Ann. Inst. Statist. Math.
37 (1985), Part A, 499-506

ON ESTIMATING A COMMON MULTIVARIATE NORMAL MEAN VECTOR

WEN-JAU CHIOU AND ARTHUR COHEN

(Received Feb. 2, 1984; revised July 31, 1984)

1. Introduction and summary

Let X_i, $i=1, 2, \cdots, n$ be a random sample of size n from a p dimensional multivariate normal distribution with mean vector μ and covariance matrix Σ_x. Let Y_i, $i=1, 2, \cdots, n$ be a random sample of size n from a p dimensional multivariate normal distribution with mean vector μ and covariance matrix Σ_y. Assume the X-sample and Y-sample are independent. Consider the problem of estimating the common mean vector μ. Let

$$\bar{X}=\Sigma X_i/n\,, \quad \bar{Y}=\Sigma Y_i/n\,, \quad A_x=\Sigma(X_i-\bar{X})(X_i-\bar{X})'\,,$$

$$A_y=\Sigma(Y_i-\bar{Y})(Y_i-\bar{Y})'\,, \quad S_x=A_x/(n-1)\,, \quad S_y=A_y(n-1)\,,$$

$$\Sigma_{\bar{x}}=\Sigma_x/n\,, \quad \text{and} \quad \Sigma_{\bar{y}}=\Sigma_y/n\,, \quad S_{\bar{x}}=S_x/n\,, \quad S_{\bar{y}}=S_y/n\,.$$

When $p=1$ a good deal is known about this problem. In that case if Σ_x, Σ_y are known then the estimator

$$T_1=(\Sigma_{\bar{y}}\bar{X}+\Sigma_{\bar{x}}\bar{Y})/(\Sigma_{\bar{x}}+\Sigma_{\bar{y}}) \tag{1.1}$$

is unbiased and has smaller variance than either $\bar{X}$ or $\bar{Y}$ uniformly in $(\mu, \Sigma_x, \Sigma_y)$. When $p=1$ and Σ_x and Σ_y are unknown then

$$T_2=(S_{\bar{y}}\bar{X}+S_{\bar{x}}\bar{Y})/(S_{\bar{x}}+S_{\bar{y}}) \tag{1.2}$$

is unbiased and has smaller variance than either $\bar{X}$ or $\bar{Y}$ as long as $n>10$. This latter result is due to Graybill and Deal [5]. Brown and Cohen [2], Cohen and Sackrowitz [3], Khatri and Shah [7] also displayed unbiased estimators with smaller variance then $\bar{X}$ and sometimes smaller than $\bar{X}$ and $\bar{Y}$ simultaneously as long as the sample size exceeded some specific integer. See Bhattacharya [1] for an update on this problem.

When $p>1$ and Σ_x and Σ_y are known consider the unbiased esti-

Key words and phrases: Common multivariate mean, unbiased estimators.
Research supported by NSF MCS-81-18564-01.

mator

$$T_3=\Sigma_{\bar{y}}(\Sigma_{\bar{x}}+\Sigma_{\bar{y}})^{-1}\bar{X}+\Sigma_{\bar{x}}(\Sigma_{\bar{x}}+\Sigma_{\bar{y}})^{-1}\bar{Y}\ , \tag{1.3}$$

which reduces to T_1 when $p=1$. It can be shown that the covariance matrix of T_3, Σ_{T_3} is such that $(\Sigma_{\bar{x}}-\Sigma_{T_3})$ and $(\Sigma_{\bar{y}}-\Sigma_{T_3})$ are both positive definite. When Σ_x and Σ_y are unknown it is logical to study the unbiased estimator

$$T_4=S_{\bar{y}}(S_{\bar{x}}+S_{\bar{y}})^{-1}\bar{X}+S_{\bar{x}}(S_{\bar{x}}+S_{\bar{y}})^{-1}\bar{Y}\ , \tag{1.4}$$

and determine whether, and for what n, are $(\Sigma_{\bar{x}}-\Sigma_{T_4})$ and $(\Sigma_{\bar{y}}-\Sigma_{T_4})$ positive semi-definite for all (Σ_x, Σ_y). The surprising result in Section 2 is that neither $\Sigma_{\bar{x}}-\Sigma_{T_4}$ nor $(\Sigma_{\bar{y}}-\Sigma_{T_4})$ is positive semi-definite for all (Σ_x, Σ_y) for any n. Evaluating unbiased estimators by the same criterion i.e. by their covariance matrices, analogues of estimators in Brown and Cohen [2], Cohen and Sackrowitz [3], Khatri and Shah [7] and others are also shown *not* to be better estimators than $\bar{X}$. We call these negative type results. In Section 3 the criterion of evaluating unbiased estimators by their covariance matrices is discussed.

2. Negative type results

By way of notation we say a covariance matrix Σ is positive semi-definite (p.s.d.) by writing $\Sigma\geqq 0$, is positive definite p.d. writing $\Sigma>0$, and is not p.s.d. by writing $\Sigma\ngeqq 0$. Consider the class of estimators

$$T_5=\bar{X}+A_1S_{\bar{x}}A_2(S_{\bar{x}}+B_1S_{\bar{y}}B_2)^{-1}C(\bar{Y}-\bar{X})\ , \tag{2.1}$$

where A_1, A_2, B_1, B_2, C are nonsingular matrices of constants. Note that when A_1, A_2, B_1, B_2, C are the identity then T_5 is the analogue of the estimator of Graybill-Deal; when $A_1=aI$, $A_2=B_1=B_2=C=I$, T_5 is the analogue of Brown-Cohen; when $B_1=cI$, $A_1=A_2=B_2=C=I$, T_5 is the analogue of Khatri-Shah. Hence all the analogues are special cases of the general form T_5.

Theorem 2.1. *The difference* $\Sigma_{\bar{x}}-\Sigma_{T_5}\ngeqq 0$ *for all* (Σ_x, Σ_y).

Proof. We give the proof for $p=2$ since it will be clear that if $p>2$ a similar but more detailed argument will work. First note that from (2.1)

$$\begin{aligned}\Sigma_{\bar{x}}-\Sigma_{T_5}=\Sigma_{\bar{x}}&-\mathrm{E}\,A_1S_{\bar{x}}A_2(S_{\bar{x}}+B_1S_{\bar{y}}B_2)^{-1}C\Sigma_{\bar{x}}\\ &-\mathrm{E}\,\Sigma_{\bar{x}}C'(S_{\bar{x}}+B_2'S_yB_1')^{-1}A_2'S_{\bar{x}}A_1'\\ &+\mathrm{E}\,A_1S_{\bar{x}}A_2(S_{\bar{x}}+B_1S_{\bar{y}}B_2)^{-1}C(\Sigma_{\bar{x}}+\Sigma_{\bar{y}})\\ &\cdot C'(S_{\bar{x}}+B_2'S_yB_1')^{-1}A_2'S_{\bar{x}}A_1'\ .\end{aligned} \tag{2.2}$$

Now let

$$A_1 S_{\bar{x}} A_2 = (1/n)U = (1/n)\begin{pmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{pmatrix}, \qquad (S_{\bar{x}} + B_1 S_{\bar{y}} B_2)^{-1} = nV = n\begin{pmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{pmatrix},$$

$$\Sigma_{\bar{x}} = (1/n)\begin{pmatrix} \sigma^2_{x_1} \rho_x \sigma_{x_1} \sigma_{x_2} \\ \sigma^2_{x_2} \end{pmatrix}, \qquad \Sigma_{\bar{y}} = (1/n)\begin{pmatrix} \sigma^2_{y_1} \rho_y \sigma_{y_1} \sigma_{y_2} \\ \sigma^2_{y_2} \end{pmatrix}$$

and consider the case where $\rho_x = \rho_y = 0$. Also let $A_i = \begin{pmatrix} a^{(i)}_{11} & a^{(i)}_{12} \\ a^{(i)}_{21} & a^{(i)}_{22} \end{pmatrix}$, $B_i = \begin{pmatrix} b^{(i)}_{11} & b^{(i)}_{12} \\ b^{(i)}_{21} & b^{(i)}_{22} \end{pmatrix}$, $i = 1, 2$, $C = \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix}$, $G = UVC\Sigma_{\bar{x}}$, $H = UVC(\Sigma_{\bar{x}} + \Sigma_{\bar{y}})C'V'U'$. From (2.2) $\Sigma_{\bar{x}} - \Sigma_{T_5} \ngeq 0$ if

$$\mathrm{E}\, h_{11} > 2\, \mathrm{E}\, g_{11} \tag{2.3}$$

for any (Σ_x, Σ_y) where h_{11} and g_{11} are elements in the first row- first column of H and G respectively. Straightforward multiplication enables (2.3) to be written as

$$\begin{aligned} \mathrm{E}\, \{ & [(\sigma^2_{x_1} + \sigma^2_{y_1})/n][c_{11}(u_{11}v_{11} + u_{12}v_{21}) + c_{21}(u_{11}v_{12} + u_{12}v_{22})]^2 \\ & + [(\sigma^2_{x_2} + \sigma^2_{y_2})/n][c_{12}(u_{11}v_{11} + u_{12}v_{21}) + c_{22}(u_{11}v_{12} + u_{12}v_{22})]^2 \} \\ & > 2\, \mathrm{E}\, (\sigma^2_{x_1}/n)[c_{11}(u_{11}v_{11} + u_{12}v_{21}) + c_{21}(u_{11}v_{12} + u_{12}v_{22})] \,. \end{aligned} \tag{2.4}$$

Let $X_i = (X_{1i}, X_{2i})'$, $Y_i = (Y_{1i}, Y_{2i})'$,

$$W_{i,x_j} = (X_{ji} - \bar{X}_j)/\sigma_{x_j}, \quad W_{i,y_j} = (Y_{ji} - \bar{Y}_j)/\sigma_{y_j}, \qquad j = 1, 2 \,.$$

Note $\sum_i (X_{ji} - \bar{X}_j)^2$ and $\sum_i (Y_{ji} - \bar{Y}_j)^2$ can be written as $\sigma^2_{x_j} \chi^2_{x_j}$ and $\sigma^2_{y_j} \chi^2_{y_j}$ respectively where $\chi^2_{x_j}$ and $\chi^2_{y_j}$ are independent χ^2 variables with $(n-1)$ degrees of freedom. Straightforward multiplication now yields the following:

$$\begin{aligned} u_{1j} = {} & a^{(1)}_{11} a^{(2)}_{1j} \sigma^2_{x_1} \chi^2_{x_1} + a^{(2)}_{12} a^{(2)}_{11} \sigma_{x_1} \sigma_{x_2} \sum_i W_{i,x_1} W_{i,x_2} \\ & + a^{(1)}_{11} a^{(2)}_{2j} \sigma_{x_1} \sigma_{x_2} \sum_i W_{i,x_1} W_{i,x_2} + a^{(1)}_{12} a^{(2)}_{2j} \sigma^2_{x_2} \chi^2_{x_2} \,, \qquad j = 1, 2 \,. \end{aligned} \tag{2.5}$$

Similarly derive v_{11}/Δ, v_{12}/Δ, v_{21}/Δ, and v_{22}/Δ

where

$$\begin{aligned} \Delta = {} & [\sigma^2_{x_1} \chi^2_{x_1} + b^{(1)}_{11} b^{(2)}_{11} \sigma^2_{y_1} \chi^2_{y_1} + b^{(1)}_{12} b^{(2)}_{11} \sigma_{y_1} \sigma_{y_2} \sum_i W_{i,y_1} W_{i,y_2} \\ & + b^{(1)}_{11} b^{(2)}_{21} \sigma_{y_1} \sigma_{y_2} \sum_i W_{i,y_1} W_{i,y_2} + b^{(1)}_{12} b^{(2)}_{21} \sigma^2_{y_2} \chi^2_{y_2}] \\ & \cdot [\sigma^2_{x_2} \chi^2_{x_2} + b^{(1)}_{21} b^{(2)}_{12} \sigma^2_{y_1} \chi^2_{y_1} + b^{(1)}_{22} b^{(2)}_{12} \sigma_{y_1} \sigma_{y_2} \sum_i W_{i,y_1} W_{i,y_2} \\ & + b^{(1)}_{21} b^{(2)}_{22} \sigma_{y_1} \sigma_{y_2} \sum_i W_{i,y_1} W_{i,y_2} + b^{(1)}_{22} b^{(2)}_{22} \sigma^2_{y_2} \chi^2_{y_2}] \\ & - [\sigma_{x_1} \sigma_{x_2} \sum W_{i,x_1} W_{i,x_2} + b^{(1)}_{21} b^{(2)}_{11} \sigma^2_{y_1} \chi^2_{y_1} + b^{(1)}_{22} b^{(2)}_{11} \sigma_{y_1} \sigma_{y_2} \sum_i W_{i,y_1} W_{i,y_2} \end{aligned} \tag{2.6}$$

$$+b_{21}^{(1)}b_{21}^{(2)}\sigma_{y_1}\sigma_{y_2}\sum W_{i,y_1}W_{i,y_2}+b_{22}^{(1)}b_{21}^{(2)}\sigma_{y_2}^2\chi_{y_2}^2]$$
$$\cdot\,[\sigma_{x_1}\sigma_{x_2}\sum W_{i,x_1}W_{i,x_2}+b_{11}^{(1)}b_{12}^{(2)}\sigma_{y_1}^2\chi_{y_1}^2+b_{12}^{(1)}b_{12}^{(2)}\sigma_{y_1}\sigma_{y_2}\sum_i W_{i,y_1}W_{i,y_2}$$
$$+b_{11}^{(1)}b_{22}^{(2)}\sigma_{y_1}\sigma_{y_2}\sum_i W_{i,y_1}W_{i,y_2}+b_{12}^{(1)}b_{22}^{(2)}\sigma_{y_2}^2\chi_{y_2}^2]\,.$$

Now let $\sigma_{x_1}\to 0$ and note from (2.5) and (2.6) that

$$\text{(2.7)}\qquad u_{11}\to a_{12}^{(1)}a_{21}^{(2)}\sigma_{x_2}^2\chi_{x_2}^2\,,\qquad u_{12}\to a_{12}^{(1)}a_{22}^{(2)}\sigma_{x_2}^2\chi_{x_2}^2\,,$$

$$\text{(2.8)}\qquad \Delta\to\Delta_1=[b_{11}^{(1)}b_{11}^{(2)}\sigma_{y_1}^2\chi_{y_1}^2+b_{12}^{(1)}b_{11}^{(2)}\sigma_{y_1}\sigma_{y_2}\sum W_{i,y_1}W_{i,y_2}$$
$$+b_{11}^{(1)}b_{21}^{(2)}\sigma_{y_1}\sigma_{y_2}\sum W_{i,y_1}W_{i,y_2}+b_{12}^{(1)}b_{21}^{(2)}\sigma_{y_2}^2\chi_{y_2}^2]$$
$$\cdot\,[\sigma_{x_2}^2\chi_{x_2}^2+b_{21}^{(1)}b_{12}^{(2)}\sigma_{y_1}^2\chi_{y_1}^2+b_{22}^{(1)}b_{12}^{(2)}\sigma_{y_1}\sigma_{y_2}\sum W_{i,y_1}W_{i,y_2}$$
$$+b_{21}^{(1)}b_{22}^{(2)}\sigma_{y_1}\sigma_{y_2}\sum W_{i,y_1}W_{i,y_2}+b_{22}^{(1)}b_{22}^{(2)}\sigma_{y_2}^2\chi_{y_2}^2]$$
$$-[b_{21}^{(1)}b_{11}^{(2)}\sigma_{y_1}^2\chi_{y_1}^2+b_{22}^{(1)}b_{11}^{(2)}\sigma_{y_1}\sigma_{y_2}\sum W_{i,y_1}W_{i,y_2}$$
$$+b_{21}^{(1)}b_{21}^{(2)}\sigma_{y_1}\sigma_{y_2}\sum W_{i,y_1}W_{i,y_2}+b_{22}^{(1)}b_{21}^{(2)}\sigma_{y_2}^2\chi_{y_2}^2]$$
$$\cdot\,[b_{11}^{(1)}b_{12}^{(2)}\sigma_{y_1}^2\chi_{y_1}^2+b_{12}^{(1)}b_{12}^{(2)}\sigma_{y_1}\sigma_{y_2}\sum W_{i,y_1}W_{i,y_2}$$
$$+b_{11}^{(1)}b_{22}^{(2)}\sigma_{y_1}\sigma_{y_2}\sum W_{i,y_1}W_{i,y_2}+b_{12}^{(1)}b_{22}^{(2)}\sigma_{y_2}^2\chi_{y_2}^2]$$

$$\text{(2.9)}\qquad u_{11}v_{11}+u_{12}v_{21}\to\{a_{12}^{(1)}a_{21}^{(2)}\sigma_{x_2}^2\chi_{x_2}^2[\sigma_{x_2}^2\chi_{x_2}^2+b_{21}^{(1)}b_{12}^{(2)}\sigma_{y_1}^2\chi_{y_2}^2$$
$$+b_{22}^{(1)}b_{12}^{(2)}\sigma_{y_1}\sigma_{y_2}\sum W_{i,y_1}W_{i,y_2}+b_{21}^{(1)}b_{22}^{(2)}\sigma_{y_1}\sigma_{y_2}\sum W_{i,y_1}W_{i,y_2}$$
$$+b_{22}^{(1)}b_{22}^{(2)}\sigma_{y_2}^2\chi_{y_2}^2]-a_{12}^{(1)}a_{22}^{(2)}\sigma_{x_2}^2\chi_{x_2}^2[b_{21}^{(1)}b_{11}^{(2)}\sigma_{y_1}^2\chi_{y_1}^2+b_{22}^{(1)}b_{11}^{(2)}\sigma_{y_1}\sigma_{y_2}\sum W_{i,y_1}W_{i,y_2}$$
$$+b_{21}^{(1)}b_{21}^{(2)}\sigma_{y_1}\sigma_{y_2}\sum W_{i,y_1}W_{i,y_2}+b_{22}^{(1)}b_{21}^{(2)}\sigma_{y_2}^2\chi_{y_2}^2]\}/\Delta_1$$

$$\text{(2.10)}\qquad u_{11}v_{12}+u_{12}v_{22}\to\{a_{12}^{(1)}a_{21}^{(2)}\sigma_{x_2}^2\chi_{x_2}^2[b_{11}^{(1)}b_{12}^{(2)}\sigma_{y_1}^2\chi_{y_1}^2+b_{12}^{(1)}b_{12}^{(2)}\sigma_{y_1}\sigma_{y_2}\sum W_{i,y_1}W_{i,y_2}$$
$$+b_{11}^{(1)}b_{22}^{(2)}\sigma_{y_1}\sigma_{y_2}\sum_i W_{i,y_1}W_{i,y_2}+b_{12}^{(1)}b_{22}^{(2)}\sigma_{y_2}^2\chi_{y_2}^2]$$
$$+a_{12}^{(1)}a_{22}^{(2)}\sigma_{x_2}^2\chi_{x_2}^2[b_{11}^{(1)}b_{11}^{(2)}\sigma_{y_1}^2\chi_{y_1}^2+b_{12}^{(1)}b_{22}^{(2)}\sigma_{y_1}\sigma_{y_2}\sum W_{i,y_1}W_{i,y_2}$$
$$+b_{11}^{(1)}b_{21}^{(2)}\sigma_{y_1}\sigma_{y_2}\sum W_{i,y_1}W_{i,y_2}+b_{12}^{(1)}b_{21}^{(2)}\sigma_{y_2}^2\chi_{y_2}^2]\}/\Delta_1\,.$$

Now observe that if $a_{12}^{(1)}\neq 0$, using (2.9) and (2.10) we see that the right hand sides of (2.4) tends to zero as $\sigma_{x_1}\to 0$ while the left hand side tends to a positive number. Thus in this case the theorem holds.

If $a_{12}^{(1)}=0$, then since A_1 is nonsingular $a_{11}^{(1)}\neq 0$. In this case

$$\text{(2.11)}\qquad u_{1j}=a_{11}^{(1)}a_{1j}^{(2)}\sigma_{x_1}^2\chi_{x_1}^2+a_{11}^{(1)}a_{2j}^{(2)}\sigma_{x_1}\sigma_{x_2}\sum W_{i,x_1}W_{i,x_2}\,.$$

Now let $\sigma_{x_2}=\sigma_{x_1}^{1/2}\to 0$, then

$$\text{(2.12)}\qquad u_{11}v_{11}+u_{12}v_{21}\to\{(a_{11}^{(1)}a_{11}^{(2)}\sigma_{x_1}^2\chi_{x_1}^2+a_{11}^{(1)}a_{21}^{(2)}\sigma_{x_1}\sigma_{x_2}\sum W_{i,x_1}W_{i,x_2})$$
$$\cdot\,[b_{21}^{(1)}b_{12}^{(2)}\sigma_{y_1}^2\chi_{y_1}^2+b_{22}^{(1)}b_{12}^{(2)}\sigma_{y_1}\sigma_{y_2}\sum W_{i,y_1}W_{i,y_2}$$
$$+b_{21}^{(1)}b_{22}^{(2)}\sigma_{y_1}\sigma_{y_2}\sum W_{i,y_1}W_{i,y_2}+b_{22}^{(1)}b_{22}^{(2)}\sigma_{y_2}^2\chi_{y_2}^2]$$
$$-(a_{11}^{(1)}a_{12}^{(2)}\sigma_{x_1}^2\chi_{x_1}^2+a_{11}^{(1)}a_{22}^{(2)}\sigma_{x_1}\sigma_{x_2}\sum W_{i,x_1}W_{i,x_2})$$
$$\cdot\,[b_{21}^{(1)}b_{11}^{(2)}\sigma_{y_1}^2\chi_{y_1}^2+b_{22}^{(1)}b_{11}^{(2)}\sigma_{y_1}\sigma_{y_2}\sum W_{i,y_1}W_{i,y_2}$$
$$+b_{21}^{(1)}b_{21}^{(2)}\sigma_{y_1}\sigma_{y_2}\sum W_{i,y_1}W_{i,y_2}+b_{22}^{(1)}b_{21}^{(2)}\sigma_{y_2}^2\chi_{y_2}^2]\}/\Delta_2\,,$$

$$
\begin{aligned}
(2.13)\quad u_{11}v_{12}+u_{12}v_{22} = &-(a_{11}^{(1)}a_{11}^{(2)}\sigma_{x_1}^2\chi_{x_1}^2+a_{11}^{(1)}a_{21}^{(2)}\sigma_{x_1}\sigma_{x_2}\sum W_{i,x_1}W_{i,x_2}) \\
&\cdot [b_{11}^{(1)}b_{12}^{(2)}\sigma_{y_1}^2\chi_{y_1}^2+b_{12}^{(1)}b_{12}^{(2)}\sigma_{y_1}\sigma_{y_2}\sum W_{i,y_1}W_{i,y_2} \\
&\quad +b_{11}^{(1)}b_{22}^{(2)}\sigma_{y_1}\sigma_{y_2}\sum W_{i,y_1}W_{i,y_2}+b_{12}^{(1)}b_{22}^{(2)}\sigma_{y_2}^2\chi_{y_2}^2] \\
&+(a_{11}^{(1)}a_{12}^{(2)}\sigma_{x_1}^2\chi_{x_1}^2+a_{11}^{(1)}a_{22}^{(2)}\sigma_{x_1}\sigma_{x_2}\sum W_{i,x_1}W_{i,x_2}) \\
&\cdot [b_{11}^{(1)}b_{11}^{(2)}\sigma_{y_1}^2\chi_{y_1}^2+b_{12}^{(1)}b_{11}^{(2)}\sigma_{y_1}\sigma_{y_2}\sum W_{i,y_1}W_{i,y_2} \\
&\quad +b_{11}^{(1)}b_{21}^{(2)}\sigma_{y_1}\sigma_{y_2}\sum W_{i,y_1}W_{i,y_2}+b_{12}^{(1)}b_{21}^{(2)}\sigma_{y_2}^2\chi_{y_2}^2]\}/\Delta_2 ,
\end{aligned}
$$

where Δ_2 is the same as Δ_1 except terms involving σ_{x_1} and σ_{x_2} are omitted. Note that (2.12), and (2.13), and $\sigma_{x_2}=\sigma_{x_1}^{1/2}\to 0$ imply that the right hand side of (2.4) tends to zero at the rate $\sigma_{x_1}^{7/2}$ whereas the left hand side tends to zero at the rate $\sigma_{x_1}^3$. This implies that parameter points exist for which (2.4) holds and thus the theorem is proved.

To prove the next negative result we need

LEMMA 2.2. $\mathrm{E}(\bar{X}-\mu|\bar{Y}-\bar{X})=-\Sigma_{\bar{x}}(\Sigma_{\bar{x}}+\Sigma_{\bar{y}})^{-1}(\bar{Y}-\bar{X})$.

PROOF. The proof is straightforward and omitted.

Now let $U=(u_{ij})$ be an $n\times n$ orthogonal matrix with nth row $u_{(n)}=(1/\sqrt{n}, 1/\sqrt{n},\cdots, 1/\sqrt{n})$. Define $U_i=\sum_{k=1}^{n} u_{ik}X_k$ so that $U_i\sim N(0, \Sigma_x)$ for $i=1, 2,\cdots, n-1$ and $U_n\sim N(\sqrt{n}\,\theta, \Sigma_x)$. Similarly define V as $n\times n$ orthogonal matrix and random vectors $V_i=\sum v_{ik}Y_k$ and so U_i, U_j, V_k, V_l are independent for $i\neq j$, $k\neq l$. Let r be an integer satisfying $6\leqq r\leqq n-2$ and define $S_x=\sum_{i=1}^{n-r-1} U_iU_i'$, $S=\sum_{i=n-r}^{n-1}(U_i+V_i)(U_i+V_i)'$. Note $S_x\sim W(p, n-r-1, \Sigma_x)$, $S\sim W(p, r, \Sigma_x+\Sigma_y)$, and S_x is independent of S. Here $W(p, q, \Sigma)$ is the Wishart distribution with parameters p, q, Σ. Consider the class of estimators

$$
(2.14)\qquad T_c=\bar{X}+cS^{-1}S_x(\bar{Y}-\bar{X}) ,
$$

for $c>0$. For $p=1$ this class of estimators was studied by Cohen and Sackrowitz [3].

THEOREM 2.3. *For any fixed $c>0$ the difference $\Sigma_{\bar{x}}-\Sigma_{T_c}\ngeq 0$ for all (Σ_x, Σ_y).*

PROOF. Again it suffices to give the proof for $p=2$. Apply Lemma 2.2, and use properties of the Wishart to find that

$$
\begin{aligned}
(2.15)\quad \Sigma_{\bar{x}}-\Sigma_{T_c}=&[c(n-r-1)/n(r-3)]\{(\Sigma_x+\Sigma_y)^{-1}\Sigma_x\Sigma_x+\Sigma_x\Sigma_x(\Sigma_x+\Sigma_y)^{-1}\} \\
&-(c^2/n)\,\mathrm{E}\,S^{-1}S_x(\Sigma_x+\Sigma_y)S_xS^{-1} .
\end{aligned}
$$

Use an identity from Haff [6] to show

$$\begin{aligned}(2.16)\quad & \mathrm{E}\, S^{-1}S_x(\Sigma_x+\Sigma_y)S_xS^{-1}\\ &=[(n-r-1)(n-r)/(r-2)(r-3)(r-5)]\\ &\quad\cdot \{\mathrm{tr}\,(\Sigma_x+\Sigma_y)^{-1}\Sigma_x(\Sigma_x+\Sigma_y)\Sigma_x\}(\Sigma_x+\Sigma_y)^{-1}\\ &\quad+[(n-r-1)(n-r)/(r-2)(r-5)]\{(\Sigma_x+\Sigma_y)^{-1}(\Sigma_x(\Sigma_x+\Sigma_y)\Sigma_x)\\ &\quad\cdot(\Sigma_x+\Sigma_y)^{-1}\}+[(n-r-1)/(r-2)(r-3)(r-5)]\\ &\quad\cdot\{[\mathrm{tr}\,\Sigma_x(\Sigma_x+\Sigma_y)][\mathrm{tr}\,(\Sigma_x+\Sigma_y)^{-1}\Sigma_x](\Sigma_x+\Sigma_y)^{-1}\}\\ &\quad+[(n-r-1)/(r-2)(r-5)][\mathrm{tr}\,\Sigma_x(\Sigma_x+\Sigma_y)]\\ &\quad\cdot(\Sigma_x+\Sigma_y)^{-1}\Sigma_x(\Sigma_x+\Sigma_y)^{-1}\,.\end{aligned}$$

Use (2.16) in (2.15) to conclude that $\Sigma_{\bar{x}}-\Sigma_{T_c}\geqq 0$ if and only if

$$\begin{aligned}(2.17)\quad & \Sigma_x\Sigma_x(\Sigma_x+\Sigma_y)+(\Sigma_x+\Sigma_y)\Sigma_x\Sigma_x\\ &>[c/(r-2)(r-5)]\{(n-r)[\mathrm{tr}\,(\Sigma_x+\Sigma_y)^{-1}\Sigma_x(\Sigma_x+\Sigma_y)\Sigma_x](\Sigma_x+\Sigma_y)\\ &\quad+(n-r)(r-3)\Sigma_x(\Sigma_x+\Sigma_y)\Sigma_x+[\mathrm{tr}\,\Sigma_x(\Sigma_x+\Sigma_y)]\\ &\quad\cdot[\mathrm{tr}\,(\Sigma_x+\Sigma_y)^{-1}\Sigma_x](\Sigma_x+\Sigma_y)+(r-3)[\mathrm{tr}\,\Sigma_x(\Sigma_x+\Sigma_y)\Sigma_x]\}\,.\end{aligned}$$

In order for (2.17) to hold clearly the element in the first row-first column on the left hand side of (2.17) must be greater than the element in the first-row first column on the right hand side of (2.17). For the case $\rho_x=\rho_y=0$ use (2.16) to find that such a condition becomes

$$\begin{aligned}(2.18)\quad 2\sigma_{x_1}^4(\sigma_{x_1}^2+\sigma_{y_1}^2)>{}&[c/(r-2)(r-5)]\{(n-r)(\sigma_{x_1}^4+\sigma_{x_2}^4)(\sigma_{x_1}^2+\sigma_{y_1}^2)\\ &+(n-r)(r-3)\sigma_{x_1}^4(\sigma_{x_1}^2+\sigma_{y_1}^2)\\ &+[\sigma_{x_1}^2(\sigma_{x_1}^2+\sigma_{y_1}^2)+\sigma_{x_2}^2(\sigma_{x_2}^2+\sigma_{y_2}^2)][\sigma_{x_1}^2(\sigma_{x_2}^2+\sigma_{y_2}^2)\\ &+\sigma_{x_2}^2(\sigma_{x_1}^2+\sigma_{y_1}^2)]/(\sigma_{x_2}^2+\sigma_{y_2}^2)\\ &+(r-3)[\sigma_{x_1}^2(\sigma_{x_1}^2+\sigma_{y_1}^2)+\sigma_{x_2}^2(\sigma_{x_2}^2+\sigma_{y_2}^2)]\sigma_{x_1}^2\}\,.\end{aligned}$$

Fix $\sigma_{x_1}, \sigma_{y_1}, \sigma_{y_2}$, and let $\sigma_{x_2}\to\infty$. It is clear that for any $c>0$ (2.18) is contradicted.

Another class of estimators studied is the vector W whose ith component is

$$(2.19)\qquad W_i=\bar{X}_i+a_i[S_{\bar{x}_i}^2/(S_{\bar{x}_i}^2+b_iS_{\bar{y}_i}^2)](\bar{Y}_i-\bar{X}_i)\,,$$

where $\bar{X}_i$ and $\bar{Y}_i$ are the ith components of $\bar{X}$ and $\bar{Y}$, and $S_{\bar{x}_i}$, $S_{\bar{y}_i}$ are the ith diagonal elements of $S_{\bar{x}}$ and $S_{\bar{y}}$. It can be shown that $\Sigma_{\bar{x}}-\Sigma_w \ngeqq 0$ for all parameter values. To show such a result again we let $p=2$ and use computations similar to those in the proof of Theorem 2.1 in writing down $\Sigma_{\bar{x}}-\Sigma_w$. This time however the correlations ρ_x and ρ_y will not be taken to be 0, and the various cases for a_i and b_i are treated separately. The parameter points for which $\Sigma_{\bar{x}}-\Sigma_w\ngeqq 0$ usually are when $\sigma_{y_1}\to\infty$. We omit the details.

3. Remarks on criterion of evaluation

Evaluating an unbiased estimator of a vector of parameters by the covariance matrix of the estimator has been done before. In fact for linear models the Gauss-Markov theorem states that the vector of least squares estimates of the regression parameter vector, $\hat{\beta}$ say, is best among linear unbiased estimators in the sense that its covariance matrix, $\Sigma_{\hat{\beta}}$ is such that $\Sigma_{\tilde{\beta}}-\Sigma_{\beta}>0$ for any other estimator $\tilde{\beta}$.

For estimating a vector of parameters a frequent criterion of evaluation is the expected sum of squared errors. A weaker criterion is to look at the vector whose ith component is the expected squared error for estimating the ith component of the vector. It is weaker in the sense that an estimator can be admissible by the vector criterion but not by the sum criterion. The criterion here is still weaker than the vector criterion. For estimating a common mean vector if the risk function were sum of squared errors, or if we studied the vector risk then the estimator $\bar{X}$ would be inadmissible by virtue of previous work: (See Brown and Cohen [2] and Cohen and Sackrowitz [4]). However with the covariance criterion the issue of admissibility of $\bar{X}$ is unresolved. The difficulty is highlighted in Section 2 of this paper.

Cohen and Sackrowitz [4] discuss decision theory formulation for vector risks. Their development is easily extended to accommodate risks more general than vector risks. The covariance criterion is an example of a risk that is more general than a vector risk. To apply the Cohen-Sackrowitz [4] development for such a criterion to be used for the problem of this paper, the parameter space would be $\Omega=\Theta\times H$ where Θ would consist of points $\theta=(\mu, \Sigma_x, \Sigma_y)$ and H would consist of points ξ where ξ would represent a direction in a p-dimesional vector space.

RUTGERS UNIVERSITY

REFERENCES

[1] Bhattacharya, C. G. (1981). Estimation of a common mean and recovery of inter-block information, Ph.D. dissertation, Indian Statistical Institute, New Delhi, India.

[2] Brown, L. D. and Cohen, A. (1974). Point and confidence estimation of a common mean and recovery of interblock information, *Ann. Statist.*, **2**, 963-976.

[3] Cohen, A. and Sackrowitz, H. B. (1974). On estimating the common mean of two normal distributions, *Ann. Statist.*, **2**, 1274-1282.

[4] Cohen, A. and Sackrowitz, H. B. (1984). Decision theory results for vector risks with applications, *Statistics and Decisions*.

[5] Graybill, F. A. and Deal, R. B. (1959). Combining unbiased estimators, *Biometrics*,

15, 543-550.

[6] Haff, L. R. (1979). An identity for the Wishart distributions with applications, *Journal of Multivariate Analysis*, **9**, 531-544.

[7] Khatri, C. G. and Shah, K. R. (1974). Estimation of location parameters from two linear models under normality, *Comm. Statist.*, B, **3**, 647-665.