J. Ind. Soc. Agril. Statist.
61(1), 2007 : 42-50

Estimation of Location and Scale Parameters of Pearson Type I Family of Distributions

N.K. Sajeevkumar, P. Yageen Thomas and Philip Samuel[1]
University of Kerala, Thiruvananthapuram
(Received : August, 2005)

SUMMARY

In this work we have derived the best linear unbiased estimator (BLUE) of the location parameter (μ) and scale parameter (σ) of Pearson Type I family of distributions, when the shape parameters p and q are known. The values of coefficients of order statistics in the BLUE's of μ and σ for $p > q$ have been explicitly derived in terms of coefficients of order statistics involved in the BLUE's of μ and σ for $p < q$. The coefficients of order statistics in the BLUEs of μ and σ, their variances and covariance for $n = 5(5)10$, $p = 2(0.5)4$, $q = 2(0.5)4$ for $p \le q$ are also evaluated. An application of the results of this paper is illustrated to a data in which lengths of fishes, is measured in centimeters (c.m.), on a catch of a trawler.

Key words : Pearson type-1 family of distributions, Beta distribution, Order statistics, Location and scale family of distributions, Estimation by order statistics.

1. INTRODUCTION

The family of distributions with probability density function (pdf) of the form

$$f(x;\ p,\ q,\ \mu,\ \alpha) = \begin{cases} \dfrac{1}{\sigma}\dfrac{1}{\beta(p,\ q)}\left(\dfrac{x-\mu}{\sigma}\right)^{p-1}\left\{1-\left(\dfrac{x-\mu}{\sigma}\right)\right\}^{q-1} & \\ 0 & \text{otherwise} \end{cases}$$

$$\text{for } \mu < x < \mu + \sigma \qquad (1)$$

where $-\infty < \mu < \infty, \sigma > 0, p > 0, q > 0$ and $\beta(p,\ q)$ is the usual complete beta function is called the Pearson Type I family of distributions (Johnson *et al.* 1995, p. 210). A distribution defined by the pdf (1) is also called as Generalized beta distribution. For convenience, we may write GBD $(p,\ q,\ \mu,\ \sigma)$ to denote the distribution defined in (1). If we put q = 1 in (1), the obtained distribution is called a Power function distribution. If X has a distribution defined by (1), then $Y = (X - \mu)/\sigma$ follows the well-known standard beta distribution with pdf given by

$$g(y : p,\ q) = \begin{cases} \dfrac{1}{\beta(p,\ q)} y^{p-1}(1-y)^{q-1} & 0 < y < 1 \\ 0 & \text{otherwise} \end{cases} \qquad (2)$$

For convenience, we may write BD(p,q) to denote the distribution defined in (2). Extensive applications of beta distribution are seen in the available literature. For example, beta distribution has been used in modelling distributions of (i) porosity/void ratio of soil (Harrop-Williams 1989), (ii) conductance of catfish retinal zones (Haynes and Yan 1990), (iii) variables affecting reproductivity of cows (McNally 1990), (iv) size of progeny in Escherchia Coli (Koppes and Grover 1992) and (v) transmission of HIV virus during a sexual contact between an infected and a susceptible individual (Wiley *et al.* 1989).

Though the problem of estimating the parameters in (1) is discussed extensively in Johnson *et al.* (1995), one may not get explicit solution for maximum likelihood estimators from likelihood equation of GBD(p, q, μ, σ) (AbouRizk *et al.* 1991 and Carnahan 1989). Moreover the procedure involved in obtaining estimators by method

[1] *University College, Thiruvananthapuram-695 034*

of moments is cumbersome (Elderton and Johnson 1969). In most real life situations only small samples could be realisable from a population and in such situations one cannot say much on reliability of estimators obtained by the method of moments or maximum likelihood procedure. These estimators are not even unbiased. In the available literature not any good finite sample estimators are seen derived and their properties analyzed for the parameters involved in (1). Hence, there is necessity to derive reasonably good finite sample estimators of the parameters involved in (1).

If p and q are known, it is clear that (1) is a member in the family of distributions, which depends only on a location parameter μ and scale parameter σ. Lloyd's (1952) method of estimating location and scale parameters of a distribution by order statistics is an extensively used method of estimation. For a survey of literature on the application of this method to various distributions, see David and Nagaraja (2003) and Balakrishnan and Cohen (1991).

In applying Lloyd's method of estimating the location and scale parameters of a distribution, one requires the expression for the first two single and product moments of order statistics arising from the standard form of the given distribution. In the available literature not any other work is seen done on the order statistics arising from the standard beta distribution BD(p, q) except the results of Thomas and Samuel (2006) regarding some recurrence relations on the single and product moments of the order statistics arising from BD(p, q).

Since Lloyd's (1952) method of estimation of μ and σ involved in (1) has not been seen derived in the available literature, in this work we use this method to obtain the small sample estimators of μ and σ involved in (1) for some known values of p and q. Cadima *et.al* (2005) have reported a data on length in centimeter (c.m.) of fishes obtained in a catch of a trawler. We have noticed that Pearson Type-1 distribution is a reasonable fit to the data and we have illustrated the methods derived in this paper to estimate μ and σ of (1) based on small sample data on lengths of fishes.

2. ESTIMATION OF LOCATION AND SCALE PARAMETERS OF A GBD(p, q, μ, σ) DISTRIBUTION

Let X be a random variable with pdf (1). Let $X_{1:n}, X_{2:n}, \ldots, X_{n:n}$ be the order statistics of a random sample of size n drawn from (1). Define $Y = (X - \mu)/\sigma$, then the pdf of Y is given by (2). Let $Y_{r:n} = (X_{r:n} - \mu)/\sigma$, r = 1, ..., n. Then $Y_{1:n}$, ..., $Y_{n:n}$ are distributed as the order statistics of a random sample of size n drawn from (2). Clearly (2) is free of μ and σ and hence the single and product moments of the order statistics $Y_{r:n}, r = 1, \ldots, n$ are also independent of μ and σ. Let

$$E(Y_{r:n}) = \alpha_{r:n} \qquad 1 \le r \le n \tag{3}$$

and

$$Cov(Y_{r:n}, Y_{s:n}) = v_{r,s:n} \qquad 1 \le r \le s \le n \tag{4}$$

in which we define $v_{r,r:n}$ as the variance of $Y_{r:n}$. If we write $\underline{X} = (X_{1:n}, \ldots, X_{n:n})$ then from (3) and (4) we have

$$E(\underline{X}) = A\theta \tag{5}$$

$$D(\underline{X}) = V\sigma^2 \tag{6}$$

where $\theta' = (\mu, \sigma)$, $A(1, \alpha)$, 1 being a column vector of n ones, $\alpha = (\alpha_{1:n}, \ldots, \alpha_{n:n})$ and $V = ((v_{r,s:n}))$.

Now using Gauss-Markov least squares theorem, the BLUE's μ^* and σ^* of μ and σ are given by (Balakrishnan and Rao 1998)

$$\mu^* = -\frac{\alpha' V^{-1}(1\alpha' - \alpha 1')V^{-1}}{\Delta}\underline{X} \tag{7}$$

$$\sigma^* = \frac{1' V^{-1}(1\alpha' - \alpha 1')V^{-1}}{\Delta}\underline{X} \tag{8}$$

with variances and covariance between the estimators given by

$$Var(\mu^*) = \frac{\alpha' V^{-1}\alpha}{\Delta}\sigma^2 \tag{9}$$

$$Var(\sigma^*) = \frac{1' V^{-1} 1}{\Delta}\sigma^2 \tag{10}$$

and

$$Cov(\mu^*, \sigma^*) = -\frac{(\alpha' V^{-1} 1)}{\Delta}\sigma^2 \tag{11}$$

where

$$\Delta = (\alpha' V^{-1}\alpha)(1' V^{-1} 1) - (\alpha' V^{-1} 1)^2 \tag{12}$$

It is to be noted that (7) and (8) can be written as

$$\mu^* = \sum_{i=1}^{n} a_{i:n} \; X_{i:n} \tag{13}$$

$$\sigma^* = \sum_{i=1}^{n} b_{i:n} \; X_{i:n} \tag{14}$$

We have computed $\alpha_{r:n}$, $1 \le r \le n$ and $V_{r,s:n}$ for $1 \le r \le s \le n$; $p = 2(0.5)4$, $q = 2(0.5)4$ for $p \le q$, $p = 1.981$, $q = 2.158$ and $n = 5(5)10$. We have verified the accuracy of the values obtained for the above moments by using the recurrence relations established by Thomas and Samuel (2006). Based on these values we have further evaluated the coefficients $a_{i:n}$, $b_{i:n}$ for i = 1, ..., n, $n = 5(5)10$, $p = 2(0.5)4$ and $q = 2(0.5)4$ with $p \le q$, $p = 1.981$, $q = 2.158$ and are presented in Table 1. The above technique of obtaining the BLUE's of μ^* and σ^* can be applied even if we have a censored sample, in which case we apply the technique with the expectation vector and covariance matrix of the vector of the realized order statistics of the sample.

It may be noted that when we put q = 1 in (1), the distribution is known as the power function distribution. The problem of estimation of the location and scale parameters involved in a power function distribution by order statistics has been considered in Kabir and Ahsanullah (1974). Hence in Table 1 we have not included the coefficients of order statistics in the estimators μ^* and σ^* for q = 1. When p = 1, q = 1, then GBD(1, 1, μ, σ) becomes a uniform distribution for which the BLUE's of μ and σ are well known.

The well-known Gupta's simplified linear unbiased estimators $\hat{\mu}$ for μ and $\hat{\sigma}$ for σ are given by Balakrishnan and Cohen (1991).

$$\hat{\mu} = \sum_{i=1}^{n} A_{i:n} \quad X_{i:n} \tag{15}$$

$$\hat{\sigma} = \sum_{i=1}^{n} B_{i:n} \quad X_{i:n} \tag{16}$$

where

$$A_{i:n} = \frac{1}{n} - \frac{\bar{\alpha}(\alpha_{i:n} - \bar{\alpha})}{\sum_{i=1}^{n}(\alpha_{i:n} - \bar{\alpha})^2} \quad \text{and}$$

$$B_{i:n} = \frac{(\alpha_{i:n} - \bar{\alpha})}{\sum_{i=1}^{n}(\alpha_{i:n} - \bar{\alpha})^2}, \quad i = 1,\ 2,\ \ldots,\ n$$

where $\bar{\alpha} = \frac{1}{n}\sum_{i=1}^{n}\alpha_{i:n}$

We have evaluated $Var(\hat{\mu})$, $Var(\hat{\sigma})$ and tabulated the values of $\frac{Var(\mu^*)}{\sigma^2}$, $\frac{Var(\sigma^*)}{\sigma^2}$, $\frac{Var(\hat{\mu})}{\sigma^2}$, $\frac{Var(\hat{\sigma})}{\sigma^2}$ and the efficiency $e_1(\mu^*/\hat{\mu})$ of μ^* relative to $\hat{\mu}$ and the efficiency $e_2(\sigma^*/\hat{\sigma})$ of σ^* relative to $\hat{\sigma}$ for $n = 5(5)10$, $p = 2(0.5)4$, $q = 2(0.5)4$ with $p \le q$ and $p = 1.981$, $q = 2.158$ and are presented in the Table 1. From Table 1 it is clear that though Gupta's linear estimates are also unbiased, our estimates μ^* and σ^* are relatively much better. It may be noted that the required estimators μ^*, σ^* for $p > q$ can be obtained from Table 1 itself. The way in which the coefficients of the order statistics in μ^* and σ^* for $p > q$ can be determined from those in Table 1 becomes clear from the results that we prove in the next Section. All the computational works involved in this paper were done using 'Mathcad'.

3. RELATIONSHIP BETWEEN THE COEFFICIENTS OF ORDER STATISTICS IN THE BLUES μ^* AND σ^* OF GBD (p, q, μ, σ) FOR p < q WITH THOSE OF p > q

Let $X_{1:n}, \ldots, X_{n:n}$ be the order statistics of a random sample of size n drawn from GBD (p, q, μ, σ). Let $Y_{1:n}, \ldots, Y_{n:n}$ be the order statistics of a random sample of size n drawn from BD(p, q). Let $\underline{Y} = (Y_{1:n}, \ldots, Y_{n:n})'$; $E(\underline{Y}) = (\alpha_{1:n}, \ldots, \alpha_{n:n})'$ and $D(\underline{Y}) = ((v_{i,j:n})) = V$. For convenience we write $\mu^*_{(p,q)}$ and $\sigma^*_{(p,q)}$ to denote the BLUEs of μ and σ involved in GBD (p, q, μ, σ). Then we have

$$\mu^*_{(p,q)} = -\frac{\alpha' V^{-1}(1\alpha' - \alpha 1')V^{-1}}{\Delta} \underline{X} \tag{17}$$

$$\sigma^*_{(p,q)} = \frac{1' V^{-1}(1\alpha' - \alpha 1')V^{-1}}{\Delta} \underline{X} \tag{18}$$

$$Var(\mu^*_{(p,q)}) = \frac{\alpha' V^{-1}\alpha}{\Delta}\alpha^2 \tag{19}$$

$$Var(\sigma^*_{(p,q)}) = \frac{1' V^{-1} 1}{\Delta}\sigma^2 \tag{20}$$

and

$$Cov(\mu^*_{(p,q)}, \sigma^*_{(p,q)}) = \frac{\alpha' V^{-1} 1}{\Delta}\ \sigma^2 \tag{21}$$

Table 1. Coefficients of $X_{r:n}$ $r = 1, 2, ..., n$ in the BLUE's $\mu^*, \sigma^*, Var(\mu^*), Var(\sigma^*), Cov(\mu^*, \sigma^*)$ and the relative efficiencies $e_1 = e_1(\mu^*/\hat{\mu}), e_2 = e_2(\sigma^*/\hat{\sigma})$, $n = 5$

p	q	Estimator	$X_{1:5}$	$X_{2:5}$	$X_{3:5}$	$X_{4:5}$	$X_{5:5}$	$Var(\mu^*)$/ $Var(\sigma^*)$	$Var(\hat{\mu})$/ $Var(\hat{\sigma})$	e_1/e_2	$Cov(\mu^*, \sigma^*)$
2	2	μ*	1.304	0.167	0.068	−0.008	−0.532	0.032	1.711	53.324	−0.0463
		σ*	−1.836	0.175	0.000	0.176	1.836	0.093	6.044	65.312	
2	2.5	μ*	1.288	0.148	0.045	−0.035	−0.446	0.023	1.534	65.422	−0.039
		σ*	−2.050	−0.136	0.078	0.298	1.810	0.097	6.757	69.442	
2	3	μ*	1.280	0.134	0.029	−0.050	−0.392	0.018	1.420	78.989	−0.034
		σ*	−2.281	−0.096	0.153	0.396	1.828	0.103	7.625	74.342	
2	3.5	μ*	1.276	0.122	0.017	−0.061	−0.354	0.014	1.339	93.860	−0.030
		σ*	−2.525	−0.056	0.226	0.493	1.861	0.108	8.616	79.916	
2	4	μ*	1.274	0.113	0.008	−0.069	−0.325	0.012	1.279	110.030	−0.027
		σ*	−2.776	−0.015	0.298	0.587	1.907	0.113	9.716	86.234	
2.5	2.5	μ*	1.344	0.230	0.090	−0.028	−0.637	0.032	2.020	62.288	−0.049
		σ*	−1.981	−0.258	0.000	0.258	1.981	0.099	7.282	73.680	
2.5	3	μ*	1.331	0.209	0.066	−0.048	−0.558	0.025	1.847	72.752	−0.043
		σ*	−2.167	−0.229	0.071	0.352	1.967	0.102	7.975	78.410	
2.5	3.5	μ*	1.325	0.193	0.048	−0.065	−0.502	0.021	1.726	84.105	−0.038
		σ*	−2.366	−0.192	0.138	0.445	1.974	0.105	8.793	83.526	
2.5	4	μ*	1.321	0.181	0.035	−0.077	−0.459	0.017	1.636	96.401	−0.035
		σ*	−2.572	−0.163	0.203	0.533	1.998	0.109	9.713	89.261	
3	3	μ*	1.388	0.282	0.106	−0.047	−0.729	0.033	2.331	71.511	−0.051
		σ*	−2.117	−0.329	0.000	0.329	2.117	0.103	8.525	82.519	
3	3.5	μ*	1.379	0.261	0.082	−0.069	−0.653	0.028	2.161	80.795	−0.046
		σ*	−2.285	−0.302	0.065	0.422	2.101	0.106	9.207	87.205	
3	4	μ*	1.373	0.245	0.064	−0.085	−0.597	0.022	2.035	90.824	−0.042
		σ*	−2.460	−0.278	0.125	0.507	2.107	0.108	9.992	92.446	
3.5	3.5	μ*	1.435	0.327	0.118	−0.069	−0.810	0.033	2.642	80.761	−0.053
		s*	−2.245	−0.396	0.000	0.396	2.245	0.107	9.769	91.436	
3.5	4	μ*	1.426	0.307	0.096	−0.089	−0.720	0.028	2.475	89.295	−0.049
		σ*	−2.397	−0.374	0.059	0.481	2.231	0.109	10.444	96.192	
4	4	μ*	1.481	0.369	0.129	−0.092	−0.885	0.033	2.954	90.092	−0.054
		σ*	−2.365	−0.460	0.000	0.460	2.365	0.120	11.016	100.515	

n = 10

p	q	Estimator	$X_{1:10}$	$X_{2:10}$	$X_{3:10}$	$X_{4:10}$	$X_{5:10}$	$X_{6:10}$	$X_{7:10}$	$X_{8:10}$	$X_{9:10}$	$X_{10:10}$	$Var(\mu^*)/Var(\sigma^*)$	$Var(\hat{\mu})/Var(\hat{\sigma})$	e_1/e_2	$Cov(\mu^*, \sigma^*)$
2	2	μ*	0.991	0.150	0.090	0.060	0.037	0.025	0.009	– 0.007	– 0.032	– 0.324	0.012	0.715	59.196	– 0.016
		σ *	– 1.315	– 0.182	– 0.097	– 0.051	– 0.012	0.012	0.051	0.097	0.183	1.315	0.032	2.458	75.870	
2	2.5	μ*	0.988	0.144	0.084	0.051	0.032	0.014	0.000	– 0.017	– 0.042	– 0.254	0.009	0.642	74.070	– 0.421
		σ *	– 1.495	– 0.186	– 0.085	– 0.026	0.009	0.057	0.093	0.153	0.253	1.227	0.035	2.745	77.930	
2	3	μ*	0.990	0.140	0.078	0.046	0.024	0.008	– 0.006	– 0.023	– 0.046	– 0.211	0.007	0.595	90.670	– 0.012
		σ *	– 1.695	– 0.191	– 0.071	– 0.007	0.045	0.089	0.139	0.203	0.310	1.178	0.038	3.092	80.690	
2	3.5	μ*	0.992	0.139	0.072	0.041	0.020	0.005	– 0.012	– 0.027	– 0.049	– 0.181	0.005	0.561	108.983	– 0.010
		σ *	– 1.897	– 0.215	– 0.049	0.018	0.074	0.122	0.182	0.250	0.363	1.152	0.041	3.488	84.650	
2	4	μ*	0.996	0.134	0.070	0.037	0.016	0.000	– 0.014	– 0.029	– 0.050	– 0.161	0.004	0.536	128.928	– 0.009
		σ *	– 2.123	– 0.199	– 0.046	0.041	0.104	0.160	0.219	0.291	0.410	1.144	0.044	3.927	89.638	
2.5	2.5	μ*	0.967	0.202	0.126	0.083	0.056	0.027	0.010	– 0.018	– 0.058	– 0.397	0.013	0.840	66.114	– 0.018
		σ *	– 1.364	– 0.259	– 0.144	– 0.073	– 0.030	0.030	0.073	0.144	0.259	1.364	0.036	2.959	81.169	
2.5	3	μ*	0.967	0.196	0.118	0.075	0.045	0.021	– 0.001	– 0.027	– 0.059	– 0.333	0.010	0.769	78.148	– 0.016
		σ *	– 1.516	– 0.268	– 0.135	– 0.056	0.004	0.058	0.117	0.193	0.293	1.307	0.038	3.238	84.716	
2.5	3.5	σ *	0.970	0.190	0.112	0.068	0.039	0.012	– 0.008	– 0.034	– 0.064	– 0.286	0.008	0.719	91.258	– 0.014
		s*	– 1.682	– 0.278	– 0.128	– 0.036	0.026	0.100	0.152	0.238	0.345	1.261	0.040	3.566	88.576	
2.5	4	σ *	0.972	0.189	0.107	0.063	0.032	0.009	– 0.014	– 0.038	– 0.067	– 0.252	0.006	0.682	105.252	– 0.013
		s*	– 1.849	– 0.293	– 0.117	– 0.021	0.058	0.123	0.193	0.279	0.394	1.234	0.042	3.935	93.125	
3	3	μ*	0.959	0.244	0.155	0.104	0.067	0.035	0.005	– 0.031	– 0.084	– 0.455	0.013	0.965	73.848	– 0.020
		σ *	– 1.414	– 0.328	– 0.186	– 0.098	– 0.033	0.033	0.098	0.186	0.328	1.414	0.039	3.461	87.792	
3	3.5	μ*	0.962	0.238	0.147	0.094	0.057	0.026	– 0.005	– 0.039	– 0.090	– 0.390	0.011	0.896	84.110	– 0.018
		σ *	– 1.551	– 0.340	– 0.180	– 0.081	– 0.004	0.063	0.138	0.230	0.377	1.350	0.041	3.736	91.561	
3	4	μ*	0.965	0.233	0.141	0.087	0.049	0.018	– 0.012	– 0.045	– 0.093	– 0.344	0.009	0.844	95.168	– 0.016
		σ *	– 1.693	– 0.353	– 0.176	– 0.065	0.018	0.094	0.173	0.271	0.420	1.310	0.042	4.051	95.915	
3.5	3.5	μ*	0.961	0.280	0.181	0.121	0.077	0.038	0.001	– 0.044	– 0.110	– 0.505	0.013	1.091	81.902	– 0.021
		σ *	– 1.465	– 0.390	– 0.225	– 0.120	– 0.039	0.039	0.120	0.225	0.390	1.465	0.042	3.964	95.099	
3.5	4	μ*	0.964	0.274	0.173	0.112	0.067	0.029	– 0.009	– 0.052	– 0.114	– 0.443	0.011	1.022	91.119	– 0.019
		σ *	– 1.587	– 0.403	– 0.221	– 0.106	– 0.015	0.068	0.156	0.265	0.434	1.409	0.043	4.235	99.119	
4	4	μ*	0.968	0.311	0.203	0.136	0.085	0.040	– 0.005	– 0.058	– 0.134	– 0.548	0.013	1.217	90.267	– 0.022
		σ *	– 1.516	– 0.446	– 0.261	– 0.141	– 0.045	0.045	0.141	0.261	0.446	1.516	0.043	4.467	102.812	

Let $Z_{1:n}, Z_{2:n}, \ldots, Z_{n:n}$ be the order statistics of a random sample of size n arising from GBD(q, p, μ, σ). Then

$$[(Z_{1:n}-\mu)/\sigma, (Z_{2:n}-\mu)/\sigma, \ldots, (Z_{n:n}-\mu)/\sigma]$$

is distributed identically as the order statistics of a random sample of size n arising from BD(q,p). Now it is clear that

$$[(Z_{1:n}-\mu)/\sigma, (Z_{2:n}-\mu)/\sigma, \ldots, (Z_{n:n}-\mu)/\sigma]$$

is distributed identically as

$$(1-Y_{n:n}, 1-Y_{n-1:n}, \ldots, 1-Y_{1:n}).$$

Consequently we have

$$E(Z_{r:n}) = \sigma(1-\alpha_{n-r+1:n}) + \mu$$

and

$$\mathrm{Cov}(Z_{r:n}, Z_{s:n}) = \mathrm{Cov}(Y_{n-s+1:n}, Y_{n-r+1:n})$$

Let $\underline{Z} = (Z_{1:n}, Z_{2:n}, \ldots, Z_{n:n})$

Thus we have

$$E(\underline{Z}) = \sigma(1-J\alpha) + 1\mu,$$

where J is the n × n matrix given by

$$J = \begin{bmatrix} 0 & 0 & \ldots & 0 & 1 \\ 0 & 0 & \ldots & 1 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 1 & 0 & \ldots & 0 & 0 \end{bmatrix}$$

Clearly we have $J = J' = J^{-1}$ and $J1 = 1$. If we write $D(\underline{X}) = \sigma^2 V$, then $D(\underline{Z}) = JVJ\sigma^2$. If we write $\mu^*_{(q,p)}$ and $\sigma^*_{(q,p)}$ to denote the BLUE's of μ and σ involved in GBD(q, p, μ, σ), then

$$\mu^*_{(q,p)} = -\frac{(1-J\alpha)'JV^{-1}J[1(1-J\alpha)'-(1-J\alpha)1']JV^{-1}J}{\Delta'}\underline{Z} \tag{22}$$

where

$$\Delta' = (1-J\alpha)'(JV^{-1}J)(1-J\alpha)(1'JV^{-1}J1) - ((1-J\alpha)'(JV^{-1}J)1)^2$$

$$= (1'V^{-1}J - \alpha'V^{-1}J)(11'V^{-1}1 - J\alpha 1'V^{-1}1) - (1'V^{-1}1 - \alpha'V^{-1}1)^2$$

$$= 1'V^{-1}1.1'V^{-1}1 - \alpha'V^{-1}1.1'V^{-1}1 - 1'V^{-1}\alpha.1'V^{-1}1 + \alpha'V^{-1}\alpha.1'V^{-1}1 - (1'V^{-1}1)^2 - (\alpha'V^{-1}1)^2 + 2.1'V^{-1}1.\alpha'V^{-1}1$$

$$= \alpha'V^{-1}\alpha.1'V^{-1}1 - (\alpha V^{-1}1)^2 \tag{23}$$

Thus using (22) and (23) we have $\Delta' = \Delta$

Further simplifying the numerator of (22) and using (23) the BLUE $\mu^*_{(q,p)}$ reduces to

$$\mu^*_{(q,p)} = \frac{1'V^{-1}(1\alpha'-\alpha 1')V^{-1}}{\Delta}(J\underline{Z}) - \frac{\alpha'V^{-1}(1\alpha'-\alpha 1')V^{-1}}{\Delta}(J\underline{Z}) \tag{24}$$

The BLUE $\sigma^*_{(q,p)}$ of σ is given by

$$\sigma^*_{(q,p)} = \frac{1'(JVJ^{-1})[1(1-J\alpha)'-(1-J\alpha)1'](JvJ)^{-1}}{\Delta}.\ \underline{Z} \tag{25}$$

Then on simplifying (25) we get

$$\sigma^*_{(q,p)} = -\frac{1'V^{-1}(1\alpha'-\alpha 1')V^{-1}}{\Delta}(J\underline{Z}) \tag{26}$$

$$\mathrm{Var}\,\mu^*_{(q,p)} = \frac{(1-J\alpha)'(JV^{-1}J)(1-J\alpha)}{\Delta}\sigma^2 = \frac{1'V^{-1}1 - 2\alpha'V^{-1}1 + \alpha'V^{-1}\alpha}{\Delta}\sigma^2 \tag{27}$$

$$\mathrm{Var}\,\sigma^*_{(q,p)} = \frac{1'JV^{-1}J1}{\Delta}\sigma^2 = \frac{1'V^{-1}1}{\Delta}\sigma^2 \tag{28}$$

and

$$\mathrm{Cov}(\mu^*_{(q,p)}, \sigma^*_{(q,p)}) = \frac{(\alpha'V^{-1}1 - 1'V^{-1}1)}{\Delta}\sigma^2 \tag{29}$$

Since in (24), $J\underline{Z} = (Z_{n:n}, Z_{n-1:n}, \ldots, Z_{1:n})$, we note that the coefficient of $Z_{r:n}$ in $\mu^*_{(q,p)}$ is equal to the sum of the coefficients of $x_{n-r+1:n}$ in $\mu^*_{(p,q)}$ and $\sigma^*_{(p,q)}$ and the coefficient of $Z_{r:n}$ in $\sigma^*_{(q,p)}$ is equal to negative value of the coefficient of $x_{n-r+1:n}$ in $\sigma^*_{(p,q)}$. Then we have proved the following theorem.

Theorem 3.1 : If $x_{1:n}$, ..., $x_{n:n}$ be the order statistics from GBD (p,q,μ,σ) and $Z_{1:n}$, ..., $Z_{n:n}$ be the order statistics arising from GBD(q,p,μ,σ) then the coefficient of $Z_{r:n}$ in the BLUE of $\mu^*_{(q,p)}$ of μ in GBD(q,p,μ,σ) is equal to the sum of the coefficients of $x_{n-r+1:n}$ in $\mu^*_{(p,q)}$ and $\sigma^*_{(p,q)}$ of the parameters μ and σ involved in GBD(p,q,μ,σ) and the coefficient of $Z_{r:n}$ in the BLUE of $\sigma^*_{(q,p)}$ of σ of GBD(q,p,μ,σ) is equal to the negative value of the coefficient of $x_{n-r+1:n}$ in $\sigma^*_{(p,q)}$ of σ involved in GBD(q,p,μ,σ). Further

$$\text{Var}(\mu^*_{(q,p)}) = \text{Var}(\mu^*_{(p,q)}) + 2\,\text{Cov}(\mu^*_{(p,q)}, \sigma^*_{(p,q)}) + \text{Var}(\sigma^*_{(p,q)})$$

$$\text{Var}(\sigma^*_{(q,p)}) = \text{Var}(\sigma^*_{(p,q)}) \text{ and}$$

$$\text{Cov}(\mu^*_{(q,p)}, \sigma^*_{(q,p)}) = -\text{Cov}(\mu^*_{(p,q)}, \sigma^*_{(p,q)}) - \text{Var}(\sigma^*_{(p,q)})$$

When the censoring is symmetric then also the above theorem can be seen to be valid.

4. AN APPLICATION

As an illustration of the theory developed, we consider the fisheries data given in Cadima *et al.* (2005). The data consist of length in c.m. of 195 fishes in a catch of a trawler. We have arranged the raw data into a frequency data as given below :

Individual total length (c.m.)	Frequency (f)
15	1
16	1
17	3
18	12
19	24
20	24
21	23
22	18
23	18
24	12
25	12
26	13
27	11
28	9
29	6
30	6
31	2
Total	195

Using this data, we check whether the given data is from the distribution defined in (1). We first consider the estimate of μ as $\hat{\mu} = 15$, the smallest value in the given data and the estimate of $\mu + \sigma$ as $\hat{\mu} + \hat{\sigma} = 31$, the largest value in the data. Therefore, the estimate of σ, is given by $\hat{\sigma} = 16$. Then considering μ and σ in (1) as determined by $\mu = 15$ and $\sigma = 16$ then one can use the results in Johnson *et al.* (1995) to obtain the estimators of p and q by using

$$\hat{p} + \hat{q} = \frac{\left(\frac{m_1 - 15}{16}\right)\left(1 - \left(\frac{m_1 - 15}{16}\right)\right)}{\left(\frac{m_2}{(16)^2}\right)} - 1 \quad (30)$$

and

$$\hat{p} = \left(\frac{m_1 - 15}{16}\right)^2 \left(1 - \left(\frac{m_2 - 15}{16}\right)\right)\left(\frac{m_2}{(16)^2}\right)^{-1} - \left(\frac{m_1 - 15}{16}\right) \quad (31)$$

where m_1 and m_2 are the mean and variance of the sample data. In this case, $m_1 = 22.656$ and $m_2 = 12.431$. Consequently the estimator of p and q are obtained as $\hat{p} = 1.981$ and $\hat{q} = 2.158$ or $\hat{p} \simeq 2$ and $\hat{q} \simeq 2$. Now using Kolmogorov-Smirnov test, we test the hypothesis that the given data is from (1). The x values, $F_0(x) = P(X \le x)$, the empirical distribution function $F^*_{195}(x)$ and the test statistic $D_n = \left|F^*_{195}(x) - F_0(x)\right|$ for each x = 15(1)31 are given below.

Clearly we have Sup $D_n = \text{Sup}\left|F^*_{195}(x) - F_0(x)\right| = 0.103$. Now for the Kolmogorov-Smirnov two-sided test, at one percent level of significance, the critical value is equal to $\frac{1.63}{\sqrt{195}} = 0.117$. Therefore, for one percent level of significance, we accept null hypothesis that given data is from a Pearson type-1 family of distribution with $p = 1.981$ and $q = 2.158$.

X	$F_0(X)$	$F^*_{195}(x)$	$D_n = \lvert F^*_{195}(x) - F_0(x) \rvert$
15	0	0.005	0.005
16	0.013	0.010	0.003
17	0.050	0.026	0.024
18	0.105	0.087	0.018
19	0.176	0.210	0.034
20	0.258	0.333	0.075
21	0.348	0.451	0.103
22	0.443	0.544	0.101
23	0.538	0.636	0.098
24	0.632	0.697	0.066
25	0.720	0.759	0.039
26	0.801	0.826	0.025
27	0.870	0.882	0.012
28	0.927	0.928	0.001
29	0.968	0.959	0.009
30	0.992	0.990	0.003
31	1	1	0

Suppose we draw a sample of size ten from the given data on the length of 195 fishes using random number table. The values are given below :

16, 21, 18, 26, 29, 19, 18, 26, 19, 19

Using these values for $p \simeq 2$, $q \simeq 2$, the BLUE of μ and σ are :

$$\mu^* = 0.990X_{1:10} + 0.146X_{2:10} + 0.085X_{3:10} + 0.057X_{4:10} + 0.035X_{5:10} + 0.018X_{6:10} + 0.008X_{7:10} - 0.012X_{8:10} - 0.038X_{9:10} - 0.290X_{10:10}$$

Similarly

$$\sigma^* = -1.376X_{1:10} - 0.180X_{2:10} - 0.093X_{3:10} - 0.042X_{4:10} - 0.002X_{5:10} + 0.022X_{6:10} + 0.067X_{7:10} + 0.115X_{8:10} + 0.217X_{9:10} + 1.268X_{10:10} = 19.463$$

with $V(\mu^*) = 0.011\sigma^2$, $V(\sigma^*) = 0.033\sigma^2$

and $Cov(\mu^*, \sigma^*) = -0.015\sigma^2$

Clearly if one is interested with the estimate of mean length of fishes in the locality with the reported catch, then one requires the BLUE of the population mean of (1) given by

$$\theta = \mu + \left(\frac{p}{p+q}\right)\sigma$$

If p and q are known as $p = 1.981$, $q = 2.158$, then

$$\theta^* = \mu^* + \left(\frac{p}{p+q}\right)\sigma^* = 12.546 + 0.479 \times 19.463 = 21.869$$

$$Var(\theta^*) = Var(\mu^*) + \left(\frac{p}{p+q}\right)^2 Var(\sigma^*) + 2\left(\frac{p}{p+q}\right)Cov(\mu^*, \sigma^*) = 0.004\,\sigma^2$$

Similarly the results of this paper are helpful to deal with several problems, when one has an evidence to see that the associated random variable is having a distribution defined by pdf given in (1).

ACKNOWLEDGEMENTS

The authors are highly grateful for the referee's valuable comments which led to the improvement of an earlier version of this paper.

REFERENCES

AbouRizk, S.M., Halpin, D.W. and Wilson, J.R. (1991). Visual iterative fitting of beta distribution. *J. Cons. Engg. Manag.*, **117**, 589-605.

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

Cadima, E.L., Caramelo, A.M., Afonso-Dias, M., Tandstad, M.O. and Laiva-Moreno, J.I. (2005). Sampling methods applied to fisheries sciences : A manual. Food and Agriculture Organization of the United Nations, Rome.

Carnahan, J.V. (1989). Maximum likelihood estimation for the 4-parameter beta distribution. *Comm. Statist. – Simula,* **18**, 513 – 516.

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

Elderton, W.P. and Johnson, N.L. (1969). *Systems of Frequency Curves*. Cambridge University Press, Cambridge.

Harrop-Williams, K. (1989). Random nature of soil porosity and related properties. *J. Engg. Mech.*, **115**, 1123-1129.

Haynes, L.W. and Yan, K.W. (1990). Single-channel measurement from the cyclic GMP- activated conductance of catfish retinal cones. *J. Physiology (London)*, **429**, 1451-1481.

Johnson, N.L., Kotz, S. and Balakrishnan, N. (1995). *Continuous Univariate Distributions*. Second edition, John Wiley and Sons, NewYork.

Kabir, A.B.M.L. and Ahsanullah, M. (1974). Estimation of the location and scale parameters of a power function distribution by linear functions of order statistics. *Comm. Statist.—Theory Methods*, **3**, 463- 467.

Koppes, L.J. and Grover, N.B. (1992). Relationship between size of parent at cell division and relative size of progeny in Escherchia Coli. *Archives Microbiology*, **157**, 402-405.

Lloyd, E.H. (1952). Least-squares estimation of location and scale parameters using order statistics. *Biometrika*, **39**, 88-95.

McNally, R.J. (1990). Maximum likelihood estimation of the parameters of the prior distributions of three variables that strongly influence reproductive performance in cows. *Biometrics*, **46**, 501-514.

Thomas, P.Y. and Samuel, P. (2006). Recurrence relation for the moments of order statistics from a beta distribution. Statistical papers (accepted for publication).

Wiley, J.A., Herschokoru, S.J. and Padiau, N.S. (1989). Heterogeneity in the probability of HIV transmission per sexual contact: The case of male-to-female transmission in penile-vaginal intercourse. *Stat. Med.*, **8**, 93-102.