Pak. J. Statist.
2006 Vol. 22 (2) pp 139-146

WEAK RECORDS OF GEOMETRIC DISTRIBUTION AND SOME CHARACTERIZATIONS

Omar Hijab
Department of Mathematics
Temple University
Philadelphia, PA, USA

and

M. Ahsanullah
Department of Mathematics
University of South Florida
Tampa,Florida,USA

ABSTRACT

Suppose $\{x_n, n \geq 1\}$ is a sequence of independent and identically distributed discrete random variables having the common distribution function F(x). Some distributional properties of the weak records under the assumption that F(x) has the geometric distribution are presented. Various properties of the weak record values and some characterizations of the geometric distribution are presented.

Key Words and Phrases Difference equation, negative binomial distribution, Equality in distribution.

1. INTRODUCTION

Let X_1, X_2, …be a sequence of independent and identically distributed random variables taking values on 0,1,….with distribution function F such that F(0) =0 and a=sup{x|F(x) <1}=∞.F(n) <1 for any n. The weak upper record times $U_w(n)$ are defined as follows:

$U_w(1)=1$
$U_w(n+1) = \min \{j> U_w(n) , X_j \geq \max (X_1, X_2, \ldots X_{j-1})\}$

and the corresponding weak upper record value is defined as $X_{U_w(n+1)}$.

The joint probability mass function (pmf) of $X_{U_w(1)}, X_{U_w(2)}, \ldots, X_{U_w(n)}$ is given by

$$P_{w,1,2,\ldots,n}(x_1, x_2, \ldots, x_n) = \left(\prod_{i=1}^{n-1} \frac{p(x_i)}{\overline{P}(x_i)}\right) \; p(x_n), \text{ for } 0 \leq x_1 \leq x_2 \leq \ldots x_n < \infty.$$

Here $P(X=x_i) = p(x_i)$, $\overline{P}(x_i) = P(X \geq x_i) = p(x_i) + p(x_i+1) + \ldots$..

The marginal pmf's of the upper weak records are given by

$P(X_{U_w(1)} = x_1) = P_{w,1}(x_1) = p(x_1),\ x_1 = 0,1,2, \ldots, \ldots$

$P(X_{U_w(2)} = x_2) = P_{w,2}(x_2) = R_{w,1}(x_2)\, p(x_2),\ x_2 = 0,1,2,\ldots.$

where

$$R_{w,1}(x_2) = \sum_{0 \le x_1 \le x_2} \frac{p(x_1)}{\bar{P}(x_1)}$$

$P(X_{U_w(n)} = x_n) = P_{w,n}(x_n) = R_{w,n-1}(x_n)\, p(x_n)$, (1.1)

where

$$R_{w,n-1}(x_n) = \sum_{0 \le x_1 \le x_2 \le \ldots x_n} \prod_{i=1}^{n-1} \frac{p(x_i)}{\vec{P}(x_i)}$$

The joint pdf of $X_{U_w(n)}$ and $X_{U_w(n-1)}$ is

$$P(X_{U_w(n)} = x_{n-1},\ X_{U_w(n-1)} = x_n)$$
$$= R_{w,n-1}(x_{n-1}) A_w(x_{n-1})\ \ p(x_n), 0 \le x_{n-1} \le x_n$$
$$= 0, \text{ otherwise,} \qquad (1.2)$$

where

$$R_{w,n-1}(x_{n-1}) \sum_{0 \le x_1 \le x_2 \le \ldots x_{n-1} < x_n} \prod_{i=1}^{n-1} \frac{p(x_i)}{\vec{P}(x_i)}, \quad A_w(x_{n-1}) = \frac{p(x_{n-1})}{\bar{P}(x_{n-1})}$$

$$P(X_{U_w(n)} = x_n \mid X_{U_w(n-1)} = x_{n-1}) = \frac{p(x_n)}{\bar{P}(x_{n-1})}, \quad 0 \le x_{n-1} \le x_n \qquad (1.3)$$

We say that the distribution function $F \in GE(p)$ (the geometric distribution) if the corresponding probability mass function (pmf) is of the following form

$P(X=k) = pq^k,\ k \in A_n,\ 0<p=1-q<1,\ A_0 = 0,1,2,\ldots$
$= 0$, otherwise. (1.3)

We choose to distinguish between GE(p) and the larger class of geometric having geometric tail. We write as $F \in p\ GT_s(p)$ if

$P(X=k) = cpq^x,\ x \in A_n$ (1.4)
$= 0$, otherwise.

Here c is such that $\sum_{j=n}^{\infty} cpq^x = 1$, *i.e.* $c = q^{-n}$.

In this paper we will present several distributional properties of $X_{U_w(n)}$ and some characterizations of GE(p) and GET(m,p).

2. DISTRIBUTIONAL PROPERTIES OF $X_{U_w(n)}$

$P(X_{U_w(1)} = x_1) = P_{w,1}(x_1) = p(x_1) = pq^{x_i}$, $x_1 = 0,1,2,\ldots$,….

$P(X_{U_w(2)} = x_2) = P_{w,2}(x_2) = R_{w,1}(x_2)\, p(x_2)$, $x_2 = 0,1,2,\ldots$.

where

$$R_{w,1}(x_2) = \sum_{0\le x_1\le x_2} \frac{p(x_1)}{\bar{P}(x_1)} = \sum_{0\le x_1\le x_2} p = (x_2+1)p \; .$$

Thus $P(X_{U_w(2)} = x_2) = P_{w,2}(x_2) = (x_2+1)p^2 q^{x_2}$, $x_2 = 0,1,2,\ldots$.

$P(X_{U_w(n)} = x_n) = P_{w,n}(x_n) = R_{w,n-1}(x_n)\, p(x_n)$,

where

$$R_{w,n-1}(x_n) = \sum_{0\le x_1\le x_2\le\ldots x_n} \prod_{i=1}^{n-1} \frac{p(x_i)}{\bar{P}(x_i)} = \frac{(x_n+1)\;(x_n+2)(x_n+3)\ldots.(x_n+n-1)}{(n-1)!}\, p^{n-1}$$

Now

$$P(X_{U_w(n)} = x_n) = P_{w,n}(x_n) = \frac{(x_n+1)\;(x_n+2)(x_n+3)\ldots.(x_n+n-1)}{(n-1)!}\, p^n q^{x_n}$$

$$= \binom{x_n+n-1}{n-1} p^n q^{x_n}\ ,\ ,\ x_n = 0,1,2,\ldots. \qquad (2.1)$$

It is evident that $X_{U_w(n)}$ has the negative binomial distribution. $X_{U_w(n)} \in NB(n,p)$. We say $X \in NB(m,p)$ if the pmf of X is as follows:

$$P(X=k) = \binom{k+n-1}{n-1} p^n q^k\ ,\ k=0,1,2,\ldots$$

Using the relation (2.10), we can write

$$X_{U_w(n)} \overset{d}{=} V_1 + V_2 + \ldots + V_n \qquad (2.2)$$

where V_i's are i.i.d. random variable with $P(V_i = x) = pq^x$, $x=0,1,2,\ldots$

The joint pmf of $X_{U_w(n)}$ and $X_{U_w(m)}$, n> m, is given by

$$P(X_{U_w(n)} = x_n, X_{U_w(m)} = x_m) = \binom{x_m+m-1}{m-1}\binom{x_n-x_m+n-m}{n-m-1} p^n q^{x_n}\ ,\ 0 \le x_m \le x_n < \infty$$

$= 0$, otherwise.

The conditional pmf of $X_{U_w(n)}$ given $X_{U_w(m)} = x_m$ is

$$P(X_{U_w(n)} = x_n \mid X_{U_w(m)} = x_m) = \binom{x_n - x_m + n - m - 1}{n-m-1} p^{n-m} q^{x_n - x_m}, \; 0 \le x_m \le x_n < \infty$$
$$= 0, \text{ otherwise.} \tag{2.3}$$

Let $Z_{n,m} = X_{U_w(n)} - X_{U_w(m)}$, then

$$P(Z_{n,m} = z \mid X_{U_w(m)} = x_m) = \binom{z+m-n-1}{n-m-1} p^{n-m} q^{z}, \; 0 \le z < \infty$$
$$= 0, \text{ otherwise.} \tag{2.4}$$

Thus $Z_{n,m}$ is independent of $X_{U_w(m)}$ and $P(Z_{n,m}=z)$ is given by (2.4).Further $Z_{n,m}$ and $X_{U_w(n-m)}$ are identically distributed. Using (2.3) we can write

$$X_{U_w(n)} \underline{\underline{d}} X_{U_w(n)} + Z_{n,m} \quad, n > m \tag{2.5}$$

Here $Z_{n,m}$ is distributed as NB(n-m,p).The following moment relations follow immediately

$E(X_{U_w(n)}) = np^{-1}q$

$Var(X_{U_w(n)}) = np^{-2}q$

$Cov(X_{U_w(n)}, X_{U_w(m)}) = Var(X_{U_w(n-m)}) = (n-m)p^{-2}q$

Let $S_{(n)} - = X_{U_w(n)} - \frac{np}{q}$, then the sequence $\{S_{(n)}, n=1,2,\ldots\}$ forms a martingale. Hence

$$E(S_{(n)} \mid S_{(1)},\ldots, S_{(n-1)}) = S_{(n-1)}. \tag{2.6}$$

3. CHARACTERIZATIONS

Using (2.6), we obtain

$$E(X_{U_w(n+1)} \mid X_{U_w(n)} = x) = \sum_{x_n = x}^{\infty} x_n\, pq^{x_n - x} = x + \frac{q}{p} \tag{3.1}$$

and

$$E(X_{U_w(n+2)} \mid X_{U_w(n)} = x) = \sum_{x_n = x}^{\infty} x_n\, pq^{x_n - x} = x + \frac{2q}{p} \tag{3.2}$$

There are several characterizations of the GE(p) distribution using the regression relations given by (2.6) and (2.7). See for details see Stepanov(1994), Aliev (1998) and Wesolowski and Ahsanullah (2001). Recently Gupta and Ahsanullah (2004) has proved the following related characterization result.

Result 1. If φ(x) is a continuous function, then $E(\varphi(X_{U_w(n+2)} \mid X_{U_w(n)} = x) = s(x)$ uniquely determines the distribution

For example if s(j) =j+b and φ(j)=j, then for k= 0,1,2,…,$P(x=k) = \frac{2}{2+b}(\frac{2}{2+b})^k$, b>0.

It is an open problem whether the regression property

$$E(X_{U_w(n+k)} \mid X_{U_w(n)} = x) = x + \frac{kq}{p}, \text{ k>2,}$$

is a characterization property of the GE(p).

We have seen that $Z_{n,m}$ and $X_{U_w(n)}$ are independent. Is it a characteristic property of the geometric distribution? As a partial answer to the question, we have the following theorem for m = n-1.

Theorem 2.1. Let $\{X_n, n \geq 1\}$ be a sequence of independent and identically distributed random variables with common distribution function F. Suppose X is concentrated on 0,1,2,…and a = {x|F(x)<1}=∞.Then $X_j \in GET(n,p)$ iff $Z_{n,n-1}$ and $X_{U_w(n-1)}$ are independent.

Proof. We have seen that $Z_{n,n-1}$ and $X_{U_w(n-1)}$ are independent if $X_j \in GE(p)$. We will proof hre the "if" part. From (1.3) , we have

$$P(Z_{n,n-1} = u \mid X_{U_w(n-1)} = x) = \frac{p(u+x)}{\bar{P}(x)}., \ 0 \leq x,u < \infty \tag{3.3}$$

Since $Z_{n,n-1}$ and $X_{U_w(n-1)}$ are independent, we can take

$$c(u) = \frac{p(u+x)}{\bar{P}(x)} = \frac{\bar{P}(u+x) - \bar{P}(u+x+1)}{\bar{P}(x)} \tag{3.4}$$

Summing both sides of (3.4) with respect to u from 0 to $u_0 - 1$ ($u_0 > 1$) and writing

$c_1(u_0) = \sum_{u=0}^{u_0-1} c(u)$ and $c_0(u_0) = 1 - c_1(u_0)$, we get

$$\bar{P}(x+u) = c_0(u) \bar{P}(x), \ x \in A_1 \text{ and } u \in A_0. \tag{3.5}$$

The general solution of (3.5) is

$\bar{P}(x) = cp^x, \quad x \in A_1$ (3.6)

where c is independent of p. Using the boundary condition, $\bar{P}(\infty) = 0$, we obtain

$$\bar{P}(x) = cp^x, \quad x \in A_1, \ 0<p<1. \tag{3.7}$$

The Theorem is proved.

We have seen that

$$X_{U_w(n+1)} \overset{d}{=} V_1 + V_2 + \ldots + V_n + 1 \tag{3.8}$$

where V_i's are i.i.d. random variable with $P(V_i = x) = pq^x$, x=0,1,2,…

From (3.8) it is evident that

$$X_{U_w(n+1)} \overset{d}{=} W_n + W \tag{3.9}$$

where W_n has the same distribution as sum of n i.i.d geometric random variable .and W is independent of W_n and has the same distribution as X_i's.

The following theorem is a characterization of GET(n,p) based on the property given in (3.9).

Theorem 2.2. Let $\{X_n, n \geq 1\}$ be a sequence of independent and identically distributed random variables with common distribution function F. Suppose X is concentrated on 0,1,2,…and a = {x|F(x)<1}=∞. Then $X_j \in$ GET(n,p) iff $X_{U_w(n+1)} \overset{d}{=} W_n + W$, where W_n has the same distribution as sum of n i.i.d geometric random variable .and W is independent of W_n and has the same distribution as X_i's.

Proof. We will give here a proof of the "if" condition.

The condition $X_{U_w(n+1)} \overset{d}{=} W_n + X$ will lead to the following functional equation

$$p(x) \sum_{0\le x_0 \le x_1 \le \ldots \le x_{n-1} \le x} \frac{p(x_0)}{\bar{P}(x_0)} \ \frac{p(x_1)}{\bar{P}(x_1)},,,, \frac{p(x_{n-1})}{\bar{P}(x_{n-1})}$$
$$= \sum_{u=0}^{x} \binom{x-u+n-1}{n-1} p^n q^{x-u} p(u). \tag{3.10}$$

Here 0<p<1, q=1-p and $\bar{P}(x) = 1-(p(0)+p(1)+\ldots+p(x-1))$.

To prove the result we will use the following two lemmas.

Lemma 1. If a$\geq$ o and b$\geq$0, then

$$\sum_{c=0}^{b+1}\binom{a+c}{a} = \binom{a+b+1}{a+1}$$

The proof is by induction on b. If b=0, this is clear, so we assume the result is true for b. Then by the basic property of binomial coefficient

$$\sum_{c=0}^{b+1}\binom{a+c}{a} = \binom{a+b+1}{a} + \sum_{c=0}^{b}\binom{a+c}{a}$$

$$= \binom{a+b+1}{a} + \binom{a+b+1}{a+1} = \binom{a+b+2}{a+1}.$$

This shows that the result is true for b+1 and completes the proof.

Lemma 2. If x $\geq$ 0 and n $\geq$ 1, then

$$\sum_{0\le x_0\le x_1\le\ldots\le x_{n-1}\le x} 1 = \binom{x+n}{n} = \sum_{u=0}^{x}\binom{x-u+n-1}{n-1}.$$

The second equality is an immediate consequence of Lemma 1. To establish the first equality, note that, by decomposing the $s_n(x)$on the left side according whether x_{n-1} =0,1,2,..,x, we obtain the recurrence formula

$s_n(x)= s_{n-1}(o)+ s_n(1)+\ldots+s_{n-1}(x)$ and $s_n(0) =1$.

By Lemma 1, $s_n(x)= \binom{x+n}{n}$ satisfies the same recursion and $\binom{n}{n}$=1. Hence the result.

Now we show that p(x) =pq^x is a solution of (3.10).If p(x) = pq^x, then $\bar{P}(x) = 1-(p(0)+p(1)+\ldots+p(x-1)) = 1-p\dfrac{1-q^x}{1-q} = q^x$ .Inserting the values of p and $\bar{P}$ into (3.10), we obtain the equality by Lemma 2.

Now we show that p(x) = pq^x is the unique solution of (3.10), and we do this by induction over x$\geq$ 0. Let p(x) be any solution of (3.10) and set f(x) =p(x) /pq^x for x$\geq$ 0.Fix an x>0 and assume p(k) =pq^k for k=0,1,2,..,x-1., then $\bar{P}(x) = q^x$. Inserting this into (3.10) and dividing by $p^{n+1}q^x$, we obtain

$$f(x) \times \sum_{0\le x_0\le x_1\le\ldots\le x_{n-1}\le x} f(x_o)\quad f(x_1)\ldots\quad f(x_{n-1}) = \binom{x+n}{n} + f(x)-1. \tag{3.11}$$

where we have used Lemma 1 again.

Because f(0)=f(1)=…=f(x-1) =1, (3.11) is a polynomial equation in one unknown f(x). Our goal is to show that f(x)=1.

If f(x) >1, then since $f(x_i) \geq 1$ and f(x) >1 in (3.11), we obtain

$$f(x)\times \sum_{0\le x_0\le x_1\le\ldots\le x_{n-1}\le x} 1 > f(x)\times \sum_{0\le x_0\le x_1\le\ldots\le x_{n-1}\le x} 1$$
$$= f(x) \binom{x+n}{n} = \binom{x+n}{n} + \binom{x+n}{n}(f(x)-1)$$
$$\geq \binom{x+n}{n} + (f(x)-1).$$

This contradicts (3.10). Hence we cannot have f(x)>1.

If f(x) < 1, then since $f(x_i) < 1$ and f(x) <1 in (3.11), we obtain

$$f(x)\times \sum_{0\le x_0\le x_1\le\ldots\le x_{n-1}\le x} 1$$
$$< f(x) \binom{x+n}{n} = \binom{x+n}{n} + \binom{x+n}{n}(f(x)-1)$$
$$\leq \binom{x+n}{n} + (f(x)-1).$$

This contradicts (3.10), hence we can't have f(x)<1. This establishes f(x)=1. By induction f(x) =1 for all $x \geq 0$.

REFERENCES

1. Aliev, F.A.(1999). New characterizations of discrete distributions through weak records. *Theory of Probability and Applications*. 44, No. 4, 756-761.
2. Gupta, R.C. and Ahsanullah, M. (2004). Some characterization results based on the conditional expectation of a function of bon-adjacent order statistics (Record Value). To appear in *Ann. Inst. Stat. Math*.
3. Stepanov, A.V. (1994). A characterization Theorem for weak records. *Theory of probability and Applications* 38, 762-764.
4. Vervaat, W. (1973). Limit Theorems for Records from Discrete Distributions. *Stochastic Processes and Their Applications*. 1, 317-33
5. Wesolowski, J. and Ahsanullah, M. (2001). Linearity of regression for non-adjacent weak records. *Statistica Sinica*, 11, 30-52