{l-1}<n} \frac{n-t_{l-1}}{\prod_{i=1}^{l-1}(t_i-1)}\left(\frac{F(y)}{F(x)}\right)^{t_{l-1}-1}(F(x))^{n-2} f(x) f(y) \\ &= \sum_{k=l}^{n-1} p_{k-1,l-1}(n-k)\left(\frac{F(y)}{F(x)}\right)^{k-1}(F(x))^{n-2} f(x) f(y), \quad y<x. \end{aligned} \tag{6}$$

Consequently,

$$f_{Y_l}(y|X_{n:n}=x) = \sum_{k=l}^{n-1} p_{k-1,l-1}\left(1-\frac{k}{n}\right)\left(\frac{F(y)}{F(x)}\right)^{k-1}\frac{f(y)}{F(x)}, \quad y<x,$$

and the proof of part (ii) is finished. □

Note that in Theorem 1, $f_{Y_l}(y|X_{n:n}=x)$, $y \geqslant x$, can also be derived from (B) and (BB) of Balakrishnan and Stepanov (2004). Furthermore, in view of Lemma 2 and (2), it is easy to check that for $l > n$ and $a < x$, $y < b$,

$$f_{Y_l}(y|X_{n:n}=x) = \sum_{k=1}^{n} p_{n,k} f_{Y_l}(y|Y_k = x).$$

Namely, the conditional distribution can be represented as a mixture of n distributions.

3. Characterizations based on conditional expectations of record values given $X_{n:n}$

In this section, we first give a lemma which can be used for proving Theorems 2–4.

Lemma 3. *Assume that f is differentiable and*

$$r'(x) = \alpha(r(x) - \beta)(r(x))^2, \quad a < x < b, \tag{7}$$

for some $\alpha > 0$ *and* $\beta > 0$. *Also assume*

there exists $a < t < b$ *such that* $r(t) \geqslant \beta$ *(condition A).*

Then $a > -\infty$, $b = \infty$ *and* $F(x) = 1 - e^{-\beta(x-a)}$, $a < x < \infty$.

Proof. First (7) can be rewritten as

$$r(u) = \beta + \frac{1}{\alpha}\frac{r'(u)}{r^2(u)}, \quad a < u < b. \tag{8}$$

Integrating both sides of (8) with respect to u over (a, x), $a < x < b$, we have

$$R(x) = \beta x - \frac{1}{\alpha}\frac{1}{r(x)} - \lim_{u \downarrow a}\left(\beta u - \frac{1}{\alpha}\frac{1}{r(u)}\right). \tag{9}$$

Suppose $a = -\infty$. Since $r(x) > 0$, $a < x < b$, and $\alpha, \beta > 0$, it yields $\lim_{u \to -\infty}(\beta u - (1/\alpha)1/r(u)) = -\infty$. This in turn implies $R(x) = \infty$, which contradicts the fact $R(x) < \infty$, $a < x < b$. Hence $a > -\infty$. It consequently follows, from (9), that $r(a^+) = \lim_{u \downarrow a} r(u) > 0$ and

$$\alpha R(x) = \alpha\beta(x - a) - \frac{1}{r(x)} + \frac{1}{r(a^+)}, \quad a < x < b. \tag{10}$$

On the other hand, from (10) and the fact $\lim_{x \uparrow b} R(x) = \infty$, it turns out $b = \infty$.

Now suppose that there exists $a < x_0 < \infty$ such that $r(x_0) \neq \beta$, we are able to derive a contradiction again. The assumption that f is differentiable implies that $r(x)$ is continuous in $x > a$, hence $r(x) \neq \beta$ for x belonging to some open interval containing x_0. Let $a_1 = \inf\{x | r(x) \neq \beta\}$ and $a_2 = \inf\{x | x > a_1, r(x) = \beta\}$. Then $a \leqslant a_1 < a_2 \leqslant \infty$ (a_2 is defined to be ∞ if $r(x) \neq \beta$ for every $x > a_1$) and $r(x) \neq \beta$, $a_1 < x < a_2$. From (7), we have

$$\frac{r'(u)}{(r(u) - \beta)(r(u))^2} = \alpha, \quad a_1 < u < a_2. \tag{11}$$

Integrating both sides of (11) with respect to u over (a_1, x), $a_1 < x < a_2$, it yields

$$\frac{1}{\beta r(x)} + \frac{1}{\beta^2}\log\left|1 - \frac{\beta}{r(x)}\right| - \lim_{u \downarrow a_1}\left(\frac{1}{\beta r(u)} + \frac{1}{\beta^2}\log\left|1 - \frac{\beta}{r(u)}\right|\right) = \alpha(x - a_1). \tag{12}$$

Now assume $a_1 > a$. The continuity of $r(x)$ implies $\lim_{u \downarrow a_1} r(u) = r(a_1) = \beta$, this in turn implies that the left side of (12) is ∞. Hence $a_1 = a$. Consequently, from (12) we conclude $r(a^+) \neq \beta$, and for $a < x < a_2$,

$$\log\left|1 - \frac{\beta}{r(x)}\right| - \log\left|1 - \frac{\beta}{r(a^+)}\right| = \beta\left(\alpha\beta(x - a) - \frac{1}{r(x)} + \frac{1}{r(a^+)}\right). \tag{13}$$

Similarly, we obtain $a_2 = \infty$. That is $r(x) \neq \beta$, for $a < x < \infty$. In view of (10) and (13), it turns out

$$\left|\frac{1}{\beta} - \frac{1}{r(x)}\right| = \left|\frac{1}{\beta} - \frac{1}{r(a^+)}\right| e^{\alpha\beta R(x)}, \quad a < x < \infty. \tag{14}$$

As $r(x)$ is continuous and $r(x) \neq \beta$, $a < x < \infty$, condition A leads to $r(x) > \beta$, $a < x < \infty$. Letting $x \to \infty$, the right side of (14) approaches ∞. This in turn implies $\lim_{x\to\infty} r(x) = 0$, which contradicts $r(x) > \beta$, $a < x < \infty$. Therefore, we conclude that there does not exist $a < x_0 < \infty$ such that $r(x_0) \neq \beta$. This results in $r(x) = \beta$, $a < x < \infty$, and $F(x) = 1 - e^{-\beta(x-a)}$, $a < x < \infty$, follows. □

In Lemma 3, suppose that condition A is replaced by the condition in which "*r* is an increasing hazard function". It can easily be seen that (7) implies that $r(x) \geqslant \beta$, $a < x < b$. Along the lines of the above proof, the assertion of Lemma 3 still follows. Another replacement for condition A is "$\lim_{x\uparrow b} r(x) > 0$" or "$\lim_{x\downarrow a} r(x) = \beta$".

As mentioned in Section 1, Nagaraja and Nevzorov (1997) proved that the independence of $Y_2 - X_{2:2}$ and $X_{2:2}$ implies that F is an exponential distribution function. In the following three theorems, we investigate the characterizations of the exponential distribution based on the independence of $Y_{l+1} - Y_l$ and $X_{2:2}$, where $l \geqslant 2$, $Y_3 - X_{2:2}$ and $X_{2:2}$, and $Y_3 - X_{3:3}$ and $X_{3:3}$, respectively.

Theorem 2. *Assume that f is differentiable and for some integer* $l \geqslant 2$,

$$E(Y_{l+1} - Y_l | X_{2:2} = x) = \eta, \quad a < x < b, \tag{15}$$

where $\eta > 0$ *is independent of x. Also assume that condition A holds with* $\beta = 1/\eta$. *Then* $a > -\infty$, $b = \infty$ *and* $F(x) = 1 - e^{-(1/\eta)(x-a)}$, $a < x < \infty$.

Proof. From (15), it can be obtained that

$$\begin{aligned}\eta(1 - F(x)) = &\int_x^b y \sum_{k=1}^{2} \frac{p_{2,k}}{(l-k)!} (R(y) - R(x))^{l-k} f(y)\,dy \\ &- \int_x^b y \sum_{k=1}^{2} \frac{p_{2,k}}{(l-k-1)!} (R(y) - R(x))^{l-k-1} f(y)\,dy, \quad a < x < b.\end{aligned} \tag{16}$$

After differentiating both sides of (16) with respect to x l times and some manipulations, we obtain

$$r'(x) = 2\eta \left(r(x) - \frac{1}{\eta}\right)(r(x))^2, \quad a < x < b.$$

Now Lemma 3 yields the conclusion. □

Theorem 3. *Assume that f is differentiable and*

$$E(Y_3 - X_{2:2} | X_{2:2} = x) = \eta, \quad a < x < b, \tag{17}$$

where $\eta > 0$ *is independent of x. Also assume that condition A holds with* $\beta = 3/(2\eta)$. *Then* $a > -\infty$, $b = \infty$ *and* $F(x) = 1 - e^{-(3/(2\eta))(x-a)}$, $a < x < \infty$.

Proof. From (17), we have

$$\int_x^b y(R(y) - R(x) + 1) f(y)\,dy = 2(x + \eta)(1 - F(x)), \quad a < x < b. \tag{18}$$

After differentiating both sides of (18) with respect to x twice, it yields

$$r'(x) = \eta \left(r(x) - \frac{3}{2\eta}\right)(r(x))^2, \quad a < x < b.$$

Again the conclusion follows from Lemma 3. □

The following Theorems 4 can be proved similar to Theorem 3.

Theorem 4. *Assume that f is differentiable and*

$$E(Y_3 - X_{3:3} | X_{3:3} = x) = \eta, \quad a < x < b, \tag{19}$$

where $\eta > 0$ *is independent of x. Also assume that condition A holds with* $\beta = 7/(6\eta)$. *Then* $a > -\infty$, $b = \infty$ *and* $F(x) = 1 - \mathrm{e}^{-(7/(6\eta))(x-a)}$, $a < x < \infty$.

In the rest of this section, we investigate some characterizations based on two conditional expectations $E(\phi(Y_{l_1}) | X_{n:n} = x)$ and $E(\phi(Y_{l_2}) | X_{n:n} = x)$, $a < x < b$, where $l_1 > l_2 \geqslant n \geqslant 1$ and ϕ is a continuous function. In view of Lemma 2, it can be seen that for $k, l \geqslant 1$ and $a < x < b$, the conditional distributions of $Y_l | X_{1:1} = x$, $Y_l | Y_1 = x$, and $Y_{k+l-1} | Y_k = x$ are identical. Hence, using our results with $n = 1$, the corresponding characterizations based on the conditional expectations $E(\phi(Y_{k+l_1-1}) | Y_k = x)$ and $E(\phi(Y_{k+l_2-1}) | Y_k = x)$, $a < x < b$, can be obtained.

Theorem 5. *Assume that for some positive integers n and i*,

$$g_{n,i}(x) = E(\phi(Y_{n+i}) | X_{n:n} = x), \quad a < x < b, \tag{20}$$

and

$$g_{n,i-1}(x) = E(\phi(Y_{n+i-1}) | X_{n:n} = x), \quad a < x < b, \tag{21}$$

where $g_{n,i}$ *is differentiable and* $g_{n,i-1}$, ϕ *are continuous functions with* $g_{n,i}(x) \neq g_{n,i-1}(x)$, *a.e. Then*

$$F(x) = 1 - \exp\left\{ - \int_a^x \frac{\mathrm{d}g_{n,i}(y)}{g_{n,i}(y) - g_{n,i-1}(y)} \right\}, \quad a < x < b. \tag{22}$$

Proof. From (20), it follows that for $a < x < b$,

$$g_{n,i}(x)(1 - F(x)) = \int_x^b \phi(y) \sum_{k=1}^{n} \frac{p_{n,k}}{(n+i-k-1)!} (R(y) - R(x))^{n+i-k-1} f(y)\, \mathrm{d}y. \tag{23}$$

Differentiating both sides of (23) with respect to x and using (21) yields

$$(1 - F(x)) g'_{n,i}(x) = (g_{n,i}(x) - g_{n,i-1}(x)) f(x), \quad a < x < b,$$

and (22) follows. □

In Theorem 5, if $n = i = 1$, condition (21) yields $g_{1,0}(x) = \phi(x)$ and condition (20) is equivalent to

$$g_{1,1}(x) = E(\phi(Y_{1+l}) | Y_l = x), \quad a < x < b, \tag{24}$$

where $l \geqslant 1$. Hence we can obtain the result reported by Gupta and Ahsanullah (2004), that is, F can be uniquely determined by (24). Some other characterizations based on conditional expectations of two adjacent record values can also be found in Nagaraja (1977, 1988) and Franco and Ruiz (1996).

In the following, we describe some applications of Theorem 5. Note that Corollary 1 with $n = i = 1$ can deduce the result of Nagaraja (1977).

Corollary 1. *Assume that for some positive integers n and i*,

$$E(Y_{n+i} | X_{n:n} = x) = c_1 x + d_1, \quad a < x < b, \tag{25}$$

and

$$E(Y_{n+i-1} | X_{n:n} = x) = c_2 x + d_2, \quad a < x < b, \tag{26}$$

where c_1, c_2, d_1 and d_2 are independent of x. Then only the following three cases are possible:

(i) $c_1 = c_2 = 1, a > -\infty, b = \infty$ *and* $F(x) = 1 - \mathrm{e}^{-\lambda(x-a)}, a < x < \infty$, *where* $\lambda = 1/(d_1 - d_2) > 0$,
(ii) $0 < c_1 < c_2 \leqslant 1, a > -\infty, b = (d_1 - d_2)/(c_2 - c_1)$ *and* $F(x) = 1 - ((b - x)/(b - a))^{\theta}, a < x < b$, *where* $\theta = c_1/(c_2 - c_1)$,
(iii) $c_1 > c_2 \geqslant 1, a > (d_1 - d_2)/(c_2 - c_1), b = \infty$ *and* $F(x) = 1 - ((x + \delta)/(a + \delta))^{\theta}, a < x < \infty$, *where* $\theta = c_1/(c_2 - c_1)$ *and* $\delta = (d_1 - d_2)/(c_1 - c_2)$.

Proof. For $n \geqslant 1, l \geqslant 1$ and $a < x < b$, it is obtained that

$$E(Y_{n+l}|X_{n:n} = x) = \int_x^b y \sum_{k=1}^{n} \frac{p_{n,k}}{(n+l-k-1)!}(R(y) - R(x))^{n+l-k-1} \frac{f(y)}{1 - F(x)}\,\mathrm{d}y$$
$$= \int_0^{\infty} R^{-1}(R(x) + u) \sum_{k=1}^{n} \frac{p_{n,k}}{(n+l-k-1)!} u^{n+l-k-1} \mathrm{e}^{-u}\,\mathrm{d}u, \tag{27}$$

where the transformation $u = R(y) - R(x)$ is used. (Note that $R'(y) = r(y) > 0, a < y < b$, then $R(y)$ is strictly increasing in (a, b) and hence the inverse function $R^{-1}(y), 0 < y < \infty$, exists.) From (27), it can be seen that $E(Y_{n+l}|X_{n:n} = x)$ is a strictly increasing function of x over (a, b). Similarly, it can also be shown that $E(Y_n|X_{n:n} = x)$ is strictly increasing in (a, b). Consequently, in view of (25) and (26), it follows $c_1 > 0$ and $c_2 > 0$.

Using Theorem 5, (25) and (26) imply

$$F(x) = 1 - \exp\left\{-\int_a^x \frac{c_1}{(c_1 - c_2)y + d_1 - d_2}\,\mathrm{d}y\right\}, \quad a < x < b. \tag{28}$$

First, consider the case $c_1 = c_2$. From (28) and the fact that $F(x)$ is non-decreasing, $0 < F(x) < 1$, $a < x < b$, and $\lim_{x \uparrow b} F(x) = 1$, it turns out that $d_1 > d_2, a > -\infty, b = \infty$ and

$$F(x) = 1 - \mathrm{e}^{-\lambda(x-a)}, \quad a < x < \infty, \tag{29}$$

where $\lambda = c_1/(d_1 - d_2)$. An expression for $E(Y_{n+i-1}|X_{n:n} = x)$ can be easily derived for the exponential distribution given in (29), and then from (26), $c_2 = 1$ follows. The proof of assertion (i) is complete.

Next consider the case $c_1 \neq c_2$. From (28), $a > -\infty, a \neq (d_1 - d_2)/(c_2 - c_1)$ and

$$F(x) = 1 - \left|\frac{x + (d_1 - d_2)/(c_1 - c_2)}{a + (d_1 - d_2)/(c_1 - c_2)}\right|^{\theta}, \quad a < x < b, \tag{30}$$

can be obtained, where $\theta = c_1/(c_2 - c_1)$. Assume $c_1 < c_2$. This implies $\theta > 0$. Using (30) and the fact that $\lim_{x \uparrow b} F(x) = 1$ yields $b = (d_1 - d_2)/(c_2 - c_1)$ and

$$F(x) = 1 - \left(\frac{b - x}{b - a}\right)^{\theta}, \quad a < x < b. \tag{31}$$

The formula of $E(Y_{n+i-1}|X_{n:n} = x)$ with respect to a distribution function as in (31) can also be easily calculated, and then from (26), it can be shown that

$$c_2 = \sum_{k=1}^{n} p_{n,k} \left(\frac{\theta}{\theta + 1}\right)^{n+i-1-k} \leqslant \sum_{k=1}^{n} p_{n,k} = 1.$$

The proof of assertion (ii) is finished. On the other hand, suppose $c_1 > c_2$. It turns out $\theta < -1$. The rest of the proof of assertion (iii) is omitted since it is similar to that of assertion (ii). □

Characterizations based on conditional expectations of non-adjacent record values were investigated by Dembińska and Wesołowski (2000), Wu and Lee (2001), Raqab (2002) and Gupta and Ahsanullah (2004). In particular, Gupta and

Ahsanullah (2004) proved that F is uniquely determined by $g(x)=E(\phi(Y_{2+l})|Y_l=x)$, $a<x<b$, where $l\geqslant 1$, and this result can also be obtained from the following theorem with $n=1$ and $i=2$.

Theorem 6. *Assume that f is continuously differentiable and for some integers* $n\geqslant 1$ *and* $i\geqslant 2$,

$$g_{n,i}(x)=E(\phi(Y_{n+i})|X_{n:n}=x), \quad a<x<b, \tag{32}$$

and

$$g_{n,i-2}(x)=E(\phi(Y_{n+i-2})|X_{n:n}=x), \quad a<x<b, \tag{33}$$

where $g_{n,i}$ *is twice continuously differentiable and* $g_{n,i-2}$, ϕ *are continuous functions. Then F can be uniquely determined.*

Proof. From (32), we have for $a<x<b$,

$$g_{n,i}(x)(1-F(x))=\int_x^b \phi(y)\sum_{k=1}^{n}\frac{p_{n,k}}{(n+i-k-1)!}(R(y)-R(x))^{n+i-k-1}f(y)\,\mathrm{d}y. \tag{34}$$

After differentiating both sides of (34) with respect to x twice and some manipulations, using (33), it yields for $a<x<b$,

$$g'_{n,i}(x)\frac{f'(x)}{f(x)}+3g'_{n,i}(x)r(x)-g''_{n,i}(x)-(g_{n,i}(x)-g_{n,i-2}(x))(r(x))^2=0. \tag{35}$$

It can be seen that (35) is essentially the same as (4.7) of Gupta and Ahsanullah (2004). Therefore, the proof is completed along the lines of the proof in Gupta and Ahsanullah (2004). □

4. Further characterizations based on record values and $X_{n:n}$

In this section, for $l,n\geqslant 1$, let $S_l=Y_l-X_{n:n}$, $S_l^-=\min\{0,S_l\}$ and $S_l^+=\max\{0,S_l\}$. It is clear that for $l\geqslant n$, $S_l^-=0$ and $S_l^+=S_l$. Balakrishnan and Stepanov (2004) recently gave characterizations of distributions based on the independence of S_2^- and $X_{n:n}$, S_2^+ and $X_{n:n}$, or S_2 and $X_{n:n}$. We now characterize F using the independence of S_l^- and $X_{n:n}$, where $l<n$.

Theorem 7. S_l^- *and* $X_{n:n}$ *are independent for some positive integers* $l<n$ *if and only if* $F(x)=\mathrm{e}^{\beta(x-b)}$, $-\infty<x<b$, *where* $\beta>0$ *and b are constants.*

Proof. The sufficiency part can be verified directly, we only prove the necessity part. For $l<n$ and $a<x<b$, the conditional distribution of S_l^- given $X_{n:n}=x$ can be derived from Theorem 1 as

$$P(S_l^-\leqslant y|X_{n:n}=x)=\begin{cases}\sum_{k=l}^{n-1}p_{k-1,l-1}\left(\frac{1}{k}-\frac{1}{n}\right)\left(\frac{F(y+x)}{F(x)}\right)^k, & y<0,\\ 1, & y\geqslant 0.\end{cases} \tag{36}$$

The independence of S_l^- and $X_{n:n}$ implies that the right side of (36) does not depend on x, i.e.

$$\sum_{k=l}^{n-1}c_k\left(\frac{F(y+x)}{F(x)}\right)^k=H(y), \quad y<0,\ a<x<b,$$

where $c_k=p_{k-1,l-1}(1/k-1/n)>0$ and $H(y)$ is some continuous function. Following the lines of the proof of Theorem 3.2 in Balakrishnan and Stepanov (2004) yields the necessity part. □

In the above theorem, if $l=n-1$, the independence condition can then be weakened. More precisely, as indicated by the following theorem with $r=0$, independence of x for $E(S_{n-1}^-|X_{n:n}=x)$ is sufficient.

Theorem 8. *Assume that for some integers* $n \geqslant 2$ *and* $r \geqslant 0$,

$$E((S_{n-1}^-)^{r+1}|X_{n:n}=x)=\eta E((S_{n-1}^-)^r|X_{n:n}=x), \quad a<x<b, \tag{37}$$

where η *is independent of* x. *Then* $\eta<0$, $a=-\infty$, $b<\infty$ *and* $F(x)=\mathrm{e}^{\beta(x-b)}$, $-\infty<x<b$, *where* $\beta=-1/(n!\eta(n-1))$ *for* $r=0$, *and* $\beta=-(r+1)/(\eta(n-1))$ *for* $r\geqslant 1$.

Proof. First, consider the case $r=0$. From (37), we have

$$\int_a^x (y-x)(F(y))^{n-2} f(y)\,\mathrm{d}y=\frac{n!\eta}{n-1}(F(x))^{n-1}, \quad a<x<b. \tag{38}$$

It is easy to see from (38) that $\eta<0$. Differentiating both sides of (38) with respect to x yields

$$f(x)=\beta F(x), \quad a<x<b, \tag{39}$$

where $\beta=-1/(n!\eta(n-1))$. This implies $F(x)=c\mathrm{e}^{\beta x}$, $a<x<b$, where $c>0$ is a constant. Using the fact that $\lim_{x\downarrow a}F(x)=0$ and $\lim_{x\uparrow b}F(x)=1$, it then turns out $a=-\infty$, $b<\infty$ and $c=\mathrm{e}^{-\beta b}$. The assertion for $r=0$ follows.

Next, consider the case $r\geqslant 1$. From (37), the following is obtained:

$$\int_a^x (y-x)^{r+1}(F(y))^{n-2} f(y)\,\mathrm{d}y=\eta\int_a^x (y-x)^{r}(F(y))^{n-2} f(y)\,\mathrm{d}y, \quad a<x<b. \tag{40}$$

We observe $\eta<0$ from (40). Differentiate both sides of (40) with respect to x $(r+1)$ times to arrive at (39) with $\beta=-(r+1)/(\eta(n-1))$. The assertion for $r\geqslant 1$ can then be obtained immediately. □

When $F(x)=\mathrm{e}^{\beta(x-b)}$, $-\infty<x<b$, where $\beta>0$ and $b<\infty$, one can deduce that

$$E((S_{n-1}^-)^2|X_{n:n}=x)=2(n!)(E(S_{n-1}^-|X_{n:n}=x))^2, \quad -\infty<x<b. \tag{41}$$

Inspired by this, we give the following characterization theorem, which also indicates that if $Var(S_{n-1}^-|X_{n:n}=x)$ is proportional to $(E(S_{n-1}^-|X_{n:n}=x))^2$, then F can be determined.

Theorem 9. *Assume that* f *is differentiable and for some integer* $n\geqslant 2$,

$$E((S_{n-1}^-)^2|X_{n:n}=x)=\eta(E(S_{n-1}^-|X_{n:n}=x))^2, \quad a<x<b, \tag{42}$$

where η *is independent of* x. *Then only the following three cases are possible*:

(i) $\eta=2(n!)$, $a=-\infty$, $b<\infty$ *and* $F(x)=\mathrm{e}^{\beta(x-b)}$, $-\infty<x<b$, *where* $\beta>0$ *is a constant*,
(ii) $n!<\eta<2(n!)$, $-\infty<a<b<\infty$ *and* $F(x)=((x-a)/(b-a))^\theta$, $a<x<b$, *where* $\theta=2(\eta-n!)/((n-1)(2(n!)-\eta))$,
(iii) $\eta>2(n!)$, $a=-\infty$, $b<\infty$ *and* $F(x)=(1+\gamma(b-x))^\theta$, $-\infty<x<b$, *where* $\theta=2(\eta-n!)/((n-1)(2(n!)-\eta))$ *and* $\gamma>0$ *is a constant*.

Proof. We know that (42) implies

$$(F(x))^{n-1}\int_a^x (y-x)^2(F(y))^{n-2} f(y)\,\mathrm{d}y=\frac{\eta(n-1)}{n!}\left(\int_a^x (y-x)(F(y))^{n-2} f(y)\,\mathrm{d}y\right)^2, \quad a<x<b. \tag{43}$$

Differentiating both sides of (43) with respect to x yields

$$f(x)\int_a^x (y-x)^2(F(y))^{n-2} f(y)\,\mathrm{d}y=\frac{2(n!-\eta)}{n!(n-1)}F(x)\int_a^x (y-x)(F(y))^{n-2} f(y)\,\mathrm{d}y, \quad a<x<b. \tag{44}$$

In view of (43) and (44), we obtain $\eta>n!$ and

$$\int_a^x (y-x)(F(y))^{n-2} f(y)\,\mathrm{d}y=\frac{2(n!-\eta)}{\eta(n-1)^2}\frac{(F(x))^n}{f(x)}, \quad a<x<b. \tag{45}$$

Again differentiating both sides of (45) with respect to x yields

$$\frac{f'(x)}{f(x)} = \left(n + \frac{\eta(n-1)}{2(n! - \eta)}\right)\frac{f(x)}{F(x)}, \quad a < x < b,$$

and this implies

$$f(x) = \beta(F(x))^{n+\eta(n-1)/(2(n!-\eta))}, \quad a < x < b, \tag{46}$$

where $\beta > 0$ is a constant.

First, consider the case $\eta = 2(n!)$. Then solving (46) yields $F(x) = c\mathrm{e}^{\beta x}$, $a < x < b$, where $c > 0$ is a constant. As $F(x)$ is a distribution function, it turns out that $a = -\infty$, $b < \infty$ and $c = \mathrm{e}^{-\beta b}$. The proof of assertion (i) is complete.

Next, consider the case $\eta \neq 2(n!)$. The general solution of (46) is

$$F(x) = (c_1 x + c_2)^{\theta}, \quad a < x < b, \tag{47}$$

where $\theta = 2(\eta - n!)/((n-1)(2(n!) - \eta))$ and $c_1 = \beta/\theta$, c_2 are constants. Assume that $n! < \eta < 2(n!)$. This, in turn, implies that $\theta > 0$ and $c_1 > 0$. Again as $F(x)$ is a distribution function, from (47), it can be seen that $a = -c_2/c_1$ and $b = (1 - c_2)/c_1$. On the other hand, suppose $\eta > 2(n!)$. Then $\theta < 0$ and $c_1 < 0$ follows. From (47), it can be shown that $a = -\infty$ and $b = (1 - c_2)/c_1$. The remaining assertions (ii) and (iii) can be obtained immediately. □

Acknowledgements

The authors are grateful to the referees for their many helpful remarks and suggestions.

References

Arnold, B.C., Balakrishnan, N., Nagaraja, H.N., 1992. A First Course in Order Statistics. Wiley, New York.
Arnold, B.C., Balakrishnan, N., Nagaraja, H.N., 1998. Records. Wiley, New York.
Balakrishnan, N., Stepanov, A., 2004. Two characterizations based on order statistics and records. J. Statist. Plann. Inference 124, 273–287.
David, H.A., Nagaraja, H.N., 2003. Order Statistics. Wiley, New York.
Deheuvels, P., 1984. The characterization of distributions by order statistics and record values—a unified approach. J. Appl. Probab. 21, 326–334.
Dembińska, A., Wesołowski, J., 2000. Linearity of regression for non-adjacent record values. J. Statist. Plann. Inference 90, 195–205.
Feller, W., 1970. An Introduction to Probability Theory and its Applications, vol. II. Wiley, New York.
Franco, M., Ruiz, J.M., 1996. On characterization of continuous distributions by conditional expectation of record values. Sankhyā A 58, 135–141.
Gupta, R.C., 1984. Relationships between order statistics and record values and some characterization results. J. Appl. Probab. 21, 425–430.
Gupta, R.C., Ahsanullah, M., 2004. Some characterization results based on the conditional expectation of a function of non-adjacent order statistic (record value). Ann. Inst. Statist. Math. 56, 721–732.
Huang, W.J., Su, J.C., 1999. On certain problems involving order statistics—a unified approach through order statistics property of point processes. Sankhyā A 61, 36–49.
Huang, W.J., Su, N.C., Su, J.C., 2007. Characterizations of the order statistics point process by the relations between its conditional moments. J. Chinese Statist. Assoc. 45, 38–54.
Nagaraja, H.N., 1977. On a characterization based on record values. Austral. J. Statist. 20, 176–182.
Nagaraja, H.N., 1988. Some characterizations of continuous distributions based on regressions of adjacent order statistics and record values. Sankhyā A 50, 70–73.
Nagaraja, H.N., Nevzorov, V.B., 1997. On characterizations based on record values and order statistics. J. Statist. Plann. Inference 63, 271–284.
Nevzorov, V.B., 2001. Records: Mathematical Theory. American Mathematical Society, Providence, RI.
Rao, C.R., Shanbhag, D.N., 1994. Choquet–Deny Type Functional Equations with Applications to Stochastic Models. Wiley, New York.
Raqab, M., 2002. Characterizations of distributions based on the conditional expectations of record values. Statist. Decision 20, 309–319.
Wu, J., Lee, W., 2001. On the characterization of generalized extreme value, power function, generalized Pareto and classical Pareto distributions by conditional expectation of record values. Statist. Papers 42, 225–242.