ISSN 2249–4553

JOURNAL OF

THE KERALA STATISTICAL ASSOCIATION

Volume 24

December 2013

JOURNAL OF THE KERALA STATISTICAL ASSOCIATION

(A Publication of the Kerala Statistical Association)

The Editors invite original research papers, typed with wide margins (format as that of Sankhya) in three copies for possible publication in this journal. Though non-members also may submit papers, at the time of the publication of their paper, the Kerala Statistical Association wishes them to be its members. Each author of this journal will be given one free copy of the volume/number of the journal in which his/her paper appears unless advance order of reprints is made.

From the papers published in each volume of this journal the best paper is selected for granting **"Professor T.S.K. Moothathu Best Paper Award"** with a fellowship.

All comunications relating to the publication of papers and subscription of the journal are to be addressed to "Manoj Chacko (Chief Editor), Department of Statistics, University of Kerala, Trivandrum - 695 581", e-mail :manojchacko02@gmail.com.

The Editorial Board of the Journal of the Kerala Statistical Association is much grateful to , P. Yageen Thomas, Kerala University, Trivandrum; Philip Samuel, University College, Trivandrum; K. Jayakumar, Calicut University; and S. Minimol, Kerala Univerity for their valuable editorial collaboration rendered to the publication of this volume of the journal.

JOURNAL OF THE KERALA STATISTICAL ASSOCIATION
Volume 24, December 2013 (Reg. No. M - 6 - 11573/80)

CONTENTS

EDITORIAL OFFICE

DEPARTMENT OF STATISTICS
UNIVERSITY OF KERALA
KARIAVATTOM, TRIVANDRUM-695 581

Journal of the Kerala Statistical Association
Vol.24, December 2013, 1-11 *ISSN 2249-4553*

Identities Emerging from Two Sample U-Statistics

Sharada V. Bhat* and I.D. Shetty
Department of Statistics
Karnatak University, Dharwad

ABSTRACT

In this paper, we derive some mathematical identities which involve combinatorial coefficients. The well known theory of two-sample U Statistics is used to derive the identities.

Key words and Phrases: *U-Statistics, two-sample location problem, combinatorial coefficients, ordered ranks.*

1 Introduction

Identities are useful in simplifying many algebraic expressions. They provide simple alternate expressions to solve complex algebraic expressions. Riordan (1968) contains many such fundamental identities, Joshi and Balakrishnan (1981) provide statistical derivations of some such identities and Baiju and Thomas (2007) describe some identities using well established theories of order statistics and U-statistics based on certain linear functions of order statistics.

In this paper, we derive some identities using two-sample U-statistics and ordered ranks. The two-sample U-statistics is described in section 2 and two-sample

*email: bhat_sharada@yahoo.com

U-Statistics from which these identities emerge are described in section 3. The identities derived from mean and variance of these U-statistics is given in section 4 and identities emerging from ordered ranks is given is section 5.

2 Definition of Two-sample U-Statistics

U-Statistics is a class of unbiased estimators of parameters of a population. They are often used as test statistics although they emphasize estimation. Randles and Wolfe (1979) describes the two-sample U-Statistics as follows:

Let $X_1, \ldots, X_m$ and $Y_1, \ldots, Y_n$ be the observations of two independent random samples drawn from cumulative distribution functions (cdf) $F(x)$ and $G(y)$ respectively. A parameter θ is said to be estimable of degree (b, d) for distributions (F, G) in a family $\boldsymbol{F}$ if b and d are the smallest sample sizes for which there exists an estimator of θ that is unbiased for every $(F, G) \in \boldsymbol{F}$. That is, there is a function $h(. \, ; \, .)$ such that $E_{F,G}\left[h(X_1, ..., X_b; Y_1, ..., Y_d)\right] = \theta$ for every $(F, G) \in \boldsymbol{F}$, where $h(. \, ; \, .)$ is called as two-sample kernel and is symmetric in it's X_i components and separately symmetric in it's Y_j components. Under these conditions a two sample U-statistic, for $m \geq b$ and $n \geq d$ has the form

$$U(X_1, ..., X_m; Y_1, ..., Y_n) = \left(\binom{m}{b}\binom{n}{d}\right)^{-1} \sum_{\alpha} h(X_{i_1}, ..., X_{i_b}, Y_{j_1}, ..., Y_{j_d}) \, , \tag{2.1}$$

where $\sum_{\alpha}$ is the collection of all subsets of $b(d)$ integers chosen without replacement from the integers {1,. . . ,m} {1,. . . ,n}.

3 Some two sample U-Statistics for location problem

Suppose $X_1, \ldots, X_m$ and $Y_1, \ldots, Y_n$ are independent random samples from populations with absolutely continuous distribution functions F(x) and G(y) having probability density functions (pdf) $f(x)$ and $g(y)$ respectively. Then the two sample location problem is to test $H_0 : F(x) \equiv G(x)$ against the alternative$H_1 : G(x) =$

$F\,(x-\theta)$, θ>0 or θ<0 or $\theta\neq$0, -∞<x<∞, that is, two distributions differ only in their location parameter.

Suppose b and d are some fixed positive integers such that $1 \leq b \leq m$ and $1 \leq d \leq n$. For testing $H_0 : \theta = 0$ against$H_1 : \theta > 0$, Shetty and Bhat (1993) proposed

$$SB_1 = [\tbinom{m}{b}\tbinom{n}{d}]^{-1} \sum_{\alpha} \phi_1\left(X_{i_1}, ..., X_{i_b}; Y_{j_1}, ..., Y_{j_d}\right), \tag{3.1}$$

where $\sum_\alpha$is sum over all $\tbinom{m}{b}\tbinom{n}{d}$possible sub samples,

$$\phi_1(X_1, ..., X_b; Y_1, ..., Y_d) = \begin{cases} 1\ , & if \quad M_1 \leq M_2 \\ 0, & otherwise \end{cases}, \tag{3.2}$$

M_1= median of $(X_1, ..., X_b)$, M_2= median of $(Y_1, ..., Y_d)$, and b and d are odd positive integers.

Shetty and Bhat (1994) proposed

$$SB_2 = [\tbinom{m}{b}\tbinom{n}{d}]^{-1} \sum_{\alpha} \phi_2\left(X_{i_1}, ..., X_{i_b}; Y_{j_1}, ..., Y_{j_d}\right) \tag{3.3}$$

and

$$SB_3 = [\tbinom{m}{b}\tbinom{n}{d}]^{-1} \sum_{\beta} \phi_3\left(X_{i_1}, ..., X_{i_d}; Y_{j_1}, ..., Y_{j_b}\right), \tag{3.4}$$

where $\sum_\beta$is sum over all $\tbinom{m}{d}\tbinom{n}{b}$possible sub samples,

$$\phi_2(X_1, ..., X_b; Y_1, ..., Y_d) = \begin{cases} 1, & if \quad X_{(b)} \leq M_2 \\ 0, & otherwise \end{cases} \tag{3.5}$$

$$\phi_3(X_1, ..., X_b; Y_1, ..., Y_d) = \begin{cases} 1, & if \quad M_3 \leq Y_{(1)} \\ 0, & otherwise \end{cases} \tag{3.6}$$

M_3= median of $(X_1, ..., X_d)$, $X_{(b)}$= maximum of $(X_1, ..., X_b)$, $Y_{(1)}$= minimum of $(Y_1, ..., Y_b)$ and d is an odd positive integer.

Shetty et al.(1997) proposed a two-sample U-Statistic with kernel being the function of sample quantiles which is given by

$$SB_4 = [\tbinom{m}{b}\tbinom{n}{d}]^{-1} \sum_{\alpha} \phi_4\left(X_{i_1}, ..., X_{i_b}; Y_{j_1}, ..., Y_{j_d}\right) \tag{3.7}$$

and

$$\phi_4(X_1,...,X_b;Y_1,...,Y_d) = \begin{cases} 1 & if \quad X_{(k_1)b} \leq Y_{(k_2)d} \\ 0 & otherwise \end{cases} \tag{3.8}$$

where,

$$\text{k}_1 = \begin{cases} b\beta, & if \quad b\beta \quad \text{is an integer} \\ [b\beta]+1, & if \quad b\beta \quad \text{is not an integer ,} \end{cases}$$

$$\text{k}_2 = \begin{cases} d\gamma, & if \quad d\gamma \quad \text{is an integer} \\ [d\gamma]+1, & if \quad d\gamma \quad \text{is not an integer ,} \end{cases}$$

β^{th} quantile of a sample size n is defined as the r^{th}order statistic, where

$$r = \begin{cases} n\beta, & if \quad n\beta \quad \text{is an integer} \\ [n\beta]+1, & if \quad n\beta \quad \text{is not an integer ,} \end{cases}$$

$X_{(k_1)b}$= k_1^{th}order statistic of $(X_1,...,X_b)$ and $Y_{(k_2)d}$=k_2^{th} order statistic of $(Y_1,....,Y_d)$. All these two sample U-Statistics are expressed in alternative forms using ordered ranks and their properties are also studied. An extensive study of these statistics is carried out in Bhat (1996).

4 Identities from mean and variance of two sample U-Statistics

In this section, we present some identities and their proofs based on mean and variance of two sample U statistics defined in section 3.

Identity 4.1.

$$\left(d!/(q!)^2\right) \sum_{i=p+1}^{b} \binom{b}{i} B\left(i+q+1, b-i+q+1\right) = 1/2 \tag{4.1}$$

for b and d being odd positive integers.

Proof.

The mean of SB_1 under H_0 is obviously 1/2.

But it is also given by

$$\mathrm{E}\,(\mathrm{SB}_1) = \int_{-\infty}^{\infty} \mathrm{F}_{M_1}\,(x)\ \ \mathrm{d}\ \mathrm{F}_{M_2}\,(x)\,,$$

where

$$F_{M_1}(x) = \sum_{i=p+1}^{b} \begin{pmatrix} b \\ i \end{pmatrix} [F(x)]^i \ \ \left[\overline{F}\,(x)\right]^{b-i},$$

$$F_{M_2}(x) = \sum_{i=q+1}^{d} \begin{pmatrix} d \\ i \end{pmatrix} [F(x)]^i \ \ \left[\overline{F}\,(x)\right]^{d-i},$$

$$p = (b-1)/2, \qquad q = (d-1)/2 \qquad \text{and} \qquad \overline{F}\,(\mathrm{x}) \ = 1 - F\,(x)\ .$$

Therefore,

$$E\,(SB_1) = \int_{-\infty}^{\infty} \sum_{i=p+1}^{b} \begin{pmatrix} b \\ i \end{pmatrix} [F(x)]^i \ \ \left[\overline{F}\,(x)\right]^{b-i}(x)\ \ \left(d!/(q!)^2\right)\ [F(x)]^q \ \ \left[\overline{F}\,(x)\right]^q dF(x)$$

$$= \sum_{\mathrm{i=p+1}}^{\mathrm{b}} \begin{pmatrix} b \\ i \end{pmatrix} \left(d!/(q!)^2\right) \int_{-\infty}^{\infty} [F(x)]^{i+q} \ \ \left[\overline{F}\,(x)\right]^{b-i+q} dF\,(x)$$

$$= \left(d!/(q!)^2\right) \sum_{i=p+1}^{b} \begin{pmatrix} b \\ i \end{pmatrix} B\,(i+q+1,\, b-i+q+1)\,,$$

where $\mathrm{B}(x,y) = \Gamma\,(x)\ \ \Gamma\,(y)\ \ /\ \Gamma(x+y)$. Therefore, we get the identity (4.1).

Identity 4.2.

$$b^2\zeta_{10}(SB_1)/\left(d^2\zeta_{01}(SB_1)\right) = 1, \qquad (4.2)$$

where

$$\zeta_{10}(SB_1) = cov\,[\phi_1(X_1, ..., X_b; Y_1, ..., Y_d)\ ,\ \phi_1(X_1, X_{b+1}, ..., X_{2b-1}; Y_1, ..., Y_{2d})]$$

and

$$\zeta_{01}(SB_1) = cov\,[\phi_1(X_1, ..., X_b; Y_1, ..., Y_d)\ ,\ \phi_1(X_{b+1}, ..., X_{2b}; Y_1, Y_{d+1}, ..., Y_{2d-1})]\,.$$

Proof. While deriving the asymptotic variance of SB_1 under the null hypothesis, certain expressions are evaluated. We have

$$\zeta_{10}(SB_1) = E\left[\phi_1(X_1,...,X_b;Y_1,...,Y_d)\ ,\ \phi_1(X_1,X_{b+1},...,X_{2b-1};Y_{d+1},....,Y_{2d})\right] - E\left[SB_1\right]^2$$

$$= \int_{-\infty}^{\infty} P^2\left(med\ (x,X_2,...,X_b) \le M_2\right) dF(x) - (1/4)\ . \qquad (4.3)$$

Also

$$P(med\ (x,X_2,...,X_b) \le M_2) = d!(b-i)![K_1 + \cdots + K_b], \qquad (4.4)$$

where

$$K_1 = P[x \le Y_1,\ \ \mathrm{X}_2 \le X_3 \le ... \le x \le ... \le X_b,\ \ \mathrm{Y}_2 \le Y_3 \le ... \le Y_1 \le ... \le Y_d],$$

x in the middle position and Y_1 in the middle position.

$$K_i = P[X_i \le Y_1,\ \ X_2 \le X_3 \le .... \le X_i \le .... \le x,\ \ \mathrm{Y}_2 \le Y_3 \le ... \le Y_1 \le .... \le Y_d,]$$

for $i = 2,...,b$ and X_i in the middle position and Y_1 in the middle position.
After evaluating the expressions for$K_1,\ldots,K_b$ and substituting in (4.4) and (4.3) we get

$$\zeta_{10}(SB_1) = (d!/(b-1)!)^2\, K\,(b,d) \Big/ \left(p!^2 q!^2 (b+d-1)\begin{pmatrix} b+d-2 \\ p+q \end{pmatrix}\right)^2, \qquad (4.5)$$

where

$$K(b,d) = \sum_{i=0}^{p+q} \begin{pmatrix} b+d-1 \\ i \end{pmatrix}^2 B(2i+1,\ 2b+2d-2i-1)$$
$$+ \sum_{i \ne i'=1}^{p+q} \sum \begin{pmatrix} b+d-1 \\ i \end{pmatrix} \begin{pmatrix} b+d-1 \\ i' \end{pmatrix} B(i+i'+1,\ \ 2b+2d-i-i'-1) - (1/4)\,.$$

Similarly,

$$\zeta_{01}(SB_1) = E[\phi_1(X_1,...,X_b;Y_1,...,Y_d)\ ,\ \phi_1(X_{b+1},...,X_{2b}\ ;\ Y_1,Y_{d+1},...,Y_{2d-1}] - (1/4)$$

$$= (b!/(d-1)!)^2 \, K\,(b,d) \,/ \left(p!^2 q!^2 (b+d-1) \begin{pmatrix} b+d-2 \\ p+q \end{pmatrix} \right)^2 . \qquad (4.6)$$

Therefore

$$\zeta_{10}(SB_1)/\zeta_{01}(SB_1) = (d!/(b-1)!)^2 \,/\, (b!/(d-1)!)^2 = (d/b)^2 . \qquad (4.7)$$

Hence, we get the identity (4.2).

Identity 4.3.

$$\zeta_{10}\,(\mathrm{SB}_2) - \zeta_{01}\,(\mathrm{SB}_3) = 0 \;, \qquad (4.8)$$

where

$$\zeta_{10}(SB_2) = cov\,[\phi_2(X_1, ..., X_b; Y_1, ..., Y_d) \;,\; \phi_2(X_1, X_{b+1}, ..., X_{2b-1}; Y_{d+1}, ..., Y_{2d})]$$

and

$$\zeta_{01}(SB_3) = cov\,[\phi_3(X_1, ..., X_b; Y_1, ..., Y_d) \;,\; \phi_3(X_{b+1}, ..., X_{2b}; Y_1, Y_{d+1} ...., Y_{2d-1})] \,.$$

Proof. Under the null hypothesis

$$E\,(SB_2) = \int_{-\infty}^{\infty} [F(x)]^b \, dF_{M_2}(x)$$

$$= \left(d!/(q!)^2\right) B\,(b+q+1,\, q+1) \qquad (4.9)$$

and

$$E\,(SB_3) = \int_{-\infty}^{\infty} \left[\overline{F}\,(x)\right]^b dF_{M_3}(x),$$

where

$$F_{M_3}(x) = \sum_{i=q+1}^{d} \begin{pmatrix} d \\ i \end{pmatrix} [F(x)]^i \left[\overline{F}\,(x)\right]^{d-i} .$$

Therefore ,

$$E(SB_3) = \left(d!/(q!)^2\right) \int_{-\infty}^{\infty} [F(x)]^q \left[\overline{F}\,(x)\right]^{b+q} dF\,(x)$$

$$= \left(d!/(q!)^2\right) B\,(q+1,\, b+q+1) \;. \qquad (4.10)$$

Since the kernel $\phi_2(.;.)$can be obtained from $\phi_3(.;.)$ by replacing(X_i)'s by $(-X_i)$'s, (Y_j)'s by $(-Y_j)$'s and interchanging labels we get$E(SB_2) = E(SB_3)$.

Also,

$$\zeta_{10}(SB_2) = E\left(\phi_2(X_1,...,X_b;Y_1,...,Y_d)\ ,\ \phi_2(X_1,X_{b+1},...,X_{2b-1};Y_{d+1},....,Y_{2d})\right) - E^2[SB_2]$$

$$= \int_{-\infty}^{\infty} P^2\left(\max\left(x,X_2,...,X_b\right) \leq M_2\right) dF(x) - E^2(SB_2) \qquad (4.11)$$

and

$$\zeta_{01}(SB_3) = E\left(\phi_3(X_1,...,X_d;Y_1,...,Y_b)\ ,\ \phi_3(X_{d+1},...,X_{2d};Y_1,Y_{b+1},....,Y_{2b-1})\right) - E^2[SB_3]$$

$$= \int_{-\infty}^{\infty} P^2\left(M_3 \leq \min\left(y,Y_2,...,Y_b\right)\right) dF(x) - E^2(SB_3). \qquad (4.12)$$

Since $E[SB_2] = E[SB_3]$ under H_0 and from symmetry, we have

$$P\left(\max\left(x,X_2,...,X_b\right) \leq M_2\right) = P\left(M_3 \leq \min\left(y,Y_2,...,Y_d\right)\right).$$

From (4.9) through (4.12) we get the identity (**??**).

Identity 4.4.

$$k\binom{d}{k}\sum_{i=k}^{b}\binom{b}{i}\ B\ (i+k,\ 2b-i-k+1) = 1/2 \qquad (4.13)$$

or

$[d!/\left((k-1)!(d-k)!\right)]\ \sum_{i=k}^{b}\binom{b}{i}\ B\ (i+k,\ 2b-i-k+1) = 1/2.$

Proof. Under the null hypothesis

$$\begin{aligned} E(SB_4) &= P(X_{(k_i)} \leq Y_{(k_2)}) \\ &= \int_{-\infty}^{\infty} P(X_{(k_i)} \leq y) dF_{Y_{(k_2)}}(y) \\ &= k_2\binom{d}{k_2}\sum_{i=k_1}^{b}\binom{b}{i}\ B\ (i+k_2,\ b+d-i-k_2+1))\ , \qquad (4.14) \end{aligned}$$

When $b = d$ and $k_1 = k_2 = k$, we have $E(SB_4) = 1/2$. Therefore we get identity (4.13).

Identity 4.5. For $i = 1,3,5,...,b+d-1$,

$$e\left(SB_1\left(b,d\right)\right)/e\left(SB_1\left(i,b+d-i\right)\right) = 1, \qquad (4.15)$$

where

$$e\left(SB_1\left(b,d\right)\right) = \left[(b!d!)\ /\ \left((p!q!)^2\ \sigma_{b,d}\right)\right]\int_{-\infty}^{\infty}\left[F\left(x\right)\right]^{p+q}\left[\overline{F}\left(x\right)\right]^{p+q}\left[f\left(x\right)\right]^2 dx,$$

$$\sigma_{b,d}^2 = \ (d!)^2\,(b!)^2\,K\,(b,d)\,/\ \left[(p!)^4\,(q!)^4\,(b+d-1)^2\left(\begin{array}{c} b+d-2 \\ p+q \end{array}\right)^2\lambda\,(1-\lambda)\right]$$

and

$$0 < \lambda = \lim_{N\to\infty}\ (m/N) < 1, \qquad N = m+n.$$

Proof. Under the null hypothesis

$$\sigma_{b,d}^2 = \left(b^2/\lambda\right)\zeta_{10}\,(SB_1) + \left(d^2/\,(1-\lambda)\right)\zeta_{01}\,(SB_1)$$

$= b^2\zeta_{10}\,(SB_1)\,/\,(\lambda\,(1-\lambda))$ or $d^2\zeta_{01}\,(SB_1)\,/\,(\lambda\,(1-\lambda))$ by identity 4.2.

It is worth to note that $e\,(SB_1\,(b,d))$ depends on $(b+d)$ and underlying distribution $F\,(x)$. Thus for $1 \le b \le m$, $1 \le d \le n$, b,d being odd positive integers, given $F\,(x)$, we get $e\,(SB_1\,(b,d)) = e\,(SB_1\,(i,b+d-i))$ for$i = 1,3,5,...,b+d-1$. Therefore, we get identity (4.15).

5 Identities based on Ordered Ranks

In this section, we present some identities based on the ordered ranks of two sample U statistics defined in section 3. Suppose that $X_{(1)} \le X_{(2)} \le .... \le X_{(m)}$ and $Y_{(1)} \le Y_{(2)} \le .... \le Y_{(n)}$ are the order statistics corresponding to X and Y sample observations respectively. Let $R_{(i)}\left(S_{(j)}\right)$ be the rank of $X_{(i)}\left(Y_{(j)}\right)$ in the joint ranking of X and Y observations. Then we have the following identities.

Identity 5.1. $\sum_{i=1}^{m}\sum_{j=0}^{q}\left(\begin{array}{c} i-1 \\ p \end{array}\right)\left(\begin{array}{c} m-i \\ p \end{array}\right)\left(\begin{array}{c} R_{(i)}-i \\ q-j \end{array}\right)\left(\begin{array}{c} n-R_{(i)}+i \\ q+1+j \end{array}\right)$

$$= \sum_{j=1}^{n}\sum_{i=0}^{p}\left(\begin{array}{c} j-1 \\ q \end{array}\right)\left(\begin{array}{c} n-j \\ q \end{array}\right)\left(\begin{array}{c} S_{(j)}-j \\ p+1+i \end{array}\right)\left(\begin{array}{c} m-S_{(j)}+j \\ p-i \end{array}\right). \qquad (5.1)$$

Proof. Choose a sub sample of size b from the X sample such that $X_{(i)}$ is the median. For a fixed i, this can be done in $\left(\begin{array}{c} i-1 \\ p \end{array}\right)\left(\begin{array}{c} m-i \\ p \end{array}\right)$ ways. Similarly choose a sub sample of size d from Ysample such that $Y_{(j)}$ is the median and is greater than $X_{(i)}$. Each such sub-sample pair results in $\phi_1(.\,;\,.) = 1$. Using the fundamental rules of counting, we get the identity (5.1).

Identity 5.2.

$$\binom{m}{b}\binom{n}{d} SB_2 = \sum_{j=1}^{n} \binom{j-1}{q}\binom{n-j}{q}\binom{S_{(j)}-j}{b}. \tag{5.2}$$

Proof. Choose a sub sample of size d from the Y sample such that $Y_{(j)}$ is the median. For a fixed j this can be done in $\binom{j-1}{q}\binom{n-j}{q}$ ways. The number of X observations less than $Y_{(j)}$ will be $\left(S_{(j)}-j\right)$. A sub sample of size b from the X observations can be chosen in $\binom{S_{(j)}-j}{b}$ ways and we get identity (5.2).

Identity 5.3.

$$\binom{m}{d}\binom{n}{b} SB_3 = \sum_{i=1}^{m} \binom{i-1}{q}\binom{m-i}{q}\binom{n-R_{(i)}+i}{b}. \tag{5.3}$$

Proof. For a fixed i, $X_{(i)}$ can be chosen as median of sub sample of size d from X observations in $\binom{i-1}{q}\binom{m-i}{q}$ ways. The number of Y observations greater than $X_{(i)}$ is $\left(n-R_{(i)}+i\right)$. A sub sample of size b from these Y observations can be chosen in $\binom{n-R_{(i)}+i}{b}$ ways. For each i, $\binom{i-1}{q}\binom{m-i}{q}\binom{n-R_{(i)}+i}{b}$ sub-sample pairs for which $\phi_3(.\,;\,.) = 1$. Then by the fundamental rule of counting, we get the identity (5.3).

Identity 5.4.

$$\sum_{i=1}^{m}\sum_{j=0}^{k_2-1} \binom{i-1}{k_1-1}\binom{m-i}{b-k_1}\binom{R_{(i)}-i}{k_2-j-1}\binom{n-R_{(i)}+i}{d-k_2+j+1}$$

$$= \sum_{j=1}^{n}\sum_{i=0}^{k_1-1} \binom{j-1}{k_2-1}\binom{n-j}{d-k_2}\binom{S_{(j)}-j}{b-k_1+i+1}\binom{m-S_{(j)}+j}{k_1-i-1}. \tag{5.4}$$

Proof. Choose a sub sample of size b from the X sample such that $X_{(i)}$ is the k_1^{th} order statistic $((b\beta)^{th}$quantile). For a fixed i, this can be done in $\binom{i-1}{k_1-1}$

$\begin{pmatrix} m-i \\ b-k_1 \end{pmatrix}$ ways. Now choose a sub sample of size d from Y sample such that $Y_{(j)}$ is the k_2^{th} order statistic $((d\gamma)^{th}$quantile) and is greater than $X_{(i)}$. Each such sub-sample pair results in $\phi_4\,(.\,;\,.) = 1$. The $Y_{(j)}$ can be selected in $\begin{pmatrix} R_{(i)}-i \\ k_2-j-1 \end{pmatrix}\begin{pmatrix} n-R_{(i)}+i \\ d-k_2+j+1 \end{pmatrix}$ ways. Thus using the fundamental rules of counting, we get the identity (5.4).

Acknowledgement: We thank the referee for his useful comments.

References

Baiju, K.V. and Thomas, P.Y. (2007). On some identities proved using Statistical mathematics, Journal of the Kerala Statistical Association 18, 34-44.

Bhat, S.V. (1996). Studies in Nonparametric Inference. An unpublished Ph.D thesis submitted to Karnatak University, Dharwad.

Joshi, P.C. and Balakrishnan, N. (1981). Applications of order Statistics in combinatorial identities, Journal of Combinatorics, Information and System Sciences 6, 271-278.

Randles, R. H. and Wolfe, D.A. (1979). Introduction to Theory of Nonparametric Statistics, John Wiley and Sons, New York.

Riordan, J. (1968). Combinatorial Identities, John Wiley and Sons, New York.

Shetty, I.D. and Bhat, S.V. (1993). Some Competitors of Mood's median test for location alternatives, Karnatak University, Journal of Science 37, 138-146.

Shetty, I.D. and Bhat, S.V. (1994). A Note on the Generalization of Mathisen's median test, Statistics and Probability Letters 19, 199-204.

Shetty, I.D, Sengupta, D. and Bhat, S. V (1997). A General class of two-sample Nonparametric tests based on sub-sample quantiles, Karnatak University, Journal of Science 41, 104-125.

Journal of the Kerala Statistical Association
Vol.24, December 2013, 12-36 *ISSN 2249-4553*

Kernel Estimation of the Past Entropy Function with dependent data

R. Maya*
Department of Statistics
University of Kerala, Thiruvananthapuram - 695 581.

ABSTRACT

The past entropy function, introduced by Di Crescenzo and Longobardi (2002), is viewed as a dynamic measure of uncertainty in past life. This measure find applications in modeling life time data. In the present work we provide non-parametric kernel type estimators for the past entropy function based on complete and censored data. Asymptotic properties of the estimators are established under suitable regularity conditions. Monte-Carlo simulation studies are carried out to compare the performance of the estimators using the mean-squared error. The methods are illustrated using real data sets.

Key words and Phrases: *Past entropy function, Residual entropy function, Kernel estimate, α-mixing, Residual life.*

1 Introduction

Recently, many researchers have shown a keen interest in the measurement of uncertainty associated with a probability distribution of particular interest in probability and statistics is the notion of entropy, introduced by Shannon (1948). If X is a random variable having an absolutely continuous distribution function F with

*email: mayasekharan@gmail.com

probability density function f, then the entropy of the random variable X is defined as

$$H(X) = H(f) = -\int_0^\infty f(x) \log f(x) dx. \tag{1.1}$$

If we consider X as the lifetime of a new unit then $H(f)$ can be viewed as a useful tool for measuring the associated uncertainty. Ebrahimi and Pellerey (1995) and Ebrahimi (1996) have introduced the concept of residual entropy in terms of a conditional measure. For a non-negative random variable X, representing the life time of a component, the residual entropy function is the Shannon's entropy associated with the random variable X given $X > t$, and is defined as

$$\begin{aligned} H(f;t) &= -\int_t^\infty \frac{f(x)}{\overline{F}(t)} \log \frac{f(x)}{\overline{F}(t)} dx \\ &= 1 - \frac{1}{\overline{F}(t)} \int_t^\infty f(x) \log h(x) dx, \ \overline{F}(t) > 0, \end{aligned} \tag{1.2}$$

where $\overline{F}(t) = P(X > t)$ denotes the survival function and $h(x) = \frac{f(x)}{\overline{F}(x)}$ is the hazard function of X. Belzunce et al. (2004) have established that if $H(f;t)$ is increasing in t then $H(f;t)$ determines the distribution uniquely. Given that an item has survived up to time t, $H(f;t)$ measures the uncertainty in its remaining life. For a discussion of the properties and applications of residual entropy we refer to Ebrahimi and Kirmani (1996), Nair and Rajesh (1998) and Asadi and Ebrahimi (2000).

It is reasonable to presume that in many realistic situations uncertainty is not necessarily related to the future but can also refer to the past. For instance, consider a system whose state is observed only at certain preassigned inspection times. If at time t, the system is inspected for the first time and it is found to be 'down', then the uncertainty relies on the past, i.e. on which instant in (0, t) it has failed. It thus seems natural to introduce a notion of uncertainty that is dual to the residual entropy, in the sense that it refers to past time and not to future time. Based on this idea Di Crescenzo and Longobardi (2002) have studied the past entropy. Further, they discussed the necessity of the past entropy, its relation with residual entropy and many interesting results. In Di Crescenzo and Longobardi (2004) a measure of discrimination based on past entropy has been studied. If X denotes the lifetime of a component/system or of living organism, then the past entropy of X at time t is defined as

$$\begin{aligned}\overline{H}(f;t) &= -\int_0^t \frac{f(x)}{F(t)} \log \frac{f(x)}{F(t)}\ dx \\ &= 1 - \frac{1}{F(t)} \int_0^t f(x) \log T(x)\ dx,\end{aligned} \tag{1.3}$$

where $T(x) = \frac{f(x)}{F(x)}$ is the reversed hazard rate of x.

It can also be used in Forensic Sciences where exact time of failure (death in case of human being) is important when at some time t the unit is found to be in failure state. Gupta and Nanda (2002) have defined generalized uncertainty of lifetime distribution by truncating the distributions above some point t. Nanda and Paul (2006) have studied some properties and applications of past entropy. Gupta (2009) established that the past entropy determines the distribution uniquely, under certain conditions.

To make valid decisions regarding the extent of uncertainty in past data, one requires a reasonable estimate of the past entropy. Motivated by this in the present paper we provide nonparametric estimators for $\overline{H}(f;t)$ using kernel type estimation for complete as well as censored data. We consider only situations where the data under study are dependent. In both situations, the underlying lifetimes are assumed to be α-mixing (see, Rosenblatt (1956)).

Definition 1.

Let $\{X_i; i \geq 1\}$ denote a sequence of random variables. Given a positive integer n, set

$$\alpha(n) = \sup_{k \geq 1} \{|P(A \cap B) - P(A)P(B)|; A \epsilon \mathfrak{F}_1^k, B \epsilon \mathfrak{F}_{k+n}^{\infty} \tag{1.4}$$

where $\mathfrak{F}_i^k$ denote the σ- field of events generated by $\{X_j; i \leq j \leq k\}$. The sequence is said to be α-mixing (strong mixing) if the mixing coefficient $\alpha(n) \to 0$ as $n \to \infty$. Among various mixing conditions, α-mixing is reasonably weak and has many practical applications.

The rest of the paper is organized as follows. In Section 2, we present estimators for $\overline{H}(f;t)$, given in (1.3), using complete and censored samples. In Section 3, we examine the consistency and asymptotic normality of the estimators. In Section 4, we evaluate the estimators for real data sets and in Section 5 a simulation study to illustrate the behavior of the estimators is undertaken.

2 Estimation

In this section, we propose nonparametric estimators for the past entropy function for complete as well as censored data sets.

2.1 Complete Samples

Let $\{X_i; 1 \leq i \leq n\}$ be a sequence of identically distributed random variables. Note that $X_i{}'s$ need not be mutually independent. A simple nonparametric estimator of $\overline{H}(f;t)$ is given as

$$\overline{H}^*(f;t) = \frac{-1}{n}\sum_{i=1}^{n} \log\left(\frac{f_n(X_i)}{F_n(t)}\right) I_{(X_i \leq t)}, \tag{2.1}$$

where $f_n(X_i) = \frac{1}{(n-1)}\sum\limits_{j\neq i}^{n} \frac{1}{b_n} K\left(\frac{X_i - X_j}{b_n}\right)$ is the kernel estimator obtained from the sample without X_i, $F_n(t)$ is either an empirical or a kernel estimator for the distribution function and

$$I_{(X_i \leq t)} = \begin{cases} 1, & \text{if } X_i \leq t \\ 0, & \text{otherwise.} \end{cases}$$

A kernel estimator of $\overline{H}(f;t)$ for the above sample is defined as

$$\overline{H}_n(f;t) = -\int_0^t \frac{f_n(x)}{F_n(t)} \log \frac{f_n(x)}{F_n(t)} dx. \tag{2.2}$$

(2.2) can also be written as

$$\overline{H}_n(f;t) = \log F_n(t) - \frac{1}{F_n(t)} \int_0^t f_n(x) \log f_n(x) dx, \tag{2.3}$$

where $f_n(x)$ is a nonparametric estimator of $f(x)$ and $F_n(t) = \int_0^t f_n(x)\,dx$ is a nonparametric estimator of distribution function $F(t)$.

The most common nonparametric density estimator of $f(x)$ is the kernel estimator given by (see, Parzen (1962), Rosenblatt (1970))

$$f_n(x) = \frac{1}{nb_n}\sum_{j=1}^{n} K\left(\frac{x - X_j}{b_n}\right), \tag{2.4}$$

where $K(x)$ is a kernel of order s with compact support satisfying the conditions

i) $K(x) \geq 0$ for all x

ii) $\int K(x)\,dx = 1$

iii) K(.) is symmetric about zero

iv) $K_n(x) = \frac{1}{b_n}K(\frac{x}{b_n})$ where $\{b_n\}$ is a bandwidth sequence of positive numbers such that $b_n \to 0$ and $nb_n \to \infty$ as $n \to \infty$ and

v) $K(.)$ satisfies Lipschitz condition, namely there exist a constant M such that $|K(x) - K(y)| \leq M|x - y|$.

Under α-mixing dependence conditions, expressions for the bias and variance of $f_n(x)$ are

$$Bias\,(f_n(x)) \backsimeq \frac{b_n^s c_s}{s!} f^{(s)}(x) \tag{2.5}$$

and

$$Var\,(f_n(x)) \backsimeq \frac{1}{nb_n} f(x) C_K, \tag{2.6}$$

where $c_s = \int\limits_{-\infty}^{\infty} u^s K(u)\ du$ and $C_K = \int\limits_{-\infty}^{\infty} K^2(u)\ du$.

2.2 Censored Samples

In reliability and life testing, due to time constraints or cost considerations the experimenter is forced to terminate the experiment after a specific period of time or after the failure of a specified number of units. In this context the underlying data will be censored. In the context of right censoring only, the lower bounds on life time will be available for some individuals and in the context of left censoring data will be recorded as the upper bound of life time for some individuals. Another common type of censoring is random censoring.

Let $\{X_i; 1 \leq i \leq n\}$ be a sequence of non-negative random variables representing the life times for n components/devices. The random variables are not assumed to be mutually independent. However they have a common unknown continuous marginal distribution function $F(x)$ with a probability density function $f(x) = F^{'}(x)$. Let the random variable X_i be censored on the right by the random variable Y_i. In this random censorship model, the censoring times Y_i are assumed to be independently and identically distributed and they are also assumed to be independent of X_i. The censoring times Y_1, Y_2, ..., Y_n have the common distribution function $P(y)$. This scheme is very common in clinical trials. In such experiments, patients enter into the study at random time points, while the experiment itself is terminated at a prespecified time. Let $Z_i = min(X_i, Y_i)$ and $\delta_i = I(X_i \leq Y_i)$, where I(.) denotes the indicator function of the event specified in parentheses. The actually observed $Z_i^{\prime}s$ have a distribution function G satisfying

$$1 - G(t) = (1 - F(t))(1 - P(t)),\ t\ \epsilon R_+ = [0, \infty).$$

Let $G^*(t) = P(Z_1 \leq t; \delta_1 = 1)$ is the corresponding sub-distribution function for the uncensored observations and $g^*(t) = f(t)(1 - P(t))$ be the corresponding sub-density. A reasonable estimator of f should behave like $\frac{g_n^*(t)}{(1-P(t))}$ where $g_n^*(t) =$

$b_n^{-1} \int\limits_{R^+} K(\frac{t-x}{b_n})\ dG_n^*(x)$ is the kernel estimator pertaining to $G_n^*(t) = \frac{1}{n} \sum\limits_{i=1}^{n} I(Z_i \le t; \delta_i = 1)$.
A nonparametric estimator for $\overline{H}(f;t)$ based on the censored data is

$$\overline{H}_*^n(f;t) = \frac{-1}{n} \sum_{i=1}^{n} \log\left(\frac{f_n(Z_i)}{F_n^*(t)}\right) I_{(Z_i \le t)}, \tag{2.7}$$

where $f_n(Z_i) = \frac{1}{(n-1)} \sum\limits_{j \ne i}^{n} \frac{1}{b_n} K\left(\frac{Z_i - Z_j}{b_n}\right)$ is the kernel estimator obtained from the sample without Z_i and $F_n^*(t)$ is a kernel estimator for the distribution function.
A non parametric estimator for past entropy function under $\alpha-$ mixing condition, in the random censorship model, is defined as

$$\overline{H}_n^*(f;t) = \log F_n^*(t) - \frac{1}{F_n^*(t)} \int_0^t f_n^*(x) \log f_n^*(x) dx, \tag{2.8}$$

where

$$f_n^*(t) = \frac{1}{b_n} \int_{R^+} \frac{K(\frac{t-x}{b_n})}{1 - P(x)}\ dG_n^*(x), \tag{2.9}$$

is a nonparametric density estimator for $f(x)$ under censoring (see, Cai (1998)), $F_n^*(t) = \int\limits_0^t f_n^*(u)\ du$, $K(x)$ is a kernel of order s with compact support which satisfies the conditions (i)-(v) in Section 2.1 and $\{b_n\}$ is a sequence of real numbers such that $b_n \to 0$ and $nb_n \to \infty$ as $n \to \infty$.
Under α-mixing dependence conditions, expressions for the bias and variance of $f_n^*(t)$ are given by (see, Cai (1998))

$$Bias\,(f_n^*(t)) \backsimeq \frac{b_n^s c_{s^+}}{s!} f^{(s)}(t) \tag{2.10}$$

and

$$Var\,(f_n^*(t)) \backsimeq \frac{1}{nb_n} \frac{f(t)}{(1 - P(t))} C_K, \tag{2.11}$$

where $c_{s^+} = \int\limits_{R^+} u^s K(u)\ du$ and $C_K = \int\limits_{-\infty}^{\infty} K^2(u)\ du$.

3 Asymptotic properties

In this section, we look in to the consistency and asymptotic normality of the estimators (2.3) and (2.8). In order to simplify the notations, define

$V_n(t) = \log F_n(t)$, $V_n^*(t) = \log F_n^*(t)$, $\quad V(t) = \log F(t)$,
$A_n(t) = \int\limits_0^t f_n(x) \log f_n(x) dx$, $A_n^*(t) = \int\limits_0^t f_n^*(x) \log f_n^*(x) dx$
and $A(t) = \int\limits_0^t f(x) \log f(x) dx$.

Theorem 3.1. *Let $K(x)$ be a kernel of order s with compact support satisfying the conditions (i)-(v) in Section 2 and $\{b_n\}$ be such that $b_n \to 0$ and $nb_n \to \infty$ as $n \to \infty$. Then*

a) $\overline{H}_n(f;t)$ is a consistent estimator of $\overline{H}(f;t)$.
b) $\overline{H}_n^*(f;t)$ is a consistent estimator of $\overline{H}(f;t)$.

Proof. a) Under α-mixing dependence conditions, we obtain the expressions for the bias and the variance of $F_n(t)$, $V_n(t)$ and $A_n(t)$ and are given by

$$Bias\,(F_n(t)) \simeq \frac{b_n^s c_s}{s!} \int\limits_0^t f^{(s)}(x)\ dx, \tag{3.1}$$

$$Var\,(F_n(t)) \simeq \frac{C_K}{nb_n} F(t), \tag{3.2}$$

$$Bias\,(V_n(t)) \simeq \frac{b_n^s c_s}{s! F(t)} \int\limits_0^t f^{(s)}(x)\ dx, \tag{3.3}$$

$$Var\,(V_n(t)) \simeq \frac{1}{nb_n} \frac{C_k}{F(t)}, \tag{3.4}$$

$$Bias(A_n(t)) \simeq \left(\frac{b_n^s}{s!} c_s\right) \int\limits_0^t (1 + log f(x)) f^{(s)}(x) dx \tag{3.5}$$

and

$$Var(A_n(t)) \backsimeq \left(\frac{1}{nb_n}C_k\right)\int_0^t (1+\log f(x))^2 f(x)dx. \tag{3.6}$$

The corresponding MSE's are given by

$$MSE\,(F_n(t)) \backsimeq \left(\frac{b_n^s c_s}{s!}\int_0^t f^{(s)}(x)\; dx\right)^2 + \frac{C_K}{nb_n}F(t), \tag{3.7}$$

$$MSE\,(V_n(t)) \backsimeq \left(\frac{b_n^s c_s}{s!F(t)}\int_0^t f^{(s)}(x)\; dx\right)^2 + \frac{1}{nb_n}\frac{C_k}{F(t)} \tag{3.8}$$

and

$$\begin{aligned} MSE\,(A_n(t)) \backsimeq & \left(\left(\frac{b_n^s}{s!}c_s\right)\int_0^t (1+logf(x))f^{(s)}(x)dx\right)^2 \\ & + \left(\frac{1}{nb_n}C_k\right)\int_0^t (1+\log f(x))^2 f(x)dx. \end{aligned} \tag{3.9}$$

From (3.7), as $n \to \infty$

$$\text{MSE}\,(F_n(t)) \to 0.$$

From (3.8), as $n \to \infty$

$$\text{MSE}\,(V_n(t)) \to 0.$$

From (3.9), as $n \to \infty$

$$\text{MSE}\,(A_n(t)) \to 0.$$

Therefore

$$\overline{H}_n(f;t) = V_n(t) - \frac{A_n(t)}{F_n(t)} \xrightarrow{p} V(t) - \frac{A(t)}{F(t)} = \overline{H}(f;t).$$

That is, $\overline{H}_n(f;t)$is a consistent estimator of $\overline{H}(f;t)$.

b) Under α-mixing dependence conditions, we obtain the expressions for the bias and the variance of $F_n^*(t)$, $V_n^*(t)$ and $A_n^*(t)$ and are given by

$$Bias\,(F_n^*(t)) \backsimeq \frac{b_n^s c_{s+}}{s!}\int_0^t f^{(s)}(u)\; du, \tag{3.10}$$

$$Var\,(F_n^*(t)) \backsimeq \frac{1}{nb_n} C_K \int_0^t \frac{f(u)}{[1-P(u)]}\ du, \tag{3.11}$$

$$Bias\,(V_n^*(t)) \backsimeq \frac{b_n^s}{s!}\frac{c_{s+}}{F(t)} \int_0^t f^{(s)}(u)\ du, \tag{3.12}$$

$$Var\,(V_n^*(t)) \backsimeq \frac{1}{nb_n}\frac{C_k}{F^2(t)} \int_0^t \frac{f(u)}{[1-P(u)]} du, \tag{3.13}$$

$$Bias(A_n^*(t)) \backsimeq \left(\frac{b_n^s}{s!} c_s^+\right) \int_0^t (1+logf(u)) f^{(s)}(u) du \tag{3.14}$$

and

$$Var(A_n^*(t)) \backsimeq \left(\frac{1}{nb_n} C_k\right) \int_0^t (1+\log f(u))^2 \frac{f(u)}{[1-P(u)]} du. \tag{3.15}$$

The corresponding MSE's are given by

$$MSE\,(F_n^*(t)) \backsimeq \left(\frac{b_n^s c_{s+}}{s!} \int_0^t\ f^{(s)}(u)\ du\right)^2 + \frac{1}{nb_n} C_K \int_0^t \frac{f(u)}{[1-P(u)]}\ du, \tag{3.16}$$

$$MSE\,(V_n^*(t)) \backsimeq \left(\frac{b_n^s}{s!}\frac{c_{s+}}{F(t)} \int_0^t f^{(s)}(u)\ du\right)^2 + \frac{1}{nb_n}\frac{C_k}{F^2(t)} \int_0^t \frac{f(u)}{[1-P(u)]} du \tag{3.17}$$

and

$$\begin{aligned} MSE\,(A_n^*(t)) \backsimeq & \left(\left(\frac{b_n^s}{s!} c_s^+\right) \int_0^t (1+logf(u)) f^{(s)}(u) du\right)^2 \\ & + \left(\frac{1}{nb_n} C_k\right) \int_0^t (1+\log f(u))^2 \frac{f(u)}{[1-P(u)]} du. \end{aligned} \tag{3.18}$$

From (3.16), as $n \to \infty$

$$\text{MSE}\,(F_n^*(t)) \to 0.$$

From (3.17), as $n \to \infty$

$$\text{MSE}\,(V_n^*(t)) \to 0.$$

From (3.18), as $n \to \infty$

$$\text{MSE}\,(A_n^*(t)) \to 0.$$

Therefore

$$\overline{H}_n^*(f;t) = V_n^*(t) - \frac{A_n^*(t)}{F_n^*(t)} \xrightarrow{p} V(t) - \frac{A(t)}{F(t)} = \overline{H}(f;t).$$

That is, $\overline{H}_n^*(f;t)$ is a consistent estimator of $\overline{H}(f;t)$. □

If $g_n(x)$ is an estimator for $g(x)$, then its MISE (Integrated mean-squared error) is given by:

$$MISE\,(g_n(x)) = E\left\{\int_{-\infty}^{\infty} (g_n(x) - g(x))^2\, w(t)dF(t)\right\},$$

where $w(.)$ is a weight function.

In the following theorem, we give expressions for the MISE of $\overline{H}_n(f;t)$ and $\overline{H}_n^*(f;t)$ as $n \to \infty$.

Theorem 3.2. *Let $K(x)$ be a kernel of order s with compact support satisfying the conditions (i)-(v) in Section 2 and $\{b_n\}$ satisfying the conditions given in Section 2. Then*

a) $\lim\limits_{n\to\infty} MISE\left(\overline{H}_n(f;t)\right) = 0.$

b) $\lim\limits_{n\to\infty} MISE\left(\overline{H}_n^*(f;t)\right) = 0.$

Proof.

$$MISE\left(\overline{H}_n(f;t)\right) = E\left\{\int_{-\infty}^{\infty} \left(\overline{H}_n(f;t) - \overline{H}(f;t)\right)^2 w(t)dF(t)\right\}. \tag{3.19}$$

$$= E\left\{\int_{-\infty}^{\infty} \left[(V_n(t) - V(t)) - \left(\frac{A_n(t)}{F_n(t)} - \frac{A(t)}{F(t)}\right)\right]^2 w(t)dF(t)\right\}$$

$$= E\left\{\int_{-\infty}^{\infty} \left(V_n(t) - V(t)\right)^2 w(t)dF(t)\right\}$$

$$+ E\left\{\int_{-\infty}^{\infty} \left(\frac{A_n(t)}{F_n(t)} - \frac{A(t)}{F(t)}\right)^2 w(t)dF(t)\right\}$$

$$- 2E\left\{\int_{-\infty}^{\infty} \left(V_n(t) - V(t)\right)\left(\frac{A_n(t)}{F_n(t)} - \frac{A(t)}{F(t)}\right) w(t)dF(t)\right\}.$$

Let H_1, H_2 and H_3 are given by

$$H_1 = E\left\{\int_{-\infty}^{\infty} \left(V_n(t) - V(t)\right)^2 w(t)dF(t)\right\}$$

$$= \int_{-\infty}^{\infty} \left[Var(V_n(t)) + Bias^2(V_n(t))\right] w(t)dF(t).$$

$$H_2 = E\left\{\int_{-\infty}^{\infty} \left(\frac{A_n(t)}{F_n(t)} - \frac{A(t)}{F(t)}\right)^2 w(t)dF(t)\right\}.$$

Using the approximation $\frac{A_n(t)}{F_n(t)} - \frac{A(t)}{F(t)} = \frac{A_n(t) - F_n(t)\frac{A(t)}{F(t)}}{F(t)}\left(1 + o_p(1)\right)$, the above equation simplifies to

$$H_2 = E\left\{\int_{-\infty}^{\infty} \left(\frac{A_n(t) - F_n(t)\frac{A(t)}{F(t)}}{F(t)}\right)^2 w(t)dF(t)\right\}.$$

Using Holder inequality, we get

$$H_2 \le \int_{-\infty}^{\infty} \left\{Var\left(A_n(t)\right) + \left[E\left(A_n(t)\right)\right]^2\right\} \frac{w(t)}{F^2(t)} dF(t)$$

$$+ \int_{-\infty}^{\infty} \left\{Var\left(F_n(t)\right)\frac{A^2(t)}{F^2(t)} + \left[E\left(F_n(t)\right)\right]^2 \frac{A^2(t)}{F^2(t)}\right\} \frac{w(t)}{F^2(t)} dF(t)$$

$$- 2\int_{-\infty}^{\infty} \frac{A(t)}{F(t)} \left[E\left(A_n(t)\right)^2\right]^{\frac{1}{2}} \left[E\left(F_n(t)\right)^2\right]^{\frac{1}{2}} \frac{w(t)}{F^2(t)} dF(t).$$

$$H_3 = 2E\left\{\int_{-\infty}^{\infty} \left(V_n(t) - V(t)\right)\left(\frac{A_n(t)}{F_n(t)} - \frac{A(t)}{F(t)}\right) w(t)dF(t)\right\}$$

$$\leq 2 \int_{-\infty}^{\infty} \left[E\left(V_n(t)-V(t)\right)^2\right]^{\frac{1}{2}} \left[E\left(\frac{A_n(t)}{F_n(t)}-\frac{A(t)}{F(t)}\right)^2\right]^{\frac{1}{2}} w(t)dF(t).$$

From (3.1), (3.2), (3.3), (3.4), (3.5) and (3.6), it follows that

$$MISE\left(\overline{H}_n(f;t)\right) \to o \ as \ n \to \infty. \tag{3.20}$$

b) The proof is similar to that of a). □

In the following theorem, we focus attention on the asymptotic normality of the estimators $\overline{H}_n(f;t)$ and $\overline{H}_n^*(f;t)$.

Theorem 3.3. *Suppose that F be continuous. Assume that K(.) satisfies the assumptions (i)-(v) of Section 2 . Then*

$$a)\,(nb_n)^{\frac{1}{2}}\left\{\frac{(\overline{H}_n(f;t)-\overline{H}(f;t))}{\sigma_{\overline{H}}}\right\} \tag{3.21}$$

has a standard normal distribution as $n\to\infty$ with
$\sigma^2_{\overline{H}} \backsimeq \frac{C_k}{nb_nF^2(t)}\int_0^t f(x)\left[\frac{A^2(t)}{F^2(t)}+1+(1+\log f(x))^2\right]dx.$

$$b)\,(nb_n)^{\frac{1}{2}}\left\{\frac{(\overline{H}^*_n(f;t)-\overline{H}(f;t))}{\sigma_{\overline{H}^*}}\right\} \tag{3.22}$$

has a standard normal distribution as $n\to\infty$ with

$\sigma^2_{\overline{H}^*} \backsimeq \frac{C_k}{nb_nF^2(t)}\int_0^t \frac{f(x)}{1-P(x)}\left[\frac{A^2(t)}{F^2(t)}+1+(1+\log f(x))^2\right]dx.$

Proof. a)

$$
\begin{aligned}
&\sqrt{nb_n}\left(\overline{H}_n(f;t)-\overline{H}(f;t)\right)\\
&\quad=\sqrt{nb_n}\left[(V_n(t)-V(t))-\left[\frac{A_n(t)}{F_n(t)}-\frac{A(t)}{F(t)}\right]\right]\\
&\quad\backsimeq\sqrt{nb_n}\left[\left(\frac{F_n(t)-F(t)}{F(t)}\right)-\left(\frac{A_n(t)F(t)-A(t)F_n(t)}{F_n(t)F(t)}\right)\right]\\
&\quad=\sqrt{nb_n}\left[\frac{(F_n(t)-F(t))}{F(t)}-\frac{F(t)\left[A_n(t)-A(t)\right]}{F_n(t)F(t)}\right]\\
&\qquad-\sqrt{nb_n}\left[\frac{A(t)\left[F_n(t)-F(t)\right]}{F_n(t)F(t)}\right].
\end{aligned}
\tag{3.23}
$$

Since $\sup_t |F_n(t)-F(t)| \to 0$ *a.s.*, (3.23) is asymptotically equal to

$$
\begin{aligned}
&\sqrt{nb_n}\left(\overline{H}_n(f;t)-\overline{H}(f;t)\right)\\
&\quad\backsimeq\sqrt{nb_n}\left[\frac{1}{F(t)}-\frac{A(t)}{F^2(t)}\right](F_n(t)-F(t))\\
&\qquad-\sqrt{nb_n}\left(\frac{A_n(t)-A(t)}{F(t)}\right)\\
&\quad\backsimeq\sqrt{nb_n}\left[\frac{1}{F(t)}-\frac{A(t)}{F^2(t)}\right]\int_0^t (f_n(u)-f(u))\ du\\
&\qquad-\frac{\sqrt{nb_n}}{F(t)}\int_0^t (f_n(u)-f(u))(1+\log f(u))\ du
\end{aligned}
\tag{3.24}
$$

Note that from Parzen (1962), $\sqrt{nb_n}\,(f_n(x)-f(x))$ is asymptotically normal with mean zero and variance σ_f^2 given in (2.6).
Now from (3.24), it is immediate that

$$
(nb_n)^{\frac{1}{2}}\left\{\frac{(\overline{H}_n(f;t)-\overline{H}(f;t))}{\sigma_{\overline{H}}}\right\}
\tag{3.25}
$$

is asymptotically normal with mean zero. The expression of variance can be obtained from (3.2), (3.4) and (3.6).

b) The proof is similar to that of *a*). □

4 Numerical illustration

Example 4.1.

To illustrate the usefulness of the proposed estimator discussed in Section2.1 with real situations, we consider the times, in months, to the first failure of 20 electric carts used for internal delivery and transportation in a large manufacturing facility (see, Zimmer et al (1998)). We use the bootstrapping procedure to find optimum value of b_n (see, Efron (1981)). The Gaussian kernel $K(z) = \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-z^2}{2}\right)$ is used as the kernel function for the estimation. At each value of t and b_n we calculate the biases and the mean-squared errors of $\overline{H}_n(f;t)$ using 250 bootstrap samples of size 20. Table 1 presents the bootstrap estimates of the biases and the mean-squared errors, in brackets, for $\overline{H}_n(f;t)$. From the Table 1 it can be seen that the optimum value of b_n is 0.2 for $0.9 < t \leq 12.6$, 0.3 for $12.6 < t \leq 16.3$, 0.4 for $16.3 < t \leq 19.3$, 0.6 for $19.3 < t \leq 22.6$, 0.7 for $22.6 < t \leq 24.8$ and 0.9 for $24.8 < t \leq 53$. In Figure 1 dark line represents the theoretical value $\overline{H}(f;t)$ and starred line represents the estimator $\overline{H}_n(f;t)$. From Figure 1, it is easy to see that for the data set considered the past entropy function is increasing with time.

Example 4.2.

For the illustration of estimation procedure discussed in Section 2.2, we consider the life times (in cycles) of 20 sodium sulphur batteries (see, Ansell and Ansell (1987)). The same method is used for calculating the bootstrap estimate of the bias and the mean-squared errors of $\overline{H}^*_n(f;t)$ as in the case of Example 4.1. The bootstrap estimates of the biases and the mean-squared errors, in brackets, for $\overline{H}^*_n(f;t)$ are given in Table 2. From the Table 2 it can be seen that the optimum value of b_n is 0.1 for $0.76 < t \leq 2.1$, 0.2 for $2.1 < t \leq 7.75$, 0.3 for $7.75 < t \leq 8.14$, 0.5 for $8.14 < t \leq 11.31$, 0.7 for $11.31 < t \leq 14.46$ and 0.9 for $14.46 < t \leq 30.77$. In Figure 2 dark line represents the theoretical value $\overline{H}(f;t)$ and starred line represents the estimator $\overline{H}^*_n(f;t)$. From Figure 2, we can say that for the given data the past entropy function is increasing with time.

5 Simulation studies

A Monte Carlo simulation study is carried out to compare the kernel estimators $\overline{H}_n(f;t)$ and $\overline{H}^*(f;t)$ in the case of complete samples and $\overline{H}_n^*(f;t)$ and $\overline{H}_*^n(f;t)$ in the case of censored samples in terms of the mean-squared error. First we consider the simulation under complete sample. The exponential distribution with parameter $\lambda = 0.6$ is used for the simulation. For the simulation under complete sample, we generated $\{X_i\}$ from AR(1) with correlation coefficient $\rho = 0.3$. The Gaussian kernel is used as the kernel function for the estimation. The estimates for various values of t $(4 < t < 4.8)$, b_n and sample sizes $n = 50$ and $n = 100$ are calculated. The ratios of mean-squared error of $\overline{H}_n(f;t)$ to that of the $\overline{H}^*(f;t)$ are computed and are given in Table 3. From a range of b_n values, we found that $b_n = 0.4$ come close to giving the smallest discrepancy between $\overline{H}_n(f;t)$ and $\overline{H}^*(f;t)$. In Figure 3 dark line represents the theoretical value $\overline{H}(f;t)$, starred line represents the estimator $\overline{H}_n(f;t)$ and dotted line represents the estimator $\overline{H}^*(f;t)$.

For the simulation under right censored sample, we generated $\{X_i\}$ from AR(1) with correlation coefficient $\rho = 0.2$ and the censoring times $\{Y_i\}$ were generated independently from $N(2,1)$. In this case also we used the Gaussian kernel for simulation. The estimates for various values of $t(2 \leq t \leq 2.8)$, b_n and sample sizes $n = 50$ and $n = 100$ are calculated. The ratios of mean-squared error of $\overline{H}_n^*(f;t)$ to that of the $\overline{H}_*^n(f;t)$ for the b_n values between 0.3 - 0.6 are computed and are given in Table 4. The ratios of mean-squared error of $\overline{H}_n^*(f;t)$ to that of the $\overline{H}_*^n(f;t)$ for $b_n = 0.6$ is very small compared to the other b_n values. In Figure 4 dark line represents the theoretical value $\overline{H}(f;t)$, starred line represents the estimator $\overline{H}_n^*(f;t)$ and dotted line represents the estimator $\overline{H}_*^n(f;t)$.

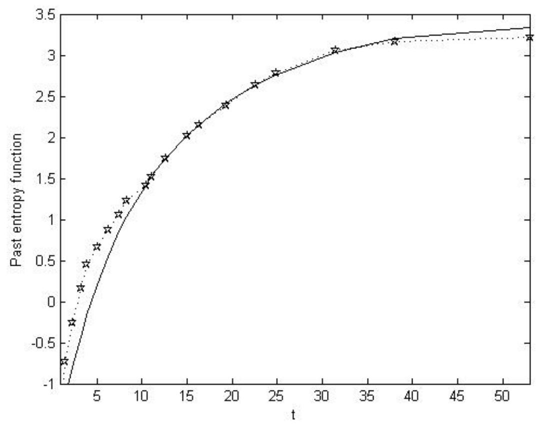

Figure 1: Plots of estimates of past entropy function for the first failure of 20 electric carts

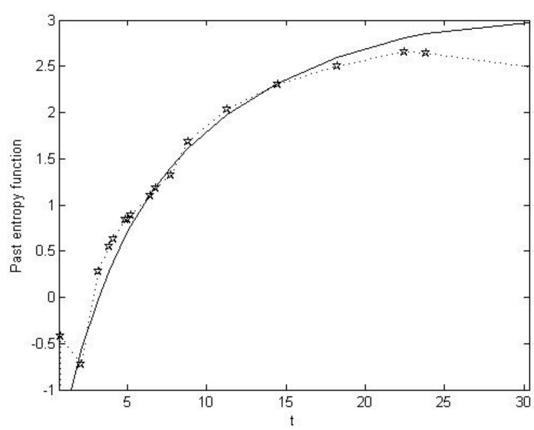

Figure 2: Plots of estimates of past entropy function for the life times of 20 sodium sulphur batteries

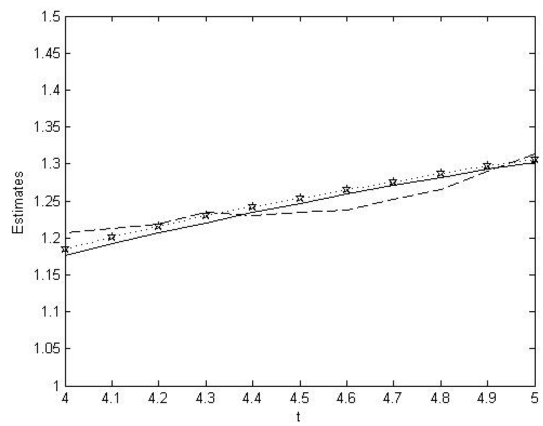

Figure 3: Plots of $\overline{H}_n(f;t)$, $\overline{H}^*(f;t)$ and $\overline{H}(f;t)$ using a simulated sample of size n=100 for $b_n = 0.4$, $\lambda = 0.6$

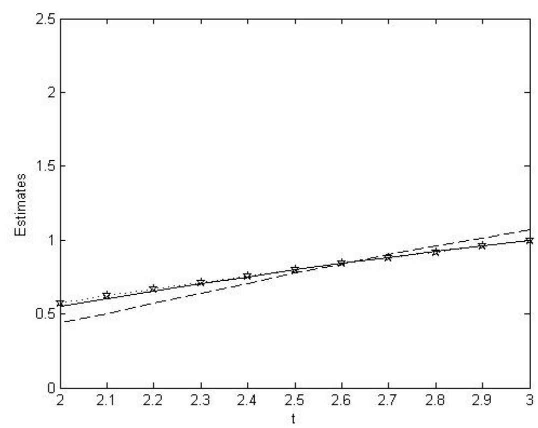

Figure 4: Plots of $\overline{H}_n^*(f;t)$, $\overline{H}_*^n(f;t)$ and $\overline{H}(f;t)$ using a simulated sample of size n=100 for $b_n = 0.6$, $\mu = 2$, $\sigma = 1$

Table 1: Bootstrap bias and mean-squared error estimates of $\overline{H}_n(f;t)$ for a real data set

	b_n							
t	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.9	0.0023	0.5532	0.8844	1.0846	1.2020	1.2609	1.2975	1.3203
	(0.2926)	(0.4403)	(0.8324)	(1.1921)	(1.4853)	(1.6067)	(1.6904)	(1.7463)
1.5	0.6078	0.9093	1.1552	1.2605	1.3597	1.4185	1.4598	1.4882
	(0.5650)	(0.9192)	(1.2871)	(1.6074)	(1.9259)	(2.0498)	(2.1491)	(2.2235)
2.3	0.5346	0.9260	1.1047	1.2274	1.3166	1.3834	1.4340	1.4725
	(0.8725)	(1.1871)	(1.4401)	(1.6560)	(1.8348)	(1.9817)	(2.1013)	(2.1978)
3.2	0.5853	0.9342	1.1175	1.2376	1.3179	1.3728	1.4117	1.4403
	(0.7449)	(1.1748)	(1.4580)	(1.6849)	(1.8515)	(1.9708)	(2.0571)	(2.1219)
3.9	0.6124	0.9964	1.1508	1.2489	1.3135	1.3569	1.3871	1.4092
	(0.6888)	(1.1554)	(1.4555)	(1.6654)	(1.8092)	(1.9068)	(1.9748)	(2.0243)
5	0.46702	0.8723	1.0389	1.1439	1.2116	1.2556	1.3099	1.3239
	(0.4823)	(0.8635)	(1.1541)	(1.3648)	(1.5114)	(1.6105)	(1.7214)	(1.7574)
6.2	0.3402	0.6974	0.8376	0.9278	0.9924	1.0413	1.0798	1.1112
	(0.3020)	(0.5770)	(0.7836)	(0.9316)	(1.0440)	(1.1323)	(1.2041)	(1.2643)
7.5	0.1942	0.5279	0.6910	0.7974	0.8721	0.9269	0.9683	1.0035
	(0.1248)	(0.3397)	(0.5243)	(0.6704)	(0.7856)	(0.9269)	(0.9516)	(1.0116)
8.3	0.1965	0.5232	0.6763	0.7705	0.8343	0.8800	0.9142	0.9407
	(0.1213)	(0.3042)	(0.4798)	(0.6102)	(0.7085)	(0.7842)	(0.8435)	(0.8909)
10.4	-0.0078	0.3083	0.4790	0.5870	0.6596	0.7103	0.7470	0.7743
	(0.0565)	(0.1166)	(0.2452)	(0.3566)	(0.4446)	(0.5123)	(0.5644)	(0.6050)
11.1	-0.0065	0.2857	0.4307	0.5210	0.5838	0.6304	0.6663	0.6947
	(0.0272)	(0.1173)	(0.2165)	(0.2983)	(0.3637)	(0.4166)	(0.4600)	(0.4961)
12.6	0.0065	0.2711	0.4051	0.4812	0.5309	0.5657	0.5910	0.6099

Continued on next page

Table 1 – continued from previous page

	b_n							
t	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
	(0.0141)	(0.0875)	(0.1796)	(0.2461)	(0.2953)	(0.3327)	(0.3615)	(0.3839)
15	-0.2251	-0.0113	0.1985	0.2833	0.3422	0.3863	0.4203	0.4468
	(0.0588)	(0.0145)	(0.0548)	(0.0954)	(0.1318)	(0.1637)	(0.1909)	(0.2136)
16.3	-0.2858	-0.0030	0.1477	0.2350	0.2945	0.3379	0.3705	0.3954
	(0.0894)	(0.0245)	(0.0365)	(0.0695)	(0.1004)	(0.1275)	(.1504)	(0.1692)
19.3	-0.4745	-0.1808	-0.0264	0.0694	0.1372	0.1882	0.2275	0.2581
	(0.2357)	(0.0494)	(0.0199)	(0.0238)	(0.0368)	(0.0524)	(0.0677)	(0.0815)
22.6	-0.6330	-0.3336	-0.1726	-0.0706	0.0026	0.0586	0.1026	0.1379
	(0.4080)	(0.1230)	(0.0432)	(0.0181)	(0.0123)	(0.0150)	(0.0214)	(0.0292)
24.8	-0.6913	-0.3843	-0.2147	-0.1062	-0.0288	0.0297	0.0749	0.1106
	(0.4870)	(0.1605)	(0.0605)	(0.0253)	(0.0140)	(0.0130)	(0.0169)	(0.0228)
31.5	-0.8917	-0.5765	-0.3966	-0.2781	-0.1914	-0.1244	-0.0711	-0.0278
	(0.8070)	(0.3483)	(0.1754)	(0.0955)	(0.0544)	(0.0329)	(0.0223)	(0.0181)
38.1	-0.9875	-0.6661	-0.4785	-0.35211	-0.2582	-0.1845	-0.1251	-0.0761
	(0.9878)	(0.4614)	(0.2498)	(0.1461)	(0.0894)	(0.0575)	(0.0399)	(0.0311)
53	-0.9854	-0.7393	-0.5461	-0.4143	-0.3154	-0.2372	-0.1736	-0.1206
	(1.0043)	(0.5629)	(0.3182)	(0.1934)	(0.1226)	(0.0807)	(0.0561)	(0.0420)

Table 2: Bootstrap bias and mean-squared error estimates of $\overline{H}_n^*(f;t)$ for a real data set

	b_n								
t	1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.76	-0.1701	0.5518	0.6866	0.7482	0.9105	1.0117	1.0768	1.1193	1.1472
	(0.0380)	(0.3064)	(0.7417)	(0.9092)	(1.0235)	(1.1234)	(1.2130)	(1.2799)	(1.3299)
0.82	0.0153	0.1799	0.4935	0.7639	0.9252	1.0277	1.0948	1.1393	1.1691
	(0.0094)	(0.8076)	(0.8673)	(0.9611)	(1.0705)	(1.1734)	(1.2606)	(1.3303)	(1.3834)
2.10	-0.1508	0.3183	0.6957	0.9149	1.0504	1.1371	1.1929	1.2281	1.2495
	(0.0587)	(0.2343)	(0.6029)	(0.9263)	(1.1642)	(1.3325)	(1.4478)	(1.5240)	(1.5714)
3.15	-0.3780	0.3057	0.6782	0.7428	0.8810	0.9716	1.0294	1.0636	1.0818
	(0.1944)	(0.1459)	(0.5032)	(0.6026)	(0.8046)	(0.9611)	(1.0724)	(1.1425)	(1.1809)
3.85	-0.4115	0.2665	0.5854	0.6760	0.7807	0.8492	0.8935	0.9212	0.9383
	(0.2279)	(0.1208)	(0.3786)	(0.4939)	(0.6329)	(0.7387)	(0.8137)	(0.8632)	(0.8942)
4.12	-0.3575	0.2448	0.5440	0.6368	0.7329	0.7951	0.8363	0.8639	0.8823
	(0.1871)	(0.1041)	(0.3265)	(0.4451)	(0.5641)	(0.6531)	(0.7176)	(0.7629)	(0.7939)
4.91	-0.4435	0.1661	0.4329	0.5530	0.6176	0.6643	0.6998	0.7268	0.7471
	(0.1871)	(0.1041)	(0.3265)	(0.4451)	(0.5641)	(0.6531)	(0.7176)	(0.7629)	(0.7939)
5.04	-0.4401	0.1309	0.4012	0.5354	0.6015	0.6483	0.6834	0.7098	0.7297
	(0.2643)	(0.0599)	(0.1849)	(0.3236)	(0.3938)	(0.4481)	(0.4911)	(0.5248)	(0.5510)
5.22	-0.3978	0.1180	0.3757	0.5163	0.5832	0.6294	0.6633	0.6885	0.7076
	(0.2091)	(0.0543)	(0.1663)	(0.3041)	(0.3720)	(0.4236)	(0.4638)	(0.4949)	(0.5191)
6.46	-0.5793	-0.0146	0.2737	0.4315	0.4953	0.5332	0.5586	0.5769	0.5906
	(0.3794)	(0.0313)	(0.0933)	(0.2081)	(0.2655)	(0.3036)	(0.3303)	(0.3499)	(0.3647)
6.78	-0.5326	-0.0154	0.2541	0.4052	0.4691	0.5075	0.5331	0.5515	0.5652
	(0.3303)	(0.0314)	(0.0840)	(0.1863)	(0.2398)	(0.2759)	(0.3015)	(0.3203)	(0.3345)
7.75	-0.6148	-0.0811	0.1868	0.3455	0.4048	0.4407	0.4651	0.4827	0.4960

Continued on next page

Table 2 – continued from previous page

	b_n								
t	1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
	(0.4281)	(0.0459)	(0.0635)	(0.1332)	(0.1714)	(0.2076)	(0.2293)	(0.2454)	(0.2578)
8.84	-0.7364	-0.1927	0.0787	0.2167	0.3184	0.3555	0.3814	0.4007	0.4157
	(0.6019)	(0.0890)	(0.0472)	(0.0784)	(0.1133)	(0.1382)	(0.1569)	(0.1713)	(0.1828)
11.31	-1.0193	-0.4619	-0.1781	-0.0260	0.0631	0.1204	0.1813	0.2117	0.2357
	(1.0966)	(0.2645)	(0.0733)	(0.0337)	(0.0298)	(0.0342)	(0.0495)	(0.0600)	(0.0691)
14.46	-1.1771	-0.5875	-0.2935	-0.1389	-0.0455	-0.0539	-0.0042	0.0334	0.0633
	(1.4036)	(0.3743)	(0.1146)	(0.0454)	(0.0274)	(0.0232)	(0.0175)	(0.0184)	(0.0176)
18.24	-1.4618	-0.9525	-0.6431	-0.4592	-0.3372	-0.2484	-0.1924	-0.1360	-0.0912
	(1.1969)	(0.9558)	(0.4599)	(0.2542)	(0.1547)	(0.1006)	(0.0518)	(0.0338)	(0.0238)
18.27	-1.5249	-0.9427	-0.6385	-0.4625	-0.3477	-0.2650	-0.1962	-0.1391	-0.0938
	(1.3631)	(0.9211)	(0.4344)	(0.2352)	(0.1386)	(0.0855)	(0.0532)	(0.0346)	(0.0243)
22.48	-1.7304	-0.1148	-0.7847	-0.5811	-0.4389	-0.3332	-0.2742	-0.2029	-0.1472
	(1.4295)	(1.2907)	(0.6567)	(0.3727)	(0.2242)	(0.1411)	(0.0945)	(0.0633)	(0.0460)
23.85	-1.8350	-1.2003	-0.8546	-0.6458	-0.5055	-0.4019	-0.3268	-0.2586	-0.2026
	(1.4978)	(1.3641)	(0.7509)	(0.4353)	(0.2732)	(0.1799)	(0.1223)	(0.0857)	(0.0628)
30.77	-2.6256	-1.6552	-1.2844	-1.0483	-0.8803	-0.7495	-0.6675	-0.5798	-0.5052
	(1.5625)	(1.4242)	(1.2883)	(1.1148)	(1.0687)	(0.8368)	(0.6163)	(0.4981)	(0.4095)

Table 3: Ratios of the mean-squared error of $\overline{H}_n(f;t)$ to that of $\overline{H}^*(f;t)$

		t								
b_n	Sample	4	4.1	4.2	4.3	4.4	4.5	4.6	4.7	4.8
0.2	n=50	0.6129	0.6451	0.6712	0.6948	0.7161	0.7340	0.7478	0.7592	0.7723
	n=100	0.2785	0.3182	0.3662	0.4235	0.4874	0.5532	0.6165	0.6748	0.7261
0.3	n=50	0.2257	0.2407	0.2560	0.2733	0.2938	0.3189	0.3494	0.3855	0.4271
	n=100	0.1598	0.1805	0.2045	0.2319	0.2629	0.2970	0.3336	0.3716	0.4095
0.4	n=50	0.1528	0.1721	0.1892	0.2061	0.2254	0.2481	0.2748	0.3058	0.3413
	n=100	0.1295	0.1453	0.1633	0.1838	0.2072	0.2335	0.2628	0.2952	0.3306
0.5	n=50	0.1556	0.1724	0.1905	0.2104	0.2323	0.2567	0.2842	0.3150	0.3494
	n=100	0.1497	0.1669	0.1860	0.2073	0.2309	0.2571	0.2858	0.3173	0.3514

Acknowledgements: The author is grateful to Dr. E. I. Abdul Sathar, Assistant Professor, Department of Statistics, University of Kerala for many of his valuable suggestions during this research. The author also wishes to express sincere thanks to Dr. B. V. Rao, Chennai Mathematical Institute and Dr. M. B. Rajarshi, University of Pune for their valuable comments for the advancement of this research. Finally, the financial support from University of Kerala is also gratefully acknowledged.

References

Asadi, M. and Ebrahimi, N. (2000). Residual entropy and its characterizations in terms of hazard function and mean residual life function, Statistics & Probability Letters 49, 263-269.

Ansell, R. O. and Ansell, J. I. (1987). Modeling the reliability of sodium sulphur cells, Reliability Engineering 17, 127-137.

Belzunce, F., Navarro, J., Ruiz, J. M. and Aguila, Y. (2004). Some results on residual entropy function, Metrica 59, 147-161.

Cai, Z. (1998). Kernel density and hazard rate estimation for censored dependent data, Journal of Multivariate Analysis 67, 23-34.

Di Crescenzo, A. and Longobardi, M. (2002). Entropy-based measure of uncertainty in past lifetime distributions, Journal of Applied Probability 39, 434-440.

Table 4: Ratios of the mean-squared error of $\overline{H}_n^*(f;t)$ to that of $\overline{H}_*^n(f;t)$

		t								
b_n	Sample	2	2.1	2.2	2.3	2.4	2.5	2.6	2.7	2.8
0.3	n=50	0.1558	0.2402	0.3572	0.5641	0.8525	1.3470	1.9956	1.9176	1.4388
	n=100	0.0596	0.1032	0.1847	0.3232	0.5891	1.4129	2.7669	1.9731	1.0698
0.4	n=50	0.0835	0.1386	0.2176	0.3516	0.5366	0.8639	1.2982	1.1960	0.8841
	n=100	0.0211	0.0424	0.0857	0.1646	0.3136	0.7762	1.4064	1.0303	0.6066
0.5	n=50	0.0367	0.0635	0.1044	0.1734	0.2670	0.4399	0.6886	0.6516	0.5071
	n=100	0.0467	0.0418	0.0586	0.0732	0.1390	0.3537	0.6653	0.5349	0.3381
0.6	n=50	0.0236	0.0309	0.0459	0.0739	0.1127	0.1890	0.3081	0.2913	0.2607
	n=100	0.0191	0.0251	0.0309	0.0651	0.0903	0.1907	0.3423	0.2924	0.1912

Di Crescenzo, A. and Longobardi, M. (2004.) A measure of discrimination between past lifetime distributions, Statistics & Probability Letters 67, 173-182.

Ebrahimi, N. (1996). How to measure uncertainty in the residual life time distribution, Sankhya A 58, 48-56.

Ebrahimi, N. and Kirmani, S. N. U. A. (1996). Some results on ordering of survival function through uncertainty, Statistics & Probability Letters 29, 167-176.

Ebrahimi, N. and Pellery, F. (1995). New partial ordering of survival functions based on the notion of uncertainty, Journal of Applied Probability 32, 202-211.

Efron, B. (1981). Censored data and the bootstrap, Journal of American Statistical Association 76, 312-319.

Gupta, R. D. and Nanda, A. K. (2002). α- and β-entropies and relative entropies of distributions, Journal of Statistical Theory and Applications 1 (3), 177-190.

Gupta, R. C. (2009). Some characterization results based on residual entropy function, Journal of Statistical Theory and Applications 8 (1), 45-59.

Nair, K. R. M. and Rajesh, G. (1998). Characterization of probability distributions using the residual entropy function, Journal of Indian Satistical Association 36, 157-166.

Nanda, A. K. and Paul, P. (2006). Some properties of past entropy and their applications, Metrika 64, 47-61.

Parzen, E. (1962). On estimation of a probability density function and mode, The Annals of Mathematical Statistics 33(3), 1065-1076.

Rosenblatt, M. (1956). A central limit theorem and a strong mixing condition, Proceedings of the National Academy of Sciences of the United States of America 42, 43-47.

Rosenblatt, M. (1970). Density estimates and Markov sequences, Nonparametric Techniques in Statistical Inference, ed. M. Puri, Oxford Cambridge University Press, 199-210.

Shannon, C. E. (1948). A mathematical theory of communication, Bell System Technical Journal, 279-423 and 623-656.

Zimmer, W. I., Keats, J. B. and Wang, F. K. (1998). The Burr xii distribution in reliability analysis, Journal of Quality Technology 30, 386-394.

Journal of the Kerala Statistical Association
Vol.24, December 2013, 37-52 *ISSN 2249-4553*

On Explicit Expressions for Single and Product Moments of Generalized Order Statistics from a New Class of Exponential Distributions and a Characterization

Jagdish Saran* and Kamal Nain
Department of Statistics
University of Delhi, Delhi -110007

ABSTRACT

In this paper, we establish exact expressions for single and product moments of generalized order statistics from a new family of exponential distributions. We also give a characterization result for this class of distributions.

Key words and Phrases: *Generalized order statistics; single and product moments; Frechet-type, Gumble and modified extreme value disributions; characterization*

1 Introduction

Generalized order statistics (gos) have been introduced and extensively studied in Kamps (1995 a,b) as a unified theoretical set-up which contains a variety of models of ordered random variables with different interpretations. Examples of such models are: Ordinary order statistics, Sequential order statistics, Progressive type II censored order statistics, Record values, kth record value and Pfeifer's records.

*email:jagdish_saran52@gmail.com

There is no natural interpretation of generalized order statistics in terms of observed random samples but these models can be effectively applied in life testing and reliability analysis, medical and life time data, and models related to software reliability analysis. The common approach makes it possible to define several distributional properties at once. The structural similarities of these models are based on the similarity of their joint density function.

Let $\{X_n, n \geq 1\}$be a sequence of absolutely continuous, independent and identically distributed random variables with cdf $F(x) = P\ (X \leq x\ )$ and pdf f(x). Assume k > 0, n$\in${2,3,...}, $\tilde{m} = (\,m_1\,,\, m_2\,,\, ...,\, m_{n-1})\in R^{n-1}$, $M_r = \sum_{j=r}^{n-1} m_j$, such that $\gamma_r = k+n-r+M_r > 0$ for all $r \in \{\,1,2,...,n-1\,\}$. Then $X\,(\,r,\, n,\, \tilde{m},\, k\,)$, r = 1, 2,...,n , are called gos if their joint pdf is given by

$$f^{X(1,n,\tilde{m},k),\, X(2,n,\tilde{m},k),\, ...,\, X(n,n,\tilde{m},k)}(\,x_1,\, x_2, ..., x_n\,) = k\left(\prod_{j=1}^{n-1}\gamma_j\right)$$

$$\left(\prod_{i=1}^{n-1}(1-F(x_i))^{m_i}\, f(x_i)\right)(1-F(x_n))^{k-1}\, f(x_n), \tag{1.1}$$

where $F^{-1}\,(0+) \ < x_1 \ \leq x_2 \leq ... \ \leq x_n < F^{-1}(1)$.

By choosing appropriate values of parameters, we get the distribution of a few very common statistics as shown in Table 1.1 given below.

The joint pdf of first r, gos is given by :

$$f^{X(1,n,\tilde{m},k),\, X(2,n,\tilde{m},k),\, ...,\, X(r,n,\tilde{m},k)}\ (\,x_1,\, x_2, ..., x_r\,) = c_{r-1}\left(\prod_{i=1}^{r-1}(1-F(x_i))^{m_i}\, f(x_i)\right)$$

$$(1-F(x_r))^{k+n-r+M_r-1}\, f(x_r), \tag{1.2}$$

where $F^{-1}\,(0+) \ < x_1 \ \leq x_2 \leq ... \ \leq x_r < F^{-1}(1)$.

We now consider two cases:

Case I: $m_1 = m_2 = ... = m_{n-1} = m$

Case II: $\gamma_i \neq \gamma_j$; $i \neq j$, i , j = 1, 2, ..., n - 1.

For case I, the gos will be denoted by X(r, n, m, k). The pdf of X(r, n, m, k) is given by

$$f^{X(r,n,m,k)}\,(\,x\,) = \frac{c_{r-1}}{(r-1)!}\ (1-F(x))^{\gamma_r - 1}\, f(\,x\,)\ g_m^{r-1}\,(F(x))\,,\quad x \in R, \tag{1.3}$$

Table 1.1:

S.No.	Choice of parameters for i = 1, 2,. . . , n	gos becomes
1	$\gamma_i = n - i + 1,\ m_1 = m_2 = \ldots = m_{n-1} = 0$ and k = 1	Joint distribution of n order statistics
2	$\gamma_i = k,\ m_1 = m_2 = \ldots = m_{n-1} = -1\,,\ k \in N$	kth record value
3	$\gamma_i = (n - i + 1)\alpha_i\,,\ \alpha_i > 0$	Sequential order statistics
4	$\gamma_i = \alpha - i + 1\,,\ \alpha > 0$	Order statistics with non integer sample size
5	$\gamma_i = \beta_i\,,\ \beta_i > 0$	Pfeifer's record values
6	$m_i \in N_o\,,\ k \in N$	Progressively type-II right censored order statistics

and the joint pdf of X(r, n, m, k) and X(s, n, m, k), $1 \leq r < s \leq n$, is given by

$$f^{X(r,\,n,m,\,k),\,X(s,\,n,m\,,k)}\,(x,\ y) = \frac{c_{s-1}}{(r-1)!\ (s-r-1)!}\,\left((1-F(x))^m f(x)\right)\ g_m^{r-1}\,(F(x))$$

$$[h_m\,(F(y)) - h_m\,(F(x))]^{\,s-r-1}\ (1-F(y))^{\gamma_s - 1} f(y)\quad,\ x < y, \tag{1.4}$$

where $c_{r-1} = \prod_{j=1}^{r} \gamma_j\,,\qquad \gamma_j = k + (n-j)(m+1)\ ,\ r = 1,\ 2, ...,\ n\,,$
$g_m(x) = \ h_m(x) - h_m(0)\,,\ \ x \in\ (0,\,1)$ and

$$h_m(x) = \begin{cases} -\ \frac{(1-x)^{m+1}}{m+1}\ , & m\ \neq -1, \\ -\ \log\,(1-x)\ , & m\ = -1. \end{cases} \tag{1.5}$$

For case II, the gos will be denoted by $X\,(r,\,n,\,\tilde{m},\,k)$. The pdf of $X\,(r,\,n,\,\tilde{m},\,k)$ is given by

$$f^{X(r,n,\tilde{m},k)}\,(x) = c_{r-1}\ f(x)\ \sum_{i=1}^{r} a_i(r)\,(1-F(x))^{\gamma_i - 1}\ ,\quad x \in R, \tag{1.6}$$

and the joint pdf of $X(r,\ n,\tilde{m},\ k)$ and $X(s,\ n,\tilde{m},k)$, $1 \leq r < s \leq n$, is given by

$$f^{X(r,\,n,\tilde{m},\,k),\,X(s,\,n,\tilde{m},k)}\quad(x,\ y) = c_{s-1}\ \left\{\textstyle\sum_{i=r+1}^{s} a_i^r(s)\left(\frac{1-F(y)}{1-F(x)}\right)^{\gamma_i}\right\}$$
$$\{\textstyle\sum_{i=1}^{r} a_i(r)\,(1-F(x))^{\gamma_i}\}\,\frac{f(x)}{1-F(x)}\ \frac{f(y)}{1-F(y)}\ , \qquad (1.7)$$

where $c_{s-1} = \prod_{j=1}^{s} \gamma_j\,, \qquad \gamma_j = k+n-j+m_j\ ,\ s = 1,\ 2, ...,\ n.$

Further, it can be proved that

(i) $a_i(r) = \prod_{j\,(\neq i)\,=1}^{r} \ (\gamma_j - \gamma_i)^{-1} \quad , \quad 1 \leq i \leq r \leq n$

(ii) $a_i^r(s) = \prod_{j\,(\neq i)\,=r+1}^{s} \ (\gamma_j - \gamma_i)^{-1} \quad , \quad r+1 \leq i \leq s \leq n$

(iii) $a_i(r) = (\gamma_{r+1} - \gamma_i)\ \ a_i(r+1)$

(iv) $c_r\ = c_{r-1}\gamma_{r+1}$

(v) $\sum_{i=1}^{r+1} a_i(r+1) = 0.$

The moments of order statistics have generated considerable interest in the recent years. Several recurrence relations and identities satisfied by single as well as product moments of order statistics have been obtained by various authors in the past. These relations help in reducing the quantum of computations involved. Joshi (1978, 1982) established recurrence relations for exponential distribution with unit mean and were further extended by Balakrishnan and Joshi (1984) for doubly truncated exponential distribution. For linear-exponential distribution, Balakrishnan and Malik (1986) derived the similar type of relations which were extended to doubly truncated linear-exponential distribution by Mohie El-Din et al. (1997) and Saran and Pushkarna (1999). Nain (2010 a, b) obtained recurrence relations for ordinary order statistics and kth record values from pth order exponential and generalized Weibull distributions, respectively.

The recurrence relations for the moments of generalized order statistics based on non identically distributed random variables were developed by Kamps (1995 a, b). Pawlas and Szynal (2001) obtained recurrence relations for single and product moments of generalized order statistics from Pareto, generalized Pareto and Burr distributions. Saran and Pandey (2004, 2009) established recurrence relations for single and product moments of generalized order statistics from linear-exponential and Burr distributions. Saran and Pandey (2011) obtained recurrence relations for marginal and joint moment generating functions of dual (lower) generalized order statistics from inverse Weibull distribution.

In this paper, we derive exact expressions for single and product moments of generalized order statistics for a new class of exponential distributions defined below in Section 2, and discuss its various particular cases. We also give a characterization result for this class of distributions. The results so obtained are generalized versions of Khan et al. (2012).

2 A New Family of Exponential Distributions

Consider a family of exponential distributions defined by the function

$$F(x) = 1 - \left(1 - e^{-\Psi(x)}\right)^{\eta}, \eta > 0 \ and \ 0 < x < \infty, \tag{2.1}$$

where$\Psi(0) = \infty$, $\Psi(\infty) = 0$ and $\Psi(x)$is monotonic in nature with inverse function $\phi(x)$, i.e. , $\Psi^{-1} = \varphi$.

The distribution has many applications in the area of equity risks, extreme floods, the amounts of large insurance losses, the size of freak waves, mutational events during evolution, large wildfires, pipeline failures due to pitting corrosion, etc. The table 2.1 given below demonstrates a few standard distributions obtained from (2.1) by choosing appropriate value of the parameter η and the function $\Psi(x)$.

The mathematical form of the distribution function, as given in (2.1), is very useful to derive the exact expressions for single and product moments of gos.

Notations

For n = 1, 2, 3,... , a > 0, b > 0, c > 0, $1 \leq r < s \leq n$, $k \geq 1 \ and \ u, v \in \{ 0, 1, 2, ... \}$,

Table 2.1:

S.No.	Choice of parameter η and the function $\Psi(x)$	Family of exponential distribution represents
1	$\eta = 1,\ \Psi(x) = \left(\frac{\beta}{x}\right)^{\alpha},\ \phi(x) = \beta\, x^{-\frac{1}{\alpha}},\ 0 < x < \infty$ *and* $\beta > 0$.	Frechet-type extreme value distribution
2	$\eta = 1,\ \Psi(x) = e^{-\lambda x},\ \phi(x) = -\frac{1}{\lambda}\log x,\ 0 < x < \infty$ *and* $\lambda > 0$.	Gumbel Extreme value distribution
3	$\eta = 1,\ \Psi(x) = e^{-(\lambda x)^{\frac{1}{\alpha}}},\ \phi(x) = \frac{1}{\lambda}\left(-\log x\right)^{\alpha},\ 0 < x < \infty$ *and* $\alpha, \lambda > 0$.	Modified extreme value Type-I distribution

we denote by

1. $$H_u\left(a,b\right) = \int_0^\infty x^u \left(1 - F(x)\right)^a f\left(x\right) g_m^b\left(F(x)\right) dx \tag{2.2}$$
2. $$H_{u\,,\,v}\left(a,b,c\right) = \int_0^\infty \int_x^\infty x^u\, y^v \left(1 - F(x)\right)^a f\left(x\right) \times \left[h_m\left(F(y)\right) - h_m\left(F(x)\right)\right]^b \left(1 - F(y)\right)^c f(y)\, dy\, dx \tag{2.3}$$
3. $$\mu^{u}_{m,\ n,\ k}\left(r\right) = E\left(\mathrm{X}\left(\mathrm{\ r,\ n,\ m,\ k\ }\right)\right)^u \tag{2.4}$$
4. $$\mu^{u,\ v}_{m,\ n,\ k}\left(r,s\right) = E\left(\left(\mathrm{X}\left(\mathrm{\ r,\ n,\ m,\ k\ }\right)\right)^u \left(\mathrm{X}\left(\mathrm{\ s,\ n,\ m,\ k\ }\right)\right)^v\right) \tag{2.5}$$
5. $$\mu^{u}_{\tilde{m},\ n,\ k}\left(r\right) = E\left[\mathrm{X}\left(\mathrm{\ r,\ n,\ }\tilde{\mathrm{m}}\mathrm{,\ k\ }\right)\right]^u \tag{2.6}$$
6. $$\mu^{u,\ v}_{\tilde{m},\ n,\ k}\left(r,s\right) = E\left[\left(\mathrm{X}\left(\mathrm{\ r,\ n,\ }\tilde{\mathrm{m}}\mathrm{,\ k\ }\right)\right)^u \left(\mathrm{X}\left(\mathrm{\ s,\ n,\ }\tilde{\mathrm{m}}\mathrm{,\ k\ }\right)\right)^v\right] \tag{2.7}$$

3 Some Auxiliary Results

In this section, we establish some results which will be useful later for deriving the main results.

Lemma 3.1 For the class of distributions defined in (2.1) and non-negative finite integers i, j, a and b,

$$H_i\left(a,b\right)=\begin{cases}\frac{1}{(m+1)^b}\sum_{d=0}^{b}\sum_{w=0}^{\infty}\begin{pmatrix} b\\ d\end{pmatrix}(-1)^d\,\frac{\beta_w(i)}{\frac{w+i}{\eta}+a+(m+1)d+1}, & m\neq -1,\\ b\,!\sum_{w=0}^{\infty}\ \frac{\beta_w(i)}{\left(\frac{w+i}{\eta}+a+1\right)^{b+1}}, & m=-1,\end{cases}$$

where $\beta_w\left(i\right)$is the co-efficient of t^{w+i} in $\left[\sum_{w=0}^{\infty}\frac{\phi_w(0)}{w!}\left(\sum_{s=1}^{\infty}\frac{t^s}{s}\right)^w\right]^i$ and $\phi=\Psi^{-1}$ as defined earlier.

Proof .

Case 1. $m\neq -1$.

Substituting $g_m^b\left(F(x)\right)=\left(\frac{1-(1-F(x))^{m+1}}{m+1}\right)^b=\frac{1}{(m+1)^b}\sum_{d=0}^{b}\begin{pmatrix} b\\ d\end{pmatrix}(-1)^d\left(1-F(x)\right)^{(m+1)d}$ in (2.2), we get

$$H_i\left(a,b\right)=\frac{1}{\left(m+1\right)^b}\sum_{d=0}^{b}\begin{pmatrix} b\\ d\end{pmatrix}(-1)^d\int_0^{\infty}x^i\left(1-F(x)\right)^{a+(m+1)d}f\left(x\right)dx.$$

Putting $t=\left(1-F(x)\right)^{\frac{1}{\eta}}$, we have

$$\begin{aligned}H_i\left(a,b\right)&=\frac{\eta}{\left(m+1\right)^b}\sum_{d=0}^{b}\sum_{w=0}^{\infty}\begin{pmatrix} b\\ d\end{pmatrix}(-1)^d\,\beta_w\left(i\right)\int_0^1 t^{w+i+(a+(m+1)d+1)\,\eta-1}\,dt\\ &=\frac{\eta}{\left(m+1\right)^b}\sum_{d=0}^{b}\sum_{w=0}^{\infty}\begin{pmatrix} b\\ d\end{pmatrix}(-1)^d\,\frac{\beta_w\left(i\right)}{w+i+\left(a+(m+1)d+1\,\right)\eta},\end{aligned}$$

which leads to the relation as stated in Lemma 3.1 for the case $m\neq -1$.

Case 2. $m=-1$

By using repeatedly the combinatorial identity (see, Ruiz, 1996)

$$\sum_{d=0}^{b}\begin{pmatrix} b\\ d\end{pmatrix}(-1)^d\ d^k=\begin{cases}0, & k=0,1,2,...,b-1,\\ (-1)^b\,b!, & k=b,\end{cases}\tag{3.1}$$

we get

$$H_i\left(a,b\right)=\lim_{m\,\to\,-1}\sum_{w=0}^{\infty}\beta_w\left(i\right)\frac{\sum_{d=0}^{b}\begin{pmatrix} b\\ d\end{pmatrix}(-1)^d\ \times\left(\frac{w+i}{\eta}+a+(m+1)d+1\right)^{-1}}{\left(m+1\right)^b}$$

$$= \sum_{w=0}^{\infty} \beta_w(i) \left(\frac{w+i}{\eta} + a + 1 \right)^{-b-1} \times \sum_{d=0}^{b} \begin{pmatrix} b \\ d \end{pmatrix} (-1)^{d+b} \times d^b$$

(by using L-Hospital Rule),

which again on using (3.1) for k = b leads to the relation as stated in Lemma 3.1 for the case $m = -1$.

Lemma 3.2 For the class of distributions (2.1) and non-negative finite integers i, j, a, b and c,

$$H_{i\,,\,j}(a,b,c) = \begin{cases} \frac{1}{(m+1)^b} \sum_{v=0}^{b} \sum_{w=0}^{\infty} \sum_{w'=0}^{\infty} \begin{pmatrix} b \\ v \end{pmatrix} (-1)^v \times \frac{\beta_w(i)}{\frac{w+w'+i+j}{\eta} + (a+\,c+b(m+1)+2)} \\ \qquad \times \frac{\beta_{w'}(j)}{\frac{w'+j}{\eta} + ((b-v)(m+1)+c+1)}, \qquad m \neq -1, \\ b! \sum_{w=0}^{\infty} \sum_{w'=0}^{\infty} \frac{\beta_w(i)}{\left(\frac{w+w'+i+j}{\eta} + a+\,c+2\right)^{b+1}} \times \frac{\beta_{w'}(j)}{\left(\frac{w'+j}{\eta} + a+1\right)^{b+1}}, \qquad m = -1, \end{cases}$$

where $\beta_w(i)$is as defined in Lemma 3.1.

Proof

Case 1: $m \neq -1$.

From (2.3), we have, for $b = 0$,

$$H_{i\,,\,j}(a,0,c) = \int_0^{\infty} \int_x^{\infty} x^i\, y^j\, (1-F(x))^a\, f(x)\, (1-F(y))^c\, f(y)\, dy\, dx$$

$$= \int_0^{\infty} x^i\, (1-F(x))^a\, f(x)\, G(x)\, dx, \tag{3.2}$$

where

$$G(x) = \int_x^{\infty} y^j\, (1-F(y))^c\, f(y)\, dy. \tag{3.3}$$

Putting $t = (1-F(y))^{\frac{1}{\eta}}$ in (3.3), we have

$$G(x) = \eta \int_0^{(1-F(x))^{\frac{1}{\eta}}} \left(\sum_{w'=0}^{\infty} \beta_{w'}(j)\ t^{j\,+\,w'} \right) t^{\eta\, c+\eta-1} dt$$

$$= \eta \sum_{w'=0}^{\infty} \beta_{w'}(j) \int_0^{(1-F(x))^{\frac{1}{\eta}}} t^{w'+j+\eta(1+\,c)-1} dt$$

$$= \sum_{w'=0}^{\infty} \beta_{w'}(j) \quad \left[\frac{(1-F(x))^{\frac{w'+j}{\eta}+c+1}}{\frac{w'+j}{\eta}+c+1} \right],$$

which on substituting in (3.2) gives,

$$H_{i\,,\,j}(a,0,c) = \eta^2 \sum_{w=0}^{\infty} \sum_{w'=0}^{\infty} \beta_w(i)\, \beta_{w'}(j) \; \frac{1}{w'+j+(c+1)\eta} \int_0^1 t^{w+w'+i+j+\eta(a+\ c+2)-1}\, dt$$

$$= \eta^2 \sum_{w=0}^{\infty} \sum_{w'=0}^{\infty} \frac{\beta_w(i)}{w+w'+i+j+(a+\ c+2)\,\eta} \times \frac{\beta_{w'}(j)}{w'+j+(c+1)\,\eta}\,. \qquad (3.4)$$

Further, on substituting the value of

$$[h_m(F(y)) - h_m(F(x))]^b = \left[\frac{(1-F(x))^{m+1} - (1-F(y))^{m+1}}{(m+1)} \right]^b$$

$$= \frac{1}{(m+1)^b} \sum_{v=0}^{b} \binom{b}{v} (-1)^v \; (1-F(x))^{v(m+1)} (1-F(y))^{(b-v)(m+1)}$$

in (2.3), we get

$$H_{i\,,\,j}(a,b,c) = \frac{1}{(m+1)^b} \sum_{v=0}^{b} \binom{b}{v} (-1)^v$$

$$\times \left[\int_0^{\infty} \int_x^{\infty} x^i\, y^j\, (1-F(x))^{a+v(m+1)} f(x)\, (1-F(y))^{(b-v)(m+1)+c} f(y)\, dy\, dx \right]$$

$$= \frac{1}{(m+1)^b} \sum_{v=0}^{b} \binom{b}{v} (-1)^v H_{i\,,\,j}(a+v(m+1),\ 0\ ,(b-v)(m+1)+c)\,,$$

(by using (3.2))

which on using (3.4) leads to the relation as stated in Lemma 3.2 for the case $m \neq -1$.

Case 2. $m = -1$.

The proof is similar to the one used in case 2 of Lemma 3.1.

4 Explicit Expressions For Single and Product Moments

Theorem 4.1 For $m_1 = m_2 = ... = m_{n-1} = m$, n = 1, 2, 3,... , $1 \leq r \leq n$, $k \geq 1$ and $u \in \{ 0, 1, 2, ... \}$,

$$\mu^{u}_{m,\ n,\ k}\,(r) = \frac{c_{\,r\,-\,1}}{(r-1)!}\, H_u\,(\gamma_r - 1,\ r-1)\,, \tag{4.1}$$

where $H_u\,(\gamma_r - 1,\ r-1)$ is as defined in Lemma 3.1.

Proof. On using (2.4) and (1.3), the *uth* order moment of $X(r, n, m, k)$ is given by

$$\mu^{u}_{m,\ n,\ k}\,(r) = \frac{c_{\,r\,-\,1}}{(r-1)!} \int_0^\infty x^u\,(1 - F(x))^{\gamma_r - 1}\, f(\,x\,)\ g_m^{r\,-\,1}\,(F(x))\,dx.$$

By using (2.2), we shall derive the relation as stated in (4.1).

Corollary 4.1 For $\gamma_i \neq \gamma_j$; $i \neq j$, $i, j = 1, 2, ..., n-1$, $n = 1, 2, 3, ...$, $1 \leq r \leq n$, $k \geq 1$ and

$$u \in \{ 0, 1, 2, ... \}\,,$$

$$\mu^{u}_{\tilde{m},\ n,\ k}\,(r) = c_{\,r\,-\,1} \sum_{i\,=\,1}^{r} a_i(r)\, H_u\,(\gamma_i - 1, 0)\,, \tag{4.2}$$

where $H_u\,(\gamma_i - 1, 0)$ is as defined in Lemma 3.1.

Proof. The proof is similar to the proof of theorem 4.1 on using (2.6) and (1.6).

Theorem 4.2 For $m_1 = m_2 = ... = m_{n-1} = m$, n = 1, 2, 3,... , $1 \leq r < s \leq n$, $k \geq 1$ and $u, v \in \{ 0, 1, 2, ... \}$,

$$\mu^{u,\ v}_{m,\ n,\ k}\,(r, s) = \frac{c_{\,s\ -1}}{(r-1)!\ (s-r-1)!\ (m+1)^{r-1}} \sum_{d=0}^{r-1} \begin{pmatrix} r-1 \\ d \end{pmatrix} (-1)^d$$

$$\times\, H_{u\,,\,v}\,(m + (m+1)d,\ s-r-1,\ \gamma_s - 1)\,, \tag{4.3}$$

where $H_{u\,,\,v}\,(m + (m+1)d,\ s-r-1,\ \gamma_s - 1)$ is as defined in Lemma 3.2.

Proof. On using (2.5) and (1.4), we have

$$\mu^{u,\ v}_{m,\ n,\ k}\,(r, s) = \frac{c_{\,s\ -1}}{(r-1)!\ (s-r-1)!} \int_0^\infty \int_x^\infty x^u\, y^v\ \left((1 - F(x))^m\, f(x) \right)\ g_m^{r\,-\,1}\,(F(x))$$

$$\times\ [h_m(F(y)) - h_m(F(x))]^{s-r-1}\,(1-F(y))^{\gamma_s - 1} f(y)\,dy\,dx. \tag{4.4}$$

Substituting

$$\begin{aligned} g_m^{r-1}(F(x)) &= \left(\frac{1-(1-F(x))^{m+1}}{m+1}\right)^{r-1} \\ &= \frac{1}{(m+1)^{r-1}} \sum_{d=0}^{r-1} \binom{r-1}{d} (-1)^d (1-F(x))^{(m+1)d} \end{aligned}$$

in (4.4), we have

$$\begin{aligned} \mu_{m,\ n,\ k}^{u,\ v}(r,s) &= \frac{c_{s\ -1}}{(r-1)!\ (s-r-1)!\ (m+1)^{r-1}} \\ & \sum_{d=0}^{r-1} \binom{r-1}{d} (-1)^d \int_0^\infty \int_x^\infty x^u\, y^v\, \Big((1-F(x))^{m+(m+1)d} f(x)\Big) \\ & \times\ [h_m(F(y)) - h_m(F(x))]^{s-r-1}\,(1-F(y))^{\gamma_s-1} f(y)\,dy\,dx. \end{aligned}$$

The relation (4.3) follows immediately on using (2.3).

Corollary 4.2 For $\gamma_i \neq \gamma_j$; $i \neq j$, $i, j = 1, 2, ..., n-1$**,** n = 1, 2, 3,... , $1 \leq r < s \leq n$,, $k \geq 1$ and $u, v \in \{\ 0, 1, 2, ...\}$,

$$\mu_{\tilde{m},n,k}^{u,v}(r,s) = c_{s-1} \sum_{i\ =\ 1}^{r} a_i(r) \sum_{j\ =\ r+1}^{s} a_j^r(s)\ H_{i\ ,\ j}(\gamma_i - \gamma_j - 1, 0, \gamma_j - 1)\,, \tag{4.5}$$

where $H_{i\ ,\ j}(\gamma_i - \gamma_j - 1, 0, \gamma_j - 1)$is as defined in Lemma 3.2.

Proof. The proof is similar to the proof of theorem 4.2 on employing (2.7) and (1.7).

5 Characterization

Let $X(r,n,m,k)$, $r = 1, 2, \ldots, n$ be the gos from a continuous type of distribution with cumulative distribution function $F(x)$ and probability density function $f(x)$. Then, in view of (1.3) and (1.4), the conditional density function of $Y = X(s,n,m,k)$ given $X(r,n,m,k) = x$, $1 \leq r < s \leq n$, is

$$f(y|x) = \sigma \left[1 - \left(\frac{1-F(y)}{1-F(x)}\right)^{m+1}\right]^{s-r-1} \left(\frac{1-F(y)}{1-F(x)}\right)^{\gamma_s - 1} \frac{f(y)}{1-F(x)}\,, \quad 0 < x < y < \infty, \tag{5.1}$$

where $\sigma = \frac{c_{s-1}}{(s-r-1)!\; c_{r-1}\; (m+1)^{s-r-1}}$.

Theorem 5.1 Let X be a non-negative, absolutely continuous type of random variable with distribution function F(x) satisfying the conditions $F(0) = 0$ and $0 < F(x) < 1$. Then a necessary and sufficient condition for

$$E\left((X(s,\; n,m\,,\,k))^i\,|X(r,\; n,m\,,\,k) = x\right) = \sum_{w=0}^{\infty} \beta_w\,(i)\, \left(1 - e^{-\psi(x)}\right)^{w+i} \times \prod_{j=1}^{s-r} \left(\frac{\gamma_{r+j}}{\frac{w+i}{\eta} + \gamma_{r+j}}\right) \tag{5.2}$$

is that

$$F\,(x) = 1 - \left(1 - e^{-\psi(x)}\right)^{\eta}\,,\; x > 0\,,\; \eta > 0,$$

where $\psi(x)$ is monotonic function satisfying $\psi o \varphi(x) = x$ for some function $\varphi(x)$.

Proof. On using (5.1), we have

$$E\left(Y^i|X\,(r,\; n,m\,,\,k) = x\right)$$
$$= \sigma \int_x^{\infty} y^i \left(1 - \left(\frac{1-F(y)}{1-F(x)}\right)^{m+1}\right)^{s-r-1} \left(\frac{1-F(y)}{1-F(x)}\right)^{\gamma_s - 1} \frac{f(y)}{1-F(x)}\,dy$$

Put $u = \frac{1-F(y)}{1-F(x)}$ in the above equation we get

$$\begin{aligned} & E\left(Y^i|X\,(r,\; n,m\,,\,k) = x\right) \\ &= \sigma \sum_{w=0}^{\infty} \beta_w\,(i)\, \left(1 - e^{-\psi(x)}\right)^{w+i} \int_0^1 u^{\frac{w+i}{\eta} + \gamma_s - 1} \left(1 - u^{m+1}\right)^{s-r-1} du, \\ &= \sigma \sum_{w=0}^{\infty} \beta_w\,(i)\, \left(1 - e^{-\psi(x)}\right)^{w+i} B\left(\frac{w+i}{\eta(m+1)} + \frac{\gamma_s}{m+1}, s-r\right) (By\ putting\ u^{m+1} = v) \\ &= \frac{c_{s-1}}{c_{r-1}} \sum_{w=0}^{\infty} \beta_w\,(i)\, \left(1 - e^{-\psi(x)}\right)^{w+i} \prod_{j=1}^{s-r} \left(\frac{w+i}{\eta} + \gamma_{r+j}\right)^{-1}, \end{aligned}$$

which, on substituting the value of $\frac{c_{s-1}}{c_{r-1}}$, leads to (5.2). This proves the sufficient condition.

Let

$$Z_r(x) = \sum_{w=0}^{\infty} \beta_w\,(i)\, \left(1 - e^{-\psi(x)}\right)^{w+i} \prod_{j\,=1}^{s-r} \left(\frac{\gamma_{r+j}}{\frac{w+i}{\eta} + \gamma_{r+j}} \right). \tag{5.3}$$

Then it implies that

$$Z_{r+1}(x) - Z_r(x) = \frac{1}{\eta\,\gamma_{r+1}} \sum_{w\,=0}^{\infty} \beta_w\,(i)\,(w+i)\, \left(1 - e^{-\psi(x)}\right)^{w+i} \prod_{j\,=1}^{s-r} \left(\frac{\gamma_{r+j}}{\frac{w+i}{\eta} + \gamma_{r+j}} \right). \tag{5.4}$$

Differentiating both sides of (5.3) with respect to x, we have

$$Z_r^{'}(x) = \frac{e^{-\psi(x)}\psi'(x)}{1 - e^{-\psi(x)}} \sum_{w=0}^{\infty} \beta_w\,(i)\,\,(w+i) \left(1 - e^{-\psi(x)}\right)^{w+i} \prod_{j=1}^{s-r} \left(\frac{\gamma_{r+j}}{\frac{w+i}{\eta} + \gamma_{r+j}} \right). \tag{5.5}$$

Using (5.4) and (5.5), we get

$$Z_r^{'}(x) = \gamma_{r+1}\,(Z_{r+1}(x) - Z_r(x))\, \frac{\eta\;\; e^{-\psi(x)}\psi'(x)}{1 - e^{-\psi(x)}}. \tag{5.6}$$

Also from (5.2), we have

$$\frac{c_{s-1}}{c_{r-1}} \int_x^{\infty} y^i\;\; \left((1-F(x))^{m+1} - (1-F(y))^{m+1}\right)^{s-r-1} (1-F(y))^{\gamma_s - 1}\, f(y)\, dy$$

$$= (s-r-1)!\,(m+1)^{s-r-1}\;\; (1-F(x))^{\gamma_{r+1}}\, Z_r(x). \tag{5.7}$$

Differentiating both sides with respect to x, we get

$$\frac{c_{s-1}}{c_{r-1}} \left[\int_x^{\infty} y^i\;\; \left((1-F(x))^{m+1} - (1-F(y))^{m+1}\right)^{s-r-2} (1-F(y))^{\gamma_s - 1}\, f(y)\, dy \right]$$

$$= \frac{(s-r-2)!\;\; (m+1)^{s-r-2}\;\, (1-F(x))^{\gamma_{r+1}-1}}{f(x)\,(1-F(x))^m} \left(\gamma_{r+1}\, f(x) Z_r(x) + (1-F(x)) Z_r^{/}(x)\right). \tag{5.8}$$

Using (5.7) in the L.H.S. of (5.8) by making appropriate changes we get

$$\frac{c_{s-1}}{c_{r-1}} \left(\frac{c_r}{c_{s-1}} (s-r-2)!\;\; (m+1)^{s-r-2}\, (1-F(x))^{\gamma_{r+2}}\, Z_{r+1}(x) \right)$$

$$= (s-r-2)!\;\; (m+1)^{s-r-2}\, (1-F(x))^{\gamma_{r+2}} \left(\gamma_{r+1}\, Z_r(x) + \frac{1-F(x)}{f(x)} Z_r^{/}(x) \right),$$

which on simplification yields

$$\gamma_{r+1}\left(Z_{r+1}(x)-Z_r(x)\right)=\frac{1-F(x)}{f(x)}Z_r^{'}(x). \tag{5.9}$$

Then, using (5.6) in (5.9) we get

$$\frac{f(x)}{1-F(x)}=\frac{\eta\ \ e^{-\psi(x)}\psi'(x)}{1-e^{-\psi(x)}}$$

This implies

$$F\left(x\right)=1-\left(1-e^{-\psi(x)}\right)^{\eta},\quad x>0\,,\,\eta>0.$$

This proves the necessary part and hence the result.

Acknowledgement: The authors wish to thank the referee for giving valuable comments which led to an improvement in the presentation of the paper

References

Balakrishnan, N. and Joshi, P.C. (1984). Product moments of order statistics from doubly truncated exponential distribution, *Naval Research Logistics Quarterly* 31, 27-31.

Balakrishnan, N. and Malik, H. J. (1986). Order statistics from linear- exponential distribution, Part I : Increasing hazard rate case, *Communication in Statistics-Theory and Methods* 15, 179-203.

Joshi, P. C. (1978). Recurrence relations between moments of order statistics from exponential and truncated exponential distributions, *Sankhya*, Ser. B 39, 362-371.

Joshi, P. C. (1982). A note on mixed moments of order statistics from exponential and truncated exponential distributions, *Journal of Statistical Planning and Inference* 6, 13-16.

Kamps, U. (1995a). *A concept of generalized order statistics*, B. G. Teubner, Stuttgart.

Kamps, U. (1995b). A concept of generalized order statistics, *Journal of Statistical Planning and Inference* 48, 1–23.

Khan, R. U., Kulshrestha, A. and Kumar, D. (2012). Exact moments of lower generalized order statistics from exponentiated Weibull distribution and its characterization, *International Journal of Statistika and Mathematika* 4, 81-89.

Mohie El-Din, M. M., Mahmoud, M. A. W., Abu-Youssef, S. E. and Sultan, K. S. (1997). Order statistics from the doubly truncated linear-exponential distribution and its characterizations, *Communication in Statistics-Simulation and Computation* 26, 281-290.

Nain, K. (2010 a). Recurrence relations for single and product moments of ordinary order statistics from pth order exponential distribution, *International Mathematical Forum* 5, 1653 – 1662.

Nain, K. (2010 b). Recurrence relations for single and product moments of kth record values from generalized Weibull distribution and a characterization, *International Mathematical Forum* 5, 1645 – 1652.

Pawlas, P. and Szynal, D. (2001). Recurrence relations for single and product moments of generalized order statistics from Pareto, Generalized Pareto and Burr distributions, *Communication in Statistics-Theory and Methods* 30, 739–746.

Ruiz, S.M. (1996). An algebraic identity leading to Wilson's Theorem, *The Mathematical Gazette* 80, 579- 582.

Saran, J. and Pandey, A. (2004). Recurrence relations for single and product moments of generalized order statistics from linear-exponential distribution, *Journal of Applied Statistical Science* 13, 323-333.

Saran, J. and Pandey, A. (2009). Recurrence relations for single and product moments of generalized order statistics from linear-exponential and Burr distributions, *Journal of Statistical Theory and Applications* 8, 383-391.

Saran, J. and Pandey, A. (2011). Recurrence relations for marginal and joint moment generating functions of dual generalized order statistics from inverse Weibull distribution, *Journal of Statistical Studies* 30, 65-72.

Saran, J. and Pushkarna, N. (1999). Moments of order statistics from doubly truncated linear- exponential distribution, *Journal of the Korean Statistical Society* 28, 279-296.

Journal of the Kerala Statistical Association
Vol.24, December 2013, 53-68 *ISSN 2249-4553*

Some Characterizations of Harris Distribution

Sandhya E*[1] and Lovely Abraham T[2]
[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) $P(s) = \frac{s}{\left(m-(m-1)s^k\right)^{\frac{1}{k}}}$, k, a positive integer and $m > 1$, is a generalization of geometric distribution on $\{1, 2, 3 \ldots\}$ 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 $\{1, 1 + k, 1 + 2k \ldots\}$. Also these probabilities coincide with that of negative binomial distribution on $\{0, 1, 2 \ldots\}$ 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 $\{a, a+k, a+2k\ldots\}$ $a>0$ integer, as a generalization of the distribution with support on $\{1, 1+k, 1+2k\ldots\}$.

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 Gnedenco 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 infinitely divisiblity.

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 $\{a, a+k, a+2k, \cdots\}$ is said to have a GPSD with parameter $\theta \geq 0$ if*

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

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

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

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

We state without proof the following theorem of random variables taking values on $S = \{a, a+k, a+2k, \ldots\}$, $a \geq 1, k \geq 1$, 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*

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

Or, equivalently

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

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

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

$Q_n(\eta)$ 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 $c_n = \frac{C^{(0)}(0)}{n!}$. $C^{(n)}(0)$ 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*

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

i.e.,

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

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

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

where $C(\eta) = \sum\limits_0^n \binom{n}{r}\eta^r = (1 + \eta)^n$

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

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*

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

putting $\lambda = \eta$,

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

where

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

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 $\{a, a + k, a + 2k, \ldots\}$ *with series function* $C(\theta) = \theta^a \left(1 - \theta^k\right)^{-\frac{a}{k}}$*, putting* $\beta = -\frac{a}{k}$ *in Sherly and Sandhya (2009)*

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

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

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

is represented by $H_a\left(\eta, k, \frac{a}{k}\right)$*. 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

$$\begin{aligned} P(X = a + rk) &= q^r p, r = 0, 1, 2, \ldots \quad \text{putting} \quad q = \eta \\ P(X = a + rk) &= \frac{\eta^r}{C(\eta)}, r = 0, 1, 2, 3 \ldots, \\ C(\eta) &= \sum \eta^r = (1 - \eta)^{-1}, \qquad (2.8) \\ \frac{d}{d\eta}P(X > a + rk) &= r(1 - \eta)^{-1}P(X = a + rk) \\ &= rC(\eta)Q_r(\eta). \end{aligned}$$

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 $P(X = x) = \dfrac{c_x\theta^x}{C(\theta)}$, $x = 0, 1, 2, \ldots$, 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

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

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

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

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 any only if the series function satisfies the following relation*

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

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*

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

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),

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

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

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

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

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

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* $u_x = \dfrac{xf(x)}{f(x-1)} = c_1x + d_1,\ x = 2, 3, 4\ldots$ *where* $c_1 = \dfrac{-a_1\eta}{b_1}$, $d_1 = \dfrac{(1+2a_1)\eta}{b_1}$.

We have the following theorem,

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

Proof. As Harris distribution belongs to GPSD we have

$$f(x) = \frac{C_x \eta^x}{C(\eta)} = \frac{\binom{a/k+x-1}{x}\eta^x}{C(\eta)}, x = 0, 1, 2\ldots$$
$$\text{and} \quad u_x = \frac{xf(x)}{f(x-1)} = \frac{xC_x\eta}{C_{x-1}}. \tag{3.8}$$

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

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

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,

$$\frac{xC_x}{C_{x-1}} = \frac{u_x}{\eta} = \frac{c_1}{\eta}x + \frac{d_1}{\eta}$$
$$= c_2(x-2) + d_2,$$

where $c_2 = \dfrac{c_1}{\eta}$, $d_2 = \dfrac{2c_1}{\eta} + \dfrac{d_1}{\eta}$.
Multiplying both sides by $(x-1)$ and rearranging

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

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

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

Also,

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

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, $\frac{\mu_{(r)}}{\mu_{(r-1)}} = c_2 r + d_2$. In Wani and Lo (1986), it is shown that it characterizes a standard generalized power series distribution by the relation

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

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*

$$V_r = \frac{\mu_{(r)}}{\mu_{(r-1)}} = c_2 r + d_2 \tag{3.11}$$

where

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

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*

$$V_r = \frac{\mu_{(r)}}{\mu_{(r-1)}} = r\left(\frac{\eta}{1-\eta}\right) + \left(\frac{a-k}{k}\right)\frac{\eta}{1-\eta} = c_2 r + d_2 \tag{3.12}$$

where

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

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

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

For Harris distribution, we have,

$$C^{(r)}(\eta) = \frac{a}{k}\,\frac{a+k}{k}\cdots\frac{a+(r-1)k}{k}(1-\eta)^{-\frac{a}{k}-r}$$
$$\mu_{(r)} = \frac{a}{k}\,\frac{a+k}{k}\cdots\frac{a+(r-1)k}{k}\,\frac{\eta^r}{(1-\eta)^r}. \tag{3.14}$$

From (3.1) and (3.2),

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

Substituting in (3.14) we get

$$\frac{\mu_{(r)}}{\mu_{(r-1)}} = \frac{\eta C^{(r)}(\eta)}{C^{(r-1)}(\eta)} \tag{3.15}$$
$$= \frac{-a_1\eta r}{a_1\eta + b_1} + \frac{(1+2a_1)\eta}{a_1\eta+b_1} = c_2 r + d_2.$$

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

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

Put $r = 2$ in V_r

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

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),

$$u_r = \frac{rf(r)}{f(r-1)} = \eta(n-r+1) = -\eta r + (n+1)\eta = c_1 r + d_1$$
$$\mu_{(r)} = n(n-1)\ldots(n-r+1)\left(\frac{\eta}{1+\eta}\right)^r$$

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

(ii) Poisson with gaps

From (2.6),

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

(iii) Geometric with gaps

From (2.8),

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

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))

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

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

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

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

Theorem 4.1. *In the class of GPSD defined on set of nonnegative integers, the p.m.f is of a Harris distribution* $H_a\left(\theta, k, \frac{1}{k}\right)$ *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* $X_1 = \{a, a+k, a+2k \cdots\}$*,* $a, k > 0$ *intergers 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* $[a, u)$*,*

$$\int_a^u f(\theta) d\theta = F(u).$$

$$\textit{If} \quad e^{F(u)} = t \sum_{x \in S} \alpha_x u^x$$

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

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 $0 < \theta < 1$

$$\begin{aligned} F(u) &= \int_0^u \frac{a}{\theta} + \frac{\theta^{k-1}}{1-\theta^k} d\theta = \ln u^a (1-u)^{-\frac{1}{k}} + t_1 \\ e^{F(u)} &= t u^a (1-u)^{-\frac{1}{k}} = t \sum \binom{1/k + x - 1}{x} u^{a+x} = t \sum \alpha_x u^{a+x} \\ \text{with} \quad & t = C(0) \\ \text{and} \quad & P(X = a + xk) = C(\theta)^{-1} c_x \theta^a (\theta^k)^x \, x \in S, \\ \text{where} \quad & c_x = t\alpha_x. \end{aligned}$$

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

$$\pi_i = \frac{ip_i}{p_0} - \sum_{j=1}^{i-1} \pi_{i-j} \frac{p_i}{p_0} \geq 0 \text{ for } i = 1, 2, \ldots \tag{5.1}$$

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

$$\pi_r = \frac{1}{k}(\theta^k)^r, \; r = 1, 2, \ldots \tag{5.2}$$

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

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

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

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

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

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*, 2^{nd} 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.

Journal of the Kerala Statistical Association
Vol.24, December 2013, 69-77 *ISSN 2249-4553*

Conditional Distributions and Reliability Analysis Associated with Marshall-Olkin Bivariate Exponential Distribution

B. Chandrasekar and Tunny Sebastian *
Department of Statistics
Loyola College, Chennai - 600034

ABSTRACT

By considering the bivariate exponential distribution due to Marshall and Olkin (1967), we derive the conditional distributions and regression equations. It is observed that the regression equations are non-linear. An alternative derivation for the moment generating function using three independent exponential distributions is also given. Assuming that the joint distribution of the component failure times is bivariate exponential, we obtain the reliability measures for two-unit standby, parallel and series systems.

Key words and Phrases: *Conditional distribution, Moment generating function, Regression, Reliability measures, Weighted average.*

1 Introduction

There is an extensive literature on the construction of bivariate exponential models. Several authors proposed different bivariate exponential models.

*email:tunnysebastian2006@yahoo.com

Marshall and Olkin (1967) proposed a multivariate exponential distribution having dependent exponential marginals which characterizes a fatal shock model. The distribution does not possess the probability density function (pdf) with respect to the Lebesgue measure in R^2. Bemis *et al.* (1972) proposed the pdf with respect to an appropriate measure and derived some properties. Paul Rajamanickam and Chandrasekar (1998) derived the asymptotic confidence interval for the steady-state availability of a one unit system with dependent structure having Marshall-Olkin bivariate exponential (MOBVE) joint distribution for the failure and repair times. Exact distributional properties of MOBVE are rare in the literature.

This paper studies the properties of MOBVE distribution through the pdf approach. Section 2 introduces the distribution and the pdf due to Bemis *et al.* (1972). The conditional distributions and the regression equations are also obtained and it is observed that the regression equations are not linear. Section 3 provides an alternative derivation for the moment generating function (mgf) using three independent exponential distributions. Section 4 derives the performance measures of two-unit standby, parallel and series systems assuming that the joint distribution of the component failure times is MOBVE. In the case of standby and parallel systems, it is observed that the system failure times are the weighted averages of exponential distributions.

2 Conditional distributions and Regression

In this section we observe that the regression equations are not linear.

The survival function of the MOBVE distribution is given by

$$\bar{F}(x,y) = \mathsf{e}^{-\lambda_1 x - \lambda_2 y - \lambda_{12}(x \vee y)}; \qquad x, y > 0; \lambda_1, \lambda_2 > 0, \lambda_{12} \geq 0.$$

Here $a \vee b = max\{a, b\}$.

Bemis *et al.* (1972) proposed a pdf as follows:

$$f(x,y) = \begin{cases} \lambda_1(\lambda_2+\lambda_{12})\mathsf{e}^{-\lambda_1 x-(\lambda_2+\lambda_{12})y}, & x < y \\ \lambda_2(\lambda_1+\lambda_{12})\mathsf{e}^{-(\lambda_1+\lambda_{12})x-\lambda_2 y}, & x > y \\ \lambda_{12}\mathsf{e}^{-(\lambda_1+\lambda_2+\lambda_{12})x}, & x = y. \end{cases} \tag{2.1}$$

***Remark* 2.1.** The pdf can be viewed as a mixture of three pdfs with the mixing proportions $\frac{\lambda_1}{\lambda}$, $\frac{\lambda_2}{\lambda}$ and $\frac{\lambda_{12}}{\lambda}$, where $\lambda = \lambda_1 + \lambda_2 + \lambda_{12}$. In fact,

$$f(x, y) = \frac{\lambda_1}{\lambda} f_1(x, y) + \frac{\lambda_2}{\lambda} f_2(x, y) + \frac{\lambda_{12}}{\lambda} f_3(x, y),$$

where

$$f_1(x, y) = \begin{cases} \lambda(\lambda_2 + \lambda_{12})\mathrm{e}^{-\lambda_1 x - (\lambda_2 + \lambda_{12})y}, & 0 < x < y < \infty \\ 0, & \text{otherwise} \end{cases},$$

$$f_2(x, y) = \begin{cases} \lambda(\lambda_1 + \lambda_{12})\mathrm{e}^{-(\lambda_1 + \lambda_{12})x - \lambda_2 y}, & 0 < y < x < \infty \\ 0, & \text{otherwise} \end{cases}$$

and

$$f_3(x, y) = \begin{cases} \lambda \mathrm{e}^{-\lambda x}, & 0 < x = y < \infty \\ 0, & \text{otherwise.} \end{cases}$$

***Theorem* 2.1.** The conditional pdf of X given $Y = y$ is

$$\begin{aligned} f(x \mid y) &= \lambda_1 \mathrm{e}^{-\lambda_1 x} I_{(0,y)}(x) + \frac{\lambda_{12}}{(\lambda_2 + \lambda_{12})} \mathrm{e}^{-\lambda_1 x} I_{\{y\}}(x) \\ &\quad + \frac{\lambda_2(\lambda_1 + \lambda_{12})}{(\lambda_2 + \lambda_{12})} \mathrm{e}^{-(\lambda_1 + \lambda_{12})x + \lambda_{12} y} I_{(y,\infty)}(x), x > 0. \end{aligned}$$

Proof. The joint pdf of (X, Y) is given in (2.1) and the pdf of Y is

$$f_2(y) = (\lambda_2 + \lambda_{12})\mathrm{e}^{-(\lambda_2 + \lambda_{12})y}, y > 0.$$

The proof follows from the fact that

$$f(x \mid y) = \frac{f(x, y)}{f_2(y)}, x > 0.$$

***Corollary* 2.1.**

$$\text{(i) } E(X \mid Y = y) = \frac{1}{\lambda_1} - \frac{\lambda \lambda_{12}}{\lambda_1(\lambda_1 + \lambda_{12})(\lambda_2 + \lambda_{12})} \mathrm{e}^{-\lambda_1 y}$$

and

$$\text{(ii) } E(Y \mid X = x) = \frac{1}{\lambda_2} - \frac{\lambda \lambda_{12}}{\lambda_2(\lambda_1 + \lambda_{12})(\lambda_2 + \lambda_{12})} \mathrm{e}^{-\lambda_2 x}.$$

Proof: From Theorem 2.1,

$$\begin{aligned}
E(X \mid Y = y) &= \int_0^\infty x f(x \mid y) dx \\
&= \int_0^y x\lambda_1 \mathrm{e}^{-\lambda_1 x} dx \\
&\quad + \int_y^\infty x \frac{\lambda_2(\lambda_1+\lambda_{12})}{(\lambda_2+\lambda_{12})} \mathrm{e}^{-(\lambda_1+\lambda_{12})x+\lambda_{12}y} dx \\
&\quad + y\frac{\lambda_{12}}{(\lambda_2+\lambda_{12})}\mathrm{e}^{-\lambda_1 y} \\
&= -y\mathrm{e}^{-\lambda_1 y} + \frac{1}{\lambda_1}(1-\mathrm{e}^{-\lambda_1 y}) \\
&\quad + \frac{\lambda_2}{(\lambda_2+\lambda_{12})}\mathrm{e}^{\lambda_{12} y}\left\{ y\mathrm{e}^{-(\lambda_1+\lambda_{12})y} + \frac{1}{(\lambda_1+\lambda_{12})}\mathrm{e}^{-(\lambda_1+\lambda_{12})y}\right\} \\
&\quad + y\frac{\lambda_{12}}{(\lambda_2+\lambda_{12})}\mathrm{e}^{-\lambda_1 y} \\
&= \frac{1}{\lambda_1} - \frac{\lambda\lambda_{12}}{\lambda_1(\lambda_1+\lambda_{12})(\lambda_2+\lambda_{12})}\mathrm{e}^{-\lambda_1 y}.
\end{aligned}$$

This proves (i). Similar arguments lead to (ii). ***Remark* 2.2.** The regression equations are

$$X = \frac{1}{\lambda_1} - \frac{\lambda\lambda_{12}}{\lambda_1(\lambda_1+\lambda_{12})(\lambda_2+\lambda_{12})}\mathrm{e}^{-\lambda_1 Y}$$

and

$$Y = \frac{1}{\lambda_2} - \frac{\lambda\lambda_{12}}{\lambda_2(\lambda_1+\lambda_{12})(\lambda_2+\lambda_{12})}\mathrm{e}^{-\lambda_2 X}.$$

The regression equations are not linear.

***Remark* 2.3.** (i) The mgf of the conditional distribution of X given $Y = y$ is

$$M(t \mid y) = \frac{\lambda_1}{(\lambda_1 - t)} + \frac{\lambda_{12} t(t-\lambda)\mathrm{e}^{-(\lambda_1 - t)y}}{(\lambda_2+\lambda_{12})(\lambda_1 - t)(\lambda_1+\lambda_{12}-t)}, t < \lambda_1.$$

(ii) The mgf of the conditional distribution of Y given $X = x$ is

$$M(t \mid x) = \frac{\lambda_2}{(\lambda_2 - t)} + \frac{\lambda_{12} t(t-\lambda)\mathrm{e}^{-(\lambda_2 - t)x}}{(\lambda_1+\lambda_{12})(\lambda_2 - t)(\lambda_2+\lambda_{12}-t)}, t < \lambda_2.$$

3 An alternative derivation for the moment generating function

Marshall and Olkin (1967) derived the mgf using pdf approach. Here we derive the mgf using three independent exponential distributions.

***Theorem* 3.1.** The mgf of (X, Y) is

$$M(t_1, t_2) = \frac{(\lambda - t_1 - t_2)(\lambda_1 + \lambda_{12})(\lambda_2 + \lambda_{12}) + \lambda_{12} t_1 t_2}{(\lambda - t_1 - t_2)(\lambda_1 + \lambda_{12} - t_1)(\lambda_2 + \lambda_{12} - t_2)}, \tag{3.1}$$

$$t_1 < \lambda_1 + \lambda_{12}, t_2 < \lambda_2 + \lambda_{12}, t_1 + t_2 < \lambda.$$

Proof. Let U_1, U_2 and U_3 be three independent exponential random variables with the parameters λ_1, λ_2 and λ_{12} respectively. Define $X = U_1 \wedge U_3$ and $Y = U_2 \wedge U_3$, where $a \wedge b = min\{a, b\}$. Then (X, Y) follows MOBVE distribution (See Barlow and Proschan(1975)).

Thus

$$\begin{aligned}
M(t_1, t_2) &= E\left\{\mathrm{e}^{t_1(U_1 \wedge U_3) + t_2(U_2 \wedge U_3)}\right\} \\
&= \int_0^\infty \int_0^\infty \int_0^\infty \mathrm{e}^{t_1(u_1 \wedge u_3) + t_2(u_2 \wedge u_3)} \lambda_1 \lambda_2 \lambda_{12} \mathrm{e}^{-\lambda_1 u_1 - \lambda_2 u_2 - \lambda_{12} u_3} du_1 du_2 du_3 \\
&= \int_0^\infty \int_{u_1}^\infty \int_{u_2}^\infty \mathrm{e}^{t_1 u_1 + t_2 u_2} \lambda_1 \lambda_2 \lambda_{12} \mathrm{e}^{-\lambda_1 u_1 - \lambda_2 u_2 - \lambda_{12} u_3} du_3 du_2 du_1 \\
&\quad + \int_0^\infty \int_{u_2}^\infty \int_{u_1}^\infty \mathrm{e}^{t_1 u_1 + t_2 u_2} \lambda_1 \lambda_2 \lambda_{12} \mathrm{e}^{-\lambda_1 u_1 - \lambda_2 u_2 - \lambda_{12} u_3} du_3 du_1 du_2 \\
&\quad + \int_0^\infty \int_{u_1}^\infty \int_{u_3}^\infty \mathrm{e}^{t_1 u_1 + t_2 u_3} \lambda_1 \lambda_2 \lambda_{12} \mathrm{e}^{-\lambda_1 u_1 - \lambda_2 u_2 - \lambda_{12} u_3} du_2 du_3 du_1 \\
&\quad + \int_0^\infty \int_{u_2}^\infty \int_{u_3}^\infty \mathrm{e}^{t_1 u_3 + t_2 u_2} \lambda_1 \lambda_2 \lambda_{12} \mathrm{e}^{-\lambda_1 u_1 - \lambda_2 u_2 - \lambda_{12} u_3} du_1 du_3 du_2 \\
&\quad + \int_0^\infty \int_{u_3}^\infty \int_{u_1}^\infty \mathrm{e}^{t_1 u_3 + t_2 u_3} \lambda_1 \lambda_2 \lambda_{12} \mathrm{e}^{-\lambda_1 u_1 - \lambda_2 u_2 - \lambda_{12} u_3} du_2 du_1 du_3 \\
&\quad + \int_0^\infty \int_{u_3}^\infty \int_{u_2}^\infty \mathrm{e}^{t_1 u_3 + t_2 u_3} \lambda_1 \lambda_2 \lambda_{12} \mathrm{e}^{-\lambda_1 u_1 - \lambda_2 u_2 - \lambda_{12} u_3} du_1 du_2 du_3 \\
&= I_1 + I_2 + I_3 + I_4 + I_5 + I_6, \text{ say.}
\end{aligned}$$

Then

$$\begin{aligned}
I_1 &= \frac{\lambda_1\lambda_2}{(\lambda_2+\lambda_{12}-t_2)(\lambda-t_1-t_2)},\\
I_2 &= \frac{\lambda_1\lambda_2}{(\lambda_1+\lambda_{12}-t_1)(\lambda-t_1-t_2)},\\
I_3 &= \frac{\lambda_1\lambda_{12}}{(\lambda_2+\lambda_{12}-t_2)(\lambda-t_1-t_2)},\\
I_4 &= \frac{\lambda_2\lambda_{12}}{(\lambda_1+\lambda_{12}-t_1)(\lambda-t_1-t_2)},\\
I_5 &= \frac{\lambda_1\lambda_{12}}{(\lambda_2+\lambda_2)(\lambda-t_1-t_2)}, \text{ and}\\
I_6 &= \frac{\lambda_2\lambda_{12}}{(\lambda_1+\lambda_{12})(\lambda-t_1-t_2)}.
\end{aligned}$$

Hence

$$\begin{aligned}
M(t_1,t_2) &= (I_1+I_3)+(I_2+I_4)+(I_5+I_6)\\
&= \frac{\lambda_1(\lambda_2+\lambda_{12})}{(\lambda_2+\lambda_{12}-t_2)(\lambda-t_1-t_2)}+\frac{\lambda_2(\lambda_1+\lambda_{12})}{(\lambda_1+\lambda_{12}-t_1)(\lambda-t_1-t_2)}+\frac{\lambda_{12}}{(\lambda-t_1-t_2)}\\
&= \frac{(\lambda-t_1-t_2)(\lambda_1+\lambda_{12})(\lambda_2+\lambda_{12})+\lambda_{12}t_1t_2}{(\lambda-t_1-t_2)(\lambda_1+\lambda_{12}-t_1)(\lambda_2+\lambda_{12}-t_2)},\\
&\qquad t_1<\lambda_1+\lambda_{12}, t_2<\lambda_2+\lambda_{12}, t_1+t_2<\lambda.
\end{aligned}$$

***Remark* 3.1.** The above expression coincides with the one given in Barlow and Proschan (1975).

4 Reliability measures

Consider a two-unit system with the failure time distributions following a MOBVE. By considering standby, parallel and series systems, we derive the system failure times and the performance measures of these systems.

4.1 Standby system

If $\lambda_{12}=0$, the failure times are independent and the results are known in the literature, (see Rau (1970)). So we assume that $\lambda_{12}>0$ throughout the section. Four cases arise according as (i) $\lambda_2-\lambda_1-\lambda_{12}\neq 0$, $\lambda_1-\lambda_2-\lambda_{12}\neq 0$ (ii) $\lambda_2-\lambda_1-\lambda_{12}\neq 0$,

$\lambda_1 - \lambda_2 - \lambda_{12} = 0$ (iii) $\lambda_2 - \lambda_1 - \lambda_{12} = 0$, $\lambda_1 - \lambda_2 - \lambda_{12} \neq 0$ (iv) $\lambda_1 = \lambda_2$. We now discuss case (i) in detail.

The system failure time is $T = X + Y$.

The mgf of T is

$$\begin{aligned} M^*(u) &= M(u,u) \\ &= \frac{(\lambda - 2u)(\lambda_1 + \lambda_{12})(\lambda_2 + \lambda_{12}) + \lambda_{12}u^2}{(\lambda - 2u)(\lambda_1 + \lambda_{12} - u)(\lambda_2 + \lambda_{12} - u)}, \text{ in view of (3.1).} \end{aligned}$$

Resolving into partial fractions we get

$$\begin{aligned} M^*(u) &= w_1 \left(1 - \frac{1}{\lambda_1 + \lambda_{12}} u\right)^{-1} + w_2 \left(1 - \frac{1}{\lambda_2 + \lambda_{12}} u\right)^{-1} + w_3 \left(1 - \frac{2}{\lambda} u\right)^{-1}, \\ & \qquad u < \lambda_1 + \lambda_{12}, u < \lambda_2 + \lambda_{12}, u < \tfrac{\lambda}{2}, \end{aligned}$$

where

$$w_1 = \frac{\lambda_2}{\lambda_2 - \lambda_1 - \lambda_{12}}, w_2 = \frac{\lambda_1}{\lambda_1 - \lambda_2 - \lambda_{12}} \text{ and } w_3 = \frac{\lambda\lambda_{12}}{(\lambda_1 - \lambda_2 + \lambda_{12})(\lambda_2 - \lambda_1 + \lambda_{12})}.$$

Thus the failure time distribution is a weighted average of three exponential distributions. Hence the reliability function and the mean time before failure (MTBF) are

$$R(t) = w_1 \mathrm{e}^{-(\lambda_1 + \lambda_{12})t} + w_2 \mathrm{e}^{-(\lambda_2 + \lambda_{12})t} + w_3 \mathrm{e}^{-\frac{\lambda}{2}t}, t > 0$$

and

$$\text{MTBF} = \frac{1}{\lambda_1 + \lambda_{12}} + \frac{1}{\lambda_2 + \lambda_{12}}.$$

***Remark* 4.1.** The mgf of the system failure time in the other three cases, in that order, are given by

$$\begin{aligned} M^*(u) &= \frac{\lambda_1}{2\lambda_{12}} \left(1 - \frac{u}{\lambda_1}\right)^{-1} + \frac{1}{2} \left(1 - \frac{u}{\lambda_1}\right)^{-2} + \frac{\lambda_{12} - \lambda_1}{2\lambda_{12}} \left(1 - \frac{u}{\lambda_1 + \lambda_{12}}\right)^{-1}, u < \lambda_1, \\ M^*(u) &= \frac{\lambda_2}{2\lambda_{12}} \left(1 - \frac{u}{\lambda_2}\right)^{-1} + \frac{1}{2} \left(1 - \frac{u}{\lambda_2}\right)^{-2} + \frac{\lambda_{12} - \lambda_2}{2\lambda_{12}} \left(1 - \frac{u}{\lambda_2 + \lambda_{12}}\right)^{-1}, u < \lambda_2, \end{aligned}$$

and

$$M^*(u) = \frac{2\lambda_1 + \lambda_{12}}{\lambda_{12}} \left(1 - \frac{2u}{2\lambda_1 + \lambda_{12}}\right)^{-1} - \frac{2\lambda_1}{\lambda_{12}} \left(1 - \frac{u}{\lambda_1 + \lambda_{12}}\right)^{-1}, u < \lambda_1 + \frac{\lambda_{12}}{2}.$$

4.2 Parallel System

In this case the failure time of the system is $T = X \vee Y$. The df of T is

$$
\begin{aligned}
P(T \leq t) &= P(X \leq t, Y \leq t) \\
&= \int_0^t \int_0^x \lambda_2(\lambda_1+\lambda_{12}) \mathrm{e}^{-\lambda_2 y-(\lambda_1+\lambda_{12})x} dydx \\
&\quad + \int_0^t \int_x^t \lambda_1(\lambda_2+\lambda_{12}) \mathrm{e}^{-\lambda_1 x-(\lambda_2+\lambda_{12})y} dydx \\
&\quad + \int_0^t \lambda_{12} \mathrm{e}^{-\lambda x} dx \\
&= \int_0^t (\lambda_1+\lambda_{12}) \mathrm{e}^{-(\lambda_1+\lambda_{12})x} \left(1-\mathrm{e}^{-\lambda_2 x}\right) dx \\
&\quad + \int_0^t \lambda_1 \mathrm{e}^{-\lambda_1 x} \left\{ \mathrm{e}^{-(\lambda_2+\lambda_{12})x} - \mathrm{e}^{-(\lambda_2+\lambda_{12})t} \right\} dx \\
&\quad + \frac{\lambda_{12}}{\lambda}\left(1-\mathrm{e}^{-\lambda t}\right) \\
&= 1-\mathrm{e}^{-(\lambda_1+\lambda_{12})t} - \frac{(\lambda_1+\lambda_{12})}{\lambda}\left(1-\mathrm{e}^{-\lambda t}\right) \\
&\quad + \frac{\lambda_1}{\lambda}\left(1-\mathrm{e}^{-\lambda t}\right) - \mathrm{e}^{-(\lambda_2+\lambda_{12})t}\left(1-\mathrm{e}^{-\lambda_1 t}\right) \\
&\quad + \frac{\lambda_{12}}{\lambda}\left(1-\mathrm{e}^{-\lambda t}\right) \\
&= 1-\mathrm{e}^{-(\lambda_1+\lambda_{12})t} - \mathrm{e}^{-(\lambda_2+\lambda_{12})t} + \mathrm{e}^{-\lambda t}.
\end{aligned}
$$

Thus the distribution of T is a weighted average of three exponential distributions. Hence the reliability function and the MTBF are

$$R(t) = \mathrm{e}^{-(\lambda_1+\lambda_{12})t} + \mathrm{e}^{-(\lambda_2+\lambda_{12})t} - \mathrm{e}^{-\lambda t}, t > 0$$

and

$$\text{MTBF} = \frac{1}{\lambda_1+\lambda_{12}} + \frac{1}{\lambda_2+\lambda_{12}} - \frac{1}{\lambda}.$$

4.3 Series System

The failure time of the system is $T = X \wedge Y$. It is well known that T is exponential with mean $\frac{1}{\lambda}$ and so

$$R(t) = \mathrm{e}^{-\lambda t}, t > 0$$

and

$$\text{MTBF } = \frac{1}{\lambda}.$$

One can also use the pdf approach to find the distribution of T.

Acknowledgement: The authors are grateful to a referee for helpful suggestions which resulted in an improvement of the presentation.

References

Barlow, R. E. and Proschan, F. (1975). Statistical Theory of Reliability and Life Testing: Probability Models. Holt, Rinehart and Winston Inc. New York.

Bemis, B. M., Bain, L. J. and Higgins, J. J. (1972). Estimation and Hypothesis Testing for the Parameters of a Bivariate Exponential Distribution Journal of the American Statistical Association 67, 927-929.

Marshall, A. W. and Olkin, I. (1967). A multivariate exponential distribution, Journal of the American Statistical Association 62, 30-44.

Paul Rajamanickam, S. and Chandrasekar, B. (1998). Confidence limits for steady state availability of a system with dependant structure for failure and repair times, Journal of Applied Statistical Science 8, 17-27.

Rau, J. G. (1970). Optimization and Probability in Systems Engineering, Van Nostrand Reinhold, New York.

The Journal of the Kerala Statistical Association is a publication of the Kerala Statistical Association and its subscription rates are given below.
India: Rs. 300/- Others Countries: U. S. $ 30/-