Statistics and Applications {ISSN 2452-7395 (online)}
Volume 18, No. 1, 2020 (New Series), pp 127-140

On Moments of Order Statistics and Some Aspects of Robustness Issues of Lindley Distribution in the Presence of Multiple Outliers

Mahesh Kumar Panda

Department of Statistics, Central University of Odisha, Koraput 763004

Received: 16 October 2019; Revised: 11 May 2020; Accepted: 23 May 2020

Abstract

This article obtains the exact expressions for the single and product moments of order statistics from one parameter Lindley distribution when multiple outliers are present in the data. Using the obtained moment relations, we compute the single and product moments (*e.g.* means, variances and covariances) of various order statistics. Next, we explore the impact of the presence of pronounced outliers on these variances and covariances, while the shape parameters have been shifted in value. We also investigate the robustness issues of the sample L-moments.

Key words: Order statistics; Outliers; Single and product moments; Covariances; Sample L-moments.

AMS Subject Classification: 62G30; 62F10

1. Introduction

In the fields of engineering, medical and biological science, the statistical analysis of lifetime data plays a significant role. In fact, the lifetime distributions are being used in different forms of investigations from the issue of survival time of manufactured items in engineering to the researches involving human diseases in biomedical sciences. In the literature, there are several statistical distributions available for modelling lifetime data. Among these distributions, the predominantly used one is exponential distribution (due to its closed form) for its survival function. The Lindley distribution belongs to the exponential family distribution and can be written as a mixture of exponential and gamma distributions. This distribution is better than exponential failure time distribution, wherein hazard rate is not unimodal or bathtub shaped [see Bakouch et al. (2012)]. The Lindley distribution, having an advantage over the exponential distribution, is due to the fact that the former possesses the increasing hazard rate and decreasing mean residual life time function (MRLF), whereas the latter one possesses constant hazard rate and MRLF. Maybe, owing to this nice property, recently many authors have paid their attention to Lindley distribution as a life time model in different perspectives [see Kumar and Jose (2018)]. Ghitany et al. (2008) showed through waiting time data that Lindley distribution provides a better model as compared to the well-known exponential distribution. This distribution also provides a better fit to competing risks lifetime data in contrast to exponential and Weibull distributions [see Mazuchelia and Achcarb (2011)]. A discrete Lindley model was introduced by Gomoz-Deniz and Calderin-Ojeda (2011) with its applications in collective risk modelling. Krishna and Kumar (2011) demonstrated that Lindley distribution might fit better than exponential, lognormal and gamma distributions in some real life problems under progressive Type-II censoring scheme. Mazucheli et al. (2019) introduced a transformed form of Lindley distribution *i.e.* unit-Lindley distribution and demonstrated that

Corresponding Author: Mahesh Kumar Panda
E-mail: mahesh2123ster@gmail.com

unit-Lindley regression could offer a better fit as compared to beta regression model by using the data of inadequate water supply and sewage in the cities of Brazil from the southeast and northeast region.

The specific area moments of order statistics has been consistently being used in other disciplines such as life testing, reliability theory, signal and image processing *etc*. In the early 70s, many researchers started working on studies of order statistics based on outlier model due to the robustness issue. An outlier in a dataset is an observation that appears to be inconsistent with the remaining observations [see Prasad et al. (2008)]. In any dataset, the presence of single or multiple outlier(s) may leads to a flawed conclusion drawn from the experiment; thus it is important to detect and handle the outlier(s) efficiently. In fact, the detection of multiple outliers in comparison to detection of a single outlier is much more difficult [see Bhar et al. (2013)].

Much of the work on order statistics in connection with robustness issue has been focused when there is one outlier present in the sample (single outlier model), but nothing much in case of multiple outliers model. Barnett and Lewis (1994) extensively discussed the topic of development on the single outlier model. Arnold and Balakrishnan (1989) obtained the density function of r^{th} order statistic as well as the joint density function of $X_{r:n}$ and $X_{s:n}$ ($1 \leq r < s \leq n$) when the sample of size is n and the sample contains an unknown single outlier. Balakrishnan (1994a) obtained the recurrence relations for the single and product moments of order statistics from right truncated exponential distribution under the multiple-outliers model. Balakrishnan ((1994b), 2007) provided many results on order statistics from multiple-outliers model and the robustness issues involved in those models. Sultan and Moshref (2014) obtained the exact expressions of order statistics for the single and product moments of order statistics from Weibull distribution under the multiple-outliers model (*i.e.* with slippage of observations).

This article derives the exact expressions of order statistics for the single and product moments of order statistics from Lindley distribution when multiple outliers are present in the data. The rest of this article is organized as follows. In Section 2, we give the preliminaries which will be used to derive the main result. In Section 3, we derive the exact expression of the single and product moments of order statistics from Lindley distribution under the multiple-outlier model. In Section 4, we obtain the L-moments of order statistics and also examine the robustness of the sample L-moments in the presence of outliers through some numerical illustrations. In Section 5, we establish some special cases. Finally, in Section 6, we sketch a conclusion of the article.

2. Preliminaries

Under the multiple outliers model set up, we assume that $x_1, x_2, \ldots, x_n$ are independent variables with $x_1, x_2, \ldots, x_{n-p}$ are $(n-p)$ independent random variables from one form of the Lindley distribution with probability density function (pdf) $f(x)$ given by

$$f(x) = \frac{\theta^2}{(1+\theta)}(1+x)e^{-\theta x},\ x \geq 0,\ \theta > 0, \tag{1}$$

while $x_{n-p+1}, x_{n-p+2},\ldots, x_n$ are the p independent random variables (*i.e.* p outliers) from another form of the Lindley distribution with pdf $g(x)$ given by

$$g(x)=\frac{\tau^2}{(1+\tau)}(1+x)e^{-\tau x},\ x\geq 0,\ \tau>0, \tag{2}$$

where θ, τ are the shape parameters of pdfs $f(x)$ and $g(x)$ respectively. We also suppose that these parameters are linked with each other by a relationship $\tau=\frac{\theta}{h},\ h\in(0,\,1)$. It can be shown that both the cumulative density functions (cdfs) are related with the corresponding pdfs by the following relationships:

$$f(x)=\frac{\theta^2}{1+\theta+\theta x}(1+x)\{1-F(x)\} \tag{3}$$

and

$$g(x)=\frac{\tau^2}{1+\tau+\tau x}(1+x)\{1-G(x)\}. \tag{4}$$

Let $x_{1:n}\leq x_{2:n}\leq\ldots\leq x_{n:n}$ denote the order statistics obtained from a sample of n observations. Then, the pdf of the r^{th} order statistic $x_{r:n}$, under the multiple-outliers model is as follows [see Childs (1996) and Sultan and Moshref (2014)]:

$$\begin{aligned} f_{r:n}[p](x)=&\sum_{s=\max(0,r-p-1)}^{\min(n-p-1,r-1)} C_1 f(x)\{F(x)\}^s\{G(x)\}^{r-s-1}\{1-F(x)\}^{n-p-s-1}\{1-G(x)\}^{p-r+s+1}\\ &+\sum_{s=\max(0,r-p)}^{\min(n-p,r-1)} C_2 g(x)\{F(x)\}^s\{G(x)\}^{r-s-1}\{1-F(x)\}^{n-p-s}\{1-G(x)\}^{p-r+s},-\infty<x<\infty \end{aligned} \tag{5}$$

where

$$C_1=\frac{(n-p)!\,p!}{s!(r-s-1)!(n-p-s-1)!(p-r+s+1)!},$$

and

$$C_2=\frac{(n-p)!\,p!}{s!(r-s-1)!(n-p-s)!(p-r+s)!}.$$

Similarly, the joint density function of the r^{th} and s^{th} order statistics $x_{r:n}$ and $x_{s:n}$ ($1\leq r< s\leq n$), under the multiple-outliers model is given by [see Childs (1994) and Sultan and Moshref (2014)]:

$$\begin{aligned} f_{r,s:n}[p](x,y)=&\sum_{j=0}^{s-r-1}\sum_{i=\max(0,s-p-j-2)}^{\min(n-p-j-2,r-1)} A_1 f(x)f(y)\{F(x)\}^i\{G(x)\}^{r-1-i}\{F(y)-F(x)\}^j\\ &\times\{G(y)-G(x)\}^{s-r-1-j}\{1-F(y)\}^{n-p-i-j-2}\{1-G(y)\}^{p-s+i+j+2}\\ &+\sum_{j=0}^{s-r-1}\sum_{i=\max(0,s-p-j-1)}^{\min(n-p-j-1,r-1)} A_2 f(x)g(y)\{F(x)\}^i\{G(x)\}^{r-1-i}\{F(y)-F(x)\}^j\\ &\times\{G(y)-G(x)\}^{s-r-1-j}\{1-F(y)\}^{n-p-i-j-1}\{1-G(y)\}^{p-s+i+j+1} \end{aligned}$$

$$+\sum_{j=0}^{s-r-1}\sum_{i=\max(0,s-p-j-1)}^{\min(n-p-j-1,r-1)} A_2 g(x)f(y)\{F(x)\}^i\{G(x)\}^{r-1-i}\{F(y)-F(x)\}^j$$
$$\times\{G(y)-G(x)\}^{s-r-1-j}\{1-F(y)\}^{n-p-i-j-1}\{1-G(y)\}^{p-s+i+j+1}$$
$$+\sum_{j=0}^{s-r-1}\sum_{i=\max(0,s-p-j)}^{\min(n-p-j,r-1)} A_3 g(x)g(y)\{F(x)\}^i\{G(x)\}^{r-1-i}\{F(y)-F(x)\}^j$$
$$\times\{G(y)-G(x)\}^{s-r-1-j}\{1-F(y)\}^{n-p-i-j}\{1-G(y)\}^{p-s+i+j}, -\infty<x<y<\infty \quad (6)$$

where

$$A_1=\frac{(n-p)!\,p!}{i!(r-1-i)!\,j!(s-r-1-j)!(n-p-i-j-2)!(p-s+i+j+2)!},$$
$$A_2=\frac{(n-p)!\,p!}{i!(r-1-i)!\,j!(s-r-1-j)!(n-p-i-j-1)!(p-s+i+j+1)!},$$

and
$$A_3=\frac{(n-p)!\,p!}{i!(r-1-i)!\,j!(s-r-1-j)!(n-p-i-j)!(p-s+i+j)!}.$$

3. Moments of Order Statistics

In this section, we obtain the exact expressions for the single and product moments of order statistics from Lindley distribution based on multiple-outliers model (based on p-outliers observations).

3.1. Single moments

In this subsection, we derive the kth moment of the r^{th} order statistics $\mu_{r:n}^{(k)}[p]$, $1\le r\le n$ under the multiple-outliers model.

Relation 1: For $1\le r\le n$, and k = 0, 1, 2,…. the k^{th} moment $\mu_{r:n}^{(k)}[p]$ is given by

$$\mu_{r:n}^{(k)}[p]=\theta^2\sum_{s=\max(0,r-p-1)}^{\min(n-p-1,r-1)} c_1\sum_{i=0}^{s}\binom{s}{i}\sum_{j=0}^{r-s-1}\binom{r-s-1}{j}\sum_{l=0}^{n-p-s+i-1}\binom{n-p-s+i-1}{l}$$
$$\times\sum_{m=0}^{p-r+s+j+1}\binom{p-r+s+j+1}{m}\sum_{q=0}^{l+m+1}\binom{l+m+1}{q}(-1)^{i+j}\frac{\theta^l\tau^m}{(1+\theta)^{n-p-s+i}(1+\tau)^{p-r+s+j+1}}$$
$$\times\frac{\Gamma(k+q+1)}{[\theta(n-p-s+i)+\tau(p-r+s+j+1)]^{k+q+1}}$$
$$+\tau^2\sum_{s=\max(0,r-p)}^{\min(n-p,r-1)} c_2\sum_{i=0}^{s}\binom{s}{i}\sum_{j=0}^{r-s-1}\binom{r-s-1}{j}\sum_{l=0}^{n-p-s+i}\binom{n-p-s+i}{l}\sum_{m=0}^{p-r+s+j}\binom{p-r+s+j}{m}$$
$$\times\sum_{q=0}^{l+m+1}\binom{l+m+1}{q}(-1)^{i+j}\frac{\theta^l\tau^m}{(1+\theta)^{n-p-s+i}(1+\tau)^{p-r+s+j+1}}$$
$$\times\frac{\Gamma(k+q+1)}{[\theta(n-p-s+i)+\tau(p-r+s+j+1)]^{k+q+1}} \quad (7)$$

Proof: For $1\le r\le n$, and k = 0, 1, 2,… and by substituting equations (3) and (4) in equation (5), we have

$$\mu_{r:n}^{(k)}[p] = \theta^2 \sum_{s=\max(0,r-p-1)}^{\min(n-p-1,r-1)} c_1 \int_0^\infty \frac{x^k(1+x)}{1+\theta+\theta x} \{F(x)\}^s \{G(x)\}^{r-s-1} \{1-F(x)\}^{n-p-s} \{1-G(x)\}^{p-r+s+1} dx$$
$$+ \tau^2 \sum_{s=\max(0,r-p)}^{\min(n-p,r-1)} c_2 \int_0^\infty \frac{x^k(1+x)}{1+\tau+\tau x} \{F(x)\}^s \{G(x)\}^{r-s-1} \{1-F(x)\}^{n-p-s} \{1-G(x)\}^{p-r+s+1} dx \qquad (8)$$

Using binomial theorem in equation (8) and subsequently expanding the same we get

$$\mu_{r:n}^{(k)}[p] = c_1 \theta^2 \sum_{s=\max(0,r-p-1)}^{\min(n-p-1,r-1)} \sum_{i=0}^{s} \sum_{j=0}^{r-s-1} (-1)^{i+j} \binom{s}{i} \binom{r-s-1}{j} \frac{1}{(1+\theta)^{n-p-s+i}(1+\tau)^{p-r+s+1+j}}$$
$$\times \int_0^\infty x^k (1+x)(1+\theta+\theta x)^{n-p-s+i-1} \left(1+\tau+\tau x\right)^{p-r+s+1+j} e^{-[(n-p-s+i)\theta+(p-r+s+1+j)\tau]x} dx$$
$$+ c_2 \tau^2 \sum_{s=\max(0,r-p)}^{\min(n-p,r-1)} \sum_{i=0}^{s} \sum_{j=0}^{r-s-1} (-1)^{i+j} \binom{s}{i} \binom{r-s-1}{j} \frac{1}{(1+\theta)^{n-p-s+i}(1+\tau)^{p-r+s+1+j}}$$
$$\times \int_0^\infty x^k (1+x)(1+\theta+\theta x)^{n-p-s+i} \left(1+\tau+\tau x\right)^{p-r+s+j} e^{-[(n-p-s+i)\theta+(p-r+s+1+j)\tau]x} dx \qquad (9)$$

Again, using binomial theorem in equation (9) and further simplifying we get the result in equation (7). □

The expression in equation (7) is used to calculate the mean and variance of the order statistics when $n = 6$, $p = 0, 1, 2$, $\theta = 1$ and $\tau = \frac{\theta}{h}$, $h = 0.1, 0.2, 0.3, 0.4, 0.5$ and are presented in Table 1 (Annexure). We can verify the results in Table 1 for the case $p = 0$, by using the well-known identity [see Arnold and Balakrishnan (1989), p. 6]

$$\sum_{i=1}^{6} \mu_{i:6} = 6E(X) = 6\frac{(\theta+2)}{\theta(\theta+1)}.$$

From Table 1, we see the following:

(1) The variance decreases as p increases.
(2) The variance is an increasing function of h for $r = 1$ and it is a decreasing function of h for $r = 5, 6$. For $r = 2, 3$ and 4 the behaviour is not consistent.
(3) For small r, the relative change in variance is more with the increase in the number of outliers from $p = 1$ to $p = 2$ for different values of h.

3.2. Product moments

In this subsection, we derive the $(k, l)^{\text{th}}$ moment of the r^{th} and s^{th} order statistics $\mu_{r,s:n}^{(k,l)}[p]$, under the multiple-outliers model.

Relation 2. For $1 \le r < s \le n$, and $k, l = 0, 1, 2, \ldots$ the product moments $\mu_{r,s:n}^{(k,l)}[p]$ is given by

$$\mu_{r,s:n}^{(k,l)}[p] = A_1 \theta^4 \sum_{j=0}^{s-r-1} \sum_{i=\max(0,s-p-j-2)}^{\min(n-p-j-2,r-1)} \sum_{b=0}^{i} \sum_{d=0}^{r-1-i} \sum_{t=0}^{j} \sum_{q=0}^{s-r-1-j} \sum_{p_1=0}^{b+j-t} \sum_{p_2=0}^{d+s-r-j-q-1} \sum_{p_3=0}^{n-p-i-j+t-2} \sum_{p_4=0}^{p-s+i+j+q+2} \sum_{l_1=0}^{p_1+p_2+1}$$

$$\times \sum_{l_2=0}^{p_3+p_4+1} \sum_{a=0}^{l+l_2} (-1)^{b+d+t+q} \binom{i}{b}\binom{r-1-i}{d}\binom{j}{t}\binom{s-r-1-j}{q}\binom{b+j-t}{p_1}$$

$$\times \binom{d+s-r-j-q-1}{p_2}\binom{n-p-i-j+t-2}{p_3}\binom{p-s+i+j+q+2}{p_4}$$

$$\times \binom{p_1+p_2+1}{l_1}\binom{p_3+p_4+1}{l_2} \frac{\theta^{p_1+p_3}\tau^{p_2+p_4}}{(1+\theta)^{b+n-p-i}(1+\tau)^{d-r+p+i+1}} \frac{(l+l_2)!}{a!}$$

$$\times \frac{\Gamma(k+l_1+a+1)}{[\theta(n-s+j+1)]^{l+l_1+1}[\theta(b+n-r+1)]^{k+l_1+a+1}}$$

$$+2A_2\,\theta^2\tau^2 \sum_{j=0}^{s-r-1} \sum_{i=\max(0,s-p-j-1)}^{\min(n-p-j-1,r-1)} \sum_{b=0}^{i}\sum_{d=0}^{r-1-i}\sum_{t=0}^{j}\sum_{q=0}^{s-r-1-j}\sum_{p_1=0}^{b+j-t}\sum_{p_2=0}^{d+s-r-j-q-1}\sum_{p_3=0}^{n-p-i-j+t-1}\sum_{p_4=0}^{p-s+i+j+q+1}$$

$$\times \sum_{l_1=0}^{p_1+p_2+1}\sum_{l_2=0}^{p_3+p_4+1} \sum_{a=0}^{l+l_2} (-1)^{b+d+t+q} \binom{i}{b}\binom{r-1-i}{d}\binom{j}{t}\binom{s-r-1-j}{q}\binom{b+j-t}{p_1}$$

$$\times \binom{d+s-r-j-q-1}{p_2}\binom{n-p-i-j+t-1}{p_3}\binom{p-s+i+j+q+1}{p_4}$$

$$\times \binom{p_1+p_2+1}{l_1}\binom{p_3+p_4+1}{l_2} \frac{\theta^{p_1+p_3}\tau^{p_2+p_4}}{(1+\theta)^{b+n-p-i}(1+\tau)^{d-r+p+i+1}} \frac{(l+l_2)!}{a!}$$

$$\times \frac{\Gamma(k+l_1+a+1)}{[\theta(n-s+j+1)]^{l+l_1+1}[\theta(b+n-r+1)]^{k+l_1+a+1}}$$

$$+A_3\,\tau^4 \sum_{j=0}^{s-r-1} \sum_{i=\max(0,s-p-j)}^{\min(n-p-j,r-1)} \sum_{b=0}^{i}\sum_{d=0}^{r-1-i}\sum_{t=0}^{j}\sum_{q=0}^{s-r-1-j}\sum_{p_1=0}^{b+j-t}\sum_{p_2=0}^{d+s-r-j-q-1}\sum_{p_3=0}^{n-p-i-j+t}\sum_{p_4=0}^{p-s+i+j+q}\sum_{l_1=0}^{p_1+p_2+1}$$

$$\times \sum_{l_2=0}^{p_3+p_4+1} \sum_{a=0}^{l+l_2} (-1)^{b+d+t+q} \binom{i}{b}\binom{r-1-i}{d}\binom{j}{t}\binom{s-r-1-j}{q}\binom{b+j-t}{p_1}$$

$$\times \binom{d+s-r-j-q-1}{p_2}\binom{n-p-i-j+t}{p_3}\binom{p-s+i+j+q}{p_4}\binom{p_1+p_2+1}{l_1}$$

$$\times \binom{p_3+p_4+1}{l_2} \frac{\theta^{p_1+p_3}\tau^{p_2+p_4}}{(1+\theta)^{b+n-p-i}(1+\tau)^{d-r+p+i+1}} \frac{(l+l_2)!}{a!}$$

$$\times \frac{\Gamma(k+l_1+a+1)}{[\theta(n-s+j+1)]^{l+l_1+1}[\theta(b+n-r+1)]^{k+l_1+a+1}} \tag{10}$$

Proof: For $1 \le r < s \le n$, and k, l = 0, 1, 2,… and by using equations (3) and (4) in equation (6), we have

$$\mu_{r,s:n}^{(k,l)}[p] = \theta^4 \sum_{j=0}^{s-r-1} \sum_{i=\max(0,s-p-j-2)}^{\min(n-p-j-2,r-1)} A_1 \int_0^\infty\int_x^\infty \frac{x^k(1+x)y^l(1+y)}{(1+\theta+\theta x)(1+\theta+\theta y)} \{F(x)\}^i\{G(x)\}^{r-1-i}\{F(y)-F(x)\}^j$$
$$\times \{G(y)-G(x)\}^{s-r-1-j}\{1-F(y)\}^{n-p-i-j-1}\{1-G(y)\}^{p-s+i+j+2}\{1-F(x)\}dxdy$$
$$+\theta^2\tau^2 \sum_{j=0}^{s-r-1} \sum_{i=\max(0,s-p-j-1)}^{\min(n-p-j-1,r-1)} A_2 \int_0^\infty\int_x^\infty \frac{x^k(1+x)y^l(1+y)}{(1+\theta+\theta x)(1+\tau+\tau y)} \{F(x)\}^i\{G(x)\}^{r-1-i}\{F(y)-F(x)\}^j$$
$$\times \{G(y)-G(x)\}^{s-r-1-j}\{1-F(y)\}^{n-p-i-j-1}\{1-G(y)\}^{p-s+i+j+2}\{1-F(x)\}dxdy$$

$$+\theta^2\tau^2\sum_{j=0}^{s-r-1}\sum_{i=\max(0,s-p-j-1)}^{\min(n-p-j-1,r-1)} A_2\int_0^\infty\int_x^\infty \frac{x^k(1+x)y^l(1+y)}{(1+\tau+\tau x)(1+\theta+\theta y)}\{F(x)\}^i\{G(x)\}^{r-1-i}\{F(y)-F(x)\}^j$$

$$\times\{G(y)-G(x)\}^{s-r-1-j}\{1-F(y)\}^{n-p-i-j}\{1-G(y)\}^{p-s+i+j+1}\{1-G(x)\}dxdy$$

$$+\tau^4\sum_{j=0}^{s-r-1}\sum_{i=\max(0,s-p-j)}^{\min(n-p-j,r-1)} A_3\int_0^\infty\int_x^\infty \frac{x^k(1+x)y^l(1+y)}{(1+\tau+\tau x)(1+\tau+\tau y)}\{F(x)\}^i\{G(x)\}^{r-1-i}\{F(y)-F(x)\}^j$$

$$\times\{G(y)-G(x)\}^{s-r-1-j}\{1-F(y)\}^{n-p-i-j}\{1-G(y)\}^{p-s+i+j+1}\{1-G(x)\}dxdy. \tag{11}$$

Using binomial theorem in equation (11) and subsequently expanding the same we get

$$\mu_{r,s:n}^{(k,l)}[p]=\theta^4\sum_{j=0}^{s-r-1}\sum_{i=\max(0,s-p-j-2)}^{\min(n-p-j-2,r-1)} A_1\sum_{b=0}^{i}\sum_{d=0}^{r-1-i}\sum_{t=0}^{j}\sum_{q=0}^{s-r-1-j}\sum_{p_1=0}^{b+j-t}\sum_{p_2=0}^{d+s-r-j-q-1}\sum_{p_3=0}^{n-p-i-j+t-2}\sum_{p_4=0}^{p-s+i+j+q+2}\sum_{l_1=0}^{p_1+p_2+1}\sum_{l_2=0}^{p_3+p_4+1}$$

$$\times(-1)^{b+d+t+q}\binom{i}{b}\binom{r-1-i}{d}\binom{j}{t}\binom{s-r-1-j}{q}\binom{b+j-t}{p_1}\binom{d+s-r-j-q-1}{p_2}$$

$$\times\binom{n-p-i-j+t-2}{p_3}\binom{p-s+i+j+q+2}{p_4}\binom{p_1+p_2+1}{l_1}\binom{p_3+p_4+1}{l_2}$$

$$\times\frac{\theta^{p_1+p_3}\tau^{p_2+p_4}}{(1+\theta)^{b+n-p-i}(1+\tau)^{d-r+p+i+1}}$$

$$\times\int_0^\infty\int_x^\infty x^{k+l_1}y^{l+l_2}e^{-x\{\theta(b+j-t+1)+\tau(d+s-r-1-j+q)\}}e^{-y\{\theta(n-p-i-j+t-1)+\tau(p-s+i+j+q+2)\}}dxdy$$

$$+\theta^2\tau^2\sum_{j=0}^{s-r-1}\sum_{i=\max(0,s-p-j-1)}^{\min(n-p-j-1,r-1)} A_2\sum_{b=0}^{i}\sum_{d=0}^{r-1-i}\sum_{t=0}^{j}\sum_{q=0}^{s-r-1-j}\sum_{p_1=0}^{b+j-t}\sum_{p_2=0}^{d+s-r-j-q-1}\sum_{p_3=0}^{n-p-i-j+t-1}\sum_{p_4=0}^{p-s+i+j+q+1}\sum_{l_1=0}^{p_1+p_2+1}\sum_{l_2=0}^{p_3+p_4+1}$$

$$\times(-1)^{b+d+t+q}\binom{i}{b}\binom{r-1-i}{d}\binom{j}{t}\binom{s-r-1-j}{q}\binom{b+j-t}{p_1}\binom{d+s-r-j-q-1}{p_2}$$

$$\times\binom{n-p-i-j+t-1}{p_3}\binom{p-s+i+j+q+1}{p_4}\binom{p_1+p_2+1}{l_1}\binom{p_3+p_4+1}{l_2}$$

$$\times\frac{\theta^{p_1+p_3}\tau^{p_2+p_4}}{(1+\theta)^{b+n-p-i}(1+\tau)^{d-r+p+i+1}}$$

$$\times\int_0^\infty\int_x^\infty x^{k+l_1}y^{l+l_2}e^{-x\{\theta(b+j-t+1)+\tau(d+s-r-1-j+q)\}}e^{-y\{\theta(n-p-i-j+t-1)+\tau(p-s+i+j+q+2)\}}dxdy$$

$$+\theta^2\tau^2\sum_{j=0}^{s-r-1}\sum_{i=\max(0,s-p-j-1)}^{\min(n-p-j-1,r-1)} A_2\sum_{b=0}^{i}\sum_{d=0}^{r-1-i}\sum_{t=0}^{j}\sum_{q=0}^{s-r-1-j}\sum_{p_1=0}^{b+j-t}\sum_{p_2=0}^{d+s-r-j-q-1}\sum_{p_3=0}^{n-p-i-j+t-1}\sum_{p_4=0}^{p-s+i+j+q+1}$$

$$\times\sum_{l_1=0}^{p_1+p_2+1}\sum_{l_2=0}^{p_3+p_4+1}(-1)^{b+d+t+q}\binom{i}{b}\binom{r-1-i}{d}\binom{j}{t}\binom{s-r-1-j}{q}\binom{b+j-t}{p_1}$$

$$\times\binom{d+s-r-j-q-1}{p_2}\binom{n-p-i-j+t-1}{p_3}\binom{p-s+i+j+q+1}{p_4}$$

$$\times\binom{p_1+p_2+1}{l_1}\binom{p_3+p_4+1}{l_2}\frac{\theta^{p_1+p_3}\tau^{p_2+p_4}}{(1+\theta)^{b+n-p-i}(1+\tau)^{d-r+p+i+1}}$$

$$\times\int_0^\infty\int_x^\infty x^{k+l_1}y^{l+l_2}e^{-x\{\theta(b+j-t)+\tau(d+s-r-j+q)\}}e^{-y\{\theta(n-p-i-j+t)+\tau(p-s+i+j+q+1)\}}dxdy$$

$$+\tau^4\sum_{j=0}^{s-r-1}\sum_{i=\max(0,s-p-j)}^{\min(n-p-j,r-1)} A_3\sum_{b=0}^{i}\sum_{d=0}^{r-1-i}\sum_{t=0}^{j}\sum_{q=0}^{s-r-1-j}\sum_{p_1=0}^{b+j-t}\sum_{p_2=0}^{d+s-r-j-q-1}\sum_{p_3=0}^{n-p-i-j+t}\sum_{p_4=0}^{p-s+i+j+q}\sum_{l_1=0}^{p_1+p_2+1}\sum_{l_2=0}^{p_3+p_4+1}$$

$$\times(-1)^{b+d+t+q}\binom{i}{b}\binom{r-1-i}{d}\binom{j}{t}\binom{s-r-1-j}{q}\binom{b+j-t}{p_1}\binom{d+s-r-j-q-1}{p_2}$$
$$\times\binom{n-p-i-j+t}{p_3}\binom{p-s+i+j+q}{p_4}\binom{p_1+p_2+1}{l_1}\binom{p_3+p_4+1}{l_2}$$
$$\times\frac{\theta^{p_1+p_3}\tau^{p_2+p_4}}{(1+\theta)^{b+n-p-i}(1+\tau)^{d-r+p+i+1}}$$
$$\times\int_0^\infty\int_x^\infty x^{k+l_1}y^{l+l_2}e^{-x\{\theta(b+j-t)+\tau(d+s-r-j+q)\}}e^{-y\{\theta(n-p-i-j+t)+\tau(p-s+i+j+q+1)\}}dxdy \qquad (12)$$

After simplifying equation (12) by evaluating the integrals using gamma function we get the required relation in equation (10).

□

Next, we have evaluated the covariance using the product moments in equation (10) with n = 6, $\theta=1$ and $\tau=\dfrac{\theta}{h}$ with p = 0, 1, 2 and tabulated in Table 2 (Annexure). The results in Table 1 can be verified for the case p = 0 by using the well-known identity [see Arnold and Balakrishnan (1989), p.10]

$$\sum_{i=1}^{5}\sum_{j=i+1}^{6}\mu_{i,j:n}=\binom{6}{2}[E(X)]^2.$$

From Table 2, we see that the covariance increases as h increases while it decreases with the increase in p values. For small r and s, the relative change in covariances is more with the increase in number of outliers from p = 1 to p = 2 for all values of h and p.

4. Robustness of the L-Moments

In this section, we discuss the issue of robustness by estimating the bias and mean square error (MSE) of sample L-moments of the population L-moments for the distribution in equation (1) under various choices of n.

According to Hosking (1990), the L-moments are basically linear functions of the data and are more robust than the usual moments when outliers are present in the data. Also, sometimes these estimators produce efficient parameter estimators as compare to maximum likelihood estimates (MLEs).

Using the expression of the first four population L-moments λ_1, λ_2, λ_3 and λ_4 [see Hosking (1990), p. 107] for the distribution in equation (1) and using equation (7) we compute the values of all the first population L-moments for p = 0, 1 and 2 which is given in Table 3 (Annexure).

The expressions of the first four sample L-moments [see Hosking (1990), p. 113] are as follows:

$$l_1=n^{-1}\sum_i x_i, \qquad (13)$$

$$l_2 = \frac{1}{2}\binom{n}{2}^{-1} \sum\sum_{i>j} (x_{i:n} - x_{j:n}), \tag{14}$$

$$l_3 = \frac{1}{3}\binom{n}{3}^{-1} \sum\sum_{i>j>k}\sum (x_{i:n} - 2x_{j:n} + x_{k:n}), \tag{15}$$

and
$$l_4 = \frac{1}{4}\binom{n}{4}^{-1} \sum\sum\sum_{i>j>k>l}\sum (x_{i:n} - 3x_{j:n} + 3x_{k:n} - x_{l:n}). \tag{16}$$

Using equations (14), (15) and (16) and the population L-moments from the Table 3 we have estimated the Bias and MSE of sample L-moments in Table 4, Table 5 and Table 6 (Annexure) for n = 10, 20 and 30 respectively. The random samples are simulated from Lindley distribution for $\theta = 1$ using the LindleyR package in R software. The bias and MSE are computed using R (based on 10,000 runs). The R code is not included but it is available upon request from the author.

From Table 4, Table 5 and Table 6 we see that

1) In general for most of the values of 'p' and 'h' the MSE decreases as 'n' increases.
2) When p = 0, l_4 has the smallest MSE among the three sample L-moments.
3) The values of bias and MSE gradually decrease with the increase in the order of the values of 'p' and 'h' *i.e.* the bias and MSE are having inverse relation with the order of the sample L-moments. Again, the relative change in bias and MSE gradually decreases with the increase in the order of the sample L-moment for different values of 'p'.

5. Special Cases

By substituting p = 0 in equation (7), it reduces to

$$\mu_{r:n}^{(k)} = \frac{\theta^2}{(1+\theta)^{n-r+1}} c_{r:n} \sum_{i=0}^{r-1}\sum_{l=0}^{n-r+i}\sum_{q=0}^{l+1} (-1)^i \frac{\theta^l}{(1+\theta)^i} \binom{r-1}{i}\binom{n-r+i}{l}\binom{l+1}{q} \frac{\Gamma(k+q+1)}{[\theta(n-r+i+1)]^{k+q+1}} \tag{17}$$

where

$$c_{r:n} = \frac{n!}{(r-1)!(n-r)!} \quad .$$

Again, replacing p = 0 in equation (10), we get

$$\mu_{r,s:n}^{(k,l)} = \frac{\theta^4}{(1+\theta)^{n-r+1}} c_{r,s:n} \sum_{b=0}^{r-1}\sum_{j=0}^{s-r-1}\sum_{p_1=0}^{b+s-r-1-j}\sum_{p_3=0}^{n+j-s}\sum_{l_1=0}^{p_1+1}\sum_{l_2=0}^{p_3+1}\sum_{a=0}^{l+l_2} (-1)^{b+j} \frac{\theta^{p_1+p_3}}{(1+\theta)^b}\binom{r-1}{b}\binom{s-r-1}{j}$$
$$\times \binom{b+s-r-j-1}{p_1}\binom{n+j-s}{p_3}\binom{p_1+1}{l_1}\binom{p_3+1}{l_2}\frac{(l+l_2)!}{a!}$$
$$\times \frac{\Gamma(k+l_1+q+1)}{[\theta(n-s+j+1)]^{l+l_2+1}[\theta(b+n-r+1)]^{k+l_1+a+1}} \tag{18}$$

where

$$c_{r,s:n} = \frac{n!}{(r-1)!(s-r-1)!\,(n-s)!}.$$

The results in equations (17) and (18) are the single and product moment of order statistics from one parameter Lindley distribution [see Sultan and AL-Thubyani (2016), p. 3 and p. 4] respectively.

6. Conclusion

In this article, we obtain the explicit form of single and product moments of order statistics from one parameter Lindley distribution when multiple outliers are present in the data. These moment relations are generalized form of the moments of order statistics of the Lindley distribution obtained by Sultan and AL-Thubyani (2016) under the multiple-outliers model set up.

The findings of the present study suggest that both the variances and covariances of order statistics of Lindley distribution decreases with the increase in the number of outliers present in the data. For small values of r (in case of single moment order statistics) and r and s (in case of product moment order statistics) the relative change in variances and covariances is comparatively more *i.e.* the smaller order statistics are more sensitive to the presence of outliers, as one would expect. While for higher values of r the variance is negatively correlated with h (scaling factor); the covariance remains positively correlated with h for all r and s, $r < s$. We also find that the bias and MSE of higher sample L-moments gradually reduced. The robustness feature of the sample L-moments is evident from the fact that the higher order sample L-moments provide more protection against the presence of pronounced outliers as the relative change in bias and MSE is reasonably less with the increase in number of outliers.

Acknowledgements

I am thankful to the anonymous reviewer for his constructive suggestions, which helped to improve the presentation of this article.

References

Arnold, B. C. and Balakrishnan, N. (1989). *Relations, Bounds and Approximations for Order Statistics. Lecture Notes in Statistics*, Vol. 53, Springer-Verlag, New York.

Bakouch, H. S., Al-Zahrani, B. M., Al-Shomarani, A. A., Marchi, V. A. A. and Louzada, F. (2012). An extended Lindley distribution. *Journal of the Korean Statistical Society*, **41**, 75-85.

Balakrishnan, N. (1994a). On order statistics from non-identical exponential random variables and some applications (with discussion). *Computational Statistics and Data Analysis*, **18**, 203-253.

Balakrishnan, N. (1994b). On order statistics from non-identical exponential random variables and some applications. *Communication in Statistics - Theory and Methods*, **23**, 3373-3393.

Balakrishnan, N. (2007). Permanents, order statistics, outliers, and robustness. *Revista Mathematica Complutense*, **20**(**1**), 7-107.

Barnett, V. and Lewis, T. (1994). *Outliers in Statistical Data*. 3rd Edition, John Wiley & Sons, Chichester.

Bhar, L. M., Gupta, V. K. and Prasad, R. (2013). Detection of outliers in designed experiments in presence of masking. *Statistics and Applications*, **11**(**1&2**), 147-160.

Childs, A. (1996). *Advances in Statistical Inference and Outlier Related Issues*. Ph. D. Thesis, Department of Mathematics and Statistics, McMaster University, Hamilton, Ontario.

Ghitany, M. E., Atieh, B. and Nadarajah, S. (2008). Lindley distribution and its application. *Mathematics and Computers in Simulation*, **78**, 493-506.

Gomoz-Deniz E., and Calderin-Ojeda, E. (2011). The discrete Lindley distribution: properties and applications. *Journal of Statistical Computation and Simulation*, **81(11)**, 1405–1416.

Hosking, J. R. M. (1990). L-moments: Analysis and estimation of distributions using linear combinations of order statistics. *Journal of the Royal Statistical Society Series B (Methodological)*, **52**(**1**), 105-124.

Krishna, H. and Kumar, K. (2011). Reliability estimation in Lindley distribution with progressively type II right censored sample. *Mathematics and Computers in Simulation*, **82**(**2**), 281-294.

Kumar, C. S., and Jose, R. (2018). On double Lindley distribution and some of its properties. *American Journal of Mathematical and Management Sciences*, **38**(**1**), 23-43.

Mazucheli, J., Menezes, A. F. B. and Chakraborty, S. (2019). On the one parameter unit-Lindley distribution and its associated regression model for proportion data. *Journal of Applied Statistics*, **46**, 700-714.

Mazuchelia, J. and Achcar, J. A. (2011). The Lindley distribution applied to competing risks lifetime data. *Computer Methods and Programs in Biomedicine*, **104**(**2**), 188-192.

Prasad, R., Nandi, P. K., Bhar, L. M. and Gupta, V. K. (2008). Outliers in multi-response experiments. *Statistics and Applications*, **6**(**1& 2**), 275-292.

Sultan, K. F. and Al-Thubyani, W. S. (2016). Higher order moments of order statistics from the Lindley distribution and associated inference. *Journal of Statistical Computation and Simulation*, **86**(**17**), 3432-3445.

Sultan, K. F. and Moshref, M. E. (2014). Moments of order statistics from Weibull distribution in the presence of multiple outliers. *Communication in Statistics -Theory and Methods*, **43**(**10-12**), 2214-2226.

ANNEXURE

Table 1: The means and variances in the presence of multiple outliers when n = 6

r	p	Mean	Var	Mean	Var	Mean	Var	Mean	Var	Mean	Var
1	0	0.2997	0.0773								
2		0.6292	0.1637								
3		1.0116	0.2774								
4		1.4885	0.4568								
5		2.1597	0.8255								
6		3.4110	2.1743								
		h = 0.1		h = 0.2		h = 0.3		h = 0.4		h = 0.5	
1	1	0.0851	0.0070	0.1448	0.0198	0.1871	0.0322	0.2179	0.0428	0.2409	0.0516
2		0.3740	0.0993	0.4132	0.0962	0.4549	0.1011	0.4930	0.1103	0.5260	0.1207
3		0.7605	0.2303	0.7781	0.2209	0.8066	0.2155	0.8402	0.2169	0.8744	0.2234
4		1.2507	0.4228	1.2567	0.4158	1.2721	0.4055	1.2961	0.3984	1.3259	0.3973
5		1.9360	0.8025	1.9376	0.7992	1.9439	0.7900	1.9575	0.7775	1.9789	0.7670
6		3.2025	2.1669	3.2027	2.1660	3.2043	2.1608	3.2093	2.1482	3.2202	2.1289
1	2	0.0492	0.0023	0.0949	0.0087	0.1355	0.0173	0.1708	0.0270	0.2012	0.0367
2		0.1338	0.0096	0.2376	0.0286	0.3196	0.0491	0.3864	0.0688	0.4425	0.0873
3		0.4645	0.1431	0.5297	0.1320	0.6050	0.1363	0.6791	0.1507	0.7481	0.1699
4		0.9590	0.3618	0.9848	0.3410	1.0322	0.3230	1.0932	0.3178	1.1606	0.3254
5		1.6622	0.7626	1.6691	0.7500	1.6902	0.7249	1.7281	0.6990	1.7811	0.6832
6		2.9492	2.1505	2.9502	2.1465	2.9557	2.1296	2.9706	2.0958	2.9995	2.0508

* *The results remain same for all values of h when p* = 0.

Table 2: The covariances in the presence of multiple outliers when n = 6

p = 0			p = 1					p = 2				
r	s	h^*	h = 0.1	h = 0.2	h = 0.3	h = 0.4	h = 0.5	h = 0.1	h =0.2	h = 0.3	h = 0.4	h = 0.5
1	2	0.0575	0.0199	0.0278	0.0331	0.0377	0.0418	0.0035	0.0101	0.0172	0.0240	0.0303
1	3	0.2005	0.0564	0.0877	0.1091	0.1266	0.1420	0.0194	0.0386	0.0590	0.0800	0.1012
2	3	0.1483	0.1006	0.1043	0.1074	0.1112	0.1162	0.0386	0.0540	0.0649	0.0756	0.0868
1	4	0.8749	0.1941	0.3036	0.3741	0.4308	0.4835	0.0465	0.0941	0.1473	0.2062	0.2695
2	4	0.6203	0.3654	0.3910	0.4159	0.4413	0.4685	0.1159	0.1887	0.2426	0.2916	0.3411
3	4	0.3949	0.3102	0.3120	0.3136	0.3165	0.3224	0.1996	0.2111	0.2208	0.2327	0.2490
1	5	4.7651	0.9196	1.4366	1.7330	1.9385	2.1201	0.1625	0.3160	0.4725	0.6448	0.8379
2	5	3.0990	1.6620	1.7596	1.8437	1.9184	1.9970	0.4198	0.6883	0.8702	1.0230	1.1769
3	5	1.9092	1.3646	1.3849	1.4162	1.4560	1.5054	0.7706	0.8494	0.9366	1.0291	1.1322
4	5	1.3947	1.1475	1.1493	1.1524	1.1577	1.1687	0.8573	0.8675	0.8815	0.9011	0.9323
1	6	34.584	6.1630	9.6334	11.450	12.436	13.080	0.8723	1.6678	2.3711	3.0249	3.6854
2	6	19.979	10.217	10.755	11.149	11.381	11.536	2.2758	3.7227	4.5735	5.1174	5.5537
3	6	11.168	7.5911	7.6829	7.8078	7.9318	8.0554	3.9772	4.3269	4.6847	5.0045	5.3116
4	6	7.7404	6.0517	6.0717	6.1271	6.2207	6.3521	4.2195	4.3019	4.4670	4.6992	4.9895
5	6	8.1472	6.9541	6.9568	6.9680	6.9931	7.0394	5.6128	5.6253	5.6636	5.7365	5.8549

* *The results remain same for all values of h when p* = 0.

Table 3: First four Population L-Moments

		λ_1	λ_2	λ_3	λ_4
$p = 0$	h*	1.5000	0.6875	0.1921	0.0978
$p = 1$	$h = 0.1$	0.1090	0.7013	0.2133	0.0856
	$h = 0.2$	0.2333	0.6591	0.2243	0.0862
	$h = 0.3$	0.3692	0.6268	0.2240	0.0911
	$h = 0.4$	0.5142	0.6056	0.2172	0.0951
	$h = 0.5$	0.6666	0.5956	0.2078	0.0971
$p = 2$	$h = 0.1$		0.0543	0.4180	0.0213
	$h = 0.2$		0.1152	0.3448	0.0716
	$h = 0.3$		0.1806	0.2852	0.1741
	$h = 0.4$		0.2489	0.2401	0.1027
	$h = 0.5$		0.3194	0.2084	0.1018

Table 4: Bias and MSE of sample L-moments l_2, l_3 and l_4 for *n* = 10

		l_2		l_3		l_4	
p	h	Bias	MSE	Bias	MSE	Bias	MSE
0	h*	0.0027	0.0528	0.0018	0.0252	0.0018	0.0187
1	0.1	0.0099	0.0533	0.0135	0.0251	0.0069	0.0189
1	0.2	0.0242	0.0550	0.0219	0.0267	0.0086	0.0193
1	0.3	0.0503	0.0566	0.0209	0.0254	0.0043	0.0186
1	0.4	0.0652	0.0570	0.0197	0.0251	0.0025	0.0185
1	0.5	0.0738	0.0571	0.0101	0.0243	0.0001	0.0182
2	0.1	0.6129	0.4279	0.1982	0.0637	0.0674	0.0228
2	0.2	0.5510	0.3575	0.1272	0.0412	0.0157	0.0190
2	0.3	0.4696	0.2725	0.0724	0.0299	0.0799	0.0248
2	0.4	0.3982	0.2101	0.0349	0.0257	0.0081	0.0181
2	0.5	0.3288	0.1591	0.0068	0.0250	0.0035	0.0250

* *The results remain same for all values of h when p* = 0.

Table 5: Bias and MSE of sample L-moments l_2, l_3 and l_4 for n = 20

		l_2		l_3		l_4	
p	*h*	Bias	MSE	Bias	MSE	Bias	MSE
0	*h**	0.0015	0.0254	0.0012	0.0107	0.0013	0.0040
1	0.1	0.0115	0.6898	0.0171	0.1965	0.0098	0.0041
1	0.2	0.0269	0.0266	0.0250	0.0115	0.0114	0.0042
1	0.3	0.0542	0.0282	0.0282	0.0113	0.0050	0.0041
1	0.4	0.0749	0.0309	0.0206	0.0114	0.0033	0.0041
1	0.5	0.0838	0.0325	0.0126	0.0109	0.0012	0.0040
2	0.1	0.6329	0.4264	0.2164	0.0580	0.0702	0.0089
2	0.2	0.5646	0.3443	0.1414	0.0310	0.0212	0.0044
2	0.3	0.4924	0.2679	0.0840	0.0178	0.0788	0.0101
2	0.4	0.4191	0.2011	0.0408	0.0122	0.0054	0.0041
2	0.5	0.3494	0.1461	0.0107	0.0106	0.0028	0.0041

** The results remain same for all values of h when p* = 0.

Table 6: Bias and MSE of sample L-moments l_2, l_3 and l_4 for n = 30

		l_2		l_3		l_4	
p	*h*	Bias	MSE	Bias	MSE	Bias	MSE
0	*h**	0.0006	0.0167	0.0009	0.0067	0.0005	0.0040
1	0.1	0.0147	0.0168	0.0197	0.0071	0.0106	0.0041
1	0.2	0.0258	0.0171	0.0288	0.0076	0.0107	0.0042
1	0.3	0.0552	0.0193	0.0294	0.0078	0.0054	0.0041
1	0.4	0.0766	0.0223	0.0230	0.0074	0.0024	0.0040
1	0.5	0.0851	0.0238	0.0139	0.0069	0.0011	0.0039
2	0.1	0.6345	0.4192	0.2202	0.0553	0.0723	0.0091
2	0.2	0.5647	0.3356	0.1467	0.0283	0.0227	0.0045
2	0.3	0.4954	0.2621	0.0881	0.0144	0.0789	0.0101
2	0.4	0.4261	0.1981	0.0425	0.0088	0.0947	0.0130
2	0.5	0.3558	0.1432	0.0127	0.0070	0.0031	0.0041

** The results remain same for all values of h when p* = 0.