Ann. Inst. Statist. Math.
36 (1984), Part A, 289-299

ASYMPTOTIC LINEAR PREDICTION OF EXTREME ORDER STATISTICS

H. N. NAGARAJA

(Received Jan. 14, 1983; revised June 23, 1983)

Summary

We consider the problem of predicting the sth order statistic using the lowest r order statistics from a large sample of size n under the assumption that the sample minimum, appropriately normalized, has a non-degenerate limit distribution as $n\to\infty$. Assuming r, s fixed and $n\to\infty$ we obtain asymptotically best linear unbiased as well as asymptotically best linear invariant predictors of the sth order statistic.

1. Introduction

The problem of predicting future order statistics based on a few observed order statistics has received a great deal of attention in life data analysis literature for quite some time (see, e.g., Nelson [12]). Most of the papers have used the finite sample theory and handled the problem for each family of distributions separately. Kaminsky and Nelson [8] have used the asymptotic theory for the sample quantiles in developing asymptotically best linear unbiased predictors (ABLUP) for order statistics. In this paper we obtain ABLUP and asymptotically best linear invariant predictors (ABLIP) based on extreme value theory.

Let $X_{1:n}\leqq X_{2:n}\leqq\cdots\leqq X_{n:n}$ be the order statistics from a random sample from a distribution with distribution function (df) F. Based on $X_{1:n},\cdots,X_{r:n}$, we would like to predict $X_{s:n}$ where r and s are small compared to n. Such a situation arises when the manufacturer of a certain product would like to predict the failure times of the products with the warranty period after observing the first few failures. For example, suppose a car is sold with a five year warranty on certain parts. After a year or two of the model year, the manufacturer would have data on the failure times of the parts failed by that time. The number of failures would be small compared to the number of cars sold.

Key words: Extreme-value distributions, best linear unbiased predictor, best linear invariant predictor.

Based on this data, he would like to predict the times of the next several failures. This would assist in estimating the cost associated with warranty as well as inventory requirements.

Suppose $X_{1:n}$, appropriately normalized, converges in law to a non-degenerate distribution with df G; that is there exist constants c_n and $d_n>0$ such that

$$\text{(1)} \qquad \mathrm{P}\left((X_{1:n}-c_n)/d_n \leqq x\right)=1-F^n(c_n+d_n x) \to G(x)\ ,$$

as $n\to\infty$ for all x in the support of G. If this happens we write $F\in \mathcal{D}(G)$. From Gnedenko [4] it is known that G can be one of the following df's (up to location and scale parameters)

$$\text{(2)} \qquad \begin{aligned} L_{1,\alpha}(x)&=\begin{cases} 1-\exp(-(-x)^{-\alpha})\ , & x<0,\ \alpha>0 \\ 1\ , & x\geqq 0 \end{cases} \\ L_{2,\alpha}(x)&=\begin{cases} 1-\exp(-x^{\alpha})\ , & x>0,\ \alpha>0 \\ 0\ , & x\leqq 0 \end{cases} \\ L_3(x)&= 1-\exp(-\exp(x))\ , \qquad -\infty<x<\infty\ . \end{aligned}$$

The norming constants c_n and d_n depend on F and G, and are available, for example in Galambos ([3], pp. 56–57). (The constant d_n in the $L_{1,\alpha}$ case should be the negative of the value given by (29) on page 56 of his book.) In the $L_{1,\alpha}$ case, $c_n=0$ whereas in the other two cases it is nonzero. If (1) holds then the limiting distribution of $(X_{k:n}-c_n)/d_n$ is that of T_k, the kth upper record value from the df G (see, e.g., Nagaraja [11]). Hence $X_{1:n}, X_{2:n}, \cdots, X_{r:n}$ behave approximately as the first r upper record values from the df $G(\mu+\sigma x)$ with $\mu=c_n$ and $\sigma=d_n$, if n is large. Thus the problem now reduces to predicting the sth upper record value using the first r record values from a location and scale parameter family if G is $L_{2,\alpha}$ or L_3, and from a scale parameter family if G is $L_{1,\alpha}$. Weissman [14] has considered the estimation of norming constants for upper extremes and used the same in the estimation of large quantiles. Ahsanullah [1] has considered the linear prediction problem for exponential distribution; that is for $L_{2,1}$.

2. Asymptotic covariance structure of extreme order statistics

The following lemma gives the representation for limiting extreme order statistics, that is for $X_{k:n}$ as $n\to\infty$ if (1) holds. It is useful in the computation of the first two moments of limiting extreme order statistics.

LEMMA 1. *Let Z_i's be independent standard exponential random* *variables and* γ *be Euler's constant* $(0.5772\cdots)$. *Then the* T_k*'s, the upper record values from* $G(\mu+\sigma x)$*, have the following representation*:

$$T_k \stackrel{d}{=} \mu+\sigma\left(\sum_{j=k}^{\infty}(1-Z_j)/j-\gamma+\sum_{j=1}^{k-1} j^{-1}\right) \qquad \text{if } G=L_3 \tag{3a}$$

$$T_k \stackrel{d}{=} \mu+\sigma\left(\sum_{j=1}^{k} Z_j\right)^{1/\alpha} \qquad \text{if } G=L_{2,\alpha} \tag{3b}$$

$$T_k \stackrel{d}{=} \mu-\sigma\left(\sum_{j=1}^{k} Z_j\right)^{-1/\alpha} \qquad \text{if } G=L_{1,\alpha}\,. \tag{3c}$$

Here $\stackrel{d}{=}$ *stands for the equality of distribution and* $\sum_{j=1}^{0} j^{-1}=0$.

PROOF. Relation (3a) follows from Hall [6] and the remaining two follow from Nagaraja [11].

Let $\alpha_i=\mathrm{E}(T_i)$ and $v_{ij}=\mathrm{Cov}(T_i, T_j)$. The following lemma exhibits these constants when G is one of the distributions in (2).

LEMMA 2. *For positive integers* i *and* j *with* $i\leqq j$,

$$\alpha_i=\sum_{l=1}^{i-1} l^{-1}-\gamma \quad \text{and} \quad v_{ij}=\sum_{l=j}^{\infty} l^{-2}\,, \qquad \text{if } G=L_3 \tag{4a}$$

$$\alpha_i=\Gamma(i+\delta)/\Gamma(i) \quad \text{and} \quad v_{ij}=\alpha_i[\{\Gamma(j+2\delta)/\Gamma(j+\delta)\}-\alpha_j]\,, \tag{4b}$$

where $\delta=\alpha^{-1}$ *and* $G=L_{2,\alpha}$

$$\alpha_i=\Gamma(i-\delta)/\Gamma(i) \quad \text{and} \quad v_{ij}=\alpha_i[\{\Gamma(j-2\delta)/\Gamma(j-\delta)\}-\alpha_j]\,, \tag{4c}$$

where $\delta=\alpha^{-1}<\min(i, j/2)$ *if* $G=L_{1,\alpha}$.

PROOF. Since $\mathrm{E}(Z_i)=1$ and $\mathrm{Var}(Z_i)=1$ and Z_i's are all mutually independent, (4a) follows directly from (3a). In the $L_{2,\alpha}$ case, since $U\equiv\sum_{l=1}^{i} Z_l$ is a Gamma $(1, i)$ variable, $\alpha_i=\mathrm{E}\,U^\delta=\Gamma(i+\delta)/\Gamma(i)$. Further $v_{ij}=\mathrm{E}\left(\sum_{l=1}^{i} Z_l\right)^\delta\left(\sum_{l=1}^{j} Z_l\right)^\delta-\alpha_i\alpha_j$ and the first term on the right side is $\mathrm{E}\,U^\delta(U+Y)^\delta$ where $Y=\sum_{l=i+1}^{j} Z_l$ is Gamma $(1, j-i)$ and U and Y are independent. Hence the first term is

$$\int_0^\infty\int_0^\infty u^\delta(u+y)^\delta\frac{e^{-u}u^{i-1}}{\Gamma(i)}e^{-y}\frac{y^{(j-i)-1}}{\Gamma(j-i)}dudy$$

$$=\frac{\Gamma(i+\delta)}{\Gamma(i)}\mathrm{E}\,(U^*+Y)^\delta \qquad \text{where } U^* \text{ is Gamma } (1, i+\delta)$$

$$=\alpha_i\,\mathrm{E}\,(W^\delta) \qquad \text{where } W \text{ is Gamma } (1, j+\delta)$$

$$=\alpha_i\Gamma(j+2\delta)/\Gamma(j+\delta)\,.$$

Hence (4b) follows. The relation (4c) follows similarly. Existence of v_{ij} in this case requires that $\delta<\min(i, j/2)$ so that the gamma functions involved are convergent.

In view of Lemma 2 it is clear that existence of $\boldsymbol{V}$, the covariance matrix of T_i's, when $G=L_{1,\alpha}$ requires that $\alpha>2$. However in all the three cases we have for $i\leqq j$, $v_{ij}=a_ib_j$. The following lemma, which has been known for quite some time (see, e.g., Greenberg and Sarhan [5], p. 757), exhibits $\boldsymbol{V}^{-1}$ for such a matrix $\boldsymbol{V}$.

LEMMA 3. *Let $\boldsymbol{V}=(v_{ij})$ be an $r\times r$ nonsingular matrix with $v_{ij}=a_ib_j$, $i\leqq j$. Then v^{ij}, the (i, j)th element of $\boldsymbol{V}^{-1}$ is given by*

$$(5)\qquad v^{ij}=\begin{cases} \dfrac{a_{i+1}b_{i-1}-a_{i-1}b_{i+1}}{(a_ib_{i-1}-a_{i-1}b_i)(a_{i+1}b_i-a_ib_{i+1})}\,, & i=j=2 \text{ to } r-1 \\ -(a_{i+1}b_i-a_ib_{i+1})^{-1}\,, & j=i+1 \text{ and } i=1 \text{ to } r-1 \\ a_2/a_1(a_2b_1-a_1b_2)\,, & i=j=1 \\ b_{r-1}/b_r(a_rb_{r-1}-a_{r-1}b_r)\,, & i=j=r \\ 0\,, & |i-j|>1\,. \end{cases}$$

and $v^{(i+1)i}=v^{i(i+1)}$.

The lemma follows by the direct manipulation of the fact that $\boldsymbol{V}\boldsymbol{V}^{-1}=\boldsymbol{I}$.

Remarks. (1) Lemma 3 shows that the elements of $\boldsymbol{V}^{-1}$ are well defined if and only if a_1 and b_r are nonzero and $a_ib_{i-1}\neq a_{i-1}b_i$ for $i=2$ to r. Hence this serves as a necessary and sufficient condition for the nonsingularity of $\boldsymbol{V}$. This is satisfied for $\boldsymbol{V}$ where the v_{ij}'s are given by $(4a-c)$.
(2) Let $\boldsymbol{\omega}'=(v_{1s}, v_{2s},\cdots, v_{rs})$ for $s\geqq r$. Then if $v_{ij}=a_ib_j$, $i\leqq j$ and $\boldsymbol{V}=(v_{ij})_{r\times r}$ is nonsingular, $\boldsymbol{\omega}\boldsymbol{V}^{-1}=(0, 0,\cdots, 0, b_s/b_r)$. This is because $\boldsymbol{\omega}'$ is (b_s/b_r) times the last row of $\boldsymbol{V}$. Kaminsky and Nelson [8] note that

$$(6)\qquad \boldsymbol{\omega}'\boldsymbol{V}^{-1}=(0, 0,\cdots, 0, q)$$

where $q=v_{ir}/v_{is}$, for $i=1$ to r, for order statistics themselves if the standardized df F belongs to the family containing exponential, power function and Pareto, and the negative versions of these distributions. Our discussion shows that (6) holds for the 'limiting' order statistics from any parent distribution provided the limit distributions exist; or in other words, it holds for upper record values from $L_{1,\alpha}$, $L_{2,\alpha}$ and L_3.

We now give a necessary and sufficient condition for a covariance matrix $\boldsymbol{V}$ to have $v_{ij}=a_ib_j$, $i\leqq j$ in terms of a general version of (6).

LEMMA 4. *Let Y_i, $1 \leqq i \leqq n$ be n random variables with $v_{ij} = \mathrm{Cov}(Y_i, Y_j)$. Let $\boldsymbol{V}_k$ be the covariance matrix of Y_i, $1 \leqq i \leqq k$ and $\boldsymbol{\omega}_k = (v_{1k}, v_{2k}, \cdots, v_{k-1,k})$. Assume that $\boldsymbol{V}_n^{-1}$ exists. Then $v_{ij} = a_i b_j$, $1 \leqq i \leqq j \leqq n$ if and only if*

$$(7) \qquad \boldsymbol{\omega}'_{k+1}\boldsymbol{V}_k^{-1} = (0, 0, \cdots, 0, b_{k+1}/b_k) , \qquad 1 \leqq k \leqq n-1 .$$

PROOF. The fact that $v_{ij} = a_i b_j$ implies (7) follows from remark (2) above. To prove the converse, note that (7) implies that $\boldsymbol{\omega}'_{k+1}\boldsymbol{V}_k^{-1}\boldsymbol{V}_k \equiv \boldsymbol{\omega}'_{k+1} = (b_{k+1}/b_k)(v_{1k}, v_{2k}, \cdots, v_{kk})$, $1 \leqq k \leqq n-1$. Putting $k=1$, we have $v_{12} = (b_2/b_1)v_{11}$; $k=2$ yields $(v_{13}, v_{23}) = b_3 b_2^{-1}(v_{12}, v_{22}) = (b_3 b_1^{-1} v_{11}, b_3 b_2^{-1} v_{22})$ and in general

$$(v_{1,k+1}, \cdots, v_{k,k+1}) = \left(\frac{b_{k+1}}{b_1} v_{11}, \frac{b_{k+1}}{b_2} v_{22}, \cdots, \frac{b_{k+1}}{b_k} v_{kk}\right) , \qquad 1 \leqq k \leqq n-1 .$$

Now defining $a_i = v_{ii} b_i^{-1}$ it follows that $v_{ij} = a_i b_j$ for $i \leqq j$.

3. General theory of linear prediction

In this section we enumerate results of interest from the general theory of linear estimation and prediction in location and scale families. Let $\boldsymbol{X}' = (T_1, T_2, \cdots, T_r)$ denote the vector of first r record values from a df $G(\mu + \sigma x)$. Then $\mathrm{E}(\boldsymbol{X}) = \mu \mathbf{1} + \sigma \boldsymbol{\alpha}$ where $\mathbf{1}' = (1, 1, \cdots, 1)$ and $\boldsymbol{\alpha}' = (\alpha_1, \cdots, \alpha_r)$ with $\alpha_k = \mathrm{E}\{(T_k - \mu)/\sigma\}$. Let $\sigma^2 \boldsymbol{V}$ be the covariance matrix of the T_i's. Then (see, for example David [2], p. 130) the best linear unbiased estimates (BLUE) of μ and σ are given by $\hat{\mu} = -\boldsymbol{\alpha}'\boldsymbol{V}^{-1}(\mathbf{1}\boldsymbol{\alpha}' - \boldsymbol{\alpha}\mathbf{1}')\boldsymbol{V}^{-1}\boldsymbol{X}/\Delta$ and $\hat{\sigma} = \mathbf{1}'\boldsymbol{V}^{-1}(\mathbf{1}\boldsymbol{\alpha}' - \boldsymbol{\alpha}\mathbf{1}')\boldsymbol{V}^{-1}\boldsymbol{X}/\Delta$ where $\Delta = (\mathbf{1}'\boldsymbol{V}^{-1}\mathbf{1})(\boldsymbol{\alpha}'\boldsymbol{V}^{-1}\boldsymbol{\alpha}) - (\mathbf{1}'\boldsymbol{V}^{-1}\boldsymbol{\alpha})^2$. Further $\mathrm{Var}(\hat{\mu}) = \sigma^2 \boldsymbol{\alpha}'\boldsymbol{V}^{-1}\boldsymbol{\alpha}/\Delta$, $\mathrm{Var}(\hat{\sigma}) = \sigma^2 \mathbf{1}'\boldsymbol{V}^{-1}\mathbf{1}/\Delta$ and $\mathrm{Cov}(\hat{\mu}, \hat{\sigma}) = -\sigma^2 \mathbf{1}'\boldsymbol{V}^{-1}\boldsymbol{\alpha}/\Delta$. If $\boldsymbol{\omega}' = (v_{1s}, v_{2s}, \cdots, v_{rs})$, where $v_{ij} = \sigma^2 \mathrm{Cov}(T_i, T_j)$, then the BLUP of T_s is (see, e.g., Kaminsky and Nelson [8], p. 146)

$$\hat{T}_s = \hat{\mu} + \alpha_s \hat{\sigma} + \boldsymbol{\omega}'\boldsymbol{V}^{-1}(\boldsymbol{X} - \hat{\mu}\mathbf{1} - \hat{\sigma}\boldsymbol{\alpha})$$

and the BLIP of T_s is (see, Kaminsky, Mann and Nelson [7], p. 525)

$$\tilde{T}_s = \hat{T}_s - \{C_{12}/(1 + C_{22})\}\hat{\sigma}$$

where

$$C_{12}\sigma^2 = \mathrm{Cov}\{\hat{\sigma}, (1 - \boldsymbol{\omega}'\boldsymbol{V}^{-1}\mathbf{1})\hat{\mu} + (\alpha_s - \boldsymbol{\omega}'\boldsymbol{V}^{-1}\boldsymbol{\alpha})\hat{\sigma}\} , \qquad C_{22}\sigma^2 = \mathrm{Var}(\hat{\sigma}) .$$

Kaminsky and Nelson [9] note that Watson's [13] result implies that $\boldsymbol{\omega}'\boldsymbol{V}^{-1}(\boldsymbol{X} - \hat{\mu}\mathbf{1} - \hat{\sigma}\boldsymbol{\alpha}) = 0$ if and only if $\boldsymbol{\omega}$ is in the column space of $(\mathbf{1}, \boldsymbol{\alpha})$. From (4a–c) it follows that $\boldsymbol{\omega}' = b_s \mathbf{1}'$ in the L_3 case and $\boldsymbol{\omega}' = b_s \boldsymbol{\alpha}'$ in the remaining two cases. However when $G = L_{1,\alpha}$ we have only a scale parameter family and hence the prediction problem is simpler. Hence we have

$$(8) \qquad \hat{T}_s=\hat{\mu}+\alpha_s\hat{\sigma}$$

when $G=L_3$ or $L_{2,\alpha}$. Also since $\boldsymbol{\omega}'\boldsymbol{V}^{-1}=(0, 0, \cdots, 0, b_s/b_r)$ we have

$$(9) \qquad \tilde{T}_s=\hat{T}_s-\{C_{12}/(1+C_{22})\}\hat{\sigma}$$

where

$$C_{12}=\{(\alpha_s-b_s\alpha_r b_r^{-1})\mathbf{1}'\boldsymbol{V}^{-1}\mathbf{1}-(1-b_s b_r^{-1})\mathbf{1}'\boldsymbol{V}^{-1}\boldsymbol{\alpha}\}/\Delta\ , \qquad C_{22}=(\mathbf{1}'\boldsymbol{V}^{-1}\mathbf{1})/\Delta\ .$$

Further, with $C_{11}\sigma^2=\mathrm{Var}\,\{(1-b_s b_r^{-1})\hat{\mu}+(\alpha_s-b_s\alpha_r b_r^{-1})\hat{\sigma}\}$,

$$(10) \quad M(\hat{T}_s)=\mathrm{E}\,(T_s-\hat{T}_s)^2=\sigma^2(v_{ss}-\boldsymbol{\omega}'\boldsymbol{V}^{-1}\boldsymbol{\omega}+C_{11})=\sigma^2(a_s b_s-a_r b_s^2 b_r^{-1}+C_{11})$$

and

$$M(\tilde{T}_s)=M(\hat{T}_s)-\{C_{12}^2/(1+C_{22})\}\sigma^2\ .$$

In the $L_{1,\alpha}$ case $\hat{\sigma}=(\boldsymbol{\alpha}'\boldsymbol{V}^{-1}\boldsymbol{X})/(\boldsymbol{\alpha}'\boldsymbol{V}^{-1}\boldsymbol{\alpha})$ (see Lloyd [10], p. 25) and $\hat{T}_s=\alpha_s\hat{\sigma}$.

We consider the detailed discussion of the three cases one by one in the next section.

4. Detailed discussion

Case (i) $G=L_3$

Here $a_i=1$, $v_{ij}=b_j=\sum_{l=j}^{\infty} l^{-2}$ and $\alpha_i=S_i-\gamma$ where $S_i=\sum_{l=1}^{i-1} l^{-1}$, $i\geqq 2$ and $S_1=0$. From these, one obtains $\mathbf{1}'\boldsymbol{V}^{-1}=(0, 0, \cdots, 0, b_r^{-1})$, $\boldsymbol{\alpha}'\boldsymbol{V}^{-1}=(-1, \cdots, -1, \alpha_r b_r^{-1}+(r-1))$, $\boldsymbol{\alpha}'\boldsymbol{V}^{-1}\boldsymbol{\alpha}=\alpha_r^2 b_r^{-1}+(r-1)$, $\mathbf{1}'\boldsymbol{V}^{-1}\boldsymbol{\alpha}=\alpha_r b_r^{-1}$, $\mathbf{1}'\boldsymbol{V}^{-1}\mathbf{1}=b_r^{-1}$ and consequently $\Delta=(r-1)b_r^{-1}$. These, on substitution yield $\hat{\mu}=T_r(1+\gamma-S_r)+(S_r-\gamma)\bar{T}_{r-1}$ and $\hat{\sigma}=T_r-\bar{T}_{r-1}$ where $\bar{T}_{r-1}=\sum_{i=1}^{r-1} T_i/(r-1)$. This fact also follows from the work of Weissman [14] on upper extremes where he essentially shows that $\hat{\mu}$ and $\hat{\sigma}$ are in fact minimum variance unbiased estimators of μ and σ respectively.

Now from (8) the BLUP of T_s is given by $\hat{T}_s=T_r+(S_s-S_r)(T_r-\bar{T}_{r-1})$. Since $C_{12}/(1+C_{22})=(\alpha_s-\alpha_r)/r=(S_s-S_r)/r$, on simplification, the BLIP of T_s reduces to $\tilde{T}_s=T_r+\{(r-1)/r\}(S_s-S_r)(T_r-\bar{T}_{r-1})$. Also one obtains $M(\hat{T}_s)=\sigma^2(b_r-b_s+(S_r-S_s)^2(r-1)^{-1})$ and $M(\tilde{T}_s)=\sigma^2(b_r-b_s+(S_r-S_s)^2\cdot r^{-1})$ so that the improvement in the mean square error (MSE) is $(S_r-S_s)^2/r(r-1)$.

Case (ii) $G=L_{2,\alpha}$

In this case the additional parameter $\delta\,(=\alpha^{-1})$ complicates the problem. First we consider the case where δ is unknown and show that no linear unbiased estimates (LUE) of the parameters exist. Then we obtain the predictors using the discussion of the preceding section, when δ is known.

LEMMA 5. *If δ is unknown no* LUE *of μ or σ or δ exists.*

PROOF. From the representation in (3b) and the fact $\mathrm{E}(Z_1+\cdots+Z_k)^\delta=\Gamma(k+\delta)/\Gamma(k)$ we have

$$\mathrm{E}\left(\sum l_k T_k\right)=\mu\sum l_k+\sigma\sum l_k\Gamma(k+\delta)/\Gamma(k) .$$

Hence for $\sum l_k T_k$ to be unbiased for μ, we should have $\sum l_k=1$ and $\sum l_k\Gamma(k+\delta)/\Gamma(k)=0$ for all δ. The latter equation can be written as

$$\sum_{k=1}^{r} m_k\delta(\delta+1)\cdots(\delta+k-1)=0 \qquad \text{for all } \delta>0 \tag{11}$$

where $m_k=l_k/\Gamma(k)$, since $\Gamma(\delta)$ is nonzero. The equation (11) being an rth degree polynomial equation, more than r solutions imply that each $m_k=0$. This in turn implies that $l_k=0$ for all k. This contradicts the fact that $\sum l_k=1$. Hence no LUE of μ exists.

Now for $\sum l_k T_k$ to be unbiased for σ, one should have $\sum l_k=0$ and

$$\sum l_k\Gamma(k+\delta)/\Gamma(k)=1 \qquad \text{for } \delta>0 . \tag{12}$$

Using (12) with δ and $\delta^*=\delta+1$, on substraction one obtains

$$\sum_{k=1}^{r}\left\{l_k\frac{\Gamma(k+\delta+1)}{\Gamma(k)}-l_k\frac{\Gamma(k+\delta)}{\Gamma(k)}\right\}=0 \qquad \text{for all } \delta>0 .$$

That is

$$l_r\frac{\Gamma(r+\delta+1)}{\Gamma(r)}+\sum_{k=1}^{r-1}(l_k-l_{k-1})\frac{\Gamma(k+\delta+1)}{\Gamma(k)}-l_1\frac{\Gamma(1+\delta)}{\Gamma(1)}=0 .$$

In other words $\sum_{j=1}^{r+1} m_j^*\Gamma(j+\delta)=0$ for all $\delta>0$. Now proceeding as before one obtains $m_j^*=0$, $j=1$ to $r+1$. This used sequentially, would give $l_j=0$ for all $j=1$ to r, contradicting (12). That is, a LUE of σ does not exist. Note that in claiming this we did not use the fact that $\sum l_k=0$; hence whether μ is known or not a LUE of σ does not exist.

Similarly one can show that no LUE of δ exists even if μ and σ are known.

If δ is known, BLUE's of μ and σ can be obtained using the general discussion considered earlier. Substantial simplification is possible since $v_{ij}=a_i b_j$, $i\leqq j$ where $a_i=\Gamma(i+\delta)/\Gamma(i)$ and $b_j=\Gamma(j+2\delta)/\Gamma(j+\delta)-a_j$.

On simplification one obtains $\Delta=\alpha_r(\mathbf{1}'\boldsymbol{V}^{-1}\mathbf{1})/b_r-b_r^{-2}$, $\hat{\mu}=\{-(T_r/b_r)+\alpha_r\mathbf{1}'\boldsymbol{V}^{-1}\cdot\boldsymbol{X}\}/b_r\Delta$ and $\hat{\sigma}=\{(\mathbf{1}'\boldsymbol{V}^{-1}\mathbf{1})T_r-\mathbf{1}'\boldsymbol{V}^{-1}\boldsymbol{X}\}/b_r\Delta$ where the elements of $\boldsymbol{V}^{-1}$ can be obtained using (5). In fact

$$v^{ii}=\frac{\Gamma(i)}{\Gamma(i+2\delta)\delta^2}[(i+\delta)^2+(i-1)(i-1+2\delta)]\ , \qquad 2\leqq i\leqq r-1$$

$$v^{i(i+1)}=v^{(i+1)i}=-\frac{\Gamma(i+1)(i+\delta)}{\delta^2\Gamma(i+2\delta)}\ , \qquad 1\leqq i\leqq r-1\ ,$$

$$v^{11}=\frac{(1+\delta)^2}{\delta^2\Gamma(1+2\delta)}\ , \qquad v^{rr}=\frac{b_{r-1}(r-1+\delta)\Gamma(r)}{b_r\delta^2\Gamma(r-1+2\delta)}$$

and $v^{ij}=0$ otherwise.

Now $\hat{T}_s=\hat{\mu}+\alpha_s\hat{\sigma}$ and $\tilde{T}_s=\hat{T}_s-\{C_{12}/(1+C_{22})\}\hat{\sigma}$ where C_{12} and C_{22} are as given in (9) with $\mathbf{1}'\boldsymbol{V}^{-1}\boldsymbol{\alpha}=b_r^{-1}$. Further the MSE's are given by (10) with

$$C_{11}=\left\{\left(1-\frac{b_s}{b_r}\right)^2\frac{\alpha_r}{b_r}+\left(\alpha_s-b_s\frac{\alpha_r}{b_r}\right)^2(\mathbf{1}'\boldsymbol{V}^{-1}\mathbf{1})-\frac{2}{b_r}\left(1-\frac{b_s}{b_r}\right)\left(\alpha_s-b_s\frac{\alpha_r}{b_r}\right)\right\}\Big/\Delta\ .$$

To illustrate the application of these formulas, some computation was carried out and the results are recorded below. The following tables give the coefficients of the T_i's for predicting T_s by $\hat{T}_s$ and $\tilde{T}_s$ as well as $M(\hat{T}_s)/\sigma^2$ and $M(\tilde{T}_s)/\sigma^2$. Table 1 gives these values for $r=5$, $s=6, 7, 10$ and Table 2 gives the coefficients and the MSE's for $r=10$, $s=11, 12, 15$. The last two rows give the coefficients of T_i's for the BLUE's of μ and σ respectively. Both the tables restrict their attention to $\delta=0.5$ and 1.5 only. The case where $\delta=1$ has been handled earlier by Ahsanullah [1] when he considered the linear prediction of record values in the two-parameter exponential distribution. So when $\delta=1$, $\hat{\mu}=(rT_1-T_r)/(r-1)$, $\hat{\sigma}=(T_r-T_1)/(r-1)$, $\hat{T}_s=((s-1)T_r-(s-r)T_1)/(r-1)$, $\tilde{T}_s=(sT_r-(s-r)T_1)/r$, $M(\hat{T}_s)=\sigma^2(s-r)(s-1)/(r-1)$ and $M(\tilde{T}_s)=\sigma^2 s(s-r)/r$.

Case (iii) $G=L_{1,\alpha}$

In this set-up we have a scale parameter family with parameters $\sigma\,(=d_n)$ and $\delta\,(=\alpha^{-1})$. As in the $L_{1,\alpha}$ case, one can show that no LUE exists either for σ or for δ when δ is unknown. If δ is known, $\hat{\sigma}=T_r/\alpha_r$ and $\hat{T}_s=\alpha_s T_r/\alpha_r$ where $\alpha_j=-\Gamma(j-\delta)/\Gamma(j)$. Note that even though application of the general theory requires that $\alpha>2$ or $\delta<1/2$, for the finite variance of $\hat{\sigma}$ one needs only $\delta<r/2$. Hence these formulas can be used as long as $\delta<r/2$. However note that in order to use these we should have $T_r<0$ since the support of $L_{1,\alpha}$ is $(-\infty, 0)$!

5. Examples

1. Suppose we are observing life data following a normal distribution. From a large random sample of size n, we have observed the first r failures $X_{1:n}, \cdots, X_{r:n}$ and would like to predict $X_{s:n}$. If n is large compared to s, we can use the formulas derived in the last section. Since the normal distribution is in $\mathcal{D}(L_3)$, the ABLUP of $X_{s:n}$ is $\hat{X}_{s:n} = X_{r:n} + (S_s - S_r)(X_{r:n} - \bar{X}_{r-1})$ and the ABLIP of $X_{s:n}$ is $\tilde{X}_{s:n} = X_{r:n} + \{(r-1)/r\}(S_s - S_r)(X_{r:n} - \bar{X}_{r-1})$ where $S_i = \sum_{j=1}^{i-1} j^{-1}$, and $\bar{X}_{r-1}$ is the mean of the first $(r-1)$ order statistics. One can also estimate the mean square errors of prediction by using $\hat{\sigma}$ as an estimate of σ. These formulas remain the same for any distribution in $\mathcal{D}(L_3)$. The lognormal distribution which is used quite often as a model in life data is one example.

2. The Weibull, gamma and beta distributions, used extensively in life testing and reliability problems, are in $\mathcal{D}(L_{2,\alpha})$. For these distributions, $c_n = \mu$, the threshold parameter and α, the shape parameter. Recently Weissman [15] has obtained confidence intervals for μ using asymptotic theory. Our discussion in the last section gives $\hat{\mu}$, asymp-

Table 1. The coefficients of T_i's when $r=5$

Predictor/estimator \ i	1	2	3	4	5	MSE/σ^2
	$\delta=0.5$					
$\hat{T}_6$	−.1440	−.0240	−.0160	−.0120	1.1960	.0548
$\tilde{T}_6$	−.1194	−.0199	−.0133	−.0010	1.1626	.0532
$\hat{T}_7$	−.2760	−.0460	−.0307	−.0230	1.3757	.1180
$\tilde{T}_7$	−.2297	−.0383	−.0255	−.0191	1.3126	.1123
$\hat{T}_{10}$	−.6220	−.1037	−.0691	−.0518	1.8466	.3420
$\tilde{T}_{10}$	−.5212	−.0869	−.0579	−.0434	1.7094	.3151
$\hat{\mu}$	1.4400	.2400	.1600	.1200	−.9600	.4800
$\hat{\sigma}$	−.6603	−.1100	−.0734	−.0550	.8987	.1521
	$\delta=1.5$					
$\hat{T}_6$	−.3571	.0179	.0071	.0036	1.3286	19.2857
$\tilde{T}_6$	−.2692	.0135	.0054	.0027	1.2477	18.1172
$\hat{T}_7$	−.7440	.0372	.0149	.0074	1.6845	52.2054
$\tilde{T}_7$	−.5564	.0278	.0111	.0056	1.5119	46.8819
$\hat{T}_{10}$	−2.0640	.1032	.0413	.0206	2.8988	262.2638
$\tilde{T}_{10}$	−1.5159	.0758	.0303	.0152	2.3946	216.8445
$\hat{\mu}$	1.1905	−.0595	−.0238	−.0119	−.0952	4.2857
$\hat{\sigma}$	−.0992	.0050	.0020	.0010	.0913	.4893

totically BLUE (ABLUE) of μ as well as ABLUP and ABLIP of $X_{s:n}$. One can use the coefficients of the T_i's given in Tables 1 and 2 as the coefficients of $X_{i:n}$'s for this purpose. One can also estimate $M(T_s)$ and $M(\tilde{T}_s)$ by replacing σ by $\hat{\sigma}$, the ABLUE of σ. The last column in the tables give MSE/σ^2 and $\hat{\sigma}$ can be obtained using the coefficients given in the tables once the sample data is known. These coefficients remain the same for any distribution in $\mathscr{D}(L_{2,\alpha})$ where α is known.

Table 2. The coefficients of T_i's when $r=10$

Predictor/estimator \ i	1 (6)	2 (7)	3 (8)	4 (9)	5 (10)	MSE/σ^2
			$\delta=0.5$			
$\hat{T}_{11}$	−.0530 (−.0029)	−.0088 (−.0025)	−.0059 (−.0022)	−.0044 (−.0020)	−.0035 (1.0853)	.0259
$\tilde{T}_{11}$	−.0488 (−.0027)	−.0081 (−.0023)	−.0054 (−.0020)	−.0041 (−.0018)	−.0033 (1.0785)	.0257
$\hat{T}_{12}$	−.1036 (−.0058)	−.0173 (−.0049)	−.0115 (−.0043)	−.0086 (−.0038)	−.0069 (1.1668)	.0534
$\tilde{T}_{12}$	−.0954 (−.0053)	−.0159 (−.0045)	−.0106 (−.0040)	−.0080 (−.0035)	−.0064 (1.1536)	.0528
$\hat{T}_{15}$	−.2438 (−.0135)	−.0406 (−.0116)	−.0271 (−.0102)	−.0203 (−.0090)	−.0163 (1.3924)	.1441
$\tilde{T}_{15}$	−.2250 (−.0125)	−.0375 (−.0107)	−.0250 (−.0094)	−.0187 (−.0083)	−.0150 (1.3621)	.1408
$\hat{\mu}$	1.0605 (.0589)	.1767 (.0505)	.1178 (.0442)	.0884 (.0393)	.0707 (−.7070)	.3535
$\hat{\sigma}$	−.3396 (−.0189)	−.0566 (−.0162)	−.0377 (−.0141)	−.0283 (−.0126)	−.0226 (.5466)	.0615
			$\delta=1.5$			
$\hat{T}_{11}$	−.1698 (.0006)	.0085 (.0004)	.0034 (.0003)	.0017 (.0002)	.0010 (1.1537)	29.7917
$\tilde{T}_{11}$	−.1472 (.0005)	.0074 (.0004)	.0029 (.0002)	.0015 (.0002)	.0008 (1.1332)	29.2638
$\hat{T}_{12}$	−.3472 (.0012)	.0174 (.0008)	.0069 (.0006)	.0035 (.0004)	.0020 1.3144	70.6449
$\tilde{T}_{12}$	−.3002 (.0011)	.0150 (.0007)	.0060 (.0005)	.0030 (.0004)	.0017 (1.2718)	68.3590
$\hat{T}_{15}$	−.9229 (.0033)	.0461 (.0022)	.0185 (.0015)	.0092 (.0011)	.0053 1.8356	273.4775
$\tilde{T}_{15}$	−.7926 (.0028)	.0396 (.0019)	.0159 (.0013)	.0079 (.0010)	.0045 (1.7176)	255.9214
$\hat{\mu}$	1.1317 (−.0040)	−.0566 (−.0027)	−.0226 (−.0019)	−.0113 (−.0014)	−.0065 (−.0247)	4.0741
$\hat{\sigma}$	−.0345 (.0001)	.0017 (.0001)	.0007 (.0001)	.0003 (.0000)	.0002 (.0312)	.2314

Acknowledgements

This work was supported by a Small Research Grant from the Graduate School of The Ohio State University.

THE OHIO STATE UNIVERSITY

REFERENCES

[1] Ahsanullah, M. (1980). Linear prediction of record values for the two parameter exponential distribution, *Ann. Inst. Statist. Math.*, **32**, 363-368.

[2] David, H. A. (1981). *Order Statistics*, Second Edition, Wiley, New York.

[3] Galambos, J. (1978). *The Asymptotic Theory of Extreme Order Statistics*, Wiley, New York.

[4] Gnedenko, B. (1943). Sur la distribution limite du terme maximum d'une serie aleatoire, *Ann. Math.*, **44**, 423-453.

[5] Greenberg, B. G. and Sarhan, A. E. (1959). Matrix inversion, its interest and application in analysis of data, *J. Amer. Statist. Ass.*, **54**, 755-766.

[6] Hall, P. (1978). Representation and limit theorems for extreme value distributions, *J. Appl. Prob.*, **15**, 639-644.

[7] Kaminsky, K. S., Mann, N. R. and Nelson, P. I. (1975). Best and simplified linear invariant prediction of order statistics in location and scale families, *Biometrika*, **62**, 525-527.

[8] Kaminsky, K. S. and Nelson, P. I. (1975a). Best linear unbiased prediction of order statistics in location and scale families, *J. Amer. Statist. Ass.*, **70**, 145-150.

[9] Kaminsky, K. S. and Nelson, P. I. (1975b). Characterization of distributions by the form of predictors of order statistics, *Statistical Distributions in Scientific Work*, (eds. Patil, G. P. et al.), Reidel Publishing Co. Dordercht-Holland, **3**, 113-115.

[10] Lloyd, E. H. (1962). Generalized least-squares theorem, *Contributions to Order Statistics* (eds. Sarhan, A. E. and Greenberg, B. G.), Wiley, New York, 20-27.

[11] Nagaraja, H. N. (1982). Record values and extreme value distributions, *J. Appl. Prob.*, **19**, 233-239,

[12] Nelson, W. (1982). *Applied Life Data Analysis*, Wiley, New York.

[13] Watson, G. S. (1972). Prediction and the efficiency of least squares, *Biometrika*, **59**, 91-98.

[14] Weissman, I. (1978). Estimation of parameters and large quantiles based on the k largest observations, *J. Amer. Statist. Ass.*, **73**, 812-815.

[15] Weissman, I. (1981). Confidence intervals for the threshold parameters, *Comm. Statist. Theor. Method*, A, **10**, 549-557.