D-optimal joint best linear unbiased prediction of order statistics

Narayanaswamy Balakrishnan\textsuperscript{1} and Ritwik Bhattacharya\textsuperscript{2}

\textsuperscript{1}Department of Mathematics and Statistics, McMaster University, Hamilton, ON L8S 4K1, Canada
\textsuperscript{2}Department of Industrial Engineering, School of Engineering and Sciences, Tecnológico de Monterrey, Querétaro 76130, México

Abstract

In life-testing experiments, it is often of interest to predict unobserved future failure times based on observed early failure times. A point best linear unbiased predictor (BLUP) has been developed in this context by Kaminsky and Nelson (1975). In this article, we develop joint BLUPs of two future failure times based on early failure times by minimizing the determinant of the variance-covariance matrix of the predictors. The advantage of applying joint prediction is demonstrated by using a real data set. The non-existence of joint BLUPs in certain setups is also discussed.

Keywords: Best linear unbiased estimate (BLUE), Best linear unbiased predictor (BLUP), Location-scale family of distribution, Lagrangian method, Order statistics, Scale family of distributions, Type-II right censored samples, Variance-covariance matrix.

1 Introduction

The issue of predicting future unobserved failure times is of great interest in reliability life-testing experiments. The aim in this case is to predict the unobserved failure times based on observed early failures as data. For instance, let us assume a continuous distribution with probability density function (pdf)

\[ \frac{1}{\sigma} f \left( \frac{x - \mu}{\sigma} \right), \]

where \( \mu \) and \( \sigma \) are the location and scale parameters, respectively. Let \( X_{1:n} < \cdots < X_{n:n} \) denote the ordered observations from (1). Suppose the first \( r \) order statistics, representing a Type-II right censored sample, are observed. Then, our main interest is in predicting the future \((n - r)\) unobserved failure times, that is, \((r + 1)\)th, \((r + 2)\)th, \cdots, \(n\)th failure times, based on the first \( r \) order statistics observed. This can also be viewed in the context
of an \( n \)-component parallel system. Based on the information on the first \( r \) components that have already failed, we may wish to predict the system failure time, which is simply the \( n \)th ordered failure time. Kaminsky and Nelson (1975) obtained the best linear unbiased predictor (BLUP) of \( X_{s:n} \) based on the observed values of \( X_{1:n} < X_{2:n} < \cdots < X_{r:n} \), where \( 1 \leq r < s \leq n \). Their predictor was developed based on the result of Goldberger (1962), who studied the best linear unbiased prediction in a generalized linear regression model. Doganaksoy and Balakrishnan (1997) showed, under Gauss-Markov model, that the best linear unbiased estimators of the model parameters remain unchanged if the predicted values of the dependent variables based on best linear unbiased predictors are treated as observed values in the best linear unbiased estimation of parameters. This property does simplify considerably the computation of point predictors. The best linear unbiased estimations, in addition to this feature, also posses some more interesting properties as shown by Balakrishnan and Rao (1997). Due to the great importance of this prediction in life-testing experiments, Nelson (2003) has presented numerous tables for the best linear unbiased prediction of order statistics for different lifetime distributions of interest.

Prediction of order statistics can also be used for detecting outliers (see Balasooriya, 1989). An interval prediction is commonly considered instead of point prediction. However, the interval prediction requires approximate methods for most models, with the exception of a few like exponential. A detailed survey of various prediction intervals can be found in Patel (1989). For a comprehensive review of prediction problems in ordered data, one may refer to Kaminsky and Nelson (1998). While there are many such works on point prediction and prediction intervals for ordered data, the joint prediction of order statistics has not been dealt with. By joint prediction, we mean the simultaneous prediction of two or more order statistics. This provides the motivation for the present work. In subsequent sections, we derive analytical expressions for the joint best linear unbiased predictors of two future order statistics by minimizing the determinant of the variance-covariance matrix of the predictors, resulting in the joint predictors being D-optimal. The advantage of joint predictors over marginal predictors is demonstrated by analyzing the design efficiency. The non-existence of joint BLUPs in certain situations is also established.

The rest of this paper is organized as follows. Analytic expressions of the joint predictors are derived in detail in Section 2. The non-existence of BLUPs in certain specific situations is established in Section 3. The proposed method is illustrated through some numerical results in Section 4. A real-
life data is analyzed to demonstrate the advantage of joint predictors over marginal predictors. Finally, some concluding remarks are made in Section 5.

2 Joint best linear unbiased predictors

In this section, we derive explicit expressions for the joint predictors of two order statistics under D-optimality criterion. For this purpose, let us denote the vector of first $r$ order statistics (observed from a life-test) from a sample of size $n$ by

$$X = (X_{1:n}, \ldots, X_{r:n})'_{r \times 1}.$$ 

We are then interested in the joint predictors $\tilde{X}_s$ and $\tilde{X}_t$ of $X_{s:n}$ and $X_{t:n}$, respectively, where $r < s < t \leq n$. Let $\alpha_i$ denote the expected value of the standardized order statistic $Z_i = X_i : n - \mu / \sigma$, $i = 1, \ldots, n$, and let us further denote

$$\alpha = (\alpha_1, \ldots, \alpha_r)'_{r \times 1}.$$ 

Also, let us denote the variance-covariance matrix of $X$ by $\sigma^2 \Sigma$, where $\Sigma$ is the $r \times r$ covariance matrix of $Z_{i:n}, i = 1, \ldots, r$. In this notation, the marginal best linear unbiased predictor $\hat{X}_s$ of $X_{s:n}$ has been derived by Kaminsky and Nelson (1975), using the results of Goldberger (1962), as

$$\hat{X}_s = \hat{\mu} + \hat{\sigma} \alpha_s + \omega'_s \Sigma^{-1}(X - \hat{\mu} 1 - \hat{\sigma} \alpha),$$ 

where $1 = (1, \ldots, 1)'_{r \times 1}$, and $\omega_s = (\omega_1, \ldots, \omega_r)'_{r \times 1}$, with $\omega_i = \text{Cov}(Z_{i:n}, Z_{s:n})$. Similarly, the marginal best linear unbiased predictor $\hat{X}_t$ of $X_{t:n}$ is exactly as in (2) with $\alpha_s$ and $\omega'_s$ being replaced by $\alpha_t$ and $\omega'_t$, respectively. In (2), $\hat{\mu}$ and $\hat{\sigma}$ are the best linear unbiased estimates (BLUEs) of $\mu$ and $\sigma$, respectively, based on $X$, given by

$$\hat{\mu} = \frac{1}{\Delta} \{(\alpha' \Sigma^{-1} \alpha)(1' \Sigma^{-1}) - (\alpha' \Sigma^{-1} 1)(\alpha' \Sigma^{-1})\} X,$$

$$\hat{\sigma} = \frac{1}{\Delta} \{(1' \Sigma^{-1} 1)(\alpha' \Sigma^{-1}) - (1' \Sigma^{-1} \alpha)(1' \Sigma^{-1})\} X,$$

$$\text{Var}(\hat{\mu}) = \frac{\alpha' \Sigma^{-1} \alpha}{\Delta} \sigma^2, \text{Var}(\hat{\sigma}) = \frac{1' \Sigma^{-1} 1}{\Delta} \sigma^2, \text{Cov}(\hat{\mu}, \hat{\sigma}) = -\frac{1' \Sigma^{-1} \alpha}{\Delta} \sigma^2,$$

$$\Delta = (1' \Sigma^{-1} 1)(\alpha' \Sigma^{-1} \alpha) - (1' \Sigma^{-1} \alpha)^2,$$ 

(3)
with \( \triangle \) being the generalized variance of BLUEs \((\hat{\mu}, \hat{\sigma})\) based on \(X\); see Balakrishnan and Cohen (1991) for pertinent details.

**Theorem 1.** The joint best linear unbiased predictors \(\hat{X}_s\) and \(\hat{X}_t\), determined by the D-optimality criterion, are of the form \(\hat{X}_s = a'X\) and \(\hat{X}_t = b'X\) in which the coefficients \(a = (a_1, \ldots, a_r)'_{r \times 1}\) and \(b = (b_1, \ldots, b_r)'_{r \times 1}\) are given by

\[
a = \frac{1}{\triangle} \begin{bmatrix}
\sum_{i=1}^{r} (\alpha_i - \alpha_s)(S_iR_1 - R_iS_1) \\
\sum_{i=1}^{r} (\alpha_i - \alpha_s)(S_iR_2 - R_iS_2) \\
\vdots \\
\sum_{i=1}^{r} (\alpha_i - \alpha_s)(S_iR_r - R_iS_r)
\end{bmatrix}
\]

and

\[
b = \frac{1}{\triangle} \begin{bmatrix}
\sum_{i=1}^{r} (\alpha_i - \alpha_t)(R_iS_1 - S_iR_1) \\
\sum_{i=1}^{r} (\alpha_i - \alpha_t)(R_iS_2 - S_iR_2) \\
\vdots \\
\sum_{i=1}^{r} (\alpha_i - \alpha_t)(R_iS_r - S_iR_r)
\end{bmatrix}
\]

where \(R_i\) and \(S_i\) are the sums of the \(i\)th rows of the matrices \(\Sigma^{-1}\) and \(\Sigma^{-1}\alpha\), respectively, and \(\triangle\) is as defined earlier in (3).

**Proof.** The BLUPs will now be derived jointly by minimizing the determinant of the variance-covariance matrix of BLUPs with respect to vectors \(a\) and \(b\).

Note that the variance-covariance matrix, say \(V\), of BLUPs \(\hat{X}_s\) and \(\hat{X}_t\) is of the form

\[
V = \sigma^2 \begin{bmatrix}
a'\Sigma a & a'\Sigma b \\
a'\Sigma b & b'\Sigma b
\end{bmatrix} = \sigma^2 \begin{bmatrix} V_{11} & V_{12} \\ V_{12} & V_{22}\end{bmatrix} \quad \text{(say)},
\]

with \(|V| = V_{11}V_{22} - V_{12}^2\) and

\[
V^{-1} = \frac{1}{|V|} \begin{bmatrix} V_{22} & -V_{12} \\ -V_{12} & V_{11}\end{bmatrix}.
\]

Note that we can leave the multiplicative factor \(\sigma^2\) in the minimization process. Moreover, while minimizing \(|V|\), we must impose four constraints due
to the unbiasedness of BLUPs and those are evidently \( a'1 = 1, a'\alpha = \alpha_s, \) \( b'1 = 1 \) and \( b'\alpha = \alpha_t \). So, the Lagrangian method can be employed to determine the optimal \( a \) and \( b \) by considering the objective function

\[
Q(a, b) = (a'\Sigma a)(b'\Sigma b) - (a'\Sigma b)^2 - 2\lambda_1(a'1 - 1) - 2\lambda_1^*(a'\alpha - \alpha_s) - 2\lambda_2(b'1 - 1) - 2\lambda_2^*(b'\alpha - \alpha_t).
\]  

(6)

Differentiating (6) with respect to \( a \) and \( b \) and equating them to vector 0 of dimension \( r \times 1 \), we obtain

\[
(b'\Sigma b)\Sigma a - (a'\Sigma b)\Sigma b - \lambda_1 1 - \lambda_1^* \alpha = 0,
\]  

(7)

\[
(a'\Sigma a)\Sigma b - (a'\Sigma b)\Sigma a - \lambda_2 1 - \lambda_2^* \alpha = 0,
\]  

(8)

respectively. Pre-multiplying (7) by \( a' \) and simplifying the resulting equation, we obtain

\[
\lambda_1 + \lambda_1^* \alpha_s = |V|.
\]  

(9)

Similarly, pre-multiplying (8) by \( b' \) and simplifying the resulting equation, we obtain

\[
\lambda_2 + \lambda_2^* \alpha_t = |V|.
\]  

(10)

Next, pre-multiplying (7) by \( b' \) and simplifying the resulting equation, we obtain

\[
\lambda_1 + \lambda_1^* \alpha_t = 0.
\]  

(11)

Finally, pre-multiplying (8) by \( a' \) and simplifying the resulting equation, we obtain

\[
\lambda_2 + \lambda_2^* \alpha_s = 0.
\]  

(12)

By solving (9) and (11), we obtain

\[
\lambda_1 = -\frac{|V|\alpha_t}{\alpha_s - \alpha_t}, \quad \lambda_1^* = \frac{|V|}{\alpha_s - \alpha_t},
\]

and similarly by solving (10) and (12), we obtain

\[
\lambda_2 = -\frac{|V|\alpha_s}{\alpha_t - \alpha_s}, \quad \lambda_2^* = \frac{|V|}{\alpha_t - \alpha_s}.
\]

Next, pre-multiplying (7) by \( \Sigma^{-1} \) and substituting for \( \lambda_1 \) and \( \lambda_1^* \), we obtain

\[
V_{22}a - V_{12}b = \frac{|V|}{\alpha_s - \alpha_t} \Sigma^{-1}(\alpha - \alpha_t 1).
\]  

(13)
Similarly, by pre-multiplying (8) by $\Sigma^{-1}$ and substituting the values of $\lambda_2$ and $\lambda_2'$, we obtain

$$V_{11}b - V_{12}a = \frac{|V|}{\alpha_t - \alpha_s} \Sigma^{-1} (\alpha - \alpha_s 1). \quad (14)$$

Note that $\alpha - \alpha_s 1 = (\alpha_1 - \alpha_s, \cdots, \alpha_r - \alpha_s)'_r \in \mathbb{R}^r$ and $\alpha - \alpha_t 1 = (\alpha_1 - \alpha_t, \cdots, \alpha_r - \alpha_t)'_r \in \mathbb{R}^r$, and let us denote them by $\alpha(s)$ and $\alpha(t)$, respectively. Thence, by writing the Eqs. (13) and (14) in a matrix form as

$$[a \ b] \begin{bmatrix} V_{22} & -V_{12} \\ -V_{12} & V_{11} \end{bmatrix} = \frac{|V|}{\alpha_t - \alpha_s} \Sigma^{-1} \begin{bmatrix} -\alpha(t) \\ -\alpha(s) \end{bmatrix} V. \quad (15)$$

The solution in (15) explicitly gives

$$a = \frac{1}{\alpha_t - \alpha_s} \{ -V_{11}(\alpha'_t \Sigma^{-1} 1) + V_{12}(\alpha'_s \Sigma^{-1} 1) \}, \quad (16)$$
$$b = \frac{1}{\alpha_t - \alpha_s} \{ -V_{11}(\alpha'_t \Sigma^{-1} \alpha) + V_{12}(\alpha'_s \Sigma^{-1} \alpha) \}. \quad (17)$$

Now, the unbiasedness conditions $a'1 = 1$ and $a'\alpha = \alpha_s$ give

$$\frac{1}{\alpha_t - \alpha_s} \{ -V_{11}(\alpha'_t \Sigma^{-1} 1) + V_{12}(\alpha'_s \Sigma^{-1} 1) \} = 1, \quad (18)$$
$$\frac{1}{\alpha_t - \alpha_s} \{ -V_{11}(\alpha'_t \Sigma^{-1} \alpha) + V_{12}(\alpha'_s \Sigma^{-1} \alpha) \} = \alpha_s. \quad (19)$$

Solving (18) and (19) for $V_{11}$ and $V_{12}$, we get

$$V_{11} = -(\alpha_t - \alpha_s) \frac{\alpha'_s \Sigma^{-1} \alpha_s}{\alpha'_t \Sigma^{-1} \alpha'_s - \alpha'_s \Sigma^{-1} \lambda_1 \alpha'_t \Sigma^{-1} \lambda_1'}, \quad (20)$$
$$V_{12} = -(\alpha_t - \alpha_s) \frac{\alpha'_t \Sigma^{-1} \alpha_s}{\alpha'_t \Sigma^{-1} \alpha'_s - \alpha'_s \Sigma^{-1} \lambda_1 \alpha'_t \Sigma^{-1} \lambda_1'}. \quad (21)$$

Now, we shall further simplify the above derived expressions. For this let us denote

$$\Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1r} \\ \sigma_{12} & \sigma_{22} & \cdots & \sigma_{2r} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{1r} & \sigma_{2r} & \cdots & \sigma_{rr} \end{bmatrix} \quad \text{and} \quad \Sigma^{-1} = \begin{bmatrix} \sigma^{11} & \sigma^{12} & \cdots & \sigma^{1r} \\ \sigma^{12} & \sigma^{22} & \cdots & \sigma^{2r} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma^{1r} & \sigma^{2r} & \cdots & \sigma^{rr} \end{bmatrix}. \quad (22)$$
Then,
\[
\Sigma^{-1} \mathbf{1} = \begin{bmatrix}
\sigma^{11} + \sigma^{12} + \cdots + \sigma^{1r} \\
\sigma^{12} + \sigma^{22} + \cdots + \sigma^{2r} \\
\vdots \\
\sigma^{1r} + \sigma^{2r} + \cdots + \sigma^{rr}
\end{bmatrix} = \begin{bmatrix} R_1 \\ R_2 \\ \vdots \\ R_r \end{bmatrix},
\]
where \( R_i = \sum_{j=1}^{r} \sigma^{ji} \) is the sum of the \( i \)th row of \( \Sigma^{-1} \), for \( i = 1, \ldots, r \), and similarly
\[
\Sigma^{-1} \mathbf{\alpha} = \begin{bmatrix}
\alpha_1 \sigma^{11} + \alpha_2 \sigma^{12} + \cdots + \alpha_r \sigma^{1r} \\
\alpha_1 \sigma^{12} + \alpha_2 \sigma^{22} + \cdots + \alpha_r \sigma^{2r} \\
\vdots \\
\alpha_1 \sigma^{1r} + \alpha_2 \sigma^{2r} + \cdots + \alpha_r \sigma^{rr}
\end{bmatrix} = \begin{bmatrix} S_1 \\ S_2 \\ \vdots \\ S_r \end{bmatrix},
\]
where \( S_i = \sum_{j=1}^{r} \alpha_j \sigma^{ji} \), for \( i = 1, \ldots, r \), is the sum of \( i \)th row of \( \Sigma^{-1} \mathbf{\alpha} \). With these notations, we have
\[
\mathbf{\alpha}'(t) \Sigma^{-1} \mathbf{1} \mathbf{\alpha}'(s) = \left( \sum_{i=1}^{r} R_i (\alpha_i - \alpha_t) \right) \begin{bmatrix} \alpha_1 - \alpha_s, \ldots, \alpha_r - \alpha_s \end{bmatrix},
\]
\[
\mathbf{\alpha}'(s) \Sigma^{-1} \mathbf{1} \mathbf{\alpha}'(t) = \left( \sum_{i=1}^{r} R_i (\alpha_i - \alpha_s) \right) \begin{bmatrix} \alpha_1 - \alpha_t, \ldots, \alpha_r - \alpha_t \end{bmatrix},
\]
and consequently,
\[
\mathbf{\alpha}'(t) \Sigma^{-1} \mathbf{1} \mathbf{\alpha}'(s) - \mathbf{\alpha}'(s) \Sigma^{-1} \mathbf{1} \mathbf{\alpha}'(t) = (\alpha_s - \alpha_t) \begin{bmatrix} \alpha_1 R - R^*, \ldots, \alpha_r R - R^* \end{bmatrix},
\]
where \( R = \sum_{i=1}^{r} R_i \) and \( R^* = \sum_{i=1}^{r} R_i \alpha_i \). Observe that \( R = \sum_{i=1}^{r} R_i = 1 \Sigma^{-1} \mathbf{1}, \quad R^* = \sum_{i=1}^{r} R_i \alpha_i = \mathbf{\alpha}' \Sigma^{-1} \mathbf{1}, \quad \sum_{i=1}^{r} S_i = 1 \Sigma^{-1} \mathbf{\alpha} = \mathbf{\alpha}' \Sigma^{-1} \mathbf{\alpha} \) and \( \sum_{i=1}^{r} \alpha_i S_i = \mathbf{\alpha}' \Sigma^{-1} \mathbf{\alpha} \). With these expressions, the denominators of (20) and (21) become
\[
(\alpha_s - \alpha_t) \begin{bmatrix} \alpha_1 R - R^*, \alpha_2 R - R^*, \ldots, \alpha_r R - R^* \end{bmatrix} \begin{bmatrix} S_1 \\ S_2 \\ \vdots \\ S_r \end{bmatrix}
\]
\[
= (\alpha_s - \alpha_t) \left( R \sum_{i=1}^{r} \alpha_i S_i - R^* \sum_{i=1}^{r} S_i \right)
\]
\[
= (\alpha_s - \alpha_t) \Delta.
\]
So, from (20) and (21) we simply obtain

\[ V_{11} = \frac{\alpha'(s) \Sigma^{-1} \alpha(s)}{\Delta} \quad \text{and} \quad V_{12} = \frac{\alpha'(t) \Sigma^{-1} \alpha(s)}{\Delta}. \]

If we use the unbiasedness conditions \( b'1 = 1 \) and \( b'\alpha = \alpha_t \) and proceed exactly as above, we obtain

\[ V_{12} = \frac{\alpha'(t) \Sigma^{-1} \alpha(s)}{\Delta} \quad \text{and} \quad V_{22} = \frac{\alpha'(t \Sigma^{-1} \alpha(t)}{\Delta}, \]

showing the uniqueness of the obtained solution. Substituting for \( V_{11} \) and \( V_{12} \) in (16), we obtain

\[ a = \frac{1}{(\alpha_t - \alpha_s) \Delta} \left\{ (\alpha'(s) \Sigma^{-1} \alpha(t) \Sigma^{-1} \alpha(s) - (\alpha'(s) \Sigma^{-1} \alpha(s)) \Sigma^{-1} \alpha(t) \right\}, \quad (22) \]

which, after some algebraic calculations, can be shown to be as given in (4). Similarly, substituting \( V_{12} \) and \( V_{22} \) in (17) and simplifying, we obtain an explicit expression for the coefficient vector \( b \) as given in (5).

\[ \square \]

**Corollary 1.** In the process of the above derivation, the variance-covariance matrix of the joint predictors is readily found to be

\[ \text{Var} \left( \begin{array}{c} \bar{X}_s \\ \bar{X}_t \end{array} \right) = \text{Var} \left( \begin{array}{c} a'X \\ b'X \end{array} \right) = \sigma^2 \begin{bmatrix} a'\Sigma a & a'\Sigma b \\ a'\Sigma b & b'\Sigma b \end{bmatrix} = \frac{\sigma^2 \Delta}{\alpha'(s) \Sigma^{-1} \alpha(s) \Sigma^{-1} \alpha(t)}, \]

where \( \Delta \) is the generalized variance as given in (3).

### 3 Non-existence of joint prediction

In this section, we discuss two cases in which the joint predictors do not exist.

#### 3.1 Prediction of more than two order statistics in location-scale family

We now formally show that the joint prediction of more than two order statistics is not possible when the lifetimes come from a location-scale family of distributions. We will first prove our claim for the joint prediction of three order statistics.

Suppose we wish to jointly predict three order statistics \( X_{s:n}, X_{t:n} \) and \( X_{u:n} \) for \( r < s < t < u \leq n \). Now, let us assume that the joint predictors
of them are of the form $\tilde{X}_s = \alpha'X$, $\tilde{X}_t = b'X$ and $\tilde{X}_u = c'X$, respectively, where $\alpha'$, $b'$ and $c'$ are all coefficient vectors of dimension $r \times 1$. By applying the Lagrangian method as done in Section 2, we will obtain

\[
\alpha' = \frac{V_{11}}{\alpha_t - \alpha_s}(\alpha_1 \Sigma^{-1} - \alpha' \Sigma^{-1}) + \frac{V_{12}}{\alpha_u - \alpha_t}(\alpha_{u1} \Sigma^{-1} - \alpha' \Sigma^{-1}) + \frac{V_{13}}{\alpha_s - \alpha_u}(\alpha_s \Sigma^{-1} - \alpha' \Sigma^{-1}),
\]

\[
b' = \frac{V_{12}}{\alpha_t - \alpha_s}(\alpha_{12} \Sigma^{-1} - \alpha' \Sigma^{-1}) + \frac{V_{22}}{\alpha_u - \alpha_t}(\alpha_{u2} \Sigma^{-1} - \alpha' \Sigma^{-1}) + \frac{V_{23}}{\alpha_s - \alpha_u}(\alpha_s \Sigma^{-1} - \alpha' \Sigma^{-1}),
\]

\[
c' = \frac{V_{13}}{\alpha_t - \alpha_s}(\alpha_{13} \Sigma^{-1} - \alpha' \Sigma^{-1}) + \frac{V_{23}}{\alpha_u - \alpha_t}(\alpha_{u3} \Sigma^{-1} - \alpha' \Sigma^{-1}) + \frac{V_{33}}{\alpha_s - \alpha_u}(\alpha_s \Sigma^{-1} - \alpha' \Sigma^{-1}),
\]

where the quantities $V_{11}$, $V_{12}$, $V_{13}$, $V_{22}$, $V_{23}$ and $V_{33}$, all of which are to be determined, are the elements of the variance-covariance matrix of the joint predictors given by

\[
\sigma^2 \begin{bmatrix} V_{11} & V_{12} & V_{13} \\ V_{12} & V_{22} & V_{23} \\ V_{13} & V_{23} & V_{33} \end{bmatrix} = \sigma^2 \begin{bmatrix} \alpha' \Sigma a & \alpha' \Sigma b & \alpha' \Sigma c \\ \alpha' \Sigma b & b' \Sigma b & b' \Sigma c \\ \alpha' \Sigma c & b' \Sigma c & c' \Sigma c \end{bmatrix}.
\]

Now, observe that the two unbiasedness conditions on $\tilde{X}_s$, namely, $\alpha' 1 = 1$ and $\alpha' \alpha = \alpha_s$ will generate two linear equations in three unknowns $V_{11}$, $V_{12}$ and $V_{13}$. The other unbiasedness conditions will similarly yield four other linear equations, and thus we end up with a system of the form

\[
\begin{bmatrix} A_{11} & A_{12} & A_{13} & 0 & 0 & 0 \\ B_{11} & B_{12} & B_{13} & 0 & 0 & 0 \\ 0 & A_{11} & 0 & A_{12} & A_{13} & 0 \\ 0 & B_{11} & 0 & B_{12} & B_{13} & 0 \\ 0 & 0 & A_{11} & 0 & A_{12} & A_{13} \\ 0 & 0 & B_{11} & 0 & B_{12} & B_{13} \end{bmatrix} \begin{bmatrix} V_{11} \\ V_{12} \\ V_{13} \\ V_{22} \\ V_{23} \\ V_{33} \end{bmatrix} = \begin{bmatrix} 1 \\ \alpha_s \\ 1 \\ \alpha_t \\ 1 \\ \alpha_u \end{bmatrix},
\]

\[
(A_{11} = \frac{\alpha_t R - R^*}{\alpha_t - \alpha_s}, A_{12} = \frac{\alpha_u R - R^*}{\alpha_u - \alpha_t}, A_{13} = \frac{\alpha_s R - R^*}{\alpha_s - \alpha_u}.
\]
Then, a straightforward matrix decomposition shows that the determinant of the coefficient matrix in (23) is zero implying the non-existence of the solution in (23). Hence, the claim. The non-existence of joint BLUPs for more than three order statistics can be established in a similar way.

3.2 Joint prediction in scale family

The non-existence of joint prediction of order statistics can also be established when the parent distribution belongs to the scale family. Assuming the joint predictors \( \hat{X}_s = a'X \) and \( \hat{X}_t = b'X \), we have two unbiasedness conditions, namely, \( a'\alpha = \alpha_s \) and \( b'\beta = \alpha_t \). Now, by applying the Lagrangian method with Lagrangian multipliers \( \lambda_1 \) and \( \lambda_2 \) for the two unbiasedness conditions \( a'\alpha = \alpha_s \) and \( b'\beta = \alpha_t \), respectively, we have two equations

\[
\begin{align*}
\lambda_1 \alpha &= (b'\Sigma b)(a'\Sigma b)\Sigma a - (a'\Sigma b)(b'\Sigma b), \\
\lambda_2 \alpha &= (a'\Sigma a)(b'\Sigma b) - (a'\Sigma b)(a'\Sigma a).
\end{align*}
\]

(24)  
(25)

Upon pre-multiplying (24) by \( a' \), we get

\[
\begin{align*}
\lambda_1 \alpha_t &= (b'\Sigma b)(b'\Sigma a) - (a'\Sigma b)(b'\Sigma b) \\
&= 0,
\end{align*}
\]

which implies \( \lambda_1 = 0 \). Similarly, pre-multiplying (25) by \( b' \) and simplifying, we get \( \lambda_2 = 0 \). We can then readily see the non-existence of the joint predictors in this case as well.

4 Numerical results

For illustrative purpose, let us consider a data set presented by Schmee and Nelson (1979). The data, assumed to follow a normal distribution, have the first \( r = 5 \) failure times to be 87.0, 92.8, 117.1, 133.6 and 138.6, and so \( X = (87.0, 92.8, 117.1, 133.6, 138.6)' \). Table 1 then presents a comparison between the marginal predictors in (2) and the joint predictors developed in the last section. Two forms of efficiency measures are defined as follows:

\[
\text{D-efficiency} = \frac{\text{Determinant of the variance-covariance matrix of joint predictors } \hat{X}_s \text{ and } \hat{X}_t}{\text{Determinant of the variance-covariance matrix of marginal predictors } \hat{X}_s \text{ and } \hat{X}_t}.
\]
Trace-efficiency = \frac{\text{Trace of the variance-covariance matrix of joint predictors } \hat{X}_s \text{ and } \hat{X}_t}{\text{Trace of the variance-covariance matrix of marginal predictors } \hat{X}_s \text{ and } \hat{X}_t}.

Then, the gain and loss in efficiency for the joint prediction can be defined as

Efficiency gain = 1 - D-efficiency

and

Efficiency loss = 1 - Trace-efficiency,

respectively. Thence, the overall gain in efficiency can be defined as

Overall efficiency gain = Efficiency gain - Efficiency loss.

Table 1 shows that, in all cases considered, Overall efficiency gain are positive indicating the advantage of using joint prediction. Further, the coefficients \( a \) and \( b \) are presented in Table 2 for three different scenarios: (i) \( s = r + 1, t = r + 2 \); (ii) \( s = r + 1, t = n \); and (iii) \( s = n - 1, t = n \), for various choices of \( n = (10, 15, 20) \) and \( r = (5, 10) \).

5 Concluding remarks

In this work, we have developed the joint best linear unbiased predictors of two unobserved order statistics based on observed order statistics through D-optimality criterion. The advantage of using point predictors over marginal predictors is demonstrated by a real data set. A possible future research problem would be to develop an optimal compound design for joint prediction based on both Trace-efficiency and D-efficiency introduced in the last section.

References

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

Balakrishnan, N. and Rao, C. R. (1997). A note on the best linear unbiased estimation based on order statistics. The American Statistician, 51:181–185.
Balasooriya, U. (1989). Detection of outliers in the exponential distribution based on prediction. *Communications in Statistics - Theory and Methods, 18*:711–720.

Doganaksoy, N. and Balakrishnan, N. (1997). A useful property of best linear unbiased predictors with applications to life-testing. *The American Statistician, 51*:22–28.

Goldberger, A. S. (1962). Best linear unbiased prediction in the generalized linear regression model. *Journal of the American Statistical Association, 57*:369–375.

Kaminsky, K. S. and Nelson, P. I. (1975). Best linear unbiased prediction of order statistics in location and scale families. *Journal of the American Statistical Association, 70*:145–150.

Kaminsky, K. S. and Nelson, P. I. (1998). Prediction of order statistics. In Balakrishnan, N. and Rao, C. R., editors, *Handbook of Statistics, Vol. 17: Order Statistics: Applications*, pages 431–450. North-Holland, Amsterdam.

Nelson, W. (2003). *Applied Life Data Analysis*. John Wiley & Sons, Hoboken, New Jersey.

Patel, J. K. (1989). Prediction intervals - a review. *Communications in Statistics - Theory and Methods, 18*:2393–2465.

Schmee, J. and Nelson, W. (1979). Predicting from early failures the last failure time of a (log) normal sample. *IEEE Transactions on Reliability, R-28*:22–28.
Table 1: Comparisons between the marginal predictors of Kaminsky and Nelson (1975) and the proposed joint predictors for the data presented in Schmee and Nelson (1979).

| Marginal predictors | Joint predictors | D-efficiency | Trace-efficiency | Overall efficiency gain |
|---------------------|------------------|--------------|-----------------|------------------------|
| $\hat{X}_{(6)} = 148.73$ | $\hat{X}_{(6)} = 148.58$ | 0.9737 | 0.9937 | 0.0200 |
| $\hat{X}_{(7)} = 158.99$ | $\hat{X}_{(7)} = 158.86$ | 0.9758 | 0.9954 | 0.0196 |
| $\hat{X}_{(8)} = 170.35$ | $\hat{X}_{(8)} = 170.25$ | 0.9777 | 0.9968 | 0.0190 |
| $\hat{X}_{(9)} = 184.37$ | $\hat{X}_{(9)} = 184.29$ | 0.9798 | 0.9980 | 0.0182 |
| $\hat{X}_{(10)} = 206.19$ | $\hat{X}_{(10)} = 206.12$ | 0.9785 | 0.9966 | 0.0181 |
| $\hat{X}_{(6)} = 148.73$ | $\hat{X}_{(6)} = 148.58$ | 0.9806 | 0.9976 | 0.0170 |
| $\hat{X}_{(7)} = 158.99$ | $\hat{X}_{(7)} = 158.86$ | 0.9829 | 0.9985 | 0.0157 |
| $\hat{X}_{(8)} = 170.35$ | $\hat{X}_{(8)} = 170.25$ | 0.9830 | 0.9983 | 0.0152 |
| $\hat{X}_{(9)} = 184.37$ | $\hat{X}_{(9)} = 184.29$ | 0.9854 | 0.9989 | 0.0135 |
| $\hat{X}_{(10)} = 206.19$ | $\hat{X}_{(10)} = 206.12$ | 0.9877 | 0.9992 | 0.0116 |
Table 2: Coefficients $a$ and $b$ for some selected joint predictors for the normal parent distribution.

| $n$ | $r|s, t$ | $a_1$ | $a_2$ | $a_3$ | $a_4$ | $a_5$ | $a_6$ | $a_7$ | $a_8$ | $a_9$ | $a_{10}$ |
|-----|---------|------|------|------|------|------|------|------|------|------|---------|
| 5   | 6, 7    | -0.1843 | -0.0322 | 0.0382 | 0.0932 | 1.0852 | -0.3088 | -0.0952 | 0.0037 | 0.0812 | 1.3191  |
|     | 6, 10   | -0.1843 | -0.0322 | 0.0382 | 0.0932 | 1.0852 | -0.8809 | -0.3849 | -0.1547 | 0.0264 | 2.3941  |
|     | 9, 10   | -0.6165 | -0.251 | -0.0815 | 0.0517 | 1.8973 |       |       |       |       |         |
| 10  | 11, 12  | -0.1696 | -0.0754 | -0.0341 | -0.0027 | 0.0239 | 0.0478 | 0.0702 | 0.0917 | 0.1130 | 0.9351  |
|     | 11, 15  | -0.4160 | -0.2266 | -0.1434 | -0.0803 | -0.0267 | 0.0215 | 0.0666 | 0.1102 | 0.1531 | 1.5415  |
|     | 14, 15  | -0.2982 | -0.1543 | -0.0912 | -0.0432 | -0.0025 | 0.0341 | 0.0683 | 0.1014 | 0.1339 | 1.2517  |
|     | 19, 20  | -1.2802 | -0.7187 | -0.4830 | -0.3093 | 3.7912 |       |       |       |       |         |
| 5   | 6, 7    | -0.1206 | -0.0361 | 0.0000 | 0.0250 | 1.1324 |       |       |       |       |         |
|     | 6, 20   | -0.2030 | -0.0846 | -0.0351 | 0.0013 | 1.3214 |       |       |       |       |         |
|     | 19, 20  | -1.5471 | -0.8758 | -0.594 | -0.3862 | 4.4032 |       |       |       |       |         |
|     | 20      | -1.2802 | -0.7187 | -0.4830 | -0.3093 | 3.7912 |       |       |       |       |         |
| 10  | 11, 12  | -0.0841 | -0.0310 | -0.0086 | 0.0080 | 0.0218 | 0.0338 | 0.0448 | 0.0550 | 0.0646 | 0.8958  |
|     | 11, 20  | -0.1168 | -0.0518 | -0.0243 | -0.0040 | 0.0129 | 0.0277 | 0.0411 | 0.0536 | 0.0654 | 0.9962  |
|     | 19, 20  | -0.5557 | -0.3308 | -0.2358 | -0.1652 | -0.1067 | -0.0554 | -0.0087 | 0.0349 | 0.0763 | 2.3470  |