Ann. Inst. Statist. Math.
Vol. 56, No. 4, 721–732 (2004)

SOME CHARACTERIZATION RESULTS BASED ON THE CONDITIONAL EXPECTATION OF A FUNCTION OF NON-ADJACENT ORDER STATISTIC (RECORD VALUE)

RAMESH C. GUPTA[1] AND MOHAMMAD AHSANULLAH[2]

[1]*Department of Mathematics and Statistics, University of Maine, Orono, ME 04469-5752, U.S.A.*
[2]*Department of Management Sciences, Rider University, Lawrenceville, NJ 08648-3099, U.S.A.*

(Received December 11, 2002; revised December 2, 2003)

Abstract. In this paper, we attempt to characterize a distribution by means of $E[\psi(X_{k+s:n}) \mid X_{k:n} = z] = g(z)$, under some mild conditions on $\psi(\cdot)$ and $g(\cdot)$. An explicit result is provided in the case of $s = 1$ and a uniqueness result is proved in the case of $s = 2$. For the general case, an expression is provided for the conditional expectation. Similar results are proved for the record values, both in the continuous as well as in the discrete case (weak records).

Key words and phrases: Adjacent order statistic, failure rate, uniqueness theorem, continuous and discrete record values.

1. Introduction

Let $X_1, X_2, \ldots, X_n$ be a random sample of size n from an absolutely continuous distribution function $F(\cdot)$ and probability density function $f(\cdot)$. Let $X_{1:n} < X_{2:n} < \cdots < X_{n:n}$ be the corresponding order statistics.

There is a vast literature on characterizing a distribution by means of the conditional expectation of $X_{k+1:n}$ (or its function) given $X_{k:n}$. More specifically, Franco and Ruiz (1999) characterized a distribution by means of

$$E[\psi(X_{k+1:n}) \mid X_{k:n} = z] = g(z),$$

under some mild conditions on $\psi(\cdot)$ and $g(\cdot)$. This general result contains some special cases. For example, Khan and Abu-Salih (1989) characterized the distribution when $g(x) = c\psi(x) + d$. Ouyang (1995) considered the case when $g(x) = h(x) + c$, where $h(\cdot)$ is differentiable and its derivative is continuous. Historically, Ferguson (1967) characterized the distribution when $\psi(x) = x$ and $g(x) = ax + b$.

The problem of characterizing a distribution by the conditional expectation of non-adjacent order statistics is rather complex. The general problem is to characterize a distribution by means of

$$E[\psi(X_{k+s:n}) \mid X_{k:n} = z] = g(z),$$

under some appropriate conditions on $\psi(\cdot)$ and $g(\cdot)$. Attempts in this direction have been made by several authors who could provide solution only in some special cases. For example Dembinska and Wesolowski (1998) characterized the distribution by means of the equation

$$E[X_{k+s:n} \mid X_{k:n} = z] = az + b.$$

They used a result of Rao and Shanbhag (1994) which deals with solutions of extended version of integrated Cauchy functional equation. The same result was proved by Lopez-Blaquez and Moreno Rebollo (1997) by using the solution of a polynomial equation, see also Franco and Ruiz (1997). A special case when $s = 2$ was considered by Wesolowski and Ahsanullah (1997). It may be remarked that Rao and Shanbhag's (1994) result is applicable only when the conditional expectation is a linear function of $X_{k:n}$. However, the following result can be easily established, following the steps of Dembinska and Wesolowski (1998).

The distribution can be characterized by means of the equation

$$E[\psi(X_{k+s:n}) \mid X_{k:n} = z] = \psi(z) + b.$$

Some other papers dealing with the characterization of distributions based on non-adjacent order statistics include Wu and Ouyang (1996, 1997), Wu (2004) and Franco and Ruiz (1997).

The present paper makes an attempt to solve the general problem described earlier. More specifically, we want to characterize a distribution by means of

$$E[\psi(X_{k+s:n}) \mid X_{k:n} = z] = g(z).$$

For this purpose, we derive an expression for $E[\psi(X_{k+s:n}) \mid X_{k:n} = z]$. For $s = 1$, we are able to obtain the expression in closed form. This is a generalization of Ouyang (1995) and unifies several results, available in the literature, for adjacent order statistic. For $s = 2$, we are able to show that the above relation determines the distribution uniquely. Before proceeding further, we present, in Section 2, an expression for $E[\psi(X_{k+s:n}) \mid X_{k:n} = z]$.

Because of the relationship between the order statistics and the record values, see Gupta (1984), one would expect similar characterization results based on the record values. The record values are defined as follows:

Let $X_1, X_2, \ldots$ be a sequence of independent identically distributed random variables with continuous distribution function $F(\cdot)$. Let us define $U(1) = 1$ and for $n > 1$

$$U(n) = \min\{k > U(n-1) : X_k > X_{U(n-1)}\}.$$

The upper record value sequence is then defined by

$$R_n = X_{U(n)}, \qquad n = 1, 2, \ldots.$$

In this connection, Nagaraja (1977, 1988) obtained a characterization result based on the linear regression of two adjacent record values. More specifically, he characterized the distribution based on the property

$$E(R_{k+1} \mid R_k = z) = az + b.$$

Other characterizations based on conditional expectation of non-adjacent record values can be seen in Raqab (2002), Wu (2004) and Wu and Lee (2001). Lopez-Blaquez and Moreno Rebollo (1997) considered this problem for non-adjacent record values under some stringent smoothness assumptions on the distribution function $F(\cdot)$. Dembinska and Wesolowski (2000) characterized the distribution by means of the relation

$$E(R_{k+s} \mid R_k = z) = az + b.$$

They used a result of Rao and Shanbhag (1994) which deals with the solution of extended version of integrated Cauchy functional equation. As pointed out earlier, Rao and Shanbhag's result is applicable only when the conditional expectation is a linear function. However, the following result can be established immitating the steps of Dembinska and Wesolowski (2000), see also Ahsanullah and Wesolowski (1998).

The distribution can be characterized by means of the equation

$$E(\psi(R_{k+s}) \mid R_k = z) = \psi(z) + b.$$

In Section 4, we shall attempt to characterize the distribution by means of the relation

$$E(\psi(R_{k+s}) \mid R_k = z) = g(z), \qquad k, s \geq 1.$$

The cases of $s = 1$ and $s = 2$ are solved, while for other cases, the problem becomes complicated because of the nature of the resulting differential equation. Section 5 contains similar results for the discrete case. Some other papers for the discrete case include Franco and Ruiz (2001), Lopez-Blaquez and Wesolowski (2001) and Wesolowski and Ahsanullah (2001).

It may be remarked that probably the results of this paper can be extended for mixtures of distributions.

2. Conditional expectation

Noting that the distribution of $X_{k+s:n} \mid X_{k:n}$ is the distribution of the s-th order statistic in a sample of size $n-k$ from a truncated distribution given by

$$G(x) = \frac{F(x) - F(z)}{\overline{F}(z)}, \qquad x > z,$$

where $\overline{F}(z) = 1 - F(z)$, we get the pdf of $X_{k+s:n}$ given $X_{k:n} = z$ as

$$= \frac{(n-k)!}{(s-1)!(n-k-s)!} \left[\frac{F(x) - F(z)}{\overline{F}(z)}\right]^{s-1} \left[\frac{\overline{F}(x)}{\overline{F}(z)}\right]^{n-k-s} \frac{f(x)}{\overline{F}(z)}.$$

This gives

$$\begin{aligned} (2.1) \qquad & E[\psi(X_{k+s:n}) \mid X_{k:n} = z] \\ & = \frac{(n-k)!}{(s-1)!(n-k-s)!} \frac{\int_z^\infty \psi(x)[\overline{F}(z) - \overline{F}(x)]^{s-1}[\overline{F}(x)]^{n-k-s} f(x)dx}{[\overline{F}(z)]^{n-k}}. \end{aligned}$$

To simplify (2.1) further, we proceed as follows:
Define

$$I_j = \frac{(n-k)!}{(s-j)!(n-k-s+j-1)!} \int_z^\infty \psi(x)[\overline{F}(z) - \overline{F}(x)]^{s-j}[\overline{F}(x)]^{n-k-s+j-1} f(x)dx.$$

This gives

$$\begin{aligned} (2.2) \qquad I_s &= \frac{(n-k)!}{(n-k-1)!} \int_z^\infty \psi(x)[\overline{F}(x)]^{n-k-1} f(x)dx \\ &= \psi(z)[\overline{F}(z)]^{n-k} + \int_z^\infty \psi'(x)[\overline{F}(x)]^{n-k} dx. \end{aligned}$$

We now assume that $j < s$. In that case

$$(2.3)\qquad \begin{aligned} I_j &= \frac{(n-k)!}{(s-j)!(n-k-s+j-1)!} \\ &\quad\times \int_z^\infty \psi(x)[\overline{F}(z)-\overline{F}(x)]^{s-j}\left(\frac{-d(\overline{F}(x))^{n-k-s+j}}{n-k-s+j}\right) \\ &= \frac{(n-k)!}{(s-j)!(n-k-s+j)!} \\ &\quad\times \Bigg[\psi(x)[\overline{F}(z)-\overline{F}(x)]^{s-j}\{-(\overline{F}(x))^{n-k-s+j}\}\,|_z^\infty \\ &\qquad + \int_z^\infty [\overline{F}(x)]^{n-k-s+j}\frac{d}{dx}\{\psi(x)[\overline{F}(z)-\overline{F}(x)]^{s-j}\}\Bigg] \\ &= \frac{(n-k)!}{(s-j)!(n-k-s+j)!} \\ &\quad\times \int_z^\infty [\overline{F}(x)]^{n-k-s+j}[\psi'(x)[\overline{F}(z)-\overline{F}(x)]^{s-j} \\ &\qquad\qquad + \psi(x)(s-j)[\overline{F}(z)-\overline{F}(x)]^{s-j+1}f(x)]dx \\ &= \frac{(n-k)!}{(s-j)!(n-k-s+j)!}\int_z^\infty [\overline{F}(x)]^{n-k-s+j}\psi'(x)[\overline{F}(z)-\overline{F}(x)]^{s-j}dx \\ &\quad + \frac{(n-k)!}{(s-j-1)!(n-k-s+j)!} \\ &\quad\times \int_z^\infty \psi(x)[\overline{F}(x)]^{n-k-s+j}[\overline{F}(z)-\overline{F}(x)]^{s-j-1}f(x)dx \\ &= \frac{(n-k)!}{(s-j)!(n-k-s+j)!} \\ &\quad\times \int_z^\infty [\overline{F}(x)]^{n-k-s+j}\psi'(x)[\overline{F}(z)-\overline{F}(x)]^{s-j}dx + I_{j+1}. \end{aligned}$$

This is a recurrence relation which will be helpful in deriving the desired expression. Noting that $E[\psi(X_{k+s:n}) \mid X_{k:n} = z] = I_1/[\overline{F}(z)]^{n-k}$, we have

$$\begin{aligned} &E[\psi(X_{k+s:n}) \mid X_{k:n} = z] \\ &\quad = \frac{(n-k)!}{(s-1)!(n-k-s+1)![\overline{F}(z)]^{n-k}} \\ &\qquad \times \int_z^\infty [\overline{F}(x)]^{n-k-s+1}\psi'(x)[\overline{F}(z)-\overline{F}(x)]^{s-1}dx + I_2. \end{aligned}$$

Using the recurrence relation (2.3) repeatedly, we can write

$$(2.4)\qquad \begin{aligned} &E[\psi(X_{k+s:n}) \mid X_{k:n} = z] \\ &\quad = \frac{1}{[\overline{F}(z)]^{n-k}}\sum_{i=1}^{s-1}\frac{(n-k)!}{i!(n-k-i)!}\int_z^\infty [\overline{F}(x)]^{n-k-i}\psi'(x)[\overline{F}(z)-\overline{F}(x)]^{i}dx \\ &\qquad + \psi(z) + \frac{1}{[\overline{F}(z)]^{n-k}}\int_z^\infty \psi'(x)[\overline{F}(x)]^{n-k}dx. \end{aligned}$$

3. The characterization results

By using (2.4), the general result

$$E[\psi(X_{k+s:n}) \mid X_{k:n} = z] = g(z)$$

becomes

$$\begin{aligned}\sum_{i=1}^{s-1} \frac{(n-k)!}{i!(n-k-i)!} \int_z^\infty [\overline{F}(x)]^{n-k-i}\psi'(x)[\overline{F}(z) - \overline{F}(x)]^i dx \\ + \int_z^\infty \psi'(x)[\overline{F}(x)]^{n-k} dx \\ = (g(z) - \psi(z))[\overline{F}(z)]^{n-k}\end{aligned} \tag{3.1}$$

or

$$\begin{aligned}&\sum_{i=1}^{s-1} \frac{(n-k)!}{i!(n-k-i)!} \sum_{j=0}^{i} [\overline{F}(z)]^j \int_z^\infty [\overline{F}(x)]^{n-k-j}\psi'(x)(-1)^{i-j} dx + \int_z^\infty \psi'(x)[\overline{F}(x)]^{n-k} dx \\ &= (g(z) - \psi(z))[\overline{F}(z)]^{n-k}.\end{aligned}$$

Differentiating the above w.r.t. z, we get

$$\begin{aligned}&\sum_{i=1}^{s-1} \frac{(n-k)!}{i!(n-k-i)!} \sum_{j=0}^{i} \Big[(-1)^{i-j-1}[\overline{F}(z)]^{n-k}\psi'(z) \\ &\qquad - j[\overline{F}(z)]^{j-1} f(z) \int_z^\infty (-1)^{i-j}[\overline{F}(x)]^{n-k-j}\psi'(x) dx\Big] \\ &\quad - [\overline{F}(z)]^{n-k}\psi'(z) \\ &= (g'(z) - \psi'(z))[\overline{F}(z)]^{n-k} - (g(z) - \psi(z))(n-k)[\overline{F}(z)]^{n-k-1} f(z).\end{aligned} \tag{3.2}$$

We now prove two characterization results.

THEOREM 3.1. *Under the assumptions stated earlier* $E[\psi(X_{k+1:n}) \mid X_{k:n} = z] =$ $g(z)$ *determines the distribution.*

PROOF. The proof can be established by using equation (3.1), for $s = 1$, and expressing $r(z)$ in terms of $g(z)$ as

$$r(z) = \frac{g'(z)}{(n-k)(g(z) - \psi(z))}, \tag{3.3}$$

where $r(z) = f(z)/\overline{F}(z)$ is the failure rate.

Thus $g(z)$ determines the distribution. □

Note that for the case $g(z) = \psi(z) + c$, see Ouyang (1995), we get

$$r(z) = \frac{\psi'(z)}{c(n-k)}. \tag{3.4}$$

For another expression, see Franco and Ruiz (1995, 1999) and Khan and Abu-Salih (1989).

Before presenting the next result, we state the following uniqueness theorem and its corollary whose proof can be found in Gupta and Kirmani (2004).

THEOREM 3.2. *Let the function f be defined and continuous in a domain $D \subset R^2$, and let f satisfy a Lipschitz condition (with respect to y) in D, namely*

$$|f(x,y_1) - f(x,y_2)| \le K|y_1 - y_2|, \qquad K > 0, \tag{3.5}$$

for every point (x, y_1) and (x, y_2) in D. Then the function $y = \phi(x)$ satisfying the initial value problem $y' = f(x,y)$ and $y(x_0) = y_0$, $x \in R$ is unique.

COROLLARY 3.1. *If f and $\frac{\partial f}{\partial y}$ are continuous in D, then the solution $y = \phi(x)$ is unique in R^2.*

We are now able to present our second characterization result.

THEOREM 3.3. *Under the assumptions stated earlier $E[\psi(X_{k+2:n}) \mid X_{k:n} = z] = g(z)$ determines the distribution.*

PROOF. For this case, (3.1) becomes

$$\begin{aligned} &-(n-k)f(z)\int_z^\infty [\overline{F}(x)]^{n-k-1}\psi'(x)dx \\ &\quad = g'(z)[\overline{F}(z)]^{n-k} - (g(z)-\psi(z))(n-k)[\overline{F}(z)]^{n-k-1}f(z). \end{aligned} \tag{3.6}$$

Differentiating (3.6) once more w.r.t. z, we get

$$\begin{aligned} f'(z)&\left[\frac{g'(z)}{f(z)}[\overline{F}(z)]^{n-k} - (n-k)[\overline{F}(z)]^{n-k-1}(g(z)-\psi(z))\right] \\ &\quad + (n-k)f(z)[\overline{F}(z)]^{n-k-1}\psi'(z) \\ &= g''(z)[\overline{F}(z)]^{n-k} - g'(z)(n-k)[\overline{F}(z)]^{n-k-1}f(z) \\ &\quad - (n-k)\{f'(z)(g(z)-\psi(z))[\overline{F}(z)]^{n-k-1} + f(z)(g'(z)-\psi'(z))[\overline{F}(z)]^{n-k-1} \\ &\qquad - f^2(z)(g(z)-\psi(z))(n-k-1)[\overline{F}(z)]^{n-k-2}\}. \end{aligned}$$

Simplifying the above equation, we get

$$\begin{aligned} \frac{f'(z)}{f(z)}g'(z) &= g''(z) - 2(n-k)g'(z)r(z) \\ &\quad + (n-k)(n-k-1)r^2(z)(g(z)-\psi(z)). \end{aligned} \tag{3.7}$$

Using the fact that

$$\frac{f'(z)}{f(z)} = \frac{r'(z)}{r(z)} - r(z),$$

equation (3.7) can be written as

$$g'(z)\left[\frac{r'(z)}{r(z)} - r(z)\right] = g''(z) - 2(n-k)g'(z)r(z) + (n-k)(n-k-1)r^2(z)(g(z)-\psi(z))$$

or

$$g'(z)\frac{r'(z)}{r(z)}+(2n-2k-1)g'(z)r(z) = g''(z)+(n-k)(n-k-1)r^2(z)(g(z)-\psi(z)). \tag{3.8}$$

This expresses $r'(z)$ as a function of $r(z)$ and the known functions. By the corollary of the uniqueness theorem, stated above, the above equation has a unique solution in $r(z)$. □

4. Record values characterization results

In this section, we shall characterize the distribution by means of conditional expectation of record values.

As explained in the introduction, let $\{R_j, j = 1, 2, \ldots\}$ be the sequence of upper record values. Then we have the following result.

THEOREM 4.1. *Let $\{X_j, j = 1, 2, 3, \ldots\}$ be a sequence of independent identically distributed random variables with absolutely continuous (w.r.t. Lebesgue measure) distribution function $F(x)$ and probability density function $f(x)$. Let $\{R_j, j = 1, 2, 3, \ldots\}$ be a sequence of upper record values. Then the condition*

$$E[\psi(R_{k+s}) \mid R_k = z] = g(z), \tag{4.1}$$

where $k, s \geq 1$, $g(z)$ is twice differentiable and $\psi(x)$ is a continuous function, determines the distribution uniquely.

PROOF. It can be verified that

$$E[\psi(R_{k+s}) \mid R_k = z] = \int_z^\infty \frac{\psi(x)[R(x)-R(z)]^{s-1}}{\overline{F}(z)} d(-\overline{F}(x)), \tag{4.2}$$

where $R(x) = -\ln \overline{F}(x)$.

Case $s = 1$. In this case, using the above two equations, we obtain

$$\int_z^\infty \psi(x)f(x)dx = g(z)\overline{F}(z). \tag{4.3}$$

Differentiating both sides of (4.3) with respect to z and simplifying, we obtain

$$r(z) = \frac{f(z)}{\overline{F}(z)} = \frac{g'(z)}{g(z)-\psi(z)}, \tag{4.4}$$

where $r(z)$ is the failure rate of the original distribution. Hence the result.

Case $s = 2$. In this case, we obtain

$$\int_z^\infty \psi(x)[R(x)-R(z)]f(x)dx = g(z)\overline{F}(z). \tag{4.5}$$

Differentiating both sides of (4.5) with respect to z, we obtain

$$-\int_z^\infty \psi(x)f(x)dx = g'(z)\frac{(\overline{F}(z))^2}{f(z)} - g(z)\overline{F}(z). \tag{4.6}$$

Differentiating both sides of (4.6) with respect to z and simplifying, we obtain

$$\psi(z)(r(z))^2 = g''(z) - 3g'(z)r(z) - g'(z)\frac{f'(z)}{f(z)} + g(z)(r(z))^2. \tag{4.7}$$

Using the relation

$$\frac{f'(z)}{f(z)} = \frac{r'(z)}{r(z)} - r(z), \tag{4.8}$$

equation (4.7) can be written as

$$g'(z)\frac{r'(z)}{r(z)} + 2g'(z)r(z) = g''(z) + (r(z))^2(g(z) - r(z)) = 0. \tag{4.9}$$

Thus $r'(z)$ has been expressed in terms of functions of $r(z)$ and the known functions. Therefore, by the corollary of the uniqueness theorem stated earlier, $r(z)$ is uniquely determined. This completes the proof.

General case. For the general case, the problem becomes more complicated because of the nature of the resulting differential equation. □

5. Characterization through weak records

Let $X_1, X_2, \dots$ be a sequence of independent identically distributed random variables taking values $0, 1, 2, \dots, N$, $N \leq \infty$ with a distribution function F such that $F(n) < 1$ for $n = 0, 1, 2, \dots$ and $E(X_1 \ln(1 + X_1)) < \infty$. Define the sequence of weak record times $V(n)$ and weak record values $X_{V(n)}$ as follows:

$$V(1) = 1, \qquad V(n+1) = \min\{j > V(n) : X_j \geq X_{V(n)}\}, \qquad n = 1, 2, \dots. \tag{5.1}$$

If we replace the sign $\geq$ by $>$ in (5.1), then we obtain record times and record values. Let

$$p_i = P(X_1 = i), \qquad q_i = p_i + p_{i+1} + \cdots + p_N.$$

Then the joint mass function of $X_{V(1)}, X_{V(2)}, \dots, X_{V(n)}$ is given by

$$P(X_{V(1)} = k_1, X_{V(2)} = k_2, \dots, X_{V(n)} = k_n) = \prod_{i=1}^{n-1} \left(\frac{p_{k_i}}{q_{k_i}}\right) p_{k_n}, \tag{5.2}$$

see Aliev (1999) for details.

We now present the following result.

THEOREM 5.1. *Let $X_1, X_2, \dots$ be a sequence of independent identically distributed random variables taking values $0, 1, 2, \dots, N$, $N \leq \infty$ with a distribution function F such that $F(n) < 1$ for $n = 0, 1, 2, \dots$ and $E(X_1 \ln(1 + X_1)) < \infty$. Let $X_{V(n)}$ be a sequence of* *weak record values as stated above. Then the condition* $E[\psi(X_{V(n+1)}) \mid X_{V(n)} = j] = g(j)$ *determines the distribution.*

PROOF. It can be verified that

$$P(X_{V(n+1)} = y \mid X_{V(n)} = x) = \frac{p_y}{q_x}, \qquad y \geq x,$$

and

$$E[\psi(X_{V(n+1)}) \mid X_{V(n)} = j] = \frac{1}{q_j} \sum_{k=j}^{N} \psi(k) p_k. \tag{5.3}$$

The right hand side of the above equation is independent of n, so we can take without loss of generality

$$E[\psi(X_{V(2)}) \mid X_{V(1)} = j] = g(j). \tag{5.4}$$

Then equation (5.3) can be written as

$$g(j) q_j = \sum_{k=j}^{N} \psi(k) p_k. \tag{5.5}$$

Taking the first order differences, we have

$$g(j) q_j - g(j+1) q_{j+1} = \psi(j) p_j. \tag{5.6}$$

Thus

$$g(j) q_j - g(j+1)(q_j - p_j) = \psi(j) p_j$$

or

$$p_j = \frac{g(j+1) - g(j)}{g(j+1) - \psi(j)} q_j. \tag{5.7}$$

Since $q_j = \frac{q_j}{q_{j-1}} \cdot \frac{q_{j-1}}{q_{j-2}} \cdots \frac{q_1}{q_0}$, $q_0 = 1$, we have

$$p_j = \frac{g(j+1) - g(j)}{g(j+1) - \psi(j)} \prod_{k=0}^{j-1} \left(\frac{q_{k+1}}{q_k} \right). \tag{5.8}$$

Also from (5.6),

$$g(j) q_j - g(j+1) q_{j+1} = \psi(j)(q_j - q_{j+1})$$

or

$$\frac{q_{j+1}}{q_j} = \frac{g(j) - \psi(j)}{g(j+1) - \psi(j)}. \tag{5.9}$$

Thus (5.8) can be written as

$$p_j = \frac{g(j+1) - g(j)}{g(j+1) - \psi(j)} \prod_{k=0}^{j-1} \left[\frac{g(k) - \psi(k)}{g(k+1) - \psi(k)} \right]. \tag{5.10}$$ □

We now present the following example.

Example 5.1. Let $g(j) = j + b$, $N = \infty$ and $\psi(j) = j$.
In this case

$$p_j = \frac{1}{1+b}\left(\frac{b}{1+b}\right)^j, \qquad j = 0, 1, 2, \ldots$$

i.e., X's are distributed as geometric.

We now present the following extension.

THEOREM 5.2. *Let $X_1, X_2, \ldots$ be a sequence of independent identically distributed random variables taking values $0, 1, 2, \ldots, N$, $N \leq \infty$ with a distribution function F such that $F(n) < 1$ for $n = 0, 1, 2, \ldots$ and $E(X_1 \ln(1 + X_1)) < \infty$. Let $X_{V(n)}$ be a sequence of weak record values as before. Then the condition $E[\psi(X_{V(n+2)}) \mid X_{V(n)} = j] = g(j)$ determines the distribution.*

PROOF. It can be verified that

$$P(X_{V(n+2)} = k \mid X_{V(n)} = j) = \frac{p_k}{q_j}\sum_{r=j}^{k}\frac{p_r}{q_r}, \qquad j \leq k \leq N. \tag{5.11}$$

The right hand side of the above equation is independent of n. So we can take without loss of generality

$$E[\psi(X_{V(3)}) \mid X_{V(1)} = j] = \frac{1}{q_j}\sum_{r=j}^{N}\frac{p_r}{q_r}\sum_{k=r}^{N}\psi(k)p_k, \tag{5.12}$$

see Wesolowski and Ahsanullah (2001).

This along with the hypothesis gives

$$\frac{g(j)q_j^2}{p_j} - \frac{g(j+1)q_jq_{j+1}}{p_j} = \sum_{k=1}^{N}\psi(k)p_k. \tag{5.13}$$

Taking the first order difference, we get

$$\frac{g(j)q_j^2 - g(j+1)q_jq_{j+1}}{p_j} - \frac{g(j+1)q_{j+1}^2 - g(j+2)q_{j+1}q_{j+2}}{p_{j+1}} = \psi(j)p_j. \tag{5.14}$$

Dividing by q_j using $r(i) = q_i/q_{i+1}$ $(r(i) - 1 = p_i/q_{i+1})$, we obtain

$$\frac{g(j)r^2(j) - g(j+1)r(j)}{r(j) - 1} - \frac{g(j+1)r(j+1) - g(j+2)}{r(j+1) - 1} = \psi(j)p_j. \tag{5.15}$$

Let $h(j) = 1/(r(j) - 1) = q_{j+1}/p_j$. We obtain on simplification, from (5.15).

$$h(j+1) = \frac{g(j) - g(j+1)}{g(j+1) - g(j+2)}(h(j) + 2) + \frac{g(j) - \psi(j)}{g(j+1) - g(j+2)}\frac{1}{h(j)}. \tag{5.16}$$

The general solution of (5.16) for $h(j)$ is difficult. However, for selected values of $g(j)$ and $\psi(j)$, solution of (5.16) for $h(j)$ can be obtained. For example if $g(j) = j + b$ and $\psi(j) = j$, then $h(j) = b/2$.

Once a solution of equation (5.16) is obtained, p_j can be obtained as follows:

Let $h^\star(j)$ be a solution of (5.16), then

$$p_j = \frac{1}{h^\star(j)} \cdot \frac{q_{j+1}}{q_j} \cdot \frac{q_j}{q_{j-1}} \cdots \frac{q_1}{q_0}, \qquad q_0 = 1.$$

Thus

$$p_j = \frac{1}{h^\star(j)} \prod_{k=0}^{j-1} \frac{h^\star(j)}{h^\star(j)+1}. \tag{5.17}$$ □

The following example illustrates the procedure.

Example 5.2. If $g(j) = j + b$, $N = \infty$ and $\psi(j) = j$, then $h^\star(j) = b/2$ and

$$p_j = \frac{2}{b+2} \left(\frac{b}{b+2} \right)^j, \qquad j = 0, 1, \ldots$$

i.e., X's are distributed as geometric.

Acknowledgements

The authors are thankful to the referees for some useful suggestions which enhanced the presentation.

References

Ahsanullah, M. and Wesolowski, J. (1998). Linearity of best predictors for non-adjacent record values, *Sankhya Series B*, **60**, 215–227.

Aliev, F. A. (1999). New characterization of discrete distributions through weak records, *Theory of Probability and Applications*, **44**(4), 756–761.

Dembinska, A. and Wesolowski, J. (1998). Linearity of regression for non-adjacent order statistics, *Metrika*, **48**, 215–222.

Dembinska, A. and Wesolowski, J. (2000). Linearity of regression for non-adjacent record values, *Journal of Statistical Planning and Inference*, **90**, 195–205.

Ferguson, T. S. (1967). On characterizing distributions by properties of order statistics, *Sankhya Series A*, **29**, 265–278.

Franco, M. and Ruiz, J. M. (1995). On characterization of continuous distributions with adjacent order statistics, *Statistics*, **26**, 375–385.

Franco, M. and Ruiz, J. M. (1997). On characterizations of distributions by expected values of order statistics and record values with gap, *Metrika*, **45**, 107–119.

Franco, M. and Ruiz, J. M. (1999). Characterization based on conditional expectation of adjacent order statistics: A unified approach, *Proceedings of the American Mathematical Society*, **127**(3), 861–874.

Franco, M. and Ruiz, J. M. (2001). Characterizations of discrete distributions based on the conditional expectations of record values, *Statistical Papers*, **42**, 101–110.

Gupta, R. C. (1984). Relationship between order statistics and record values and some characterization results, *Journal of Applied Probability*, **21**, 433–438.

Gupta, R. C. and Kirmani, S. N. U. A. (2004). Some characterization of distributions by functions of failure rate and mean residual life, *Communications in Statistics*: *Special issue on characterization of distributions* (to appear).

Khan, A. H. and Abu-Salih, M. S. (1989). Characterizations of probability distributions by conditional expectation of order statistics, *Metron*, **47**, 171–181.

Lopez-Blaquez, F. and Moreno Rebollo, J. L. (1997). A characterization of distributions based on linear regression of order statistics and record values, *Sankhya Series A*, **59**, 311–323.

Lopez-Blaquez, F. and Wesolowski, J. (2001). Discrete distribution for which the regression of the first record on the second is linear, *TEST*, **10**, 121–131.

Nagaraja, H. N. (1977). On a characterization based on record values, *Australian Journal of Statistics*, **19**, 70–73.

Nagaraja, H. N. (1988). Some characterizations of continuous distributions based on regressions of adjacent order statistics and record values, *Sankhyā Series A*, **50**, 70–73.

Ouyang, L. Y. (1995). Characterization through the conditional expectation of a function of one order statistic relative to an adjacent one, *Sankhyā Series A*, **57**, 500–503.

Rao, C. R. and Shanbhag, D. N. (1994). *Choquet-Deny Type Functional Equations with Applications to Stochastic Models*, John Wiley, New York.

Raqab, M. Z. (2002). Characterization of distributions based on the conditional expectations of record values, *Statistics and Decisions*, **20**, 309–319.

Wesolowski, J. and Ahsanullah, M. (1997). On characterizing distributions by linearity of regression for order statistics, *Australian Journal of Statistics*, **39**, 69–78.

Wesolowski, J. and Ahsanullah, M. (2001). Linearity of regression for non-adjacent weak records, *Statistica Sinica*, **11**, 39–52.

Wu, J. W. (2004). On characterizing distributions by conditional expectations of functions of non-adjacent record values, *Journal of Applied Statistical Sciences*, **13**(1) (to appear).

Wu, J. W. and Lee, W. C. (2001). On the characterization of generalized extreme value, power function, generalized Pareto and classical Pareto distributions by conditional expectation of record values, *Statistical Papers*, **42**, 225–242.

Wu, J. W. and Ouyang, L. Y. (1996). On characterizing distributions by conditional expectations of the functions of order statistics, *Metrika*, **43**, 135–147.

Wu, J. W. and Ouyang, L. Y. (1997). Characterizations based on conditional expectation of non-adjacent order statistics, *The International Journal of Information and Management Sciences*, **8**(3), 45–53.