Statistica Sinica **11**(2001), 39-52

LINEARITY OF REGRESSION FOR NON-ADJACENT WEAK RECORDS

Jacek Wesołowski and Mohammad Ahsanullah

Politechnika Warszawska and Rider University

Abstract: A complete characterization of the family of distributions with linearity of regression for future non-adjacent weak records with spacing equal to two is given. Also, in the adjacent case, weak-record versions of some known characterization results for regular records are presented.

Key words and phrases: Beta-negative binomial distribution, difference equations, geometric distribution, linearity of regression, negative hypergeometric distribution, weak records.

1. Introduction

Linearity of regressions for observations preserving some type of ordering has been of interest since Ferguson (1967), where all continuous distributions satisfying such a property for adjacent order statistics were identified. An analogous problem in the continuous case for adjacent record statistics was solved in Nagaraja (1977). However no major progress for non-adjacent order or record statistics in the continuous case was made until the 1990's - consult the monographs by Arnold, Balakrishnan and Nagaraja (1992, 1998). Both problems have been finally settled in Dembińska and Wesołowski (1998, 2000).

The discrete case for adjacent order statistics was investigated thoroughly in Nagaraja (1988). He considerably improved earlier results of Kirmani and Alam (1980) (see also the comments in Rao and Shanbhag (1986)), and raised some serious reservations about the validity of linearity of regression for non-adjacent order statistics in the discrete case. No progress has been done in this area (consult for instance Nagaraja (1992)), except for the recent paper by López-Blázquez and Salamanca-Miño (1998) which characterizes the geometric law for specially chosen adjacent order statistics.

Characterizations based on properties of records for discrete distributions are commonly concentrated not on the distributions themselves but rather on their tails. In an early paper Srivastava (1979) proved that geometric tail distributions are the only ones that exhibit the constancy of regression for the difference of adjacent records (for a further development and comments see Ahsanullah and Holland (1984), Rao and Shanbhag (1986) and Nagaraja, Sen and Srivastava (1989)). Korwar (1984) extended this result by considering a more general condition of linearity of regression, and characterized additionally negative hypergeometric tail distributions. The third distribution, claimed to be determined in Korwar (1984), has a bounded support, and in this case the statement of the problem is meaningless. This fact has been noted only recently in Arnold, Balakrishnan and Nagaraja (1998), which can be consulted for a wider review of the subject (an alternative reference is Ahsanullah (1995)). A characterization given by Balakrishnan and Balasubramanian (1995), using higher conditional moments, will be discussed in details in the sequel. Related problems were considered in Stepanov (1990), Huang and Li (1993) and Huang and Su (1999).

Vervaat (1973) modified the concept of records for discrete distributions to weak records. Let $X_1, X_2, \ldots$ be a sequence of independent observations on X and let $U(1) = 1, \quad U(n) = \inf\{j > U(n-1) : \ X_j \geq X_{U(n-1)}\}, \quad n = 2, 3, \ldots$. Then the rv $R_n = X_{U(n)}$ is called the n-th weak (upper) record for $(X_n)_{n=1,2,\ldots}$. Recall that for regular records the second inequality in the definition of $U(n)$ is sharp and consequently no ties among regular records are allowed. This results in a meaningful definition of regular records for discrete distributions only in the case where X has unbounded support. However weak records are also defined for distributions with bounded discrete supports. Observe that if X is of the continuous type, then weak and regular records are essentially indistinguishable.

Since ties are permitted using weak records, characterizations of the whole distributions, not only of the tails, can be obtained. To some extent this was exploited for linearity of regression for adjacent weak records in Stepanov (1994). A generalization to any regression function in the adjacent case was given in Aliev (1998). However these papers suffer from unnecessary assumption of unboundedness of the support.

The main aim of the present paper is to provide a complete characterization of the discrete distributions with linearity of regression for future weak records with spacing equal to two. The family includes geometric and negative hypergeometric distributions of the first (beta-binomial) and second (beta-negative binomial) type. This is done in Section 3 together with derivation of the general difference equation for any regression function, while in Section 2 direct results are given. In Section 4 we give some comments concerning the case of adjacent weak records, completing the results of Stepanov (1994), Aliev (1998), Balakrishnan and Balasubramanian (1995) and Nagaraja, Sen and Srivastava (1989).

Observe, adopting for instance the discussion in Nagaraja (1988), that the problems of linearity of regressions reviewed above can be rephrased in a more applied language as: identifying families of distributions with the property that the best unbiased predictor of the order (record, weak record) statistic, based on another order (record, weak record) statistic, is equal to its best linear unbiased predictor.

Finally let us point out that there are efforts to build a universal approach for treating distributional properties of order statistics and records together - see Deheuvels (1984), Kamps (1995) or Huang and Su (1999), which concentrated almost exclusively on absolutely continuous distributions. A fairly recent contribution by Asadi, Rao and Shanbhag (1999) brings a possible unified approach without restricting to absolutely continuous distributions. However such approaches fail for linearity of regression for discrete distributions. For order statistics it does not hold except for some special cases with spacing equal to one (see for instance Nagaraja (1988) or López-Blázquez and Salamanca-Miño (1998)), while this property holds in general for weak records with spacing equal to one or two, as observed for instance in the present paper.

Throughout this paper, assume that $supp(X) = \{0, \ldots, N\}$, $N \leq \infty$, where if $N = \infty$, then $N - k = \infty$, $k = 0, 1, \ldots$.

2. Regressions of Weak Records for Selected Distributions

In this section we study discrete distributions that have a property of linearity of regression of weak records with spacing equal to two, i.e. $E(R_{n+2}|R_n) = aR_n + b$, for some real numbers a and b and any natural number n. To this end, first we introduce some basic distributional properties of weak records.

The joint probability mass function (pmf) for the first n weak records can easily be determined as

$$P(R_1 = k_1, \ldots, R_n = k_n) = \left(\prod_{i=1}^{n-1} \frac{p_{k_i}}{q_{k_i}} \right) p_{k_n}, \tag{1}$$

for any $0 \leq k_1 \leq \cdots \leq k_n \leq N$ (if $N = \infty$ the last inequality is, obviously, sharp), where $p_k = P(X = k)$, $q_k = \sum_{j=k}^{N} p_j = P(X \geq k)$, $k \in \{0, \ldots, N\}$. The conditional distribution of R_{n+2} given R_n is defined by $P(R_{n+2} = k|R_n = l) = P(R_3 = k|R_1 = l) = \frac{p_k}{q_l} \sum_{r=l}^{k} \frac{p_r}{q_r}$, $l \leq k \leq N$. Hence, the regression of R_{n+2} on R_n (if it exists) can be computed (upon changing the order of summation) by

$$E(R_{n+2}|R_n = l) = \frac{1}{q_l} \sum_{r=l}^{N} \frac{p_r}{q_r} \sum_{k=r}^{N} k p_k = s(l), \quad l \in \{0, \ldots, N\}. \tag{2}$$

Now let us define the distributions we are going to work with.

(1) Denote by $ge(p)$ the geometric distribution defined by the pmf $p_j = P(X = j) = pq^j$, $j = 0, 1, \ldots$, where $p = 1 - q \in (0, 1)$. Then $q_j = q^j$, $j = 0, \ldots$.

(2) Denote by $nh_I(\alpha,\beta,N)$ the negative hypergeometric distribution of the first type (also known as beta-binomial) defined by the pmf

$$p_j = \binom{\alpha+j-1}{j}\binom{\beta-\alpha+N-j}{N-j}\binom{\beta+N}{N}^{-1}, \quad j=0,\dots,N,$$

where α and β are real numbers such that $\beta+1>\alpha>0$, and N is a positive integer - see Johnson, Kotz and Kemp (1992). Specifically for the $nh_I(1,\beta,N)$ distribution it follows that $\beta>0$ and

$$p_j = \binom{\beta+N-j-1}{N-j}\binom{\beta+N}{N}^{-1}, \quad j=0,\dots,N.$$

Consequently

$$q_j = \binom{\beta+N}{N}^{-1}\sum_{k=0}^{N-j}\binom{\beta+k-1}{k} = \binom{\beta+N-j}{N-j}\binom{\beta+N}{N}^{-1}, \quad j=0,\dots,N.$$

(3) Denote by $nh_{II}(\alpha,\beta,\gamma)$ the negative hypergeometric distribution of the second type (also known as beta-negative binomial), defined by the pmf

$$p_j = \frac{\gamma}{\gamma+j}\binom{\beta}{\gamma}\binom{\alpha+j-1}{j}\binom{\alpha+\beta+j}{\gamma+j}^{-1}, \quad j=0,1,\dots,$$

where α, β and γ are real numbers such that $\alpha>0$, $\beta+1>\gamma>0$ - see Johnson, Kotz and Kemp (1992).

Specifically for the $nh_{II}(1,\beta,\gamma)$ distribution with $\beta>\gamma$ (this additional restriction implies existence of the first moment), it follows that

$$p_j = \frac{\gamma}{\gamma+j}\binom{\beta}{\gamma}\binom{\beta+j+1}{\gamma+j}^{-1}, \quad j=0,1,\dots.$$

Then by the summation rule for $nh_{II}(1,\beta+j,\gamma+j)$,

$$q_j = \gamma\binom{\beta}{\gamma}\sum_{i=0}^{\infty}(\gamma+j+i)^{-1}\binom{\beta+j+1+i}{\gamma+j+i}^{-1} = \frac{\gamma}{\gamma+j}\binom{\beta}{\gamma}\binom{\beta+j}{\gamma+j}^{-1}, \quad j=0,1,\dots.$$

It is shown in the Appendix that for the above three distributions the relation $E(R_{n+2}|R_n)=aR_n+b$ a.s. holds, where
(1) for the $ge(p)$ distribution, $a=1$, $b=2q/p$;
(2) for the $nh_I(1,\beta,N)$ distribution, $a=\beta^2(\beta+1)^{-2}\in(0,1)$, $b=N(2\beta+1)(\beta+1)^{-2}$;

(3) for the $nh_{II}(1,\beta,\gamma)$ distribution and $\delta = \beta - \gamma > 0$, $a = (1+\delta^{-1})^2 > 1$, $b = \gamma(2\delta+1)\delta^{-2}$.

The main result of this paper, considered in the next section, lies in proving that these are the only distributions with such a property. It is possible that linearity of regression, $E(R_{n+k}|R_n) = aR_n + b$, holds only for these three distributions for any $k = 1, 2, \ldots$. Here we checked it for $k = 2$ (for the case $k = 1$ see the comments in Sections 1 and 4). While the direct result for any k seems to be within reach, the converse looks rather difficult and possibly requires new techniques.

3. Characterization by Linearity of Regression of R_{n+2} on R_n

Denote for any $l = 0, \ldots, N$, $s(l) = E(R_{n+2}|R_n = l)$. Then by (2) one gets $s(l)q_l = \sum_{r=l}^N \frac{p_r}{q_r} \sum_{k=r}^N kp_k$, $l = 0, \ldots, N$. Hence, upon taking the first order differences, it follows that

$$s(l)\frac{q_l^2}{p_l} - s(l+1)\frac{q_l q_{l+1}}{p_l} = \sum_{k=l}^N kp_k, \quad l = 0, \ldots, N-1. \tag{3}$$

Again taking the difference we have

$$\frac{s(l)q_l^2 - s(l+1)q_l q_{l+1}}{p_l} - \frac{s(l+1)q_{l+1}^2 - s(l+2)q_{l+1}q_{l+2}}{p_{l+1}} = lp_l, \quad l=0, \ldots, N-2.$$

Now divide both sides of the above equation by q_{l+1}. Denoting $r(l) = q_l/q_{l+1}$, we obtain

$$\frac{s(l)r^2(l) - s(l+1)r(l)}{r(l)-1} - \frac{s(l+1)r(l+1) - s(l+2)}{r(l+1)-1} = l[r(l)-1], \quad l = 0, \ldots, N-2.$$

Now for $h(l) = 1/(r(l)-1)$ it follows that

$$h(l+1) = \frac{s(l) - s(l+1)}{s(l+1) - s(l+2)}(h(l)+2) + \frac{s(l) - l}{s(l+1) - s(l+2)}\frac{1}{h(l)}, \quad l=0, \ldots, N-2.$$

The general solution of the above equation for h seems to be difficult to derive. However in the linear case, i.e. for $s(l) = al + b$, $l = 0, \ldots, N$, where a and b are some real numbers, it simplifies to

$$h(l+1) = h(l) + 2 - \frac{(a-1)l + b}{a}\frac{1}{h(l)}, \quad l = 0, \ldots, N-2. \tag{4}$$

On the other hand, from (3) we see that

$$\begin{aligned} lq_l \le \sum_{k=l}^N kp_k &= (al+b)\frac{q_l(p_l + q_{l+1})}{p_l} - (al+a+b)\frac{q_l q_{l+1}}{p_l} \\ &= (al+b)q_l - a\frac{q_l q_{l+1}}{p_l}, \quad l = 0, \ldots, N-1. \end{aligned}$$

Since $h(l) = q_{l+1}/p_l$, $l = 0, \ldots, N-1$, it follows that

$$(a-1)l + b > ah(l), \quad l = 0, \ldots, N-1. \tag{5}$$

In this case we are able to solve (4) and, consequently, to obtain the characterization result.

Theorem 1. *Let X be a rv with a nondegenerate distribution concentrated on $\{0, \ldots, N\}$, $N \leq \infty$. Assume that*

$$E(R_{n+2}|R_n) = aR_n + b, \quad a.s., \tag{6}$$

for some real numbers a and b. Then $a > 0$, $b > 0$, and only the following cases are possible:

(i) $0 < a < 1, b/(1-a)$ is a positive integer and $X \sim nh_I(1, \sqrt{a}/(1-\sqrt{a}),$ $b/(1-a))$; (7)

(ii) $a = 1$ and $X \sim ge(b/(2+b))$; (8)

(iii) $a > 1$ and $X \sim nh_{II}(1, (b + \sqrt{a} + 1)/(a-1), b/(a-1))$. (9)

Proof. Observe first that since $s(l) = E(R_{n+2}|R_n = l)$ cannot be a strictly decreasing or constant function on $supp(X)$, a has to be positive. Since X is a nondegenerate rv, $b = s(0)$ also has to be positive.

Observe also, from the obvious inequality $R_n \leq R_{n+2}$ a.s. and the fact that the supports of X and weak records coincide, that $N = \inf\{l = 0, 1, \ldots : s(l) = l\}$. Consequently

$$N = \inf\{l = 0, 1, \ldots : al + b = l\} = \begin{cases} b/(1-a), & \text{if} \quad 0 < a < 1, \\ \\ \infty, & \text{if} \quad a \geq 1. \end{cases} \tag{10}$$

Case $0 < a < 1$.

Define $\beta = \sqrt{a}/(1-\sqrt{a}) > 0$. Observe that, by (10), $N = b/(1-a)$ is a positive integer and $supp(X) = \{0, \ldots, b/(1-a)\}$. Consequently equation (4) can be rewritten as

$$h(l+1) = h(l) + 2 - \frac{2\beta+1}{\beta^2}\frac{N-l}{h(l)}, \quad l = 0, \ldots, N-2. \tag{11}$$

Take first $h(0) = N/\beta$. Then, by (11), $h(l) = \frac{N-l}{\beta}$, $l = 0, \ldots, N-1$. Hence $r(l) = \frac{h(l)+1}{h(l)} = \frac{N-l+\beta}{N-l}$, $l = 0, \ldots, N-1$. Now by the definition of r one easily gets, since $P(X \geq 0) = 1$,

$$q_l = \binom{\beta + N - l}{N - l}\binom{\beta + N}{N}^{-1}, \quad l = 0, \ldots, N.$$

Hence $X \sim nh_I(1, \beta, N)$ and (7) holds.

Assume now that $h(0) > N/\beta$. Then by (11) it follows inductively that $h(l) > (N-l)/\beta$ for any $l = 0, \ldots, N-1$. Consequently $h(N-1) > 1/\beta$. Similarly, $h(0) < N/\beta$ implies that $h(N-1) < 1/\beta$.

Substituting $l = N-1$ in (3) with linear s, we get $(N-a)\frac{q_{N-1}^2}{p_{N-1}} - N\frac{q_{N-1}q_N}{p_{N-1}} = (N-1)p_{N-1} + Np_N = Nq_{N-1} - p_{N-1}$. Hence $a(p_{N-1}+p_N)^2 = p_{N-1}^2$ and since, by definition, $h(N-1) = q_N/p_{N-1}$ (observe that $q_N = p_N$), we get $h(N-1) = 1/\beta$. Conseqently the assumption $h(0) \neq N/\beta$ is contradicted.

Case $a = 1$.

By (10), $supp(X) = \{0, 1, \ldots\}$, and (4) takes a simple form

$$h(l+1) = h(l) + 2 - b/h(l), \quad l = 0, 1, \ldots \tag{12}$$

First assume that $h(0) = b/2$. Then (12) implies that $h(l) = b/2$ for all $l = 0, 1, \ldots$. Hence, coming back to r, one gets $q_l/q_{l+1} = (b+2)/b$, $l = 0, 1, \ldots$, and finally $q_l = (b/(b+2))^l$, $l = 0, 1, \ldots$, which gives (8).

Take $h(0) < b/2$. Then (12) implies that the sequence $(h(l))_{l=0,1,\ldots}$ is decreasing. Since it is bounded from below by 0, $\alpha = \lim_{l\to\infty} h(l)$ exists. Consequently, by passing to the limit (as $l \to \infty$) on both sides of (12), it follows that $\alpha = b/2$, which is a contradiction.

Consider now the case $h(0) > b/2$. Then by (12), $(h(l))_{l=0,1,\ldots}$ is increasing and, since it is bounded from above by b/a (see (5)), it follows that the limit $\alpha = \lim_{l\to\infty} h(l)$ exists. Again $\alpha = b/2$, which contradicts the assumption $h(0) > b/2$.

Case $a > 1$.

Define $\gamma = b/(a-1)$, $\beta = (b + \sqrt{a} + 1)/(a-1)$, $\delta = \beta - \gamma > 0$. Then $N = \infty$ by (10), and (4) takes the shape

$$h(l+1) = h(l) + 2 - \frac{2\delta+1}{(\delta+1)^2}\frac{\gamma+l}{h(l)}, \quad l = 0, 1, \ldots \tag{13}$$

Define a sequence $(t(l))_{l=0,1,\ldots}$ by $t(l) = (\delta+1)h(l)/(\gamma+l)$, $l = 0, 1, \ldots$. Then (13) can be rewritten as

$$t(l+1) = \frac{(\gamma+l)t(l) + 2(\delta+1) - (2\delta+1)/t(l)}{\gamma+l+1}, \quad l = 0, 1, \ldots. \tag{14}$$

Define a sequence of functions $(f_l)_{l=0,1,\ldots}$ on $(0, \infty)$ by

$$f_l(t) = \frac{(\gamma+l)t + 2(\delta+1) - (2\delta+1)/t}{\gamma+l+1}, \quad t > 0,\ l = 0, 1, \ldots.$$

Observe that f_l is strictly increasing, $f_l(1) = 1$, $f_l(2\delta+1) = 2\delta+1$ for any $l = 0, 1, \ldots$. Hence for $A = [1, 2\delta+1]$ we have $f_l(A) = A$, $l = 0, 1, \ldots$. Further we can determine the set $A_l = \{t : f_l(t) \geq t\}$, for any $l = 0, 1, \ldots$, by solving the inequality $2(\delta+1) - (2\delta+1)/t \geq t$ in $(0, \infty)$. We easily get $A_l = [1, 2\delta+1] = A$ independent of l.

Suppose $t(0) = 1$. Since (4) can be rewritten as $t(l+1) = f_l(t(l))$, $l = 0, 1, \ldots$, and 1 is a fixed point of f_l for any $l = 0, 1, \ldots$, one gets $h(l) = (\gamma + l)/(\delta + 1)$, $l = 0, 1, \ldots$, and consequently $q_l/q_{l+1} = (\gamma + \delta + l + 1)(\gamma + l)$, $l = 0, 1, \ldots$.

Finally

$$q_l = \frac{\gamma}{\gamma + l}\binom{\gamma+\delta}{\gamma}\binom{\gamma+\delta+l}{\gamma+l}^{-1}, \quad l = 0, 1, \ldots,$$

the negative hypergeometric distribution given at (9).

Now we have to show that other choices for $t(0)$ are impossible. Suppose $t(0) \in (1, 2\delta+1] \subset A$. Then $t(l+1) = f_l(t(l)) \in [t(l), 2\delta+1]$, $l = 0, 1, \ldots$ and consequently the sequence $(t(l))_{l=0,1,\ldots}$ is nondecreasing and bounded from above by $2\delta+1$. Hence it converges for $l \to \infty$. Denote its limit by α. Rewrite (14) as

$$l[t(l+1) - t(l)] + \gamma[t(l+1) - t(l)] + t(l+1) = 2(\delta+1) - \frac{2\delta+1}{t(l)}, \quad l = 0, 1, \ldots. \tag{15}$$

Since $(t(l))_{l=0,1,\ldots}$ converges, $\lim_{l\to\infty} l[t(l+1) - t(l)] = 0$. Consequently (15) implies that $\alpha = 2\delta+1$ since the other solution $\alpha = 1$ contradicts the assumption $t(0) > 1$. To see that $\alpha = 2\delta+1$ is also contradictory, take $\epsilon > 0$ such that $\delta > \epsilon$ and $t(l) > 2\delta + 1 - \epsilon$ for $l > L$ sufficiently large. Then $h(l) > l + \frac{2\delta+1-\epsilon}{\delta+1}\gamma + \frac{\delta-\epsilon}{\delta+1}l$, $l > L$, which contradicts (5) since $\delta - \epsilon > 0$.

Let $t(0) > 2\delta+1$. Since $f_l((2\delta+1, \infty)) = (2\delta+1, \infty)$, $l = 0, \ldots$, $2\delta+1 < t(l+1) < t(l)$, $\forall\, l = 0, 1, \ldots$. Consequently $\alpha = \lim_{l\to\infty} t(l)$ exists, and by (15) $\alpha = 1$ or $\alpha = 2\delta+1$, both contradictory as was observed above.

Take finally $0 < t(0) < 1$. Then, similarly, $(t(l))_{l=0,1,\ldots}$ is a decreasing sequence bounded from below by 0. Again (15) yields possible limits of 1 or $2\delta+1$, and both the possibilities yield contradictions.

4. The Adjacent Case - Comments and Complements

4.1. On Stepanov (1994) and Aliev (1998)

In the adjacent case, essentially resolved in Stepanov (1994) and Aliev (1998), the formulas for the conditional pmf and regression are much simpler: for any $n = 1, 2, \ldots$

$$P(R_{n+1} = k | R_n = l) = P(R_2 = k | R_1 = l) = p_k/q_l, \quad k \geq l, \tag{16}$$

and consequently $E(R_{n+1}|R_n = l) = q_l^{-1}\sum_{k=l}^{N} kP(X = k) = s_1(l)$, $l = 0, \ldots, N$. Here we do not follow the approach of these two papers. Taking the first differences in the above identity we get

$$\frac{q_l}{q_{l+1}} = \frac{s_1(l+1) - l}{s_1(l) - l}, \quad l = 0, \ldots, N-1.$$

Hence it follows immediately that

$$p_l = \frac{s_1(l+1) - s_1(l)}{s_1(l+1) - l} \prod_{j=0}^{l-1} \frac{s_1(j) - j}{s_1(j+1) - j}, \quad l = 0, \dots, N. \tag{17}$$

A version of this formula was obtained in Aliev (1998) in the case of unbounded support. Similarly one can approach the problem by considering $E(H(R_{n+1})|R_n = l) = s_H(l)$ for some functions H. Observe that, as in the non-adjacent case, we have $N = inf\{l = 0, 1, \dots : s_1(l) = l\}$. Consequently the main result of Aliev (1998) about unique determination of the distribution of X by the function s_1 holds without the restriction of unboundedness of support.

If it is assumed that $s_1(l) = al + b$, $l = 0, \dots, N$, then the following complete version of Stepanov's (1994) result can be easily derived from (17).

Theorem 2. *Assume that X has a non-degenerate distribution and $E(R_{n+1}|R_n) = aR_n + b$, a.s., where a and b are some real constants. Then $a > 0$, $b > 0$, and one of the following holds:*

(i) $0 < a < 1$, *$b/1-a$ is a positive integer and $X \sim nh_I(1, a/(1-a), b/(1-a))$;*
(ii) $a = 1$ *and $X \sim ge(1/(1+b))$;*
(iii) $a > 1$ *and $X \sim nh_{II}(1, (b+1)/(a-1), b/(a-1))$.*

The case (i) is missing in Stepanov (1994).

4.2. On Balakrishnan and Balasubramanian (1995)

These authors considered the problem of charactrizing discrete distributions with the property $E((\tilde{R}_2 - \tilde{R}_1)^2|\tilde{R}_1) = const$, a.s., where $\tilde{R}_n$ denotes the regular nth record. They claim to obtain characterization of the geometric distribution, while essentially the distribution is a geometric tail distribution of the form $\alpha\delta_0 + (1-\alpha)ge(p)$, for any $\alpha \in [0, 1]$, where δ_0 denotes the unit mass concentrated at 0. Also in that paper a lot of effort is devoted to proving that the distribution of X has unbounded support (a necessary condition for introducing regular records). A much simpler derivation of this fact in the case of weak records will be given in the sequel. Observe that geometric tail distributions can be characterized by $E((\tilde{R}_{n+1} - \tilde{R}_n)^2|\tilde{R}_n) = const$, a.s.; only then the distribution has the shape $\alpha_0\delta_0 + \cdots + \alpha_{n-1}\delta_{n-1} + \alpha_n ge(p)$, where δ_k is a unit mass concentrated at k, $k = 0, \dots, n-1$, and $\alpha_0, \dots, \alpha_n$ are positive numbers summing to one.

Here we study an analogoue of Balakrishnan and Balasubramanian's (1995) result for weak records. First we observe that if $X \sim ge(p)$ then $E((R_{n+1} - R_n)^2|R_n = l) = \sum_{k=l}^{\infty}(k-l)^2 q^k p/q^l = 2(q/p)^2 + q/p = c$, where $q = 1-p$. Now we derive the converse of this statement.

Theorem 3. *Assume that X has a discrete non-degenerate distribution concentrated on $\{0,\ldots,N\}$ such that $E((R_{n+1}-R_n)^2)<\infty$. If*

$$E((R_{n+1}-R_n)^2|R_n)=c, \quad a.s., \tag{18}$$

where c is a real number, then $c>0$ and $X\sim ge\left(\frac{\sqrt{8c+1}-3}{2(c-1)}\right)$.

Proof. Assume $N<\infty$. Then $E((R_{n+1}-R_n)^2|R_n=N)=0$, which is possible only in the case $N=0$ and $P(X=0)=1$. Since we consider only non-degenrate distributions, $N=\infty$. By (16) and (18) we have $\sum_{k=l}^{\infty}(k-l)^2p_k=cq_l$, $l=0,1,\ldots$. Now, repeating the argument from Balakrishnan and Balasubramanian (1995), i.e., taking the difference operator with respect to l on the above equation twice, one arrives at $cp_l-(2c+1)p_{l+1}+(c-1)p_{l+2}=0$, $l=0,1,\ldots$. (Balakrishnan and Balasubramanian (1995) derived an analoguous equation: $(c-1)p_l-(2c+1)p_{l+1}+cp_{l+2}=0$, $l=1,2,\ldots$) Since this is a linear difference equation of the second order, we first solve its characteristic equation $(c-1)r^2-(2c+1)r+c=0$, which has solutions $r_{1,2}=\frac{2c+1\pm\sqrt{8c+1}}{2(c-1)}$. Since $r_1>1$ and $0<r_2<1$ we must have $p_l=\alpha r_2^l$, $l=0,1,\ldots$. Since $\sum_{l=0}^{\infty}\alpha r_2^l=1$ we get $\alpha=1-r_2=\frac{\sqrt{8c+1}-3}{2(c-1)}\in(0,1)$. Observe that $c<1$ is allowed.

4.3. On Nagaraja, Sen and Srivastava (1989)

These authors characterized geometric tail distributions for X by the condition $E(\tilde{R}_{n+2}-\tilde{R}_{n+1}|\tilde{R}_n)=b$, a.s. Here a version of this result for weak records is given.

Observe first that if $X\sim ge(p)$, $E(R_{n+2}-R_{n+1}|R_n)=E(R_{n+2}|R_n)-E(R_{n+1}|R_n)=(R_n+2q/p)-(R_n+q/p)=q/p$, a.s. Now we consider a converse of this observation.

Theorem 4. *Let X be a non-degenerate rv concentrated on $\{0,\ldots,N\}$ and $E(R_{n+2})<\infty$. Assume that*

$$E(R_{n+2}-R_{n+1}|R_n)=c, \quad a.s. \tag{19}$$

Then $X\sim ge(1/(1+c))$.

Proof. As in the proof of Theorem 3, it follows that if $N<\infty$ then $c=0$, which is contradicts the non-degeneracy assumption. Hence $supp(X)=\{0,1,\ldots\}$.

By the formulas derived in Section 3 and in the section above, it follows that (19) can be rewritten in the form $cq_l=\sum_{r=l}^{\infty}\frac{p_r}{q_r}\sum_{k=r}^{\infty}kp_k-\sum_{k=l}^{\infty}kp_k$, $l=0,1,\ldots$. Taking differences one gets $\sum_{k=l}^{\infty}kp_k=(c+l)q_l$, $l=0,1,\ldots$. Again taking differences it follows that $p_l(1+c)=q_l$, $l=0,1,\ldots$, which implies $(1+c)p_{l+1}=cp_l$, $l=0,1,\ldots$. The latter identity immediately yields the final result.

Acknowledgement

The authors are indebted to the editors, especially to the former Chair-Editor, Dr. Ching-Shui Cheng, and referees, whose comments significantly improved the exposition of the paper. A careful reading of the draft version by A. Dembińska helped to clarify some arguments.

Appendix. The Direct Results

Here we present details of the computations leading to the results of Section 2.

(1) For the $ge(p)$ distribution we have $p_j/q_j = p$, $j = 0, 1, \ldots$. Now (2) implies that for any $l = 0, 1, \ldots$ $E(R_{n+2}|R_n = l) = \frac{p^2}{q^l}\sum_{r=l}^{\infty}\sum_{k=r}^{\infty} kq^k$, and routine technique leads to $E(R_{n+2}|R_n = l) = l + 2q/p$, $l = 0, 1, \ldots$.

(2) For the $nh_I(1, \beta, N)$ distribution we have $p_j/q_j = \beta/(\beta + N - j)$, $j = 0, \ldots, N$. Applying (2) one gets, for any $l = 0, \ldots, N$,

$$E(R_{n+2}|R_n = l) = \binom{\beta + N - l}{N - l}^{-1} \sum_{r=l}^{N} \frac{\beta}{\beta + N - r} \sum_{k=r}^{N} k \binom{\beta + N - k - 1}{N - k}.$$

Observe that

$$\begin{aligned}\sum_{k=r}^{N} k \binom{\beta + N - k - 1}{N - k} &= \sum_{i=0}^{N-r} (N - i) \binom{\beta + i - 1}{i} \\ &= N \sum_{i=0}^{N-r} \binom{\beta + i - 1}{i} - \sum_{i=1}^{N-r} i \binom{\beta + i - 1}{i} \\ &= N \sum_{i=0}^{N-r} \binom{\beta + i - 1}{i} - \beta \sum_{j=0}^{N-r-1} \binom{\beta + j}{j}.\end{aligned}$$

Now by summation rules for $nh_I(\beta, \beta, N - r)$ and $nh_I(\beta + 1, \beta + 1, N - r - 1)$, respectively, for $r = 0, \ldots, N - 1$ and for $r = N$ directly, it follows that

$$\sum_{k=r}^{N} k \binom{\beta + N - k - 1}{N - k} = \frac{N + r\beta}{\beta + 1} \binom{\beta + N - r}{N - r}.$$

Consequently

$$E(R_{n+2}|R_n = l) = \frac{\beta}{(\beta + 1)} \binom{\beta + N - l}{N - l}^{-1} \sum_{r=l}^{N} \frac{N + r\beta}{\beta + N - r} \binom{\beta + N - r}{N - r}.$$

But

$$\sum_{r=l}^{N} \frac{N + r\beta}{\beta + N - r} \binom{\beta + N - r}{N - r} = \sum_{r=l}^{N} \frac{N + r\beta}{\beta} \binom{\beta + N - r - 1}{N - r}$$

$$= \left(\frac{N}{\beta} + \frac{N+l\beta}{\beta+1}\right)\binom{\beta+N-l}{N-l}, \quad l = 0,\dots,N.$$

Finally one gets $E(R_{n+2}|R_n = l) = (\beta/(\beta+1))^2 l + (2\beta+1)N/(\beta+1)^2$, $l = 0,\dots,N$.

(3) For the $nh_{II}(1,\beta,\gamma)$ distribution with $\beta > \gamma$, it follows that $p_j/q_j = (\beta-\gamma+1)/(\beta+j+1)$, $j = 0,1,\dots$. For any $r = 0,1,\dots$ we have

$$\sum_{k=r}^{\infty} kP(X=k)$$

$$= \sum_{j=0}^{\infty} \frac{\gamma(r+j)}{\gamma+r+j}\binom{\beta}{\gamma}\binom{\beta+r+j+1}{\gamma+r+j}^{-1}$$

$$= \frac{\gamma(r-1)}{\gamma+r}\binom{\beta}{\gamma}\binom{\beta+r}{\gamma+r}^{-1}\sum_{j=0}^{\infty}\frac{\gamma+r}{\gamma+r+j}\binom{\beta+r}{\gamma+r}\binom{\beta+r+j+1}{\gamma+r+j}^{-1}$$

$$+\frac{\gamma}{\gamma+r}\binom{\beta}{\gamma}\binom{\beta+r-1}{\gamma+r}^{-1}\sum_{j=0}^{\infty}\frac{\gamma+r}{\gamma+r+j}\binom{2+j-1}{j}\binom{\beta+r-1}{\gamma+r}\binom{2+(\beta+r-1)+j}{\gamma+r+j}^{-1}$$

$$= \frac{\gamma(r-1)}{\gamma+r}\binom{\beta}{\gamma}\binom{\beta+r}{\gamma+r}^{-1} + \frac{\gamma}{\gamma+r}\binom{\beta}{\gamma}\binom{\beta+r-1}{\gamma+r}^{-1}$$

$$= \frac{\gamma[r(\beta-\gamma+1)+\gamma]}{(\gamma+r)(\beta-\gamma)}\binom{\beta}{\gamma}\binom{\beta+r}{\gamma+r}^{-1},$$

which follows by summation rules for the $nh_{II}(1,\beta+r,\gamma+r)$ and $nh_{II}(2,\beta+r-1,\gamma+r)$ distributions, respectively.

Plugging the above expression into (2), one gets for any $l = 0,1,\dots$,

$$E(R_{n+2}|R_n = l)$$

$$= \frac{1}{q_l}\sum_{r=l}^{\infty}\frac{(\beta-\gamma+1)\gamma[r(\beta-\gamma+1)+\gamma]}{(\beta+r+1)(\gamma+r)(\beta-\gamma)}\binom{\beta}{\gamma}\binom{\beta+r}{\gamma+r}^{-1}$$

$$= \frac{1}{q_l}\binom{\beta}{\gamma}\left[\frac{\beta-\gamma+1}{\beta-\gamma}\sum_{r=l}^{\infty} r\frac{\gamma}{\gamma+r}\binom{\beta+r+1}{\gamma+r}^{-1} + \frac{\gamma}{\beta-\gamma}\sum_{r=l}^{\infty}\frac{\gamma}{\gamma+r}\binom{\beta+r+1}{\gamma+r}^{-1}\right]$$

$$= \frac{1}{q_l}\left[\frac{(\beta-\gamma+1)[l(\beta-\gamma+1)+\gamma]}{(\beta-\gamma)^2}q_l + \frac{\gamma}{\beta-\gamma}q_l\right].$$

Finally $E(R_{n+2}|R_n = l) = \left(1+\frac{1}{\delta}\right)^2 l + \frac{2\delta+1}{\delta^2}\gamma$, $l = 0,1,\dots$, where $\delta = \beta-\gamma > 0$.

References

Ahsanullah, M. (1995). *Record Statistics.* Nova Science Publishers, Commack, New Jersey.

Ahsanullah, M. and Holland, B. (1984). Record values and the geometric distribution. *Statist. Hefte* **25**, 319-327.

Aliev, F. A. (1998). Characterizations of distributions through weak records. *J. Appl. Statist. Sci.* **8**, 13-16.

Arnold, B. C., Balakrishnan, N. and Nagaraja, H. N. (1992). *A First Course in Order Statistics.* Wiley, New York.

Arnold, B. C., Balakrishnan, N. and Nagaraja, H. N. (1998). *Records.* Wiley, New York.

Asadi, M., Rao, C. R. and Shanbhag, D. N. (1999). Some unified characterization results on generalized Pareto distributions. *Preprint*, 1-23.

Balakrishnan, N. and Balasubramanian, K. (1995). A characterization of geometric distribution based on records. *J. Appl. Statist. Sci.* **2**, 277-282.

Deheuvels, P. (1984). The characterization of distributions by order statistics and record values - a unified approach. *J. Appl. Probab.* **21**, 326-334. (Correction. *Ibidem* **22**, 997.)

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

Dembińska, A. and Wesołowski, J. (2000). Linearity of regression for non-adjacent records. *J. Statist. Plann. Inference.* To appear.

Ferguson, T. S. (1967). On characterizing distributions by properties of order statistics. *Sankhyā Ser. A* **29**, 265-278.

Huang, W.-J. and Li, S.-H. (1993). Characterization results based on record values. *Statist. Sinica* **3**, 583-599.

Huang, W.-J. and Su, J.-S. (1999). On certain problems involving order statistics - a unified approach through order statistics property of point processes. *Sankhā Ser. A* **61**, 36-49.

Johnson, N. L., Kotz, S. and Kemp, A. (1992). *Discrete Univariate Distributions.* Wiley, New York.

Kamps, U. (1995). *A Concept of Generalized Order Statistics.* B.G. Teubner, Stuttgart.

Kirmani, S. N. U. A. and Alam, S. N. (1980). Characterization of the geometric distribution by the form of a predictor. *Comm. Statist. Theory Methods* **9**, 541-547.

Korwar, R. M. (1984). On characterizing distributions for which the second record value has a linear regression on the first. *Sankhyā Ser. B* **46**, 108-109.

López-Blázquez, F. and Salamanca-Miño, B. (1998). A characterization of the geometric distribution. *Statist. Papers* **39**, 231-236.

Nagaraja, H. N. (1977). On a characterization based on record values. *Austral. J. Statist.* **19**, 70-73.

Nagaraja, H. N. (1988). Some characterizations of discrete distributions based on linear regressions of adjacent order statistics. *J. Statist. Plann. Inference* **20**, 65-75.

Nagaraja, H. N. (1992). Order statistics from discrete distributions. *Statist.* **23**, 189-216.

Nagaraja, H. N., Sen, P. and Srivastava, R. C. (1989). Some characterizations of geometric tail distributions based on record values. *Statist. Papers* **30**, 147-155.

Rao, C. R. and Shanbhag, D. N. (1986). Recent results on characterization of probability distributions: a unified approach through extensions of Deny's theorem. *Adv. Appl. Probab.* **18**, 660-678.

Srivastava, R. C. (1979). Two characterizations of the geometric distribution by record values. *Sankhyā Ser. B* **40**, 276-278.

Stepanov, A. V. (1990). Characterizations of a geometric class of distributions. *Theory Probab. Math. Statist.* **41**, 133-136.

Stepanov, A. V. (1994). A characterization theorem for weak records. *Theory Probab. Appl.* **38**, 762-764.

Vervaat, W. (1973). Limit theorems for records from discrete distributions. *Stochastic Process. Appl.* **1**, 317-334.

Wydział Matematyki i Nauk Informacyjnych, Politechnika Warszawska, Plac Politechniki 1, 00-661 Warszawa, Poland.

E-mail: wesolo@alpha.mini.pw.edu.pl

Department of Management Sciences, Rider University, Lawrenceville, NJ 08648-3099, U.S.A.

E-mail: ahsan@rider.edu

(Received February 1999; accepted April 2000)