Sankhyā : The Indian Journal of Statistics
2005, Volume 67, Part 4, pp 715-721

On Characteristic Properties of the Uniform Distribution

G. Arslan
Başkent University, Turkey
M. Ahsanullah
University of South Florida, USA
I.G. Bayramoglu (Bairamov)
İzmir University of Economics, Turkey

Abstract

In this paper two characterizations of the uniform distribution using record values will be considered. The first characterization is based on the relation $X_{U(m)} - X_{U(m-1)} \stackrel{d}{=} X_{L(m)}$, $m > 1$, where $X_{U(m)}$ and $X_{L(m)}$ denote the m-th upper and lower record values, respectively. The second characterization involves the second record range.

AMS (2000) *subject classification.* Primary 62E10, 62G30.
Keywords and phrases. Characterization, uniform distribution, Choquet-Deny Theorem, records, record range.

1 Introduction

Let $X_1, X_2, ..., X_n$ be a set of independent and identically distributed (iid) random variables and denote the corresponding order statistics by $X_{1:n}, X_{2:n}, ..., X_{n:n}$. For various properties of order statistics one may refer to Ahsanullah and Nevzorov (2002), Arnold et al. (1992) and David (1981).

Let $\{X_i, i = 1, 2, ...\}$ be an infinite sequence of iid random variables. The lower record values of this sequence can be defined in the following way. Let $Y_1 = X_1$ and $Y_n = min\{X_1, ..., X_n\}$ for $n > 1$. Then $X_j, j > 1$ is called a *lower record value* of the sequence $\{X_i\}$ if $Y_j < Y_{j-1}$. Upper record values are defined similarly. We will denote by $X_{L(m)}$ and $X_{U(m)}$ the m-th lower and upper record values, respectively. More details on records can be found in Ahsanullah (1995), Arnold et al. (1998) and Nevzorov (2001), among others.

Several characterization results involving spacings can be found in the literature. For example, Puri and Rubin (1970) investigated the relation $X_{1:1} \stackrel{d}{=} X_{2:2} - X_{1:2}$ and showed all possible distributions satisfying this relation. Gather (1989) considered a more general relation; $X_{j-i:n-i} \stackrel{d}{=} X_{j:n} - X_{i:n}$. Fisz (1958), on the other hand, showed that the independence of $X_{2:2} - X_{1:2}$ and $X_{1:2}$ is a characteristic property of the exponential distribution. There are also several characterization results involving spacings of record statistics. Tata (1969), for example, showed that the independence of $X_{U(2)} - X_{U(1)}$ and $X_{U(1)}$ is a characteristic property of the exponential distribution. Bairamov and Aliev (1998) showed that the sequence $\left\{E\left(X_{U(n)} - X_{U(n-1)}\right), n \geq N\right\}$, for some $N \geq 0$, characterizes the distribution function of X_1. Ahsanullah (1991), on the other hand, showed that if $E\left(X_{U(n)} - X_{U(m)}\right) \stackrel{d}{=} E\left(X_{U(n-m)}\right), n > m$, then the random variable X_1 is exponentially distributed. The purpose of this paper is to investigate the relation $X_{U(m)} - X_{U(m-1)} \stackrel{d}{=} X_{L(m)}$ and to present another characterization based on the second record range for characterizing the uniform distribution.

To prove our main results we need a variant of the Choquet-Deny Theorem (Fosam and Shanbhag (1997)). Therefore we restate the theorem for ready reference.

THEOREM 1.1. *Let p be a positive integer and A be a non-empty subset of $[0,\infty)^p \setminus \{\boldsymbol{0}\}$ with the property that $\boldsymbol{x} \in A$ implies $[\boldsymbol{0},\boldsymbol{x}] \setminus \{\boldsymbol{0}\} \subset A$, where $\boldsymbol{0} = (0,\ldots,0)$. Also, let for each $x \in A$, $B_{\boldsymbol{x}} = [\boldsymbol{0},\boldsymbol{x}] \setminus (\{\boldsymbol{x}\} \bigcup \{\boldsymbol{0}\})$, and $\{\mu_{\boldsymbol{x}} : \boldsymbol{x} \in A\}$ be a family of probability measures (on $\mathbf{R}$) such that for each $\boldsymbol{x}$, $\mu_{\boldsymbol{x}}$ is concentrated on $B_{\boldsymbol{x}}$. Then a continuous real-valued function H on A such that $H(\boldsymbol{x})$ has a limit as $\|\boldsymbol{x}\|$ tends to $0+$, satisfies*

$$H(\boldsymbol{x}) = \int_{B_{\boldsymbol{x}}} H(\boldsymbol{x}-\boldsymbol{y})\mu_{\boldsymbol{x}}(d\boldsymbol{y}), \ \boldsymbol{x} \in A$$

if and only if it is identically equal to a constant.

The following corollary of Theorem 1.1 is used in the proofs of the results given in the next section.

COROLLARY 1.1. *Let $A = (0,\beta)$ and for each $x \in A$, $B_x = (0, \beta - x)$. Also, let $\{\mu_x : x \in A\}$ be a family of probability measures as in Theorem 1.1. Then a continuous real-valued function H on A such that $H(x)$ has a limit as x tends to $\beta-$, satisfies*

$$H(x) = \int_{B_x} H(x+y)\mu_x(dy), \ x \in A$$

if and only if it is identically equal to a constant.

Let $\{X_i, i = 1, 2, ...\}$ be an infinite sequence of iid nonnegative random variables. If the iid sequence of random variables is distributed uniformly on $(0, \beta)$, that is $X_i \sim \mathrm{U}(0, \beta)$, then it can be shown that

$$X_{U(m)} - X_{U(m-1)} \stackrel{d}{=} X_{L(m)}\ ,\ m > 1. \tag{1.1}$$

See, for example, Bairamov and Eryılmaz (2001)

We begin with a special case of relation (1.1); the case when $m = 2$. We will write $F \in \mathbf{F}_P$ if F is a member of the following family:

$$F(x; \alpha, \beta, \gamma) = 1 - \beta^{-\gamma} (\alpha + \beta - x)^{\gamma}\ ,$$

where $\alpha < x < \alpha + \beta,\ 0 < \beta,\ 0 \leq \alpha,\ 0 < \gamma.$

Let $\{X_i, i = 1, 2, ...\}$ be an infinite sequence of iid nonnegative random variables with absolutely continuous distribution function $F \in \mathbf{F}_P$. Then it can be shown that

$$X_{U(2)} - X_{U(1)} \stackrel{d}{=} X_{L(2)},$$

if and only if $F \sim \mathrm{U}(0, \beta)$.

2 Results

Under the assumption of symmetry, we have the following theorem.

THEOREM 2.1. *Let $\{X_i, i = 1, 2, ...\}$ be an infinite sequence of iid nonnegative random variables with absolutely continuous distribution function F. In addition assume that the random variables are symmetric about $\beta/2$. Then*

$$X_{U(m)} - X_{U(m-1)} \stackrel{d}{=} X_{L(m)}$$

for a fixed $m \geq 2$, it follows that $F \sim U(0, \beta)$.

PROOF. Denoting the probability density function (pdf) of $X_{U(m)} - X_{U(m-1)}$ by f_m, we have

$$f_m(x) = \int_0^{\beta - x} \frac{[R(y)]^{m-2}}{\Gamma(m-1)} r(y) f(x+y) dy,\ x \in A,$$

where $R(y) = -\ln[1 - F(y)]$, $r(y) = \frac{dR}{dy}$ and $A = (0, \beta)$. On the other hand the pdf of $X_{L(m-1)}$ is given by

$$f_{X_{L(m)}} = f(x) \frac{[H(x)]^{m-1}}{\Gamma(m)}\ ,$$

where $H(x) = -\ln F(x)$. Equating the last two relations, we obtain

$$f(x)\frac{[H(x)]^{m-1}}{\Gamma(m)} = \int_0^{\beta-x} \frac{[R(y)]^{m-2}}{\Gamma(m-1)} r(y) f(x+y)dy, \ x \in A.$$

Since the random variables are symmetric about $\beta/2$, we have

$$H(x) = -\ln F(x) = -\ln\left(\bar{F}(\beta - x)\right) = R(\beta - x),$$

where $\bar{F}(x) = 1 - F(x)$. Hence, on simplification,

$$f(x) = \int_0^{\beta-x} \frac{d\,[R(y)]^{m-1}}{[R(\beta-x)]^{m-1}} f(x+y) \tag{2.1}$$

is obtained.

Let μ_x be defined by

$$\mu_x(B) = \int_{B \cap B_x} \frac{d\,[R(y)]^{m-1}}{[R(\beta-x)]^{m-1}},$$

where $B_x = (0, \beta - x)$. Now, equation (2.1) can be written as

$$f(x) = \int_{B_x} f(x+y)\mu_x(dy), \ x \in A.$$

Applying Corollary 1.1 it follows that $f(x)$ must be a constant on A and that $F \sim \mathrm{U}(0, \beta)$. □

The second result is based on a relation involving the second record range. Before stating the result some basic information on record ranges will be reviewed.

Let $\{X_i, i = 1, 2, ...\}$ be a sequence of iid random variables with a common distribution function $F(x)$. Assume that $F(x)$ is absolutely continuous (with respect to Lebesgue measure) with pdf $f(x)$.

Now suppose that R_n^l and R_n^s are the largest and the smallest observations, respectively, when observing the n-th lower or upper record of the sequence $\{X_i, i = 1, 2, ...\}$. One may also interpret R_n^l and R_n^s as the current upper and lower record values of the $\{X_i, i = 1, 2, ...\}$ sequence when the n-th record of any kind (either lower or upper record) is observed. The n-th *record range*, then, is defined as $W_n^r = R_n^l - R_n^s$, $n > 1$, and the joint pdf of R_n^l and R_n^s is given by (see Arnold et al., 1998, p. 275)

$$f_{R_n^l, R_n^s}(x, y) = \frac{2^{n-1}}{(n-2)!}\left[-\ln\left(\bar{F}(y) + F(x)\right)\right]^{n-2} f(x)f(y),$$
$$-\infty < x < y < \infty.$$

The pdf of the n-th record range $f_{W_n^r}$ of W_n^r is given by

$$f_{W_n^r}(w) = \int_{-\infty}^{\infty} \frac{2^{n-1}}{(n-2)!}\left[-\ln\left(\bar{F}(w+u) + F(u)\right)\right]^{n-2} f(w+u)f(u)du.$$

For the uniform distribution with $F(x) = x$, $0 < x < 1$, we have

$$f_{W_n^r}(w) = \frac{2^{n-1}}{(n-2)!}(1-w)\left[-\ln(1-w)\right]^{n-2}, \; 0 < w < 1.$$

Thus W_n^r is distributed as the $(n-1)$-th, $(n \geq 2)$ upper record value from a sequence of iid random variables distributed like the minimum of two independent uniform $U(0,1)$ random variables. In the next theorem it will be shown that this is a characteristic property for the uniform distribution when $n = 2$.

THEOREM 2.2. *Suppose that $\{X_i, i = 1, 2, ...\}$ is a sequence of iid absolutely continuous random variables with a common distribution function $F(x)$. In addition assume that the random variables X_i are symmetric about 1/2. Then the following two statements are equivalent.*

(a) X_1 *has the uniform distribution with* $F(x) = x$, $0 < x < 1$,

(b) W_2^r *has the pdf* $2\overline{F}(x)f(x)$, $0 < x < 1$.

PROOF. It is easy to show that (a) $\Rightarrow$ (b). We will prove here that also (b) $\Rightarrow$ (a) holds.

Since

$$f_{W_2^r}(x) = \int_0^{1-x} 2f(x+y)f(y)dy,$$

from the assumption of the theorem we have

$$2\bar{F}(x)f(x) = \int_0^{1-x} 2f(x+y)f(y)dy, \; 0 < x < 1,$$

or

$$f(x) = \frac{1}{\bar{F}(x)} \int_0^{1-x} f(x+y)f(y)dy, \; 0 < x < 1.$$

This equation can be written as

$$f(x) = \int_{B_x} f(x+y)\mu_x(dy),$$

where

$$\mu_x(B) = \int_{B\cap B_x} \frac{dF(y)}{\bar{F}(x)},$$

and $B_x = (0, 1-x)$. Now, applying Corollary 1.1 it follows that $f(x)$ must be a constant on (0,1), implying that $F(x) = x$, $0 < x < 1$. □

REMARK 2.1. It is an open problem to characterize the uniform distribution by the distributional properties of W_n^r for $n > 2$.

Acknowledgements. The authors thank the referees for comments that helped to improve the presentation of this paper.

References

AHSANULLAH, M. (1991). Some characteristic properties of the record values from the exponential distribution. *Sankhyā Ser. B*, **53**, 403-408.

AHSANULLAH, M. (1995). *Record Statistics.* Nova Science Publishers, New York.

AHSANULLAH, M. and NEVZOROV, V. (2002). *Ordered Random Variables.* Nova Science Publishers, New York.

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.

BAIRAMOV, I.G. and ALIEV, F.A. (1998). On the characterization of distributions through order statistics and record values. *J. Appl. Statist. Sci.*, **7**, 249-253.

BAIRAMOV, I.G. AND ERYILMAZ, S. (2001). On properties of statistics connected with minimal spacing and record exceedances. *Appl. Statist. Sci.*, **V**, 245-254.

DAVID, H.A. (1981). *Order Statistics.* Wiley, New York.

FISZ, M. (1958). Characterization of some probability distributions. *Skand. Aktuarietidskr.*, **41**, 65-70.

FOSAM, E.B. and SHANBHAG, D.N. (1997). Variants of the Choquet-Deny theorem with applications. *J. Appl. Probab.*, **34**, 101-106.

GATHER, U. (1989). On a characterization of the exponential distribution by properties of order statistics. *Statist. Probab. Lett.*, **7**, 93-96.

NEVZOROV, V.B. (2001). *Records: Mathematical Theory.* Translations of Mathematical Monographs **194**. American Mathematical Society, Providence, R.I., USA.

Puri, S. and Rubin, H. (1970). A characterization based on the absolute difference of two iid random variables. *Ann. Math. Statist.*, **41**, 2113-2122.

Tata, M.N. (1969). On outstanding values in a sequence of random variables. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, **12**, 9-20.

G. Arslan
Başkent University
Department of Statistics
& Computer Science
06530 Ankara, Turkey
E-mail: guvenca@baskent.edu.tr

M. Ahsanullah
University of South Florida
Department of Mathematics
Tampa, Florida, 33620, USA
E-mail: ahsan@mail.cas.usf.edu

I.G. Bayramoglu (Bairamov)
İzmir University of Economics
35330 İzmir, Turkey
E-mail: ismihan.bayramoglu@izmirekonomi.edu.tr

Paper received June 2005; revised September 2005.