Electronic Journal of Applied Statistical Analysis
EJASA, Electron. J. App. Stat. Anal.
http://siba-ese.unisalento.it/index.php/ejasa/index
e-ISSN: 2070-5948
DOI: 10.1285/i20705948v14n2p373

The extended Farlie-Gumbel-Morgenstern bivariate Lindley distribution: Concomitants of order statistics and estimation

By Irashad et al.

Published: 20 November 2021

Electronic Journal of Applied Statistical Analysis
Vol. 14, Issue 02, November 2021, 373-388
DOI: 10.1285/i20705948v14n2p373

The extended Farlie-Gumbel-Morgenstern bivariate Lindley distribution: Concomitants of order statistics and estimation

M.R. Irshad*[a], R. Maya[b], Amer Ibrahim Al-Omari[c], S.P. Arun[d], and Ghadah A. Alomani[e]

[a] *Department of Statistics,, CUSAT, Kerala.*
[a] *Department of Statistics,, Govt. College for Women,, Trivandrum-695 014.*
[c] *Department of Mathematics, Faculty of Science, Al al-Bayt University, Mafraq, Jordan*
[d] *Kerala University Library, Research Centre, University of Kerala,, Trivandrum-695 034*
[e] *Department of Mathematical Sciences, College of Science,, Princess Nourah bint Abdulrahman University,, P.O. Box 84428, Riyadh 11671, Saudia Arabia*

Published: 20 November 2021

The ranked set sampling (*RSS*) is a new sampling method alternative to the simple random sampling (SRS). In this work, we develop the theory of concomitant of order statistics (COS) ascending from the extended Farlie-Gumbel-Morgenstern bivariate Lindley distribution (EFGMBLD). Also, we have discussed the problem of estimating the parameters related with the distribution of the variable interest, Y, based on the *RSS* defined by ordering the marginal observations on an auxiliary variable X, provided that (X, Y) follows an EFGMBLD. When the association parameters corresponding to Y are known, we have derived two estimators, viz., an unbiased estimator based on Stoke's *RSS* and the best linear unbiased estimator (BLUE) based on the Stoke's *RSS*. The BLUE and the unbiased estimators are also compared based on simulation study.

keywords: Concomitants of order statistics; Ranked set sampling; Extended Farlie-Gumbel-Morgenstern bivariate Lindley distribution..

*Corresponding author: alomari_amer@yahoo.com

ISSN: 2070-5948
http://siba-ese.unisalento.it/index.php/ejasa/index

1 Introduction

The study of COS open is the key for analysis of data withdrawn from bivariate model in a theoretical as well as applied perspective. Suppose $(X_1, Y_1), (X_2, Y_2), \cdots$ is a sequence of random variables (rvs) which are independent and identically distributed (iid) with a joint cumulative distribution function (cdf) $F(x, y), x, y \in R$. Also, let $F(x)$ and $F(y)$ are the cdf's (marginal) X and Y, respectively. If the sample values on the marginal random variable X are ordered as $X_{1:m}, X_{2:m}, \cdots, X_{m:m}$, then the related Y rv in an ordered pair with X equals to $X_{j:m}$ symbolised by $Y_{[j:m]}$ and is known as the concomitant of the j^{th} order statistic $Y_{[j:m]}$. Based on a random sample of size m coming from a bivariate distribution, the COS are $Y_{[1:m]}, Y_{[2:m]}, \cdots, Y_{[m:m]}$. Most of the theoretical developments in the COS till 2003 are available in David and Nagaraja (2003).

The main applications of COS varies over biological selection problems, engineering, and development of structural methods and so on. The COS ascending from the Farlie-Gumbel-Morgenstern (FGM) family are elucidated by Scaria and Nair (1999). Beg and Ahsanullah (2008) investigated concomitants for generalized order statistics selected from FGM distributions. Recently, Maya et al. (2021) thoroughly studied certain inferential aspects of FGM bivariate Bilal distribution using COS.

One of the main applications of COS is in the *RSS*. The *RSS* technique was first developed by McIntyre (1952) in order to estimate the population mean of pasture yields. McIntyre's idea of ranking is possible whenever it can be done easily by some inexpensive method.

For recent developments in *RSS*, one can refer to some references as Al-Omari and Al-manjahie (2021), Al-Omari (2021) for maximum likelihood estimation in location-scale families using varied L RSS, Hassan et al. (2021) for stress-strength reliability for the generalized inverted exponential distribution using median RSS, Al-Omari and Abdallah (2021), Jemain et al. (2007), Benchiha and Al-Omari (2021) for generalized quasi Lindley distribution, Al-Omari and Haq (2019), Mahdizadeh and Zamanzade (2021), Mahdizadeh and Zamanzade (2021), Terpstra and Miller (2006), Haq et al. (2016a), Koyuncu and Al-Omari (2021) for generalized robust-regression-type estimators under different RSS methods, and Haq et al. (2016b). In some practical problems, the study variable say Y, is intricate or cost to measure, while an auxiliary variable X related with Y can be easily measure or ordered exactly. In this case, Stokes (1977) developed another scheme of *RSS*, which is as follows: randomly select m independent bivariate sets, each of size m. From the first set of m units, select the variate Y associated with minimum ordered X for actual measure. From the m units in the second set, the Y variate associated with the second minimum X is selected. This process is continued until the Y associated with the largest X from the last set is measured. The measurements on the Y variate of the new set of m units chosen by the above method gives a *RSS* as suggested by Stokes (1977). Let $X_{(j:m)j}$ be the measured observation on the variable X from the chosen unit from the j^{th} set, then $Y_{[j:m]j}$ denotes to the via measurement based on Y, the study variable on this unit and hence $Y_{[j:m]j}$, for $j = 1, 2, \cdots, m$ are the *RSS* units. Here, $Y_{[j:m]j}$ is the concomitant of the j^{th} order statistic OS obtained via the j^{th}

sample. Takahasi and Wakimoto (1968) offered that for the degenerate distribution the efficiency (Eff) of the *RSS* with respect to the SRS is

$$1 \leq Eff\left(\bar{X}_{RSS}, \bar{X}_{SRS}\right) = \frac{\operatorname{Var}\left(\bar{X}_{SRS}\right)}{\operatorname{Var}\left(\bar{X}_{RSS}\right)} \leq \frac{m+1}{2},$$

where

$$\bar{X}_{RSS} = \frac{1}{m}\sum_{i=1}^{m} X_{i(i:m)}$$

and

$$\operatorname{Var}\left(\bar{X}_{RSS}\right) = \frac{\sigma^2}{m} - \frac{1}{m^2}\sum_{i=1}^{m}\left(\mu_{(i:m)} - \mu\right)^2.$$

The ith order statistics, has the cdf and pdf, respectively, defined by

$$F_{(i:m)}(x) = \begin{pmatrix} m \\ i \end{pmatrix} \int_0^{F(x)} \tau^{i-1}(1-\tau)^{m-i} d\tau$$

and

$$f_{(i:m)}(x) = \begin{pmatrix} m \\ i \end{pmatrix} [1-F(x)]^{m-i} f(x)[F(x)]^{i-1}.$$

Irshad et al. (2019) discussed the problem of estimating the parameter of FGM bivariate Lindley distribution (FGMBLD) by the *RSS*.

Morgenstern (1956) suggested a family of new bivariate distribution functions $F(x, y)$, its representation is

$$F(x, y) = \{1 + \theta[1 - F(x)][1 - F(y)]\}F(x)F(y), \tag{1}$$

where the dependence parameter θ is constrained to be in the closed interval $[-1, 1]$. The family of bivariate distributions with distribution function $F(x, y)$ as given in (1) is also called in the literature as FGM family of bivariate distributions.

Johnson and Kotz (1977) introduced another generalization of family, so-called extended FGM (EFGM) family, with dependence parameters ρ and θ. The EFGM distribution has a bivariate cdf is

$$\begin{gathered} H(x, y) = \{1 + \theta\overline{F}(y)\overline{F}(x) + \rho F(x)F(y)\overline{F}(y)\overline{F}(x)\}F(x)F(y), \\ |\theta| \leq 1, -\theta - 1 \leq \rho \leq \tfrac{1}{2}[3 - \theta + \left(9 - 6\theta - 3\theta^2\right)^{1/2}], \end{gathered} \tag{2}$$

where $\overline{F} = 1 - F$.

The corresponding bivariate pdf is given by

$$\begin{aligned} h(x, y) \quad &= \quad f(x)f(y)\left\{1 + \theta[1 - 2F(x)][1 - 2F(y)] + \right. \\ &\left. + \rho[2 - 3F(x)][2 - 3F(y)]F(x)F(y)\right\}, \end{aligned} \tag{3}$$

where $f(x)$ and $f(y)$ are the pdf's, respectively of X and Y. Let $\rho = 0$ in (2), to get the joint cdf of the model given in (1). The maximum value of the correlation coefficient between X and Y having the cdf given in (2) is 0.5072, that is greater than the maximum value corresponding to the distribution related to the FGM family which is equal to 0.3333.

Hence, the EFGM distribution is commonly used with maximal values of the correlation coefficient between the component rvs related to higher dimension parameter space. Due to the flexibility of EFGM distribution compared to the FGM distribution, in this paper we consider an important member of EFGM family, which is known as EFGM bivariate Lindley distribution (EFGMBLD).

The distribution theory of COS obtained from the EFGMBLD is modified and discussed in Section 2.1. In Section 3, we provide an unbiased estimator of the parameter of the study variate contained in the EFGMBLD via Stoke's *RSS*. The BLUE of this parameter based on the observations of Stoke's *RSS* are derived and given in Section 4. Also, the efficiency values of the BLUE with respect to the unbiased estimator are presented. Section 5 is devoted for concluding remarks.

2 EFGM bivariate Lindley distribution

An EFGM bivariate distribution with univariate Lindley distribution as marginal are known as the EFGMBLD. The EFGMBLD has the joint pdf $h(x, y)$ given in (4) by substituting the pdf's $f(x) = \frac{1}{2\sigma_1}\left(1 + \frac{x}{\sigma_1}\right)e^{-\frac{x}{\sigma_1}}$, $f(y) = \frac{1}{2\sigma_2}\left(1 + \frac{y}{\sigma_2}\right)e^{-\frac{y}{\sigma_2}}$ and cdf's $F(x) = 1 - \left(1 + \frac{x}{2\sigma_1}\right)e^{-\frac{x}{\sigma_1}}$, $F(y) = 1 - \left(1 + \frac{y}{2\sigma_2}\right)e^{-\frac{y}{\sigma_2}}$ of two univariate Lindley distributions in (3) as

$$
\begin{aligned}
h(x,y) = {} & \frac{0.5}{\sigma_1}\left(1+\frac{x}{\sigma_1}\right)e^{-\frac{x}{\sigma_1}}\frac{0.5}{\sigma_2}\left(1+\frac{y}{\sigma_2}\right)e^{-\frac{y}{\sigma_2}}\{1+ \\
& \theta\left[2\left(1+\frac{0.5x}{\sigma_1}\right)e^{-\frac{x}{\sigma_1}}-1\right]\left[2\left(1+\frac{0.5y}{\sigma_2}\right)e^{-\frac{y}{\sigma_2}}-1\right]+ \\
& \rho\left[3\left(1+\frac{x}{2\sigma_1}\right)e^{-\frac{x}{\sigma_1}}-1\right]\left[3\left(1+\frac{0.5y}{\sigma_2}\right)e^{-\frac{y}{\sigma_2}}-1\right] \\
& \times\left[1-\left(1+\frac{0.5x}{\sigma_1}\right)e^{-\frac{x}{\sigma_1}}\right]\left[1-\left(1+\frac{y}{2\sigma_2}\right)e^{-\frac{y}{\sigma_2}}\right]\bigg\},
\end{aligned}
\tag{4}
$$

where $x, y > 0; \sigma_1, \sigma_2 > 0; |\theta| \leq 1; -\theta-1 \leq \rho \leq \frac{1}{2}[3-\theta+(9-6\theta-3\theta^2)^{\frac{1}{2}}]$ and is zero, elsewhere.

The matching cdf is

$$\begin{aligned} H(x,y) &= \left[1-\left(1+\frac{0.5x}{\sigma_1}\right)e^{-\frac{x}{\sigma_1}}\right]\left[1-\left(1+\frac{0.5y}{\sigma_2}\right)e^{-\frac{y}{\sigma_2}}\right]\{1+\theta \\ &\times\left[1+\frac{0.5x}{\sigma_1}\right]e^{-\frac{x}{\sigma_1}}\left[1+\frac{0.5y}{\sigma_2}\right]e^{-\frac{y}{\sigma_2}} \\ &+\rho\left[1-\left(1+\frac{0.5x}{\sigma_1}\right)e^{-\frac{x}{\sigma_1}}\right]\left[1-\left(1+\frac{0.5y}{\sigma_2}\right)e^{-\frac{y}{\sigma_2}}\right] \\ &\times\left[\left(1+\frac{0.5x}{\sigma_1}\right)e^{-\frac{x}{\sigma_1}}\left(1+\frac{0.5y}{\sigma_2}\right)e^{-0.5y}\right]\right\}. \end{aligned} \quad (5)$$

For $\rho = 0$, the EFGMBLD leads to the FGMBLD (see, Maya et al. , 2018 and Irshad et al., 2019).

Clearly,

$$E(X) = 1.5\sigma_1, \quad Var(X) = \frac{7}{9}\sigma_1^2,$$

$$E(Y) = 1.5\sigma_2, \quad Var(Y) = \frac{7}{9}\sigma_2^2.$$

If we make the transformation,

$$W = \frac{X}{\sigma_1} \quad \text{and} \quad Z = \frac{Y}{\sigma_2}, \quad (6)$$

the standard EFGMBLD has the joint pdf as

$$\begin{aligned} h^*(w,z) = 0.25e^{-w-z}(1+z)(1+w)\{1+ \\ \theta\left[2e^{-w}(1+0.5w)-1\right]\left[2e^{-z}(1+0.5z)-1\right]+ \\ \rho\left[3e^{-w}(1+0.5w)-1\right]\left[3e^{-z}(1+0.5z)-1\right] \\ \times\left[1-e^{-w}(1+0.5w)\right]\left[1-e^{-z}(1+0.5z)\right]\}. \end{aligned} \quad (7)$$

It is clear that the variables W and Z have the standard univariate Lindley distribution as a marginal functions with pdf's are given by, respectively:

$$f_W(w) = \begin{cases} 0.5\,(1+w)\,e^{-w}, & if \quad w > 0, \\ 0, & otherwise. \end{cases}$$

and

$$f_Z(z) = \begin{cases} 0.5\,(1+z)\,e^{-z}, & if \quad z > 0. \\ 0, & otherwise. \end{cases} \quad (8)$$

2.1 Distribution theory of COS from EFGMBLD

The general theory of COS arising from the EFGM family are derived by Philip (2011). This section introduces the distribution theory of COS obtained from the EFGMBLD given in (4).

Let (X_i, Y_i) and (W_i, Z_i), be random samples of size m each arising from the EFGMBLD and standard EFGMBLD with pdfs given by (4) and (7), respectively. Let $Z_{[j:m]}$ denotes the concomitant to the j^{th} order statistic $W_{j:m}$ arising from (7). Then, the joint pdf $h^*_{[j,s:m]}(z_1, z_2)$ of $Z_{[j:m]}$ and $Z_{[s:m]}$ and the pdf $h^*_{[j:m]}(z)$ of $Z_{[j:m]}$ and are given by (see Philip (2011).

For $1 \leq j \leq m$,

$$\begin{aligned} h^*_{[j:m]}(z) &= h_Z(z)\left\{1 + \theta\frac{(m-2j+1)}{2(m+1)}\left[h_{1:2}(z) - h_{2:2}(z)\right] + \right. \\ &\left. +\rho\frac{(2m-3j+1)j}{3(m+1)(m+2)}\left[h_{2:3}(z) - h_{3:3}(z)\right]\right\}. \end{aligned} \tag{9}$$

For $1 \leq j < s \leq m$,

$$\begin{aligned} h^*_{[j,s:m]}(z_1, z_2) = &\, h_Z(z_1)h_Z(z_2) + \theta\frac{(m-2j+1)}{2(m+1)}\left[h_{1:2}(z_1) - h_{2:2}(z_1)\right]h_Z(z_2) \\ &+ \theta\frac{(m-2s+1)}{2(m+1)}\left[h_{1:2}(z_2) - h_{2:2}(z_2)\right]h_Z(z_1) \\ &+ \rho\frac{(2m-3j+1)j}{3(m+1)(m+2)}\left[h_{2:3}(z_1) - h_{3:3}(z_1)\right]h_Z(z_2) \\ &+ \rho\frac{(2m-3s+1)s}{3(m+1)(m+2)}\left[h_{2:3}(z_2) - h_{3:3}(z_2)\right]h_Z(z_1) \\ &+ \frac{\theta^2}{4}\left\{\frac{m-2s+1}{m+1} - \frac{2j(m-2s)}{(m+1)(m+2)}\right\} \\ &\times\left[h_{1:2}(z_1) - h_{2:2}(z_1)\right]\left[h_{1:2}(z_2) - h_{2:2}(z_2)\right] \\ &+ \frac{\theta\rho}{6}\left\{-\frac{2j}{m+1} + \frac{[3j(j+1) + 4j(m-s+1)]}{(m+1)(m+2)} - \frac{6j(j+1)(m-s+1)}{(m+1)(m+2)(m+3)}\right\} \\ &\times\left[h_{2:3}(z_1) - h_{3:3}(z_1)\right]\left[h_{1:2}(z_2) - h_{2:2}(z_2)\right] \\ &+ \frac{\theta\rho}{6}\left\{-\frac{s}{m+1} + \frac{[2j(s+1) + 3s(m-s+1)]}{(m+1)(m+2)} - \frac{6j(s+1)(m-s+1)}{(m+1)(m+2)(m+3)}\right\} \\ &\times\left[h_{1:2}(z_1) - h_{2:2}(z_1)\right]\left[h_{2:3}(z_2) - h_{3:3}(z_2)\right] \\ &+ \frac{\rho^2}{9}\left\{-\frac{2j(s+1)}{(m+1)(m+2)} + \frac{[3j(j+1)(s+2) + 6j(s+1)(m-s+1)]}{(m+1)(m+2)(m+3)}\right. \\ &\left. - \frac{9j(j+1)(s+2)(m-s+1)}{(m+1)(m+2)(m+3)(m+4)}\right\} \\ &\times\left[h_{2:3}(z_1) - h_{3:3}(z_1)\right]\left[h_{2:3}(z_2) - h_{3:3}(z_2)\right], \end{aligned} \tag{10}$$

where $h_{1:2}(.)$, $h_{2:2}(.)$, $h_{2:3}(.)$ and $h_{3:3}(.)$ in (9) and (10) are the pdf's of OS $W_{1:2}, W_{2:2}, W_{2:3}$ and $W_{3:3}$, respectively. Computing the values $h_{1:2}(.)$, $h_{2:2}(.)$, $h_{2:3}(.)$ and $h_{3:3}(.)$ and substituting these values in equations (9) and (10), we obtain $h^*_{[j:m]}(z)$ and $h^*_{[j,s:m]}(z_1, z_2)$ as

$$\begin{aligned} h^*_{[j:m]}(z) &= 0.5(1+z)e^{-z} + \theta\frac{(m-2j+1)}{2(m+1)}(1+z)e^{-z}\left[2\left(1+0.5z\right)e^{-z}-1\right] \\ &+ \rho\frac{(2m-3j+1)j}{2(m+1)(m+2)}(1+z)e^{-z}\left[1-\left(1+0.5z\right)e^{-z}\right]\left[3\left(1+0.5z\right)e^{-z}-1\right], \\ &\text{where } z>0; |\theta|\le 1; -\theta-1\le\rho\le 0.5[3-\theta+(9-6\theta-3\theta^2)^{0.5}]. \end{aligned} \tag{11}$$

and

$$\begin{aligned} h^*_{[j,s:m]}(z_1,z_2) &= 0.5(1+z_1)e^{-z_1}0.5(1+z_2)e^{-z_2} \\ &+\theta\frac{(m-2j+1)}{2(m+1)}(1+z_1)e^{-z_1}\left[2\left(1+0.5z_1\right)e^{-z_1}-1\right]0.5(1+z_2)e^{-z_2} \\ &+\theta\frac{(m-2s+1)}{2(m+1)}(1+z_2)e^{-z_2}\left[2\left(1+0.5z_2\right)e^{-z_2}-1\right]0.5(1+z_1)e^{-z_1} \\ &+\rho\frac{(2m-3j+1)j}{4(m+1)(m+2)}(1+z_1)(1+z_2)e^{-z_1-z_2} \\ &\times\left[1-\left(1+0.5z_1\right)e^{-z_1}\right]\left[3\left(1+0.5z_1\right)e^{-z_1}-1\right] \\ &+\rho\frac{(2m-3s+1)s}{4(m+1)(m+2)}(1+z_1)(1+z_2)e^{-z_1-z_2} \\ &\times\left[1-\left(1+0.5z_2\right)e^{-z_2}\right]\left[3\left(1+0.5z_2\right)e^{-z_2}-1\right] \\ &+\frac{\theta^2}{4}\left\{\frac{m-2s+1}{m+1}-\frac{2j(m-2s)}{(m+2)(m+1)}\right\} \\ &\times(1+z_1)(1+z_2)e^{-z_1-z_2}\left[2\left(1+0.5z_1\right)e^{-z_1}-1\right]\left[2\left(1+0.5z_2\right)e^{-z_2}-1\right] \\ &+\frac{\theta\rho}{4}\left\{-\frac{2j}{m+1}+\frac{[3j(j+1)+4j(m-s+1)]}{(m+1)(m+2)}-\frac{6j(j+1)(m-s+1)}{(m+3)(m+2)(m+1)}\right\} \\ &\times(1+z_1)(1+z_2)e^{-z_1-z_2}\left[1-\left(1+0.5z_1\right)e^{-z_1}\right]\left[3\left(1+0.5z_1\right)e^{-z_1}-1\right] \\ &\times\left[2\left(1+0.5z_2\right)e^{-z_2}-1\right] \\ &+\frac{\theta\rho}{4}\left\{-\frac{s}{m+1}+\frac{[2j(s+1)+3s(m-s+1)]}{(m+2)(m+1)}-\frac{6j(s+1)(m-s+1)}{(m+3)(m+2)(m+1)}\right\} \\ &\times(1+z_1)(1+z_2)e^{-z_1-z_2}\left[2\left(1+\frac{z_1}{2}\right)e^{-z_1}-1\right] \\ &\times\left[1-\left(1+0.5z_2\right)e^{-z_2}\right]\left[3\left(1+0.5z_2\right)e^{-z_2}-1\right] \\ &+\frac{\rho^2}{9}\left\{-\frac{2j(s+1)}{(m+1)(m+2)}+\frac{[3j(j+1)(s+2)+6j(s+1)(m-s+1)]}{(m+3)(m+2)(m+1)}\right. \\ &\left.-\frac{9j(j+1)(s+2)(m-s+1)}{(m+2)(m+1)(m+3)(m+4)}\right\} \\ &\times\frac{9}{4}(1+z_1)(1+z_2)e^{-z_1-z_2}\left[1-\left(1+0.5z_1\right)e^{-z_1}\right]\left[3\left(1+0.5z_1\right)e^{-z_1}-1\right] \\ &\times\left[1-\left(1+0.5z_2\right)e^{-z_2}\right]\left[3\left(1+0.5z_2\right)e^{-z_2}-1\right], \\ &w,z>0. \end{aligned} \tag{12}$$

Some basic properties of the standard univariate Lindley distribution given in (8), are

given below.
The k^{th} moment of the standard univariate Lindley given in (8)is obtained as

$$\mu^{(k)} = E(Z^k) = \frac{1}{2}\left(\Gamma{k+1} + \Gamma{k+2}\right). \tag{13}$$

The k^{th} moment of the j^{th} order statistics of Z when $m = 2$ is given by

$$E(Z^k_{j:2}) = \frac{2^{k+1}}{(2-j)!(j-1)!} \int_0^\infty t^k(1+2t)\left(\tfrac{1-(t+1)e^{-2t}}{t+1}\right)^j \frac{(t+1)^2 e^{-2t(3-j)}}{1-(t+1)e^{-2t}} dt.$$

For $k = 1, 2$ and $j = 1, 2$ we have

$$E(Z_{1:2}) = z_{1:2} = \tfrac{13}{16},$$

$$E(Z_{2:2}) = z_{2:2} = \tfrac{35}{16},$$

$$E(Z^2_{1:2}) = z^{(2)}_{1:2} = \tfrac{19}{16},$$

and

$$E(Z^2_{2:2}) = z^{(2)}_{2:2} = \tfrac{109}{16}.$$

The k^{th} moment of the jth order statistics of Z when $m = 3$ is given by,

$$E(Z^k_{j:3}) = \frac{3\left(2^{k+1}\right)}{(3-j)!(j-1)!} \int_0^\infty t^k(1+2t)\left(\tfrac{1-(t+1)e^{-2t}}{t+1}\right)^j \frac{(t+1)^3 e^{-2t(4-j)}}{1-(t+1)e^{-2t}} dt.$$

For $k = 1, 2$ and $j = 2, 3$ we have

$$E(Z_{2:3}) = z_{2:3} = \tfrac{565}{432},$$

$$E(Z_{3:3}) = z_{3:3} = \tfrac{1135}{432},$$

$$E(Z^2_{2:3}) = z^{(2)}_{2:3} = \tfrac{3113}{1296},$$

and

$$E(Z^2_{3:3}) = z^{(2)}_{3:3} = \tfrac{11687}{1296}.$$

Then, based on (11), we have

$$E(Z_{[j:m]}) = z_{[j:m]} = \frac{3}{2} - \theta\left[\frac{m-2j+1}{m+1}\right]\left(\frac{11}{16}\right) - \rho\left[\frac{(2m-3j+1)j}{(m+2)(m+1)}\right]\left(\frac{95}{216}\right), \tag{14}$$

$$E(Z_{[j:m]}^2) = z_{[j:m]}^{(2)}$$
$$= 4 - \theta\left[\frac{m-2j+1}{m+1}\right]\left(\frac{45}{16}\right) - \rho\left[\frac{(2m-3j+1)j}{(m+2)(m+1)}\right]\left(\frac{1429}{648}\right)$$

and by (12), we compute the product moment of $Z_{[j:m]}$ and $Z_{[s:m]}$ as, for $1 \leq j \leq s \leq m$,

$$\begin{aligned}
E\left(Z_{[j:m]}Z_{[s:m]}\right) &= z_{j,s:m} \\
&= \frac{9}{4} - \theta\left[\frac{m-(j+s)+1}{m+1}\right]\frac{33}{16} - \rho\left[\frac{(2m+1)(j+s)-3(j^2+s^2)}{3(m+1)(m+2)}\right]\frac{95}{48} \\
&+ \frac{\theta^2}{4}\left[\frac{m-2s+1}{m+1} - \frac{2j(m-2s)}{(m+2)(m+1)}\right]\left(\frac{11}{8}\right)^2 \\
&+ \frac{\theta\rho}{6}\left[\frac{-(2j+s)}{m+1} + \frac{(3j+2s+5)j+(4j+3s)(m-s+1)}{(m+2)(m+1)}\right]\frac{1045}{576} \\
&- \frac{\theta\rho}{6}\left[\frac{6j(j+s+2)(m-s+1)}{(m+1)(m+2)(m+3)}\right]\frac{1045}{576} \\
&+ \frac{\rho^2}{9}\left[\frac{-2j(s+1)}{(m+2)(m+1)} + \frac{3j(j+1)(s+2)+6j(s+1)(m-s+1)}{(m+3)(m+2)(m+1)}\right. \\
&\left. - \frac{9j(j+1)(s+2)(m-s+1)}{(m+2)(m+1)(m+3)(m+4)}\right]\left(\frac{95}{72}\right)^2.
\end{aligned}$$

For $1 \leq j \leq m$, the variance of $Z_{[j:m]}$, is obtained as

$$\begin{aligned}
Var(Z_{[j:m]}) &= 4 - \theta\left[\frac{m-2j+1}{m+1}\right]\left(\frac{45}{16}\right) - \rho\left[\frac{(2m-3j+1)j}{(m+2)(m+1)}\right]\left(\frac{1429}{648}\right) \\
&- \left\{\frac{3}{2} - \theta\left[\frac{m-2j+1}{m+1}\right]\left(\frac{11}{16}\right) - \rho\left[\frac{(2m-3j+1)j}{(m+2)(m+1)}\right]\left(\frac{95}{216}\right)\right\}^2.
\end{aligned} \tag{15}$$

For $1 \leq j \leq s \leq m$, the covariance between $Z_{[j:m]}$ and $Z_{[s:m]}$ is obtained as ,

$$
\begin{aligned}
Cov\left(Z_{[j:m]}, Z_{[s:m]}\right) = \\
&= \frac{9}{4} - \theta\left[\frac{m-(j+s)+1}{m+1}\right]\frac{33}{16} - \rho\left[\frac{(2m+1)(j+s)-3(j^2+s^2)}{3(m+1)(m+2)}\right]\frac{95}{48} \\
&+ \frac{\theta^2}{4}\left[\frac{m-2s+1}{m+1} - \frac{2j(m-2s)}{(m+2)(m+1)}\right]\left(\frac{11}{8}\right)^2 \\
&+ \frac{\theta\rho}{6}\left[\frac{-(2j+s)}{m+1} + \frac{(3j+2s+5)j+(4j+3s)(m-s+1)}{(m+1)(m+2)}\right]\frac{1045}{576} \\
&- \frac{\theta\rho}{6}\left[\frac{6j(j+s+2)(m-s+1)}{(m+1)(m+2)(m+3)}\right]\frac{1045}{576} \\
&+ \frac{\rho^2}{9}\left[\frac{-2j(s+1)}{(m+2)(m+1)} + \frac{3j(j+1)(s+2)+6j(s+1)(m-s+1)}{(m+3)(m+2)(m+1)}\right. \\
&\left.- \frac{9j(j+1)(s+2)(m-s+1)}{(m+2)(m+1)(m+3)(m+4)}\right]\left(\frac{95}{72}\right)^2. \\
&- \left\{1.5 - \theta\left[\frac{m-2j+1}{m+1}\right]\left(\frac{11}{16}\right) - \rho\left[\frac{(2m-3j+1)j}{(m+2)(m+1)}\right]\left(\frac{95}{216}\right)\right\} \\
&\times \left\{1.5 - \theta\left[\frac{m-2s+1}{m+1}\right]\left(\frac{11}{16}\right) - \rho\left[\frac{(2m-3s+1)j}{(m+2)(m+1)}\right]\left(\frac{95}{216}\right)\right\}.
\end{aligned} \tag{16}
$$

If we define the constants $\varphi_{j:m}, \eta_{j,j:m}$ and $\eta_{j,s:m}$ by the terms on the right side of equations (14), (15) and (16), then Equations (14)-(16) can be

$$E(Z_{[j:m]}) = \varphi_{j:m}, Var(Z_{[j:m]}) = \eta_{j,j:m}, 1 \leq j \leq m$$

and

$$Cov(Z_{[j:m]}, Z_{[s:m]}) = \eta_{j,s:m}, 1 \leq j < s \leq m,$$

respectively. From the substitution given in equation (6), we have

$$X_i = \sigma_1 W_i \quad \text{and} \quad Y_i = \sigma_2 Z_i, \quad \text{for} \quad i = 1, 2, \cdots, m.$$

Thus, to the j^{th} order statistics $Y_{[j:m]}$, $j = 1, 2, \cdots, m$ follows the EFGMBLD, the means and variances of the COS are

$$
\begin{aligned}
E(Y_{[j:m]}) =& \sigma_2 E(Z_{[j:m]}) \\
&= \sigma_2 \varphi_{j:m}
\end{aligned} \tag{17}
$$

and

$$
\begin{aligned}
Var(Y_{[j:m]}) =& \sigma_2^2 Var(Z_{[j:m]}) \\
&= \sigma_2^2 \eta_{j,j:m}.
\end{aligned} \tag{18}
$$

The COS $Y_{[j:m]}$ and $Y_{[s:m]}$, for $1 \leq j < s \leq m$, have a covariance as

$$
\begin{aligned}
Cov(Y_{[j:m]}, Y_{[s:m]}) =& \sigma_2^2 Cov(Z_{[j:m]}, Z_{[s:m]}) \\
&= \sigma_2^2 \eta_{j,s:m}.
\end{aligned} \tag{19}
$$

It is clear that the constants included in $\varphi_{j:m}, \eta_{j,j:m}$ and $\eta_{j,s:m}$ are known for known values of θ and ρ.

3 Unbiased estimator of the parameter σ_2 using RSS

The observations based on COS are correlated, which leads one to evaluate the variance and covariance of COS to use them for inference problems. However, in case of Stoke's *RSS* scheme, the number of units to be selected is definite and there exists no correlation between one observation to another as they are drawn from different samples so that handling the observations in *RSS* for inferential problem will be very easy. Also, estimation using Stoke's *RSS* can be effectively applied when it is difficult to quantified the study variate Y but the auxiliary variable is correlated with X which can be quantified easily. Hence, the problem of estimating the parameter of the study variable of the suggested distribution is considered here using the observations based on Stoke's *RSS* scheme.

Let (X, Y) be a bivariate rv which follows the EFGMBLD with pdf given in (4). Assume that m sets each of size m are drawn from the distribution defined in (4). By ranking the X measured observations from the j^{th} set, considering $X_{(j:m)j}$ as the j^{th} in the set, then $Y_{[j:m]j}$, where $j = 1, 2, \cdots, m$ is the actual measurement to the Y characteristic of the observation whose X value is $X_{(j:m)j}$.

It is of interest to note here that $Y_{[j:m]j}$ has the same distribution of the concomitant of the j^{th} OS of a sample of size m selected from the distribution given in (4).

Using the means and variances of COS obtained from the EFGMBLD, then $Y_{[j:m]j}$ for $1 \le j \le m$ has the means and variances given by

$$E(Y_{[j:m]j}) = \sigma_2 \left\{ 1.5 - \theta \left[\frac{m - 2j + 1}{m + 1} \right] \left(\frac{11}{16} \right) - \rho \left[\frac{(2m - 3j + 1)j}{(m + 2)(m + 1)} \right] \left(\frac{95}{216} \right) \right\} \tag{20}$$

and

$$\begin{aligned} Var(Y_{[j:m]j}) =& \sigma_2^2 \left\{ 4 - \theta \left[\frac{m - 2j + 1}{m + 1} \right] \left(\frac{45}{16} \right) - \rho \left[\frac{(2m - 3j + 1)j}{(m + 2)(m + 1)} \right] \left(\frac{1429}{648} \right) \right. \\ & \left. - \left\{ 1.5 - \theta \left[\frac{m - 2j + 1}{m + 1} \right] \left(\frac{11}{16} \right) - \rho \left[\frac{(2m - 3j + 1)j}{(m + 1)(m + 2)} \right] \left(\frac{95}{216} \right) \right\}^2 \right\}. \end{aligned} \tag{21}$$

Since the two measurements $Y_{[j:m]j}$ and $Y_{[s:m]s}$ $(j \neq s)$ of Y are based on two different groups, then we obtain

$$Cov(Y_{[j:m]j}, Y_{[s:m]s}) = 0, \quad j \neq s. \tag{22}$$

The next theorem proposes an unbiased estimator of σ_2 contained in (4).

Suppose (X, Y) has a EFGMBLD. Let $Y_{[j:m]j}$ for $, j = 1, 2, ..., m$ be the *RSS* units based on Y and the ranking is on X. Then, the estimator

$$\sigma_2^* = \frac{2}{3m} \sum_{j=1}^{m} Y_{[j:m]j} \tag{23}$$

is an unbiased estimator of σ_2 and with variance

$$Var\left(\sigma_2^*\right) = \frac{\sigma_2^2}{m}\left\{\frac{16}{9} - \frac{4}{9m}\right.$$
$$\left.\sum_{j=1}^{m}\left\{\frac{3}{2} - \theta\left[\frac{m-2j+1}{m+1}\right]\left(\frac{11}{16}\right) - \rho\left[\frac{(2m-3j+1)j}{(m+2)(m+1)}\right]\left(\frac{95}{216}\right)\right\}^2\right\}. \tag{24}$$

Proof. Taking expectations on equation (23), we get

$$E(\sigma_2^*) = \frac{2}{3m}\sum_{j=1}^{m} E\left(Y_{[j:m]j}\right) \tag{25}$$

Substituting (20) in (25), we get

$$E(\sigma_2^*) = \frac{2\sigma_2}{3m}\sum_{j=1}^{m}\left\{\frac{3}{2} - \theta\left[\frac{m-2j+1}{m+1}\right]\left(\frac{11}{16}\right) - \rho\left[\frac{(2m-3j+1)j}{(m+2)(m+1)}\right]\left(\frac{95}{216}\right)\right\}.$$

Since

$$\sum_{j=1}^{m}(m-2j+1) = 0 \quad \text{and} \quad \sum_{j=1}^{m}(2m-3j+1)j = 0, \tag{26}$$

we get

$$E(\sigma_2^*) = \sigma_2.$$

Hence, the variance of σ_2^* is

$$Var(\sigma_2^*) = \frac{4}{9m^2}\sum_{j=1}^{m} Var(Y_{[j:m]j}). \tag{27}$$

Applying (21) and (26) in (27) and simplifying, we get

$$Var\left(\sigma_2^*\right) = \frac{\sigma_2^2}{m}\left\{\frac{16}{9} - \frac{4}{9m}\right.$$
$$\left.\sum_{j=1}^{m}\left\{\frac{3}{2} - \theta\left[\frac{m-2j+1}{m+1}\right]\left(\frac{11}{16}\right) - \rho\left[\frac{(2m-3j+1)j}{(m+2)(m+1)}\right]\left(\frac{95}{216}\right)\right\}^2\right\}.$$

Thus, the theorem is proved.

4 BLUE of the parameter σ_2 of EFGMBLD using RSS

Here, a good estimator $\tilde{\sigma_2}$ of σ_2 is developed by finding the BLUE assuming that the parameters θ and ρ are known.
Assume that m sets of size m each are follow the EFGMBLD and $\mathbf{Y}_{[m]} = (Y_{[1:m]1}, Y_{[2:m]2}, \cdots, Y_{[m:m]})'$

is the column vector of COS taken from (4). It is obvious that $Y_{[j:m]j}$ has the same distribution as that of $Y_{[j:m]}$, the concomitant of the j^{th} OS. Hence, with reference to Equation (17), the mean vector of $\mathbf{Y}_{[m]}$ is

$$E(\mathbf{Y}_{[m]}) = \sigma_2 \varphi, \tag{28}$$

where φ=$(\varphi_{1:m}, \varphi_{2:m}, \cdots, \varphi_{m:m})'$. From equations (18) and (19), the dispersion matrix of $\mathbf{Y}_{[m]}$ can be given as

$$D[\mathbf{Y}_{[m]}] = \sigma_2^2 \mathbf{H}, \tag{29}$$

where $\mathbf{H} = diag(\eta_{1,1:m}, \eta_{2,2:m}, \cdots, \eta_{m,m:m})$.

For known values of θ and ρ, then based on (28) and (29) a generalized Gauss-Markov setup can be defined and then the BLUE of σ_2 is given by

$$\tilde{\sigma_2} = (\varphi'\mathbf{H}^{-1}\varphi)^{-1}\varphi'\mathbf{H}^{-1}\mathbf{Y}_{[m]}$$

with variance given by

$$Var(\tilde{\sigma_2}) = \frac{\sigma_2^2}{\varphi'\mathbf{H}^{-1}\varphi}.$$

On simplifying, we get

$$\tilde{\sigma_2} = \frac{\sum_{j=1}^{m} \frac{\varphi_{j:m}}{\eta_{j,j:m}}}{\sum_{j=1}^{m} \frac{\varphi_{j:m}^2}{\eta_{j,j:m}}} Y_{[j:m]j} \tag{30}$$

and

$$Var(\tilde{\sigma_2}) = \frac{\sigma_2^2}{\sum_{j=1}^{m} \frac{\varphi_{j:m}^2}{\eta_{j,j:m}}}. \tag{31}$$

From (30), we have $\tilde{\sigma_2}$ is a linear functions of the ranked set sample observations $Y_{[j:m]j}$, for $j = 1, 2, \cdots, m$ and hence $\tilde{\sigma_2}$ can be formed as $\tilde{\sigma_2} = \sum_{j=1}^{m} a_j Y_{[j:m]j}$, where

$$a_j = \frac{\frac{\varphi_{j:m}}{\eta_{j,j:m}}}{\sum_{j=1}^{m} \frac{\varphi_{j:m}^2}{\eta_{j,j:m}}}.$$

Next, we have evaluated the efficiency $e(\tilde{\sigma_2}/\sigma_2^*) = \frac{Var(\sigma_2^*)}{Var(\tilde{\sigma_2})}$ with $m = 2, 3, ..., 10$; $\theta = -1, -0.75, -0.50, 0.25$; $\rho = 0.5, 1$. The results are given in Table 1. Based on Table 1, it is observed that the numerical values of $e(\tilde{\sigma_2}/\sigma_2^*) = \frac{Var(\sigma_2^*)}{Var(\tilde{\sigma_2})}$ are greater than unity for all values of θ, ρ and m and increases with increasing the sample size m. Also, it is observed that for any negative value of θ, with fixed value of ρ, the $e(\tilde{\sigma_2}/\sigma_2^*)$ increases as m increases.

Table 1: The $e(\tilde{\sigma_2}/\sigma_2^*)$ for some selected ρ, m, and θ

m	ρ	θ			
		-1	-0.75	-0.50	0.25
2		1.00066	1.00062	1.00052	1.00026
3		1.00124	1.00102	1.00082	1.00031
4		1.00171	1.00126	1.00109	1.00036
5		1.00202	1.00145	1.00123	1.00045
6	0.5	1.00219	1.00166	1.00132	1.00039
7		1.00237	1.00176	1.00145	1.00046
8		1.00261	1.00190	1.00156	1.00052
9		1.00269	1.00190	1.00164	1.00047
10		1.00286	1.00198	1.00169	1.00052
2		1.00199	1.00184	1.00165	1.00120
3		1.00348	1.00312	1.00271	1.00153
4		1.00451	1.00402	1.00347	1.00168
5		1.00533	1.00464	1.00395	1.00177
6	1.0	1.00593	1.00519	1.00436	1.00190
7		1.00637	1.00560	1.00463	1.00194
8		1.00914	1.00589	1.00488	1.00201
9		1.00726	1.00615	1.00503	1.00212
10		1.00740	1.00632	1.00533	1.00232

5 Concluding remarks

The distribution theory of COS selected from EFGMBLD is developed. This development further provides necessary statistical foundation to formulate *RSS* strategies for a population random variable following a EFGMBLD. Finally, according to Stoke's *RSS* scheme, we have derived some estimators of the parameter related with the variable of primary interest. As a future work, the authors recommend to estimate the EFGMBLD parameters using other variations of *RSS*, see Al-Omari (2011), Zamanzade and Al-Omari (2016), and Haq et al. (2015).

Acknowledgments This research was funded by the Deanship of Scientific Research at Princess Nourah bint Abdulrahman University through the Fast-track Research Funding Program.

References

Al-Omari, A.I. (2011). Estimation of mean based on modified robust extreme ranked set sampling. *Journal of Statistical Computation and Simulation*, 81(8), 1055-1066.

Al-Omari, A.I. (2021). Maximum likelihood estimation in location-scale families using varied L ranked set sampling. *RAIRO Operations Research*, 55, 2759-2771.

Al-Omari, A.I. and Abdallah, M.S. (2021). Estimation of the distribution function using moving extreme and minimax ranked set sampling. *Communications in Statistics-Simulation and Computation*, https://doi.org/10.1080/03610918.2021.1891433.

Al-Omari, A.I. and Almanjahie, I.M. (2021). New improved ranked set sampling design with an application to real data. *Computers, Materials and Continua*, 67(2), 1503-1522.

Al-Omari, A.I. and Haq, A. (2019). A new sampling method for estimating the population mean. *Journal of Statistical Computation and Simulation*, 89(11), 1973-1985.

Beg, M.I. and Ahsanullah, M. (2008). Concomitants of generalized order statistics from Farlie–Gumbel–Morgenstern distributions. *Statistical Methodology*, 5(1), 1-20.

Benchiha, S. and Al-Omari, A.I. (2021). Generalized quasi Lindley distribution: theoretical properties, estimation methods and applications. *Electronic Journal of Applied Statistical Analysis*, 14(1), 167-196.

David, H.A. and Nagaraja, H.N. (2003). Order Statistics. Third edition. John Wiley and Sons, New York..

Jemain, A.A., Al-Omari, A.I. and Ibrahim, K. (2007). Multistage median ranked set sampling for estimating the population median. *Journal of Mathematics and Statistics*, 3(2), 58-64.

Haq, A., Brown, J., Moltchanova, E. and Al-Omari, A.I. (2015). Varied L ranked set sampling scheme. *Journal of Statistical Theory and Practice*, 9(4), 741-767.

Haq, A., Brown, J., Moltchanova, E. and Al-Omari, A.I. (2016a). Best linear unbiased and invariant estimation in location-scale families based on double ranked set sampling. *Communications in Statistics-Theory and Methods*, 45(1), 225-48.

Haq, A., Brown, J., Moltchanova, E. and Al-Omari, A.I. (2016b). Paired double ranked set sampling. *Communications in Statistics-Theory and Methods*, 45(1), 2873-2889.

Hassan, A.S., Al-Omari, A.I. and Nagy, H.F. (2021). Stress-strength reliabilbity for the generalized inverted Exponential distribution using MRSS. *Iranian Journal of Science and Technology, Transactions A: Science*, 45(2), 641-659.

Johnson N.L. and Kotz, S. (1977). On some generalized Farlie-Gumbel-Morgenstern distributions-II: Regression, correlation and further generalizations. *Communications in Statistics-Theory and Methods*, 6, 485-496.

Koyuncu, N. and Al-Omari, A.I. (2021). Generalized robust-regression-type estimators under different ranked set sampling. *Mathematical Sciences*, 15, 29-40.

Irshad, M.R., Maya, R. and Arun S.P. (2019). Estimation of a parameter of Morgenstern type bivariate Lindley distribution by ranked set sampling. *iSTATISTIK: Journal of the Turkish Statistical Association*, 12(1-2), 25-34.

Mahdizadeh,M. and Zamanzade, E. (2021). Smooth estimation of the area under the ROC curve in multistage ranked set sampling. *Statistical Papers*, 62, 1753-1776.

Mahdizadeh,M. and Zamanzade, E. (2021). New estimator for the variances of strata in ranked set sampling. *Soft Computing*, 25(13), 8007-8013.

Maya, R., Irshad, M.R. and Shibu, D.S. (2018). Concomitants of record values arising from the Morgenstern type bivariate Lindley distribution. *Afrika Statistika*, 13(4), 1865-1875.

Maya, R., Irshad, M.R. and Arun, S.P. (2021). Farlie-Gumbel-Morgenstern bivariate Bilal Distribution and its inferential aspects using concomitants of order statistics. *Journal of Probability and Statistical Science*, 19(1), 1-20.

McIntyre, G.A. (1952). A method for unbiased selective sampling using ranked sets. *Australian Journal of Agricultural Research* 3, 385-390.

Morgenstern, D. (1956). Einfache beispiele zweidimensionaler verteilungen. *Mittelingsblatt fur Mathematische Statistik*, 8, 234-235.

Philip, A. (2011). Concomitants of order statistics from an extended Farlie-Gumbel-Morgenstern distribution. *Journal of the Kerala Statistical Association*, 22, 5-20.

Scaria, J. and Nair, N.U. (1999). On concomitants of order statistics from Morgenstern family. *Biometrical Journal*, 41, 483-489.

Stokes, S.L. (1977). Ranked set sampling with concomitant variables. *Communications in Statistics-Theory and Methods*, 6, 1207-1211.

Takahasi, K. and Wakimoto, K. (1968). On the unbiased estimates of the population mean based on the sample stratified by means of ordering. *Annals of the Institute of Statistical Mathematics*, 20, 1-31.

Terpstra, J.T. and Miller, Z.A. (2006). Exact inference for a population proportion based on a ranked set sample. *Communications in Statistics: Simulation and Computation*, 35(1), 19-26.

Zamanzade, E. and Al-Omari, A.I. (2016). New ranked set sampling for estimating the population mean and variance. *Hacettepe Journal of Mathematics and Statistics*, 45(6), 1891-1905.