STATISTICA, anno LXXIV, n. 3, 2014

ESTIMATION OF THE LOCATION PARAMETER OF DISTRIBUTIONS WITH KNOWN COEFFICIENT OF VARIATION BY RECORD VALUES.

N. K. Sajeevkumar [1]

Department of Statistics, Govt. College Kariavattom, Trivandrum-695581, Kerala, India

M. R. Irshad

Department of Statistics, Govt. College Kariavattom, Trivandrum-695581, Kerala, India

1. Introduction

Estimation of population parameters are considered by several statisticians when prior information such as coefficient of variation, kurtosis or skewness is known. The use of prior information in inference is well established in the Bayesian arena of statistical methodology. In some instances prior information can be incorporated in classical models as well. Searls and Intarapanich (1990) derived an estimator of variance when the kurtosis of the sampled population is known. In some of the problems of biological and physical sciences, situations where the scale parameter is proportional to the location parameter are seen reported in the literature, then knowing the proportionality constant is equivalent to knowing the population coefficient of variation.

Situations where coefficient of variation is known, do occur in practice. In clinical chemistry batches of some substance (chemicals) are to be analysed, after sufficient batches of the substances are analyzed, their coefficient of variation will be known. In biological experiments, it is customary to conduct multi-locational trials. When the results of a few centres are available, the coefficient of variation is known and it can be used for inferential purpose of an experiment to be conducted in a new location. In environmental studies the mean represents the average concentration level of a particular chemical or pollutant and the standard deviation is directly proportional to the mean, in this case coefficient of variation is known either from past studies or physical characteristics of the environmental setup (see Guo and Pal, 2003).

For all these situations the interest for statisticians comes from the fact that, the inferential procedures instead of getting simplified becomes more complex. The property of completeness enjoyed by the statistics is no longer holds in this situation, therefore the standard theory of uniformly minimum variance unbiased

[1] Corresponding Author. E-mail: sajeevkumarnk@gmail.com

estimation (UMVUE) is not applicable in this case. For details see, Kunte (2000), Guo and Pal (2003), Bhat and Rao (2011), Hedayat et al. (2011), Kagan and Malinovsky (2013), Fu et al. (2013) and Khan (2013) and the references therein. In the case of location scale family of distributions, order statistics play a key role in the development of efficient estimators. Thomas and Sajeevkumar (2003) considered the problem of estimation of the mean of normal distribution with known coefficient of variation using order statistics. Also Sajeevkumar and Thomas (2005) discussed about the estimation of the mean of the logistic distribution with known coefficient of variation by order statistics. Recently Sajeevkumar and Irshad (2011, 2013a) discussed about the estimation of the location parameter of the exponential distribution with known coefficient of variation by order statistics. Estimation of the mean of double exponential and normal distributions with known coefficient of variation by U-statistics are respectively discussed by Sajeevkumar and Irshad (2012) and Sajeevkumar and Irshad (2013b).

The concept of record values introduced by Chandler (1952) has gained momentum in a theoretical perspective as well as in terms of its applications. Also the study of record values in many ways parallels the study of order statistics, indeed they are inextricably related. This motivates the authors to take up the study based on record values. Record values and associated statistics are of great importance in several real life problems involving weather, economic and sports data. The statistical study of record values started with Chandler (1952) and has now spread in different directions. Resnick (1973) and Shorrock (1973) documented the asymptotic theory of records. Glick (1978) provides a survey of the literature on records. For a detailed discussion on the developments in the theory and applications of record values, see Arnold et al. (1998) and Ahsanullah (1995).

Let $X_1, X_2, ...$ be a sequence of independent observations arising from a population. An observation X_j will be called an upper record value (or simply a record) if its value exceeds that of all previous observations. Thus X_j is a record if $X_j > X_i$, $\forall\ i < j$. The first observation X_1 is taken as the initial record R_1. The next record R_2 is the observation following R_1 which is greater than R_1 and so on. The records $R_1, R_2, \cdots$ as defined above are sometimes referred to as the sequence of upper records. An analogous definition deals with lower record values. But we are not of interest in the present paper and hence whenever we use the term record values in this work it means upper record values.

Let $X_1, X_2, \cdots$ be a sequence of independent observations arising from a population with absolutely continuous cumulative distribution function (cdf) $F_X(x)$ and pdf $f_X(x)$. If we write R_n to denote the n^{th} upper record value, then its pdf is given by (see Arnold et al. 1998, p.10),

$$f_{R_n}(x) = \frac{1}{(n-1)!}[-\log\{1 - F_X(x)\}]^{n-1} f_X(x). \tag{1}$$

The joint pdf of the m^{th} and n^{th} upper record values R_m and R_n for $m < n$ is given by (see Arnold et al. 1998, p.11),

$$f_{R_{m,n}}(x_1, x_2) = \frac{1}{(m-1)!(n-m-1)!}[-\log\{1 - F_X(x_1)\}]^{m-1} \frac{f_X(x_1)}{1 - F_X(x_1)}$$

$$\times[-\log\{1 - F_X(x_2)\} + \log\{1 - F_X(x_1)\}]^{n-m-1} f_X(x_2), x_1 < x_2. \quad (2)$$

The rest of the paper is organized as follows. In section 2, we discuss the general technique of estimating the location parameter of a distribution with known coefficient of variation by record values. Results of section 2, are directly applied in normal, logistic and exponential distributions are discussed in sections 3, 4 and 5 respectively. Real-life data have been used to illustrate the application of the results obtained and is given in section 6.

2. Estimation of the location parameter when the scale parameter is proportional to the location parameter using record values

In this section we consider the family $\mathcal{G}$ of all absolutely continuous distributions which depend on a location parameter μ and a scale parameter σ, such that $\sigma = c\mu$. Throughout this work we assume that c is known. Then any distribution belongs to $\mathcal{G}$ has a pdf of the form,

$$f(x; \mu, c\mu) = \frac{1}{c\mu} f_0\left(\frac{x-\mu}{c\mu}\right), \quad \mu > 0, c > 0, x \in R. \quad (3)$$

If in the pdf defined in (3), μ is the mean and $c\mu$ is the standard deviation, then c is the known coefficient of variation. In real life situations, coefficient of variation is meaningful, only when it is positive. In standard case our pdf defined in (3), reduces to the usual standard form of location scale family of distributions.

Let $X_{U(1)}, X_{U(2)}, \cdots, X_{U(n)}$ be the first n upper record values arising from (3) and let $Y_{U(1)}, Y_{U(2)}, \cdots, Y_{U(n)}$ be the corresponding upper record values arising from the standard form of the pdf defined in (3). Let us now denote $E(Y_{U(n)})$ by α_n, $Var(Y_{U(n)})$ by $\beta_{n,n}$, $E(Y_{U(m)}Y_{U(n)})$ by $\alpha_{m,n}$, $Cov(Y_{U(m)}, Y_{U(n)})$ by $\beta_{m,n}$ and $E(Y^2_{U(n)})$ by α^2_n. Let $\mathbf{X}_{(U)} = (X_{U(1)}, X_{U(2)}, \cdots, X_{U(n)})'$ be the vector of first n upper record values arising from (3) and let $\mathbf{Y}_{(U)} = (Y_{U(1)}, Y_{U(2)}, \cdots, Y_{U(n)})'$ be the corresponding vector of first n upper record values arising from the standard form of (3). Now we derive the best linear unbiased estimator (BLUE) of μ involved in (3), and is given by the following theorem.

Theorem 1. *Suppose $X_{U(1)}, X_{U(2)}, \cdots, X_{U(n)}$ are the first n upper record values arising from (3) . Let $Y_{U(1)}, Y_{U(2)}, \cdots, Y_{U(n)}$ be the corresponding upper record values arising from the standard form of the distribution defined in (3). Let $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \cdots, \alpha_n)'$, $\boldsymbol{B} = ((\beta_{i,j}))$, $1 \le i \le j \le n$, be the vector of means and dispersion matrix respectively of $\boldsymbol{Y}_{(U)} = (Y_{U(1)}, Y_{U(2)}, \cdots, Y_{U(n)})'$. Then the BLUE $\tilde{\mu_1}$ of the parameter μ is given by,*

$$\tilde{\mu_1} = \frac{(c\boldsymbol{\alpha} + \mathbf{1})'\mathbf{B}^{-1}}{(c\boldsymbol{\alpha} + \mathbf{1})'\mathbf{B}^{-1}(c\boldsymbol{\alpha} + \mathbf{1})}\mathbf{X}_{(u)} \quad (4)$$

and

$$Var(\tilde{\mu_1}) = \frac{c^2\mu^2}{(c\boldsymbol{\alpha} + \mathbf{1})'\mathbf{B}^{-1}(c\boldsymbol{\alpha} + \mathbf{1})}, \quad (5)$$

where $\mathbf{1} = (1, 1, \cdots, 1)'$.

PROOF. Given $X_{U(1)}, X_{U(2)}, \cdots, X_{U(n)}$ are first n upper record values arising from (3).

$$\text{Let} \quad E(Y_{U(i)}) = \alpha_i, \quad i = 1, 2, \cdots, n,$$

$$Var(Y_{U(i)}) = \beta_{i,i}, \quad i = 1, 2, \cdots, n$$

and

$$Cov(Y_{U(i)}, Y_{U(j)}) = \beta_{i,j}, \quad 1 \leq i \leq j \leq n.$$

Then we have,

$$\frac{X_{U(i)} - \mu}{c\mu} \stackrel{d}{=} Y_{U(i)}, \quad i = 1, 2, \cdots, n.$$

Therefore,

$$E(X_{U(i)}) = (c\alpha_i + 1)\mu, \qquad i = 1, 2, \cdots, n, \tag{6}$$

$$Var(X_{U(i)}) = c^2\mu^2\beta_{i,i} \tag{7}$$

and

$$Cov(X_{U(i)}, X_{U(j)}) = c^2\mu^2\beta_{i,j}. \tag{8}$$

Using (6) to (8) one can also write,

$$E(\mathbf{X}_{(U)}) = (c\boldsymbol{\alpha} + \mathbf{1})\mu \tag{9}$$

and

$$D(\mathbf{X}_{(U)}) = \mathbf{B}c^2\mu^2. \tag{10}$$

Then by generalized Gauss-Markov setup, the BLUE $\tilde{\mu_1}$ of the parameter μ is given by,

$$\tilde{\mu_1} = \frac{(c\boldsymbol{\alpha} + \mathbf{1})'\mathbf{B}^{-1}}{(c\boldsymbol{\alpha} + \mathbf{1})'\mathbf{B}^{-1}(c\boldsymbol{\alpha} + \mathbf{1})}\mathbf{X}_{(U)}$$

and

$$Var(\tilde{\mu_1}) = \frac{c^2\mu^2}{(c\boldsymbol{\alpha} + \mathbf{1})'\mathbf{B}^{-1}(c\boldsymbol{\alpha} + \mathbf{1})}.$$

Thus the theorem is proved.

Clearly $\tilde{\mu_1}$ can be written as a linear function of $X_{U(i)}$ as,

$$\tilde{\mu_1} = \sum_{i=1}^{n} a_i X_{U(i)}, \text{where } a_i, i = 1, 2, \cdots, n \text{ are constants.}$$

For comparison purpose, we take two other linear unbiased estimators of μ. For that considering μ as the location parameter of (3), a linear unbiased estimator of μ based on the upper record values is given by (same argument that of David 1981, pp: 128-131 and Balakrishnan and Cohen 1991, pp: 80-82),

$$\hat{\mu_1} = -\frac{\boldsymbol{\alpha}'\mathbf{B}^{-1}(\mathbf{1}\boldsymbol{\alpha}' - \boldsymbol{\alpha}\mathbf{1}')\mathbf{B}^{-1}}{(\boldsymbol{\alpha}'\mathbf{B}^{-1}\boldsymbol{\alpha})(\mathbf{1}'\mathbf{B}^{-1}\mathbf{1}) - (\boldsymbol{\alpha}'\mathbf{B}^{-1}\mathbf{1})^2}\mathbf{X}_{(U)} \tag{11}$$

and

$$Var(\hat{\mu_1}) = \frac{(\boldsymbol{\alpha}'\mathbf{B}^{-1}\boldsymbol{\alpha})c^2\mu^2}{(\boldsymbol{\alpha}'\mathbf{B}^{-1}\boldsymbol{\alpha})(\mathbf{1}'\mathbf{B}^{-1}\mathbf{1}) - (\boldsymbol{\alpha}'\mathbf{B}^{-1}\mathbf{1})^2}. \tag{12}$$

Clearly the estimator considered in (11) is independent of c. But its variance containing c.

Also by considering $c\mu$ as the scale parameter of (3), a linear unbiased estimator of $c\mu$ based on upper record values is given by (same argument that of David 1981, pp: 128-131 and Balakrishnan and Cohen 1991, pp: 80-82),

$$T_1 = \frac{\mathbf{1}'\mathbf{B}^{-1}(\mathbf{1}\boldsymbol{\alpha}' - \boldsymbol{\alpha}\mathbf{1}')\mathbf{B}^{-1}}{(\boldsymbol{\alpha}'\mathbf{B}^{-1}\boldsymbol{\alpha})(\mathbf{1}'\mathbf{B}^{-1}\mathbf{1}) - (\boldsymbol{\alpha}'\mathbf{B}^{-1}\mathbf{1})^2}\mathbf{X}_{(U)} \tag{13}$$

and

$$Var(T_1) = \frac{(\mathbf{1}'\mathbf{B}^{-1}\mathbf{1})c^2\mu^2}{(\boldsymbol{\alpha}'\mathbf{B}^{-1}\boldsymbol{\alpha})(\mathbf{1}'\mathbf{B}^{-1}\mathbf{1}) - (\boldsymbol{\alpha}'\mathbf{B}^{-1}\mathbf{1})^2}. \tag{14}$$

From (14) we can obtain another linear unbiased estimator μ_1^* of μ, and is given by,

$$\mu_1^* = \frac{\mathbf{1}'\mathbf{B}^{-1}(\mathbf{1}\boldsymbol{\alpha}' - \boldsymbol{\alpha}\mathbf{1}')\mathbf{B}^{-1}}{c\left[(\boldsymbol{\alpha}'\mathbf{B}^{-1}\boldsymbol{\alpha})(\mathbf{1}'\mathbf{B}^{-1}\mathbf{1}) - (\boldsymbol{\alpha}'\mathbf{B}^{-1}\mathbf{1})^2\right]}\mathbf{X}_{(U)} \tag{15}$$

and

$$Var(\mu_1^*) = \frac{(\mathbf{1}'\mathbf{B}^{-1}\mathbf{1})\mu^2}{(\boldsymbol{\alpha}'\mathbf{B}^{-1}\boldsymbol{\alpha})(\mathbf{1}'\mathbf{B}^{-1}\mathbf{1}) - (\boldsymbol{\alpha}'\mathbf{B}^{-1}\mathbf{1})^2}. \tag{16}$$

The results derived in this section are directly applied in normal, logistic and exponential distributions are discussed in sections 3, 4 and 5 respectively.

3. Estimation of the mean of the normal distribution with known coefficient of variation by using record values

A continuous random variable X is said to have the normal distribution with location parameter μ and scale parameter $c\mu$, if its pdf is given by,

$$f(x;\mu,c\mu) = \frac{1}{c\mu\sqrt{2\pi}}\exp\left\{\frac{-(x-\mu)^2}{2c^2\mu^2}\right\},\ \mu > 0, c > 0, x \in R. \tag{17}$$

We will write $N(\mu, c\mu)$ to denote the normal distribution defined in (17). The mean and variance of the above distribution are given by $E(X) = \mu$ and $Var(X) = c^2\mu^2$, where c is the known coefficient of variation. The above mentioned normal model with known coefficient of variation is useful in environmental studies where μ represents mean concentration level of a particular chemical or pollutant (in air or water), and the standard deviation $c\mu$ is directly proportional to mean concentration level with the proportionality constant c being known from either past studies or physical characteristics of the environmental setup (see Guo and Pal 2003). For other applications of this distribution see also Gleser and Healy (1976).

The record values and associated inference arising from the two parameter normal distribution $N(\mu, \sigma)$ are discussed by Balakrishnan and Chan (1998). The

TABLE 1
Coefficients of $X_{U(i)}$ in the BLUE $\tilde{\mu_1}$ for n=2(1)10

c	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9	a_{10}
0.2	0.09	0.76								
	0.08	0.06	0.66							
	0.06	0.05	0.04	0.59						
	0.05	0.04	0.03	0.03	0.56					
	0.05	0.04	0.03	0.02	0.02	0.53				
	0.04	0.03	0.02	0.02	0.02	0.02	0.50			
	0.04	0.03	0.02	0.02	0.01	0.01	0.01	0.49		
	0.04	0.03	0.02	0.02	0.01	0.01	0.01	0.01	0.47	
	0.03	0.03	0.02	0.02	0.01	0.01	0.01	0.01	0.01	0.46
0.4	-0.03	0.76								
	-0.02	0.02	0.62							
	-0.01	0.01	0.01	0.54						
	-0.01	0.01	0.01	0.01	0.49					
	-0.01	0.01	0.01	0.01	0.01	0.46				
	-0.01	0.01	0.01	0.01	0.01	0.01	0.43			
	-0.01	0.01	0.01	0.01	0.01	0.01	0	0.40		
	-0.01	0.01	0.01	0.01	0	0	0	0	0.38	
	-0.01	0.01	0.01	0.01	0	0	0	0	0	0.37

means, variances and convariances of the record values arising from the standard form of $N(\mu, \sigma)$ are also available in this paper. Using the results of Balakrishnan and Chan (1998) and also using the results based on upper record values given in section 2, we have evaluated the coefficients of $X_{U(i)}$ in the BLUE $\tilde{\mu_1}$, $Var(\hat{\mu}_1)$ defined in (12), $Var(\mu_1^*)$ defined in (16), $Var(\tilde{\mu_1})$ defined in (5) for the parameter μ involved in (17) with the help of MATHCAD and MATHEMATICA software's. We have also evaluated the numerical values of the efficiency $e_1 = \frac{Var(\hat{\mu}_1)}{Var(\tilde{\mu_1})}$ of $\tilde{\mu_1}$ relative to $\hat{\mu_1}$ and the numerical values of the efficiency $e_2 = \frac{Var(\mu_1^*)}{Var(\tilde{\mu_1})}$ of $\tilde{\mu_1}$ relative to μ_1^* for $c = 0.20, 0.40$ and $n = 2(1)10$ and are given in *TABLE 2*. From this table, it may be noted that in all the cases our estimator $\tilde{\mu_1}$ is much better than that of estimators $\hat{\mu_1}$ and μ_1^*.

4. Estimation of the mean of the logistic distribution with known coefficient of variation by using record values

The logistic distribution is a well-known and widely used statistical distribution because of its simplicity and its historical importance as a growth curve (see Erkelens 1968). It has several important applications in biological, actuarial, industrial and engineering fields. Some applications of order statistics from the logistic distribution in the fields of life-testing and reliability studies have been mentioned by Lawless (1982). Also logistic distribution has a shape similar to that of normal distribution, which makes it simpler and also profitable on suitable occasions to replace the normal distribution by the logistic distribution to simplify the anal-

TABLE 2

$V_1 = \frac{Var(\hat{\mu}_1)}{\mu^2}$, $V_2 = \frac{Var(\mu_1^*)}{\mu^2}$, $V_3 = \frac{Var(\tilde{\mu_1})}{\mu^2}$, *the efficiency* e_1 *of* $\tilde{\mu_1}$ *relative to* $\hat{\mu}_1$ *and the efficiency* e_2 *of* $\tilde{\mu_1}$ *relative to* μ_1^*

c	n	V_1	V_2	V_3	e_1	e_2
0.20	2	0.04000	0.72165	0.02216	1.80505	32.56543
	3	0.03897	0.34485	0.01635	2.38349	21.09174
	4	0.03790	0.22274	0.01332	2.84535	16.72222
	5	0.03692	0.16305	0.01142	3.23292	14.27758
	6	0.03604	0.12793	0.01008	3.57539	12.69147
	7	0.03526	0.10486	0.00909	3.87899	11.53575
	8	0.03457	0.08862	0.00830	4.16506	10.67711
	9	0.03394	0.07662	0.00768	4.41927	9.97656
	10	0.03337	0.06738	0.00715	4.66713	9.42378
0.40	2	0.16000	0.72165	0.06726	2.37883	10.72926
	3	0.15589	0.34485	0.04388	3.55264	7.85893
	4	0.15160	0.22274	0.03308	4.58283	6.73337
	5	0.14767	0.16305	0.02679	5.51213	6.08623
	6	0.14418	0.12793	0.22650	6.36556	5.64812
	7	0.14106	0.10486	0.01969	7.16404	5.32555
	8	0.13826	0.08862	0.01747	7.91414	5.07270
	9	0.13576	0.07662	0.01573	8.63064	4.87095
	10	0.13348	0.06738	0.01433	9.31472	4.70133

ysis without too great discrepancies in the respective theories. Sajeevkumar and Thomas (2005) considered the problem of estimation with known coefficient of variation in logistic model using order statistics. Hence in this section we consider the problem of estimation of the mean of logistic distribution with known coefficient of variation by record values.

A continuous random variable X is said to have the logistic distribution with location parameter μ and scale parameter $c\mu$, if its pdf is given by (see Sajeevkumar and Thomas, 2005),

$$f(x;\mu,c\mu) = \frac{\pi}{\sqrt{3}} \frac{\exp\left\{\frac{-\pi}{\sqrt{3}}\left(\frac{x-\mu}{c\mu}\right)\right\}}{c\mu\left[1+\exp\left\{\frac{-\pi}{\sqrt{3}}\left(\frac{x-\mu}{c\mu}\right)\right\}\right]^2},\ \mu > 0, c > 0, x \in R. \tag{18}$$

We will write $LD(\mu, c\mu)$ to denote the logistic distribution defined in (18). The mean and variance of the above distribution are given by $E(X) = \mu$ and $Var(X) = c^2\mu^2$, where c is the known coefficient of variation. Applications of record values arising from usual two parameter logistic distribution are elucidated by (see Arnold et al. 1995, pp: 65-68). Using those results, first we have evaluated the means, variances and covariances of the record values arising from the standard form of the distribution defined in (18) with the help of MATHCAD and MATHEMATICA software's. Using those values and also using the results based on record values given in section 2, we have evaluated the coefficients of $X_{U(i)}$ in the BLUE $\tilde{\mu_1}$, $Var(\hat{\mu}_1)$ defined in (12), $Var(\mu_1^*)$ defined in (16), $Var(\tilde{\mu_1})$ defined in (5) for the

TABLE 3
Coefficients of $X_{U(i)}$ in the BLUE $\tilde{\mu_1}$ for n=2(1)10

c	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9	a_{10}
0.2	0.06	0.80								
	0.13	0.24	0.45							
	0.13	0.22	0.11	0.31						
	0.13	0.21	0.11	0.06	0.25					
	0.13	0.21	0.11	0.06	0.03	0.21				
	0.13	0.20	0.11	0.06	0.03	0.01	0.19			
	0.13	0.19	0.10	0.06	0.03	0.01	0.01	0.18		
	0.13	0.19	0.10	0.05	0.03	0.01	0.01	0	0.18	
	0.12	0.19	0.09	0.05	0.03	0.01	0.01	0	0	0.17
0.4	-0.14	0.84								
	-0.03	0.21	0.46							
	-0.02	0.19	0.07	0.34						
	-0.01	0.17	0.07	0.04	0.28					
	-0.01	0.16	0.07	0.04	0.02	0.24				
	0	0.15	0.06	0.04	0.02	0.01	0.22			
	0	0.14	0.06	0.04	0.02	0.01	0.01	0.21		
	0	0.14	0.06	0.03	0.02	0.01	0	0	0.20	
	0	0.13	0.05	0.03	0.02	0.01	0	0	0	0.19

parameter μ involved in (18) with the help of MATHCAD and MATHEMATICA software's. Also we have evaluated the numerical values of the efficiency $e_1 = \frac{Var(\hat{\mu}_1)}{Var(\tilde{\mu_1})}$ of $\tilde{\mu_1}$ relative to $\hat{\mu_1}$ and the numerical values of the efficiency $e_2 = \frac{Var(\mu_1^*)}{Var(\tilde{\mu_1})}$ of $\tilde{\mu_1}$ relative to μ_1^* for $c = 0.20, 0.40$ and $n = 2(1)10$ and are given in *TABLE 4*. From this table, it may be noted that in all the cases our estimator $\tilde{\mu_1}$ is much better than that of estimators $\hat{\mu_1}$ and μ_1^*.

5. Estimation of the location parameter of the exponential distribution with known coefficient of variation by using record values

A distribution occupying a commanding position, especially in life-testing problems, is the two-parameter exponential distribution $E(\mu, \sigma), \sigma > 0$. In the context of life-length studies, the location parameter μ and the scale parameter σ respectively represent the minimum guaranteed life and the average excess life of an equipment or a system. The parameters μ and σ are functionally independent and the statistical inference about these parameters make use of the existence of complete minimal sufficient statistics. This brings about a substantial simplification in the inferential problems. There, however, exist situations where the average life σ depends on the minimum guaranteed life μ and the functionally independent nature of the parameters no longer hold, resulting in the loss of optimal properties of the statistics. For instance, in testing for time to breakdown of an insulating fluid under high test voltages, it is observed that high voltages quickly

TABLE 4

$V_1 = \frac{Var(\hat{\mu}_1)}{\mu^2}$,$V_2 = \frac{Var(\mu_1^*)}{\mu^2}$, $V_3 = \frac{Var(\tilde{\mu_1})}{\mu^2}$, *the efficiency* e_1 *of* $\tilde{\mu_1}$ *relative to* $\hat{\mu}_1$ *and the efficiency* e_2 *of* $\tilde{\mu_1}$ *relative to* μ_1^*

c	n	V_1	V_2	V_3	e_1	e_2
0.20	2	0.04000	0.51480	0.02598	1.53965	19.81524
	3	0.03984	0.35423	0.02298	1.73368	15.41471
	4	0.03894	0.25903	0.02163	1.80028	11.97550
	5	0.03796	0.20374	0.02084	1.82150	9.77639
	6	0.03712	0.16816	0.02027	1.83128	8.29600
	7	0.03644	0.14341	0.01980	1.84040	7.24293
	8	0.03590	0.12515	0.01938	1.85243	6.45769
	9	0.03547	0.11111	0.01900	1.86684	5.84789
	10	0.03512	0.09994	0.01864	1.88412	5.36159
0.40	2	0.16000	0.51480	0.07701	2.07765	6.68485
	3	0.15936	0.35423	0.06348	2.51040	5.58018
	4	0.15576	0.25903	0.05666	2.74903	4.57166
	5	0.15184	0.20374	0.05242	2.89660	3.88668
	6	0.14848	0.16816	0.04929	3.01238	3.41165
	7	0.14577	0.14341	0.04673	3.11941	3.06891
	8	0.14362	0.12515	0.04453	3.22524	2.81046
	9	0.14188	0.11111	0.04258	3.33208	2.60944
	10	0.14047	0.09994	0.04082	3.44121	2.44831

yield breakdown data and the linear relationship between the minimum time to breakdown and the average excess time to breakdown seem to hold (see Nelson 1990, p.129).

In this situation, the two-parameter model reduces to a one-parameter model $E(\mu, a\mu)$, where a is known and positive. The interest for theoretical statisticians comes from the fact that in the reduced model $E(\mu, a\mu)$, the inferential procedures instead of getting simplified becomes more complex. The property of completeness enjoyed by the statistics in the case of $E(\mu, \sigma)$ distribution no longer holds. Hence it is not possible to obtain the unique minimum variance unbiased estimator by the use of Rao-Blackwell theorem. For some early works based on $E(\mu, a\mu)$ distribution, see Ghosh and Razmpour (1982) and Handa et al. (2002).

A continuous random variable X is said to follow the exponential distribution with location parameter μ and scale parameter $a\mu$, if its pdf is given by,

$$f(x; \mu, a\mu) = \begin{cases} \frac{1}{a\mu} \exp\left\{-\frac{x-\mu}{a\mu}\right\}, \; x \geq \mu, \mu > 0, a > 0 \\ 0, \; \text{otherwise.} \end{cases} \tag{19}$$

In this case it can be shown that $E(X) = \mu + a\mu = (a+1)\mu$, $Var(X) = a^2\mu^2$ and coefficient of variation $c = \frac{a}{a+1}$ is a known positive constant for the known value of a. The results based on record values arising from the two parameter exponential distribution $E(\mu, \sigma)$ are available in (see Arnold et al. 1998, p.134). Using the results available in (see Arnold et al. 1998, p.134), the linear unbiased estimator

$\hat{\mu_1}$ given in (11), corresponding to the parameter μ involved in (19) reduces to,

$$\hat{\mu_1} = \frac{1}{n}\left[(n+1)R_0 - R_n\right], \tag{20}$$

where R_0 and R_n denote the 1^{st} and n^{th} record values. And

$$Var(\hat{\mu_1}) = \frac{n+1}{n}a^2\mu^2. \tag{21}$$

Also using the results available in (see Arnold et al. 1998, p.134), the linear unbiased estimator μ_1^* given in (15), corresponding to the parameter μ involved in (19) reduces to,

$$\mu_1^* = \frac{1}{an}(R_n - R_0) \tag{22}$$

and

$$Var(\mu_1^*) = \frac{\mu^2}{n}. \tag{23}$$

By using the results available in (see Arnold et al. 1998, p.134), we have found out the following results,

$$\boldsymbol{\alpha}'\mathbf{B}^{-1}\boldsymbol{\alpha} = n+1, \quad \boldsymbol{\alpha}'\mathbf{B}^{-1}\mathbf{1} = 1, \quad \mathbf{1}'\mathbf{B}^{-1}\mathbf{1} = 1,$$
$$\boldsymbol{\alpha}'\mathbf{B}^{-1} = (0,0,\cdots,1) \text{ and } \mathbf{1}'\mathbf{B}^{-1} = (1,0,\cdots,0).$$

Using the above results, the BLUE corresponding to (4) for the parameter μ involved in (19) reduces to,

$$\tilde{\mu_1} = \frac{2+n(n+1)a}{2[a^2(n+1)+2a+1]}R + \frac{a}{a^2(n+1)+2a+1}R_n,$$

that is,

$$\tilde{\mu_1} = g_1 R + g_2 R_n, \tag{24}$$

where $g_1 = \frac{2+n(n+1)a}{2[a^2(n+1)+2a+1]}$, $g_2 = \frac{a}{a^2(n+1)+2a+1}$ and $R = R_0 + R_1 + \cdots + R_n$. And its variance reduces to,

$$Var(\tilde{\mu_1}) = \frac{a^2\mu^2}{a^2(n+1)+2a+1}. \tag{25}$$

The main advantage of this results given in (24) and (25) is that, one can obtain the BLUE and its variance of the location parameter μ of the exponential distribution with known coefficient of variation by record values without knowing the values of means, variances and covariances of the record values arising from the standard form of the distribution defined in (19). We have evaluated the coefficients of R, R_n in the BLUE $\tilde{\mu}_1$ defined in (24), $Var(\hat{\mu}_1)$ defined in (21), $Var(\mu_1^*)$ defined in (23), $Var(\tilde{\mu}_1)$ defined in (25), the efficiency $e_1 = \frac{Var(\hat{\mu}_1)}{Var(\tilde{\mu_1})}$ of $\tilde{\mu_1}$ relative to $\hat{\mu_1}$ and the efficiency $e_2 = \frac{Var(\mu_1^*)}{Var(\tilde{\mu_1})}$ of $\tilde{\mu_1}$ relative to μ_1^* for $n = 2(1)10$ and $c = 0.15(0.05)0.30$ and are given in *TABLE 5* and *TABLE 6*. From this table, it may be noted that, in all the cases our estimator $\tilde{\mu_1}$ is much better than that of estimators $\hat{\mu_1}$ and μ_1^*.

TABLE 5

Coefficients of R, coefficients of R_n in the BLUE $\tilde{\mu}_1$, $V_1 = \frac{Var(\hat{\mu}_1)}{\mu^2}$, $V_2 = \frac{Var(\mu_1^*)}{\mu^2}$, $V_3 = \frac{Var(\tilde{\mu_1})}{\mu^2}$, the efficiency e_1 of $\tilde{\mu_1}$ relative to $\hat{\mu}_1$ and the efficiency e_2 of $\tilde{\mu_1}$ relative to μ_1^*

c	n	g_1	g_2	V_1	V_2	V_3	e_1	e_2
0.15	2	1.057	0.122	0.046	0.500	0.022	2.170	23.223
	3	1.393	0.119	0.042	0.333	0.021	1.970	15.812
	4	1.833	0.117	0.039	0.250	0.021	1.886	12.112
	5	2.369	0.115	0.037	0.200	0.020	1.848	9.891
	6	2.996	0.112	0.036	0.167	0.019	1.833	8.409
	7	3.708	0.110	0.036	0.143	0.019	1.831	7.349
	8	4.502	0.108	0.035	0.125	0.019	1.837	6.554
	9	5.372	0.106	0.034	0.111	0.019	1.849	5.939
	10	6.314	0.104	0.034	0.100	0.018	1.865	5.443
0.20	2	1.037	0.148	0.094	0.500	0.037	2.531	13.499
	3	1.429	0.143	0.083	0.333	0.036	2.333	9.334
	4	1.931	0.138	0.078	0.250	0.034	2.266	7.251
	5	2.533	0.133	0.075	0.200	0.033	2.250	6.000
	6	3.226	0.129	0.073	0.167	0.032	2.260	5.166
	7	4.000	0.125	0.071	0.143	0.031	2.286	4.572
	8	4.848	0.121	0.070	0.125	0.030	2.320	4.125
	9	5.765	0.118	0.069	0.111	0.029	2.361	3.778
	10	6.743	0.114	0.069	0.100	0.029	2.406	3.500

TABLE 6

Coefficients of R, coefficients of R_n in the BLUE $\tilde{\mu}_1$, $V_1 = \frac{Var(\hat{\mu}_1)}{\mu^2}$, $V_2 = \frac{Var(\mu_1^*)}{\mu^2}$, $V_3 = \frac{Var(\tilde{\mu_1})}{\mu^2}$, the efficiency e_1 of $\tilde{\mu_1}$ relative to $\hat{\mu}_1$ and the efficiency e_2 of $\tilde{\mu_1}$ relative to μ_1^*

c	n	g_1	g_2	V_1	V_2	V_3	e_1	e_2
0.25	2	1.000	0.167	0.166	0.500	0.056	2.999	8.999
	3	1.421	0.158	0.148	0.333	0.053	2.815	6.333
	4	1.950	0.150	0.139	0.250	0.050	2.778	5.000
	5	2.571	0.143	0.133	0.200	0.048	2.800	4.200
	6	3.273	0.136	0.130	0.167	0.045	2.852	3.667
	7	4.043	0.130	0.127	0.143	0.043	2.920	3.286
	8	4.875	0.125	0.125	0.125	0.042	2.999	2.999
	9	5.760	0.120	0.123	0.111	0.040	3.087	2.778
	10	6.692	0.115	0.122	0.100	0.038	3.178	2.600
0.30	2	0.949	0.178	0.276	0.500	0.076	3.612	6.556
	3	1.378	0.165	0.245	0.333	0.071	3.456	4.703
	4	1.904	0.154	0.230	0.250	0.066	3.469	3.778
	5	2.510	0.145	0.220	0.200	0.062	3.551	3.222
	6	3.182	0.136	0.214	0.167	0.058	3.667	2.852
	7	3.908	0.129	0.210	0.143	0.055	3.802	2.588
	8	4.680	0.122	0.207	0.125	0.052	3.949	2.389
	9	5.492	0.116	0.204	0.111	0.050	4.105	2.235
	10	6.337	0.110	0.202	0.100	0.047	4.265	2.111

6. Real life example

Roberts (1979) has given the one-hour average concentration of sulpher dioxide (in pphm) from Long Beach, California, for the years 1956 to 1974. From this data, we observe the upper record values for the months July and August are respectively given by (14, 18, 37) and (21, 25, 26, 40, 55). Also the upper record values for the month of September to be (33, 38). For analysing the data of September based on upper record values, we have only two record values. Obviously we know that, for analysing a data based on only two values, it is not enough to get good result. This is a major drawback of record values, that is, statistical inference based on records arise due to the fact that the occurrences of record data are very rare in practical situations and moreover the expected waiting time is infinite for every record after the first. In this case, if any prior information can be incorporated, the result will be improved. The coefficient of variation of record values, for the months July and August are respectively given by 0.43 and 0.40 respectively. The coefficient of variation is a stable measure of dispersion and thus does not change quite rapidly over the years. Here also we can observe that the coefficient of variation is constant and is approximately equal to 0.40. This is a situation where one can make use of the past data to have a knowledge regarding coefficient of variation.

A simple plot of the upper record values of each month against the expected values arising from usual standard normal distribution given in Table 1 of Balakrishnan and Chan (1998) indicate a very strong correlation (that is, the correlation coefficient between expected values and record values of each months as high as 0.90). Hence the assumption that the record values have come from a normal distribution $N(\mu, c\mu)$ is quite reasonable. From Table 1, we then determine the best linear unbiased estimator (BLUE) of μ to be,

$$\begin{aligned}\tilde{\mu} &= -0.03 \times 33 + 0.76 \times 38\\ &= 27.89\end{aligned}$$

and

$$Var(\tilde{\mu}) = 0.06726.$$

Once again without loss of generality assume that the record values for the month of September arising from a logistic distribution $LD(\mu, c\mu)$. From Table 3, we then determine the best linear unbiased estimator (BLUE) of μ to be,

$$\begin{aligned}\tilde{\mu} &= -0.14 \times 33 + 0.84 \times 38\\ &= 27.30\end{aligned}$$

and

$$Var(\tilde{\mu}) = 0.07701.$$

ACKNOWLEDGEMENTS

The authors are highly thankful for some of the helpful comments of the referee. Also authors express their gratefulness to UGC India for supporting this work by providing fund in the form of major research project.

REFERENCES

M. Ahsanullah (1995). *Record Statistics*, Nova Science Publishers, Commack, New York.

B. C. Arnold, N. Balakrishnan, H. N. Nagaraja (1998). *Records*, John Wiley and Sons, New York.

N. Balakrishnan, P. S. Chan (1998). *On the normal record values and associated inference*, Statistics and Probability Letters, 39, pp. 73-80.

N. Balakrishnan, A. C. Cohen (1991). *Order Statistics and Inference: Estimation Methods*, San Diego, Academic Press.

K. Bhat, K. A. Rao (2011). *Inference for Normal Mean with known Coefficient of Variation: Comparison using Simulation and Real Examples*, Lap Lambert Academic Publishing GmbH and Co. KG.

K. N. Chandler (1982). *The distribution and frequency of record values*, Journal of Royal Statistical Society, Series B, 14, pp. 220-228.

H. A. David (1981). *Order Statistics*, Second Edition, John Wiley and Sons, New York.

J. Eekelens (1968). *A method of calculation for logistic curve*, Statistica Neerlandica, 22, pp. 213-217.

Y. Fu, H. Wang, A. Wong (2013). *Inference for the normal mean with known coefficient of variation*, Open Journal of Statistics, 3, pp. 45-51.

M. Ghosh, A. Razmpour (1982). *Estimating the location parameter of an exponential distribution with known coefficient of variation*, Calcutta Statistical Association Bulletin, 31, pp. 137-150.

L. J. Gleser, J. D. Healy (1976). *Estimating the mean of a normal distribution with known coefficient of variation*, Journal of the American Statistical Association, 71, pp. 977-981.

N. Glick (1978). *Breaking records and breaking boards*, Amer. Math. Monthly., 85, pp. 2-26.

H. Guo, N. Pal (2003). *On a normal mean with known coefficient of variation*, Calcutta Statistical Association Bulletin, 54, pp. 17-29.

B. R. Handa, N. S. Kambo, C.D. Ravindran (2002). *Testing of scale parameter of the exponential distribution with known coefficient of variation: conditional approach*, Communications in Statistics-Theory and Methods, 31, pp. 73-86.

S. Hedayat, B. K. Sinha, W. Zhang. (2011). *Some aspects of inference on a normal mean with known coefficient of variation*, International Journal of Statistical Science, 11, pp. 159-181.

A. M. Kagan, Y. Malinovsky (2013). *On the Nile problem by Sir Ronald Fisher*, Electronic Journal of Statistics, 7, pp. 1968-1982.

R. A. Khan (2013). *A remark on estimating the mean of a normal distribution with known coefficient of variation*, Statistics: A Journal of Theoretical and Applied Statistics, pp. 1-7.

S. Kunte (2000). *A Note on consistent maximum likelihood estimation for $N(\theta, \theta^2)$ family*, Calcutta Statistical Association Bulletin, 50, PP. 325-328.

J. E. Lawless (1982). *Statistical Models and Methods of Lifetime Data*, John Wiley and Sons, New York.

W. B. Nelson (1990). *Accelerated Testing: Statistical Models, Test Plans and Data Analysis*, John Wiley and Sons, New York.

S. I. Resnick (1973). *Limit laws for record values*, Stochastic Processes and Their Applications, 1, pp. 67-82.

E. M. Roberts (1979). *Review of statistics of extreme values with applications to air quality data*, Journal of the Air Pollution Control Association, 29, pp. 733-740.

N. K. Sajeevkumar, P. Y. Thomas (2005). *Estimating the mean of logistic distribution with known coefficient of variation by order statistics*, Recent Advances in Statistical Theory and Applications, ISPS proceedings 1, pp. 170-176.

N. K. Sajeevkumar, M. R. Irshad (2011). *Estimating the parameter μ of the exponential distribution with known coefficient of variation using censored sample by order statistics*, IAPQR Transactions, 36, pp. 155-169.

N. K. Sajeevkumar, M. R. Irshad (2012). *Estimation of the mean of the double exponential distribution with known coefficient of variation by U-Statistics*, Journal of the Kerala Statistical Association, 23, pp. 61-71.

N. K. Sajeevkumar, M. R. Irshad (2013a). *Estimating the parameter μ of the exponential distribution with known coefficient of variation by order statistics*, Aligarh Journal of Statistics, 33, pp. 22-32.

N. K. Sajeevkumar, M. R. Irshad (2013b). *Estimation of the mean of the normal distribution with known coefficient of variation by U-Statistics*, IAPQR, Transactions, 38, pp. 51-65.

D. T. Searls, P. Intarapanich (1990). *A note on an estimator for the variance that utilizes kurtosis*, The American Satistician, 44, pp. 295-296.

R .W. Shorrock (1973). *Record values and inter-record times*, Journal of Applied Probability, 10, pp. 543-555.

P. Y. Thomas, N. K. Sajeevkumar (2003). *Estimating the mean of normal distribution with known coefficient of variation by order statistics*, Journal of the Kerala Statistical Association, 14, pp. 26-32.

SUMMARY

Estimation of the location parameter of distributions with known coefficient of variation by record values.

In this article, we derived the Best Linear Unbiased Estimator (BLUE) of the location parameter of certain distributions with known coefficient of variation by record values. Efficiency comparisons are also made on the proposed estimator with some of the usual estimators. Finally we give a real life data to explain the utility of results developed in this article.

Keywords: Record values; Normal distribution; Logistic distribution; Exponential distribution; Best linear unbiased estimator