

SOME CHARACTERIZATIONS OF HARRIS DISTRIBUTION

SANDHYA E^{*1} and LOVELY ABRAHAM T²

¹ Department of Statistics, Prajyoti Niketan College, Pudukkad - 680301.

² Department of Statistics, St.Thomas College, Pala - 686574

ABSTRACT

The fact that Harris distribution is a member of generalized power series distribution gives rise to the characterizations based on the sensitivity of the reliability function, mean and variance, infinite divisibility and series function. Some other distributions with gaps in this support are also characterized.

Key words and Phrases: *Harris distribution, generalized power series distribution, infinite divisibility, standard power series distribution.*

1 Introduction

Harris family of distributions defined by its probability generating function (p.g.f) $P(s) = \frac{s}{(m-(m-1)s^k)^{\frac{1}{k}}}$, k , a positive integer and $m > 1$, is a generalization of geometric distribution on $\{1, 2, 3 \dots\}$ to which it reduces when $k = 1$. The speciality of this distribution is that its atoms are k integers apart or the probabilities are concentrated on the points $\{1, 1+k, 1+2k \dots\}$. Also these probabilities coincide with that of negative binomial distribution on $\{0, 1, 2 \dots\}$ with parameters $\frac{1}{m}$ and $\frac{1}{k}$. The distribution was first introduced by Harris (1948) in connection with a simple branching process where a particle either splits into $k+1$ identical particles or

^{*}email:esandhya@hotmail.com

remains the same during a short time interval Δt . We can also consider a Harris distribution with support in $\{a, a + k, a + 2k \dots\}$ $a > 0$ integer, as a generalization of the distribution with support on $\{1, 1 + k, 1 + 2k \dots\}$.

Harris distribution plays a key role in schemes with random sums in general and in branching processes and time series models in particular. (See, Satheesh (2002), Satheesh and Nair (2002a,b)). Satheesh and Sandhya (2005) have developed a time series model that has an inherent N -sum structure where N is Harris distributed. Satheesh and Nair (2004) have shown that a Harris-sum of Harris distribution is again Harris and they used this property to generalize Marshal-Olkin parametrizing scheme. Sandhya *et al.* (2008) have presented the Harris distribution as an appropriate marketing distribution for a specific manufacturing unit.

Again, parallel to the notion of infinite divisibility and geometric infinite divisibility generalizations of these concepts have also been worked out in many directions for example, see Gnedenko and Korolev (1996) and Stutel and Van Harn (2004). It has been shown in Sandhya *et.al* (2008) that Harris distribution is infinitely divisible.

It has also been shown that Harris distribution is a member of generalized power series distribution (GPSD), defined in Patil (1962), and Johnson *et al.* (1992). Sherly and Sandhya (2009) use this for the estimation of natural parameter and its k^{th} power. Also UMP and UMPU tests of the population means as well as the sampling distribution of the test statistics are discussed.

In this paper we characterize GPSD in the context of sensitivity of the reliability function corresponding to a system working in a random environment. Section 2 deals with the characterization of GPSD which includes all power series distributions (PSD) with gaps in their support. In section 3 Harris distribution is characterized (i) through a differential equation based on the ratio of two successive derivative of the series function (ii) through the ratio of two probabilities associated with two successive values in the range of the random variable and (iii) through the ratio of two consecutive factorial moments. Again Harris distribution is characterized based on its mean and variance and these results are included in section 4. Finally in section 5 Harris distribution is characterized based on infinite divisibility.

2 Characterization of GPSD

First we define a GPSD (See Patil (1962) and Johnson *et al.* (1992))

Definition 2.1. A random variable X on $\{a, a+k, a+2k, \dots\}$ is said to have a GPSD with parameter $\theta \geq 0$ if

$$Q_n(\theta) = P(X = a + nk) = \frac{c_n \theta^{a+nk}}{C(\theta)}, \quad n = 0, 1, 2, \dots, \quad (2.1)$$

Also $c_n > 0$ and $C(\theta) = \sum_{n=0}^{\infty} c_n \theta^{a+nk}$, $a > 0$, $k > 0$ are integers.

Without loss of generality we assume that $C(0) = 1$, writing in terms of η , where $\eta = \theta^k$ and dividing both numerator and denominator by θ^a we get

$$Q_n(\eta) = P(X = a + nk) = \frac{c_n \eta^n}{C(\eta)}, \quad n = 0, 1, 2, \dots$$

$$\text{and } C(\eta) = \sum_{n=0}^{\infty} c_n \eta^n, \quad Q(0) = 0. \quad (2.2)$$

We state without proof the following theorem of random variables taking values on $S = \{a, a+k, a+2k, \dots\}$, $a \geq 1, k \geq 1$, integers. The proof follows on the same line as the proof of theorem 2 in Sreehari and Pathak (1999). The advantage of this theorem is that it includes all power series distributions(PSD) with gaps of interval k in their support.

Theorem 2.1. If X has a GPSD given by (2.2) then

$$\frac{d}{d\eta} P(X > a + rk) = \frac{C^{(1)}(\eta)}{C(\eta)} P(X \leq a + rk) - \sum_{j=0}^{r-1} \frac{C^{(j+1)}(0)}{C^{(j)}(\eta)} Q_j(\eta) \quad r = 0, 1, 2, \dots \quad (2.3)$$

Or, equivalently

$$\frac{d}{d\eta} \sum_{j=0}^r Q_j(\eta) C(\eta) = C(\eta) \sum_{j=0}^{r-1} \frac{C^{(j+1)}(0)}{C^{(j)}(0)} Q_j(\eta). \quad (2.4)$$

Conversely assume that, if X is a non negative random variable with the probability distribution

$$P(X = a + nk) = Q_n(\eta), \quad \eta \geq 0 \quad \text{and} \quad Q_0(0) = 1,$$

$Q_n(\eta)$ satisfies (2.3) or (2.4) for some power series $C(\eta)$ positive, finite, differentiable function and $C(0) = 1$. Then X has a GPSD given by (2.2), where $c_n = \frac{C^{(n)}(0)}{n!}$. $C^{(n)}(0)$ denotes the values of the n^{th} derivative of $C(\eta)$ at $\eta = 0$, R.H.S of (2.4) is interpreted as zero for $r = 0$.

Remark 2.1.

(i) If X has binomial distribution with gaps, we have the probability mass function (p.m.f) as

$$P(X = a + rk) = \binom{n}{r} p^r q^{n-r}, 0 < p < 1, q = 1 - p, r = 0, 1, 2, \dots$$

i.e.,

$$P(X = a + rk) = \frac{\binom{n}{r} (\theta^k)^r}{(1 + \theta^k)^n}, r = 0, 1, 2, \dots, n$$

where $\theta^k = \frac{p}{1-p}$ by putting $\eta = \theta^k$ we have,

$$P(X = a + rk) = \frac{\binom{n}{r} \eta^r}{C(\eta)}, r = 0, 1, 2, \dots, n,$$

where $C(\eta) = \sum_0^n \binom{n}{r} \eta^r = (1 + \eta)^n$

$$P(X = a + rk) = \binom{n}{r} \left(\frac{\eta}{1 + \eta} \right)^r \left(\frac{1}{1 + \eta} \right)^{n-r}. \quad (2.5)$$

Then the equation (2.3) reduces to

$$\frac{d}{d\eta} P(X > a + rk) = \frac{n-r}{1+\eta} P(X = a + rk).$$

(ii) If X has Poisson distribution with gaps it has p.m.f

$$P(X = a + rk) = \frac{e^{-\lambda} \lambda^r}{r!}, r = 0, 1, 2, \dots$$

putting $\lambda = \eta$,

$$P(X = a + rk) = \frac{\eta^r}{C(\eta)r!}, r = 0, 1, 2, \dots$$

where

$$C(\eta) = \sum \frac{\eta^r}{r!} = e^\eta. \quad (2.6)$$

For Poisson with gaps

$$\frac{d}{d\eta}P(X > a + rk) = P(X = a + rk).$$

(iii) We consider generalized Harris distribution with support on $\{a, a + k, a + 2k, \dots\}$ with series function $C(\theta) = \theta^a (1 - \theta^k)^{-\frac{a}{k}}$, putting $\beta = -\frac{a}{k}$ in Sherly and Sandhya (2009)

$$P(X = a + rk) = \frac{\binom{\frac{a}{k} + r - 1}{r} \theta^a (\theta^k)^r}{\theta^a (1 - \theta^k)^{-\frac{a}{k}}}, \quad a > 0, k > 0 \text{ are integers}$$

Putting $\theta^k = \eta$ and canceling common terms the p.m.f reduces to

$$P(X = a + rk) = \frac{\binom{(a/k) + r - 1}{r} \eta^r}{(1 - \eta)^{-\frac{a}{k}}}, \quad 0 \leq \eta \leq 1 \quad r = 0, 1, 2, \dots, \quad (2.7)$$

is represented by $H_a(\eta, k, \frac{a}{k})$. Here $C(\eta) = (1 - \eta)^{-\frac{a}{k}} = \sum \binom{(a/k) + r - 1}{r} \eta^r$. Then

$$\frac{d}{d\eta}P(X > a + rk) = \frac{a + rk}{k(1 - \eta)}P(X = a + rk).$$

(iv) If X has a geometric distribution with gaps its p.m.f

$$P(X = a + rk) = q^r p, \quad r = 0, 1, 2, \dots \quad \text{putting } q = \eta$$

$$P(X = a + rk) = \frac{\eta^r}{C(\eta)}, \quad r = 0, 1, 2, 3, \dots,$$

$$C(\eta) = \sum \eta^r = (1 - \eta)^{-1}, \quad (2.8)$$

$$\begin{aligned} \frac{d}{d\eta}P(X > a + rk) &= r(1 - \eta)^{-1}P(X = a + rk) \\ &= rC(\eta)Q_r(\eta). \end{aligned}$$

3 Characterization in terms of series function

Wani and Lo (1986) characterized a subclass of PSD known as standard power series distribution (i) through a differential equation based on the ratio of two successive derivatives of the series function (ii) through the ratio of two probabilities associated with two successive values in the range of the random variable and (iii) through the

ratio of two successive factorial moments which includes familiar discrete distributions binomial, Poisson, logarithmic series distributions and negative binomial.

Consider the PSD with p.m.f $P(X = x) = \frac{c_x \theta^x}{C(\theta)}$, $x = 0, 1, 2, \dots$, the series function is $C(\theta) = \sum c_x \theta^x$. If $C(\theta)$ is the series function of a standard power series distribution then

$$\frac{C^{(x)}(\theta)}{C^{(x+1)}(\theta)} = a_x \theta + b_x, x = 0, 1, 2 \dots \quad (3.1)$$

$$a_x = \frac{a_1}{1 - (x-1)a_1}, b_x = \frac{b_1}{1 - (x-1)a_1} \text{ with} \quad (3.2)$$

$$a_1 = 1 - \frac{C'(0)C^{(3)}(0)}{\{C''(0)\}^2}, b_1 = \frac{C'(0)}{C''(0)}.$$

Now we state the following theorems (without proof) for the subclass of GPSD parallel to the theorems given in Wani and Lo (1986).

Theorem 3.1. *The GPSD is a standard generalized power series distribution if and only if the series function satisfies the following relation*

$$\frac{C^{(1)}(\eta)}{C^{(2)}(\eta)} = a_1 \eta + b_1, \quad (3.3)$$

with the value of a_1 determines the exact functional form of the standard generalized power series distribution.

This leads to the following characterizations of Harris distribution,

Theorem 3.2. *The Harris distribution is a generalized standard power series distribution if and only if the series function satisfies the following relation*

$$\frac{C^{(1)}(\eta)}{C^{(2)}(\eta)} = a_1 \eta + b_1 = \frac{-k}{a+k} \eta + \frac{k}{a+k}, \text{ here } a_1 = -b_1 \text{ and } -1 < a_1 < 0. \quad (3.4)$$

Proof. The necessary part is clear. We can prove the sufficient part from (3.4), $\frac{C^{(1)}(\eta)}{C^{(2)}(\eta)} = \frac{k}{a+k}(1-\eta)$. Put $C^{(1)}(\eta) = V(\eta)$, then $\frac{V^{(1)}(\eta)}{V(\eta)} = \frac{1}{\frac{k}{a+k}(1-\eta)}$, integrating both sides with respect to η we get $V(\eta) = k_1(1-\eta)^{\frac{-a+k}{k}}$, integrating again with respect to η we get

$$C(\eta) = -\frac{k_1}{-a/k}(1-\eta)^{\frac{-a}{k}} + k_2.$$

From the condition $C(0) = 1$ we get, $k_1 = \frac{a}{k}$ and $k_2 = 0$

$$C(\eta) = (1 - \eta)^{-\frac{a}{k}}.$$

Remark 3.1. *Following the same lines as in Wani and Lo (1986).*

(i) $0 < a_1 \leq 1$ for a binomial with gaps $C(\eta) = (1 + \eta)^n$ from (2.5),

$$\frac{C^{(1)}(\eta)}{C^{(2)}(\eta)} = \frac{1}{n-1}\eta + \frac{1}{n-1}. \quad (3.5)$$

(ii) $a_1 = 0$ yields a Poisson with gaps, $C(\eta) = e^\eta$ from (2.6).

$$\frac{C^{(1)}(\eta)}{C^{(2)}(\eta)} = 1 = a_1\eta + b_1. \quad (3.6)$$

(iii) $a_1 = -\frac{1}{2}$ for a geometric with gaps distribution $C(\eta) = (1 - \eta)^{-1}$ from (2.8).

$$\frac{C^{(1)}(\eta)}{C^{(2)}(\eta)} = \frac{1}{2}\eta + \frac{1}{2}. \quad (3.7)$$

In Dubey (1976) and Ord (1967) it is observed that power series distribution with p.m.f $f(x)$ the ratio $u_x = xf(x)/f(x-1) = c_1x + d_1$ is a linear function of x . Later in Wani and Lo (1986) it is shown that it is the characteristic property of standard power series distribution when $c_1 = \frac{-a_1\theta}{b_1}$, $d_1 = \frac{(1+2a_1)\theta}{b_1}$. These suggest the following theorem for GPSD.

Theorem 3.3. *A GPSD with probability function $f(x)$ is a generalized standard power series distribution if and only if $u_x = \frac{xf(x)}{f(x-1)} = c_1x + d_1$, $x = 2, 3, 4, \dots$ where $c_1 = \frac{-a_1\eta}{b_1}$, $d_1 = \frac{(1+2a_1)\eta}{b_1}$.*

We have the following theorem,

Theorem 3.4. *The Harris distribution with probability function $f(x) = \frac{(\frac{a}{k} + x - 1)\eta^x}{(1-\eta)^{-a/k}}$ is a standard generalized power series distribution if and only if $u_x = \frac{xf(x)}{f(x-1)} = \eta x + (\frac{a}{k} - 1)\eta = c_1x + d_1$, $x = 2, 3, 4, \dots$*

Proof. As Harris distribution belongs to GPSD we have

$$\begin{aligned} f(x) &= \frac{C_x \eta^x}{C(\eta)} = \frac{\left(\frac{a}{k} + x - 1\right) \eta^x}{C(\eta)}, x = 0, 1, 2, \dots \\ \text{and } u_x &= \frac{x f(x)}{f(x-1)} = \frac{x C_x \eta}{C_{x-1}}. \end{aligned} \quad (3.8)$$

Now we have,

$$\frac{C_x}{C_{x-1}} = \frac{1 - \frac{k}{a+k}(x-2)}{\frac{k}{a+k}x}.$$

Substituting in (3.8) we get

$$\begin{aligned} u_x &= \frac{\frac{k}{a+k} \eta x + \left(1 - 2 \frac{k}{a+k}\right) \eta}{\frac{k}{a+k}} \\ &= \eta x + \left(\frac{a}{k} - 1\right) \eta = c_1 x + d_1. \end{aligned}$$

Conversely suppose that $u_x = c_1 x + d_1 = \eta x + \left(\frac{a}{k} - 1\right) \eta$. Then from (3.8) we have,

$$\begin{aligned} \frac{x C_x}{C_{x-1}} &= \frac{u_x}{\eta} = \frac{c_1}{\eta} x + \frac{d_1}{\eta} \\ &= c_2(x-2) + d_2, \end{aligned}$$

where $c_2 = \frac{c_1}{\eta}$, $d_2 = \frac{2c_1}{\eta} + \frac{d_1}{\eta}$.

Multiplying both sides by $(x-1)$ and rearranging

$$(x-1)C_{x-1} = C_x x(x-1)d_3 + C_{x-1}(x-1)(x-2)c_3, \quad (3.9)$$

where $c_3 = \frac{-c_2}{d_2}$, $d_3 = \frac{1}{d_2}$.

Now,

$$C^{(1)}(\eta) = \sum_{x=1}^{\infty} x C_x \eta^{x-1} = \sum_{x=2}^{\infty} (x-1) C_{x-1} \eta^{x-2} \quad (3.10)$$

Also,

$$\begin{aligned} C^{(2)}(\eta) &= \sum_{x=2}^{\infty} x(x-1) C_x \eta^{x-2} \\ &= \sum_{x=3}^{\infty} (x-1)(x-2) C_{x-1} \eta^{x-3}. \end{aligned}$$

Substituting (3.9) in (3.10) we get

$$C^{(1)}(\eta) = (c_3\eta + d_3) C'''(\eta)$$

$$\frac{C^{(1)}(\eta)}{C^{(2)}(\eta)} = \frac{-k}{a+k}\eta + \frac{k}{a+k}.$$

This implies that Harris distribution is standard generalized power series distribution.

Ottestad (1939) observed that the ratio of factorial moments of order r and $r-1$ is a linear function of r that is, $\frac{\mu_{(r)}}{\mu_{(r-1)}} = c_2r + d_2$. In Wani and Lo (1986), it is shown that it characterizes a standard generalized power series distribution by the relation

$$c_2 = \frac{-a_1\theta}{a_1\theta + b_1}, \quad d_2 = \frac{(1 + 2a_1)\theta}{a_1\theta + b_1}.$$

So we can state the following theorem for GPSD,

Theorem 3.5. *A GPSD is a standard generalized power series distribution if and only if*

$$V_r = \frac{\mu_{(r)}}{\mu_{(r-1)}} = c_2r + d_2 \quad (3.11)$$

where

$$c_2 = \frac{-a_1\eta}{a_1\eta + b_1}, \quad d_2 = \frac{(1 + 2a_1)\eta}{a_1\eta + b_1}.$$

So we have the following characterization for the Harris distribution is,

Theorem 3.6. *The Harris distribution is a standard generalized power series distribution if and only if*

$$V_r = \frac{\mu_{(r)}}{\mu_{(r-1)}} = r \left(\frac{\eta}{1-\eta} \right) + \left(\frac{a-k}{k} \right) \frac{\eta}{1-\eta} = c_2r + d_2 \quad (3.12)$$

where

$$c_2 = \frac{-a_1\eta}{a_1\eta + b_1}, \quad d_2 = \frac{(1 + 2a_1)\eta}{a_1\eta + b_1}.$$

Proof. The r^{th} factorial moment of GPSD

$$\mu_{(r)} = \frac{\eta^r C^{(r)}(\eta)}{C(\eta)} \quad (\text{from Johnson et al. (1992)}). \quad (3.13)$$

For Harris distribution, we have,

$$\begin{aligned} C^{(r)}(\eta) &= \frac{a}{k} \frac{a+k}{k} \cdots \frac{a+(r-1)k}{k} (1-\eta)^{-\frac{a}{k}-r} \\ \mu_{(r)} &= \frac{a}{k} \frac{a+k}{k} \cdots \frac{a+(r-1)k}{k} \frac{\eta^r}{(1-\eta)^r}. \end{aligned} \quad (3.14)$$

From (3.1) and (3.2),

$$\frac{C^{(r)}(\eta)}{C^{(r-1)}(\eta)} = 1/a_r\eta + b_r \text{ and } a_r = \frac{a_1}{1-(r-1)a_1}, \quad b_r = \frac{b_1}{1-(r-1)a_1}$$

Substituting in (3.14) we get

$$\begin{aligned} \frac{\mu_{(r)}}{\mu_{(r-1)}} &= \frac{\eta C^{(r)}(\eta)}{C^{(r-1)}(\eta)} \\ &= \frac{-a_1\eta r}{a_1\eta + b_1} + \frac{(1+2a_1)\eta}{a_1\eta + b_1} = c_2r + d_2. \end{aligned} \quad (3.15)$$

Also from (3.14)

$$V_r = \frac{\mu_{(r)}}{\mu_{(r-1)}} = \left(\frac{\eta}{1-\eta} \right) r + \left(\frac{a-k}{k} \right) \frac{\eta}{1-\eta}.$$

Conversely suppose that

$$V_r = \frac{\mu_{(r)}}{\mu_{(r-1)}} = \left(\frac{\eta}{1-\eta} \right) r + \left(\frac{a-k}{k} \right) \frac{\eta}{1-\eta} = c_2r + d_2.$$

Put $r = 2$ in V_r

$$\begin{aligned} \frac{C^{(2)}(\eta)}{C^{(1)}(\eta)} &= \frac{2}{1-\eta} + \frac{a-k}{k} \frac{1}{(1-\eta)} \\ \frac{C^{(1)}(\eta)}{C^{(2)}(\eta)} &= \frac{-k}{a+k}\eta + \frac{k}{a+k}. \end{aligned}$$

Now we find u_r and V_r ratios for some standard generalized power series distributions.

Remark 3.2. (i) *Binomial with gaps*

From (2.5),

$$\begin{aligned} u_r &= \frac{rf(r)}{f(r-1)} = \eta(n-r+1) = -\eta r + (n+1)\eta = c_1r + d_1 \\ \mu_{(r)} &= n(n-1)\cdots(n-r+1) \left(\frac{\eta}{1+\eta} \right)^r \end{aligned}$$

$$V_r = \frac{\mu_r}{\mu_{r-1}} = \left(\frac{-\eta}{1+\eta} \right) r + (n+1) \frac{\eta}{1+\eta} = c_2 r + d_2.$$

(ii) *Poisson with gaps*

From (2.6),

$$\begin{aligned} u_r &= \eta = c_1 r + d_1 \\ \mu_{(r)} &= \eta^r \\ V_r &= \frac{\mu_r}{\mu_{r-1}} = \eta = c_2 r + d_2. \end{aligned}$$

(iii) *Geometric with gaps*

From (2.8),

$$\begin{aligned} u_r &= \eta = c_1 r + d_1 \\ \mu_{(r)} &= r! \frac{\eta^r}{(1-\eta)^r} \\ V_r &= \frac{\mu_r}{\mu_{r-1}} = r\eta(1-\eta)^{-1} = c_2 r + d_2. \end{aligned}$$

4 Characterization based on mean and variance

Here we characterize Harris distribution in terms of its mean and variance, similar to a characterization of generalized negative binomial distribution as given in Ashanullah (1991). Let X follows a power series distribution with p.m.f define by (See, Ashanullah (1991))

$$P(X = x) = \begin{cases} \frac{b_x g(\theta)^x}{C(\theta)}, & x \in A_s \\ 0, & \text{otherwise.} \end{cases} \quad (4.1)$$

where $b_x > 0$, $C(\theta) = \sum_{x \in A_s} b_x g(\theta)^x$ and A_s is a subset of integers $\{s, s+1, s+2, \dots\}$, $s \rightarrow 0$.

Also mean and variance of r.v X with p.m.f (4.1)) is given by

$$\begin{aligned} \mu &= \frac{g(\theta)}{g'(\theta)} \frac{C'(\theta)}{C(\theta)} = \alpha_1(\theta) \text{ (say),} \\ \sigma^2 &= \eta \frac{d\mu}{d\theta} = \mu' \frac{g(\theta)}{g'(\theta)} = \alpha_2(\theta) \text{ (say).} \end{aligned} \quad (4.2)$$

Theorem 4.1. *In the class of GPSD defined on set of nonnegative integers, the p.m.f is of a Harris distribution $H_a(\theta, k, \frac{1}{k})$ iff*

$$\begin{aligned}\mu &= \alpha_1(\eta) = a + \frac{\theta^k}{1 - \theta^k} \\ \sigma^2 &= \alpha_2(\eta) = \frac{k \theta^k}{(1 - \theta^k)^2}, 0 < \theta < 1.\end{aligned}\tag{4.3}$$

Proof. The necessary part can be easily shown. We prove the sufficiency part.

From (4.3) we get,

$$\begin{aligned}\mu' &= \frac{d\mu}{d\theta} = \frac{k \theta^k}{\theta(1 - \theta^k)^2} \\ \alpha_2(\theta) &= \frac{\mu' g(\theta)}{g'(\theta)} = \alpha_1'(\theta) \cdot \frac{g(\theta)}{g'(\theta)} \\ \frac{g'(\theta)}{g(\theta)} &= \frac{\alpha_1'(\theta)}{\alpha_2(\theta)} = \frac{1}{\theta}.\end{aligned}\tag{4.4}$$

Hence $g(\theta) = \theta + c_1$, c_1 is arbitrary constant.

With the initial condition $g(0) = 0$, $c_1 = 0$.

Next is a characterization of Harris distribution based only on its mean. For a similar discussion see Ashanullah (1992).

Theorem 4.2. *If mean of GPSD defined on the subset of integers X_1 , where $X_1 = \{a, a + k, a + 2k, \dots\}$, $a, k > 0$ integers is $E(X) = a + \frac{\theta^k}{1 - \theta^k} = \theta \cdot f(\theta)$ and the parameter θ lies in the interval $b < \theta < c$, $-\infty < b < c < \infty$, where $f(\theta)$ is integrable over a finite interval $[a, u]$,*

$$\begin{aligned}\int_a^u f(\theta) d\theta &= F(u). \\ \text{If } e^{F(u)} &= t \sum_{x \in S} \alpha_x u^x\end{aligned}$$

$$\text{then } c_x = t \alpha_x \quad \text{and} \quad t^{-1} = \sum_{x \in S} \alpha_x (\theta^k)^x = C(\theta).$$

Proof. As Harris distribution belongs to GPSD,

$$\begin{aligned}
 E(X) &= \theta \cdot \frac{C'(\theta)}{C(\theta)} = a + \frac{\theta^k}{1 - \theta^k} \text{ from (4.3)} \\
 &= \theta \cdot \left(\frac{a}{\theta} + \frac{\theta^k}{(1 - \theta^k)\theta} \right) = \theta \cdot f(\theta) \\
 f(\theta) &= \frac{a}{\theta} + \frac{\theta^{k-1}}{(1 - \theta^k)} \\
 F(u) &= \int_a^u f(\theta) d\theta, b < \theta < c
 \end{aligned}$$

As we know that $0 < \theta < 1$

$$\begin{aligned}
 F(u) &= \int_0^u \frac{a}{\theta} + \frac{\theta^{k-1}}{1 - \theta^k} d\theta = \ln u^a (1 - u)^{-\frac{1}{k}} + t_1 \\
 e^{F(u)} &= t u^a (1 - u)^{-\frac{1}{k}} = t \sum \binom{1/k + x - 1}{x} u^{a+x} = t \sum \alpha_x u^{a+x}
 \end{aligned}$$

with $t = C(0)$

and $P(X = a + xk) = C(\theta)^{-1} c_x \theta^a (\theta^k)^x \quad x \in S,$

where $c_x = t \alpha_x$.

5 Characterization of Harris distribution based on infinite divisibility

Katti (1967) has shown that a necessary and sufficient condition for a r.v X to be infinitely divisible (i.d) is that

$$\pi_i = \frac{i p_i}{p_0} - \sum_{j=1}^{i-1} \pi_{i-j} \frac{p_j}{p_0} \geq 0 \text{ for } i = 1, 2, \dots \quad (5.1)$$

Theorem 5.1. A r.v X has the Harris distribution with parameters $H_a(\theta, k, \frac{1}{k})$ with p.g.f $\frac{C(\theta s)}{C(\theta)}$ iff

$$\pi_r = \frac{1}{k} (\theta^k)^r, \quad r = 1, 2, \dots \quad (5.2)$$

Proof. For Harris distribution $H_a(\theta, k, \frac{1}{k})$ we have,

$$P[X = a + rk] = \frac{\binom{1/k+r-1}{r}(\theta^k)^r}{C(\theta)}, r = 0, 1, 2, \dots$$

$$\pi_1 = \frac{p_1}{p_0} = \frac{1}{k} \theta^k$$

$$\pi_2 = \frac{2p_2}{p_0} - \pi_1 \frac{p_1}{p_0} = \frac{1}{k} (\theta^k)^2 \quad \text{and so on.}$$

$$\text{In general } \pi_r = \frac{1}{k} (\theta^k)^r, r = 1, 2, \dots$$

Conversely suppose (5.2) is true. Then from (5.1),

$$p_1 = \pi_1 p_0 = \binom{1/k}{1} \theta^k p_0$$

$$2p_2 = \binom{1/k+1}{2} (\theta^k)^2 p_0$$

$$3p_3 = \binom{1/k+2}{3} (\theta^k)^3 p_0$$

In general

$$\begin{aligned} rp_r &= \sum_{j=0}^{r-1} \pi_{r-j} p_j \\ &= (\theta^k)^r p_0 \sum_{j=0}^{r-1} \binom{1/k+j-1}{j} \\ &= \binom{1/k+r-1}{r-1} (\theta^k)^r p_0 \\ p_r &= \binom{1/k+r-1}{r} (\theta^k)^r p_0. \end{aligned}$$

Hence the proof.

Acknowledgement: The authors are highly thankful to the referee for some valuable comments.

References

- Ashanulla, M; Two characteristic properties of the generalized negative binomial distribution, *Biometrical Journal*, 33, 7, (1991), 867–864.
- Ashanulla, M; On some characteristic properties of one parameter power series distribution, *Pakistan Journal of Statistics*, Vol.8 A, (1992), 45–49.

- Dubey, S.D; Graphical tests for discrete distributions, *American Statistician*, 20, (1976), 23–25.
- Gnedenco, B.V and Korolev, V.Yu; *Random Summation-Limit theorems and Applications*, CRC Press Boca Raton, 1996.
- Harris, T.E; Branching Processes, *Annals of Mathematical Statistics*, 19, (1948), 474–494.
- Johnson, N. L., Kotz, S and Kemp, A.W; Univariate Discrete Distributions, *Wiley Series in Probability and Mathematical Statistics*, 2nd edition, (1992).
- Katti, S.K; Infinite divisibility of integer valued random variables, *Annals of Mathematical Statistics*, 38, (1967), 1306–1308.
- Patil, G.P; Certain properties of generalized power series distribution, *Annals of the institute of Statistical Mathematics*, 14, (1962), 179–182.
- Ord, J.K; Graphical methods for a class of discrete distributions, *Journal of Royal Statistical Society Series A*, 130, (1967), 232–238.
- Ottestad, P; On the use of the factorial moments in the study of discontinuous frequency distributions, *Skandinavisk Aktuarietidskrift.*, 22, (1939), 22–31.
- Sandhya, E., Sherly, S and Raju, N; Harris family of discrete distribution, *Some Innovations in Statistics, special volume in honor of Professor T S K, Moothathu*, University of Kerala, Trivadram, India, (2008), 57–72.
- Satheesh, S; Aspects of randomization in infinitely divisible and max-infinitely divisible laws, *Probability and Statistical Models*, 1, (2002), 7–6.
- Satheesh, S and Nair, N.U; A Note on maximum and minimum stability of certain distributions, *Calcutta Statistical Association Bulletin*, 53, (2002a) 249–252.
- Satheesh, S and Nair, N. U; Some classes of distributions on the non-negative lattice, *Journal of Indian Statistical Association*, 40, (2002b), 41–58.
- Satheesh, S and Nair, N. U; On the stability of geometric extremes, *Journal of Indian Statistical Association*, 42, (2004), 99–109.

- Satheesh, S and Sandhya, E; Semi-self decomposable laws and related processes, *Journal of Indian Statistical Association*, 43, (2005), 157–166.
- Sherly, S and Sandhya, E; Harris distribution as a member of Generalized power series family, *Journal of the Kerala Statistical Association*, 20, (2009), 12–22.
- Sreehari, M and Pathak, A.G; Sensitivity of a reliability function and a related characterization, *Stochastic modeling and application*, Vol. 2, No.2, (1999), 13–18.
- Steutel, F.W and Van Harn, K; *Infinite Divisibility of Probability Distributions on the Real Line*, Mercel Dekker, Inc. New York-Basel, 2004.
- Wani K. Jaganath and Hing Po Lo; Selecting a power series distribution for goodness of fit, *Canadian Journal of Statistics*, 14, No.4: (1986), 347–353.