On record values of discrete distributions

Fazil Alioglu Aliev
Baskent University, Faculty of Engineering,
Department of Industrial Engineering, Ankara, Turkey
e-mail: falioglu@baskent.edu.tr

Abstract

Let $X_1, X_2, ...$ be a sequence of independent and identically distributed random variables taking on values $0, 1, ...$ with distribution function F such that $F(n) < 1$ for any n, and let $EX_1 \ln(1 + X_1) < \infty$. Let $X_{U(n)}$ be the n-th record value. In this paper we show that $E(X_{U(n+2)} - X_{U(n)} \mid X_{U(n)} = l) = c$, for some $n > 0$ and all $l \geq n - 1$ if and only if X_1 has the distribution of the form

$p_k = P(X_1 = k) = (1 - p_0)\left(1 - \frac{2}{c}\right)\left(\frac{2}{c}\right)^{k-1}, k = 1, 2, 3, ...$

($p_0 = P(X_1 = 0)$ is arbitrary $0 < p_0 < 1$). Our result is generalization of the theorems Srivastava (1979), Korwar (1984) and Aliev (2000).

Key Words: records, characterization of discrete distributions, characterization of geometric distribution.

1. INTRODUCTION

A lot of papers in the field of records are devoted to characterizations of distributions via records (see references below and references in the book Ahsanullah (1995) and in the paper Nevzorov (1987) among many others). Great interest in records exists because we often come across with them in our everyday life in such a way that singling out and fixing of record values.

Let $X_1, X_2, ...$ be a sequence of independent and identically distributed random variables (r.v.$'$s) taking on values $0, 1, ...$ with distribution function F such that $F(n) < 1$ for any n, and let $EX_1 \ln(1 + X_1) < \infty$. Define the sequence of weak record times $L(n)$, record times $U(n)$, weak record values $X_{L(n)}$ and record values $X_{U(n)}$ as follows:

$$\begin{aligned} L(1) &= 1, \quad L(n+1) = \min\{j > L(n) : X_j \geq X_{L(n)}\}, \quad n = 1, 2, ..., \\ U(1) &= 1, \quad U(n+1) = \min\{j > U(n) : X_j > X_{U(n)}\}, \; n = 1, 2, ... \end{aligned}$$

The following are well-known:

1) X_1 has a distribution of the form $P\{X_1 \geq m\} = \prod_{i=1}^{m} \frac{\alpha+(i-1)\beta}{1+\alpha+i\beta}$ for some $\alpha > 0, \beta \geq 0$, and $m = 1, 2, ...,$ if and only if $E(X_{L(n+1)} - X_{L(n)} \mid X_{L(n)} = l) = \alpha + \beta l$, for some $n \geq 1$ and for all $l = 0, 1, ...$ (Stepanov (1994)). If $\beta = 0$ this result corresponds to the geometric distribution;

2) if $\{S_i\}_{i=0}^{\infty}$ is any sequence of positive numbers such that $S_i > S_{i-1} - 1$ for all i and $\prod_{i=1}^{\infty} \frac{S_i}{1+S_i} = 0$ then X_1 has distribution of the form $P(X_1 \geq m) = \prod_{i=1}^{m} \frac{S_{i-1}}{1+S_i}$ for all $m = 1, 2, ...$ if and only if $E(X_{L(n+1)} - X_{L(n)} \mid X_{L(n)} = l) = S_l$ for some $n \geq 1$ and for all $l = 0, 1, ...$ (Aliev (1998)). In the case of $S_l = \alpha + \beta l$ from this result implies the above result of Stepanov (1994);

3) if $\{S_i\}_{i=0}^{\infty}$ is any sequence of positive finite numbers such that $S_i > S_{i-1} - 1$ for all i and if there exists $F(x)$ such that $E(X_{L(n+2)} - X_{L(n)} \mid X_{L(n)} = l) = S_l$ for some $n \geq 1$ and all $l = 0, 1, ...$ then $F(x)$ is unique(Aliev (1999)). There are some examples of characterization theorems for known distributions is this paper and it is easy to extend these characterizations for other distributions. Note also that the characterization problem using $E\{X_{L(n+k)} - X_{L(n)} \mid X_{L(n)} = l\}$ for $k > 2$ is steel open;

4) X_1 has a distribution of the form $p_m = (1-p_0)\left(\frac{\alpha-1}{\alpha}\right)^{m-1}\frac{1}{\alpha}$ for some $0 < p_0 < 1$ and $m = 1, 2, ...,$ if and only if $E(X_{U(n+1)} - X_{U(n)} \mid X_{U(n)} = l) = \alpha$, for some $n \geq 1, \alpha > 1$ and for all $l = 0, 1, ...$ (Srivastava (1979));

5) X_1 has a distribution of the form $p_1 = (1-p_0)\frac{1+\beta}{\alpha+\beta}$, $p_k = (1-p_0)\frac{1+\beta}{\alpha+k\beta}\prod_{j=1}^{k-1}\left(1 - \frac{1+\beta}{\alpha+j\beta}\right)$ for some $0 < p_0 < 1$ and $m = 2, 3, ...,$ if and only if $E(X_{U(n+1)} - X_{U(n)} \mid X_{U(n)} = l) = \alpha + \beta l$, for some $n \geq 1$, $\alpha > 1$, $\beta \geqslant 0$ and for all $l = 0, 1, ...$ (Korwar (1984));

6) if $\{S_i\}_{i=0}^{\infty}$ is any sequence of positive numbers such that $S_k > 1$, $S_{k+1} > S_k - 1$ for all k and $\prod_{i=0}^{\infty} \frac{S_i - 1}{S_i} = 0$ then X_1 has distribution of the form $p_1 = (1-p_0)\left(1 - \frac{S_0-1}{S_1}\right)$, $p_k = (1-p_0)\prod_{j=1}^{k-1}\frac{S_{j-1}-1}{S_j}\left(1 - \frac{S_{k-1}-1}{S_k}\right)$, $k = 2, 3, ...$ for arbitrary $0 < p_0 < 1$ $(p_j = P(X_1 = j))$ if and only if $E(X_{U(n+1)} - X_{U(n)} \mid X_{U(n)} = l) = S_l$, for some $n > 0$ and all $l \geq n-1$ (Aliev (2000)). This result is generalization of the theorems Srivastava (1979) and Korwar (1984).

Similar results for continuous distributions are obtained in the papers of Nagaraja (1977), Blaquez and Rebollo (1997), Ahsanullah and Wesolowski (1998) and Demblinka and Wesolowski (1999).

2. MAIN RESULT

Theorem. Let $X_{U(n)}$ be the n-th record value and $c > 0$. A necessary and sufficient condition for a r.v. X_1 to have distribution of the form

$$p_k = P(X_1 = k) = (1 - p_0)\left(1 - \frac{2}{c}\right)\left(\frac{2}{c}\right)^{k-1}, k = 1, 2, 3, ...$$

$$(p_0 = P(X_1 = 0) \text{ is arbitrary } 0 < p_0 < 1)$$

is

$$E(X_{U(n+2)} - X_{U(n)} \mid X_{U(n)} = l) = c, \text{ for some } n > 0 \text{ and all } l \geq n - 1.$$

Remark 1. The characterization problems when the regression $E\{X_{U(n+k)} - X_{U(n)} \mid X_{U(n)} = s\}$ is not constant for $k = 2$ and the case $k > 2$ are steel open.

Remark 2. The proof of the theorem is based on the technique of the paper Aliev (2000) and may be extended for the case $k = 2$, when $E\{X_{U(n+k)} - X_{U(n)} \mid X_{U(n)} = s\}$ is not constant.

References

[1] Ahsanullah, M. (1995). *Record Statistics,* Nova Science Publishers, Inc.,Commack, NY.

[2] Ahsanullah, M. and Holland, B. (1984). Record values and the geometric distribution. *Statistische Helfe*, 25, 319-327.

[3] Ahsanullah, M. and Wesolowski, J. (1998). Linearity of best predictors for non-adjacent record values. *Sankhya*, 60, B, 221-227.

[4] Aliev, F.A. (1998). Characterization of discrete distributions through weak records. *Journal of Applied Statistical Science*, 8, N. 1, 13-16.

[5] Aliev, F.A. (1999). New characterization of discrete distributions through weak records. *Theory Prob. Appl.*, 44, No 4, 756-761.

[6] Aliev, F.A. (2000). Characterization of discrete distributions through record values. *Proc. Int. Conference on Statistics in 21st Century,* University of Maine, Orono, Maine, USA.

[7] Blaquez, F.L. and Rebollo J.L.M. (1997). A characterization of distributions based on linear regression of order statistics and record values. *Sankhya*, 59, A, 311-323.

[8] Dembinska A, Wesolowski J. (2000). Linearity of regression for non-adjacent record values. *Journal of Stat. Plan. and Infer.*, 90, (2) 195-205.

[9] Kirmani, S.N.U.A. and Beg, M.I. (1984). On characterization of distributions by expected records. *Sankhya*, 49, A, 463-465.

[10] Korwar, R.M. (1984). On characterizing distributions for which the second record value has a linear regression of the first. *Sankhya*, 46, B, 108-109.

[11] Nagaraja, H.N. (1977). On characterization based on record values. *Australian Journal of Statistics*, 19, 70-73.

[12] Nagaraja, H.N. (1988). Some characterizations of continuous distributions based on regressions of adjacent order statistics and record values. *Sankhya*, A, 50, 70-73.

[13] Nevzorov, V.B. (1987). Records. *Theory Prob. Appl.*, 32, 201-228.

[14] Stepanov, A.V. (1994). A characterization theorem for weak records. *Theory Prob. Appl.*,39, 762-764.

[15] Srivastava, R.C. (1979). Two characterization of the geometric distribution by record values. *Sankhya*, 40, B, 276-278.