The Annals of Mathematical Statistics
1972, Vol. 43, No. 2, 455-467

ESTIMATION AND TESTING FOR DIFFERENCES IN MAGNITUDE OR DISPLACEMENT IN THE MEAN VECTORS OF TWO MULTIVARIATE NORMAL POPULATIONS

BY CHARLES H. KRAFT,[1] INGRAM OLKIN,[2] AND CONSTANCE VAN EEDEN[1]

Université de Montréal, Stanford University and Université de Montréal

0. Summary. This paper is concerned with the determination of whether two multivariate normal mean vectors differ by a multiplicative factor or by a multiplicative factor plus a displacement along a specified direction. These considerations arise, for example, in the evaluation of drugs, where two drugs affecting the same symptoms are considered equivalent if they can be made equal by a change in dosage. Maximum likelihood estimators and likelihood ratio tests for the relevant parameters are considered.

1. Introduction. The model considered stems from an evaluation of drugs, in which x_i denotes the effect of a drug on symptom i, $i = 1, \cdots, p$. Two drugs with measurements $(x_1, \cdots, x_p)$ and $(y_1, \cdots, y_p)$ can be considered to be equivalent if $Ex_i = cEy_i$, $i = 1, \cdots, p$, since by a change of dosage the effects may be made equal. In some cases, equivalence may be defined by $Ex_i = cEy_i + d$, $i = 1, \cdots, p$. The parameters c and d are called magnitude and displacement parameters, respectively. The basic problem is to estimate the relevant parameters, and to determine whether two drugs are equivalent.

Models which contain magnitude and displacement may be generated in various ways. We may view x and y as arising from a model in which $x_{i\alpha} = \mu_i + \varepsilon_{i\alpha}$, $y_{i\beta} = \nu_i + \delta_{i\beta}$, $i = 1, \cdots, p$, $\alpha = 1, \cdots, N_1$, $\beta = 1, \cdots, N_2$. Here $(\varepsilon_{1\alpha}, \cdots, \varepsilon_{p\alpha})$ and $(\delta_{1\beta}, \cdots, \delta_{p\beta})$ are independently distributed having a common multivariate normal distribution with means 0 and covariance matrix Σ. The ε's and δ's denote measurement error and are from different subjects. Thus the model becomes $\mathscr{L}(x_{1\alpha}, \cdots, x_{p\alpha}) = \mathscr{N}(\mu, \Sigma)$, $\mathscr{L}(y_{1\beta}, \cdots, y_{p\beta}) = \mathscr{N}(\nu, \Sigma)$, and we wish to test $H_1: \mu = c\nu$ or $H_2: \mu = c\nu + de$, where $e = (1, \cdots, 1)$, versus the alternative hypothesis that μ and ν are unrestricted. When the subjects are the same, the x's and y's are correlated, and a more complicated model is required to take this dependence into account. Alternatively, we

Received December 1968.

[1] Work supported in part by the National Science Foundation under Grant Number GP-3816 and by the AFOSR under Contract Number 88565, while authors were at the University of Minnesota.

[2] Work supported in part by the National Science Foundation under Grant Number GP-6681.

may suspect that the y's behave as the x's, except for a multiplicative factor, i.e., $\mathscr{L}(x) = \mathscr{L}(cy)$. Consequently, we may wish to test $H: \mu = c\nu$, $\Sigma_1 = c^2\Sigma_2$, where now Σ_1 (or Σ_2) is the covariance matrix of x (or y), against general alternatives. Our concern in the present paper is with the former model. Likelihood ratio tests (LRT) for various hypotheses and maximum likelihood estimators (MLE) of the parameters when the hypotheses prevail are derived.

In the case of a single bivariate distribution, Paulson (1942) obtained a confidence interval for the ratio of means. This type of problem arises in sampling theory when using ratio estimates. The model considered has a closer connection with classification procedures, as indicated in Section 2.

Cochran (1943) provides a comprehensive study of the comparison of different scales of measurement. In the above context his concern is with testing that scales are (a) equivalent, (b) differ by a constant, or are (c) linearly related, i.e., (a) $\mu_i = \nu_i$, (b) $\mu_i = \nu_i + d$, (c) $\alpha\mu_i + \beta\nu_i = d$. Tests for each case are discussed in some detail. However it should be noted that this model assumes repeated measurements on the same set of subjects. The models we consider —though quite similar—complement those of Cochran (1943); the testing procedures will follow along similar lines, though the development is quite different. In the present paper the emphasis is, to a great extent, on estimating the parameters of the model, which was not considered for (a)–(c).

If we define $M = \binom{\mu}{\nu}$, then the hypothesis $\mu = c\nu$ is equivalent to the hypothesis that the rank of M is one. This problem has been considered in detail by Anderson (1951), who provided the LRT of this hypothesis. Note that the hypothesis $\mu_i = c\nu_i + d$ $(1 \leqq i \leqq p)$ is not a hypothesis on the rank of M.

2. Preliminaries. If the two p-variate normal populations are denoted by $\mathscr{N}(\mu, \Sigma)$ and $\mathscr{N}(\nu, \Sigma)$, the hypotheses of interest are

$$H_1: \mu_1 = c\nu_1, \cdots, \mu_p = c\nu_p ,$$
$$H_2: \mu_1 = c\nu_1 + d, \cdots, \mu_p = c\nu_p + d ,$$

where c and d are unknown numbers. In each case the alternatives are $H: -\infty < \mu_i < \infty$, $-\infty < \nu_i < \infty$, $i = 1, \cdots, p$. We are also concerned with testing H_1 against H_2.

In order to apply the results of Wald (1943) to obtain asymptotic distributions, it is convenient to make the following assumptions on the families of populations $\{\mathscr{N}(\mu, \Sigma)\}$, $\{\mathscr{N}(\nu, \Sigma)\}$. For some $\varepsilon, B, 0 < \varepsilon < B < \infty$, and $\eta < 1$,

(a) $0 < \varepsilon < \sigma_{ii} < B , \qquad i = 1, \cdots, p ,$

(b) $|\sigma_{ij}| < \eta(\sigma_{ii}\sigma_{jj})^{\frac{1}{2}} , \qquad i \neq j ,$

(c) $|\nu_p| > \varepsilon$ for H_1, $|\nu_p - \nu_{p-1}| > \varepsilon$ for H_2, and $|\mu_i| < B$, $|\nu_i| < B$, $i = 1, \cdots, p$,

where $\Sigma = (\sigma_{ij})$.

In practice the implications of these conditions are that none of the variances and not all of the means ν_i (not all of the differences $\nu_i - \nu_j$) be too close to zero and that no two of the variables be (linearly) redundant.

Independent samples of size N_1 and N_2 are taken from populations 1 and 2. Let $X = (x_{ij}) : p \times N_1$ and $Y = (y_{ij}) : p \times N_2$ denote the matrices of observations, and define $\bar{x} = (\bar{x}_1, \cdots, \bar{x}_p)$, $\bar{y} = (\bar{y}_1, \cdots, \bar{y}_p)$ to be the vectors of means, $S^{(1)} = (s_{ij}^{(1)})$, and $S^{(2)} = (s_{ij}^{(2)})$ the product moment matrices, and $S = S^{(1)} + S^{(2)}$. Then $(\bar{x}, \bar{y}, S)$ is a sufficient statistic for (μ, ν, Σ). Thus we have the reduction to the model: $\bar{x}, \bar{y}, S$ are independently distributed with

$$\mathscr{L}(\bar{x}) = \mathscr{N}(\mu, \Sigma/N_1), \qquad \mathscr{L}(\bar{y}) = \mathscr{N}(\nu, \Sigma/N_2), \qquad \mathscr{L}(S) = \mathscr{W}(\Sigma; p, n),$$

where $n = N - 2$, $N = N_1 + N_2$, and $\mathscr{W}(\Sigma; p, n)$ denotes the Wishart distribution with n degrees of freedom and parameters p and $ES = n\Sigma$. The joint density of $\bar{x}, \bar{y}, S$ is

$$(2.1) \qquad p(\bar{x}, \bar{y}, S; \mu, \nu, \Lambda) = K(S)|\Lambda|^{N/2} \exp -\tfrac{1}{2}[N_1(\bar{x} - \mu)\Lambda(\bar{x} - \mu)' + N_2(\bar{y} - \nu)\Lambda(\bar{y} - \nu)' + \operatorname{tr} \Lambda S],$$

where $\Lambda = \Sigma^{-1}$, $K(S) = |S|^{(n-p-1)/2}[(2\pi)^p 2^{np/2}\pi^{p(p-1)/4} \prod_1^p \Gamma((n - i + 1)/2)]^{-1}$.

Note that the problem of testing $\mu = c\nu$ against general alternatives is left invariant under the transformation $(\bar{x}, \bar{y}, S) \to (\bar{x}A, \bar{y}A, A'SA)$, for nonsingular matrices A. The maximal invariant under this group is

$$(t_{11}, t_{22}, t_{12}) = (\bar{x}S^{-1}\bar{x}', \bar{y}S^{-1}\bar{y}', \bar{x}S^{-1}\bar{y}').$$

In the classification problem in which there are two multivariate normal populations $\mathscr{N}(\mu, \Sigma)$ and $\mathscr{N}(\nu, \Sigma)$, the classification statistics of Wald and of Anderson are functions of the maximal invariant. The joint distribution of these statistics has been obtained by Sitgreaves (1952) under the hypothesis that the two mean vectors are proportional, i.e., $\mu = c_1\delta$, $\nu = c_2\delta$. Thus, we could use the joint distribution of (t_{11}, t_{22}, t_{12}) to obtain the LRT. However, it appears to be more troublesome to start at this stage rather than with the original variables.

3. MLE of the magnitude and displacement parameters. When $\mu_i = c\nu_i + d$, $i = 1, \cdots, p$, the joint density (2.1) becomes

$$(3.1) \qquad p(\bar{x}, \bar{y}, S; c, d, \nu, \Lambda) = K(S)|\Lambda|^{N/2} \exp -\tfrac{1}{2}[N_1(\bar{x} - c\nu - de)\Lambda(\bar{x} - c\nu - de)' + N_2(\bar{y} - \nu)\Lambda(\bar{y} - \nu)' + \operatorname{tr} \Lambda S],$$

where $e = (1, \cdots, 1)$. To obtain the MLE, first maximize with respect to ν. Setting the derivative of (3.1) equal to zero yields the vector equation

$$cN_1(\bar{x} - c\nu - de)\Lambda + N_2(\bar{y} - \nu)\Lambda = 0.$$

Since the logarithm of (3.1) is a strictly concave function of ν, the maximizer is

$$\hat{\nu} = (cN_1\bar{x} + N_2\bar{y} - N_1cde)/(c^2N_1 + N_2).$$

A direct substitution of $\hat{\nu}$ in (3.1) yields

$$(3.2)\quad \max p(\bar{x}, \bar{y}, S; c, d, \nu, \Lambda)$$
$$= K(S)|\Lambda|^{N/2} \exp -\tfrac{1}{2}\left[\frac{N_1 N_2(\bar{x} - c\bar{y} - de)\Lambda(\bar{x} - c\bar{y} - de)'}{c^2 N_1 + N_2} + \operatorname{tr} \Lambda S\right].$$

Maximization of (3.2) with respect to $\Lambda > 0$ yields (e.g., see Anderson (1958), Lemma 3.2.2)

$$\Lambda = N\left(S + \frac{N_1 N_2(\bar{x} - c\bar{y} - de)'(\bar{x} - c\bar{y} - de)}{c^2 N_1 + N_2}\right)^{-1}$$
$$= NS^{-1}\left(I + \frac{N_1 N_2(\bar{x} - c\bar{y} - de)'(\bar{x} - c\bar{y} - de)S^{-1}}{c^2 N_1 + N_2}\right)^{-1},$$

so that

$$(3.3)\quad \max_{\nu,\Lambda} p(\bar{x}, \bar{y}, S; c, d, \nu, \Lambda)$$
$$= \frac{K(S)|NS^{-1}|^{N/2} e^{-\frac{1}{2}pN}}{\{1 + [N_1 N_2(\bar{x} - c\bar{y} - de)S^{-1}(\bar{x} - c\bar{y} - de)']/(c^2 N_1 + N_2)\}^{N/2}}.$$

In this form we need to minimize

$$(3.4)\quad g(c, d) = \frac{(\bar{x} - cy - de)S^{-1}(\bar{x} - c\bar{y} - de)'}{c^2 N_1 + N_2}$$

with respect to c and d if both the magnitude and displacement parameters are present. When only the magnitude parameter is present we consider $g(c, 0)$.

Although in the present context $c > 0$, so that we have to maximize over a restricted range, we do not take account of this restriction in the derivation of the tests. The reason for this is that the procedures become quite complicated, whereas if $\hat{c}$ were negative, we would wish to reexamine the basic model.

3.1. MLE *of the magnitude parameter.* For simplicity of notation, set

$$(3.5)\quad T = \begin{pmatrix} t_{11} & t_{12} \\ t_{12} & t_{22} \end{pmatrix} = \begin{pmatrix} \bar{x}S^{-1}\bar{x}' & \bar{x}S^{-1}\bar{y}' \\ \bar{y}S^{-1}\bar{x}' & \bar{y}S^{-1}\bar{y}' \end{pmatrix},$$

so that we have to minimize

$$g(c, 0) = \frac{t_{11} - 2ct_{12} + c^2 t_{22}}{(c^2 N_1 + N_2)}.$$

Setting $g' = 0$ yields $N_1 t_{12} c^2 + (N_2 t_{22} - N_1 t_{11})c - N_2 t_{12} = 0$. Because $t_{12} \neq 0$ with probability 1, there are two roots:

$$\frac{N_1 t_{11} - N_2 t_{22} \pm [(N_1 t_{11} - N_2 t_{22})^2 + 4N_1 N_2 t_{12}^2]^{\frac{1}{2}}}{2N_1 t_{12}}.$$

Since $\hat{c} = \pm\infty$ is not a solution, and $\frac{1}{2}(\hat{c}^2 N_1 + N_2)^2 g''(\hat{c}) = N_2 t_{22} - N_1 t_{11} + 2\hat{c}t_{12}N_1 = \pm[(N_1 t_{11} - N_2 t_{22})^2 + 4N_1 N_2 t_{12}^2]^{\frac{1}{2}}$, it follows that the upper sign must hold, namely,

$$\hat{c} = \frac{N_1 t_{11} - N_2 t_{22} + [(N_1 t_{11} - N_2 t_{22})^2 + 4N_1 N_2 t_{12}^2]^{\frac{1}{2}}}{2N_1 t_{12}} . \tag{3.6}$$

For later reference we note that

$$\begin{aligned} N_1 N_2 g(\hat{c}, 0) &= N_2 t_{22} - \frac{t_{12} N_2}{\hat{c}} \\ &= \tfrac{1}{2} \{N_1 t_{11} + N_2 t_{22} - [(N_1 t_{11} - N_2 t_{22})^2 + 4N_1 N_2 t_{12}^2]^{\frac{1}{2}}\} . \end{aligned} \tag{3.7}$$

3.2. *Asymptotic distribution of the* MLE *of the magnitude parameter.* We first obtain the joint asymptotic distribution of t_{11}, t_{12}, t_{22}, from which it is relatively straightforward to obtain the asymptotic distribution of $\hat{c}$.

Recall from (3.5) that $T \equiv T(\bar{x}, \bar{y}, S)$ so that T evaluated at the mean $T(\mu, \nu, n\Sigma)$ is

$$\tau = \begin{pmatrix} \tau_{11} & \tau_{12} \\ \tau_{12} & \tau_{22} \end{pmatrix} = \begin{pmatrix} \mu\Sigma^{-1}\mu' & \mu\Sigma^{-1}\nu' \\ \nu\Sigma^{-1}\mu' & \nu\Sigma^{-1}\nu' \end{pmatrix} . \tag{3.8}$$

To obtain the covariance matrix of the asymptotic distribution we note that

$$N_1 \operatorname{Cov}(\bar{x}) = N_2 \operatorname{Cov}(\bar{y}) = \Sigma , \qquad \operatorname{Cov}(s_{ij}, s_{kl}) = n(\sigma_{ij}\sigma_{jl} + \sigma_{il}\sigma_{jk}) .$$

If g_1 and g_2 are functions of $\bar{x}$, $\bar{y}$, and S, then the covariance of the asymptotic distribution is given by

$$\begin{aligned} &\operatorname{Cov}(g_1, g_2) \\ &\quad = \textstyle\sum_{i,j} \frac{\partial g_1}{\partial \bar{x}_i} \frac{\partial g_2}{\partial \bar{x}_j} \operatorname{Cov}(\bar{x}_i, \bar{x}_j) + \sum_{i,j} \frac{\partial g_1}{\partial \bar{y}_i} \frac{\partial g_2}{\partial \bar{y}_j} \operatorname{Cov}(\bar{y}_i, \bar{y}_j) \\ &\quad\quad + \textstyle\sum_{i \leq j, k \leq l} \frac{\partial g_1}{\partial s_{ij}} \frac{\partial g_2}{\partial s_{kl}} \operatorname{Cov}(s_{ij}, s_{kl}) \\ &\quad \equiv \left(\frac{\partial g_1}{\partial \bar{x}}\right) \frac{\Sigma}{N_1} \left(\frac{\partial g_2}{\partial \bar{x}}\right)' + \left(\frac{\partial g_1}{\partial \bar{y}}\right) \frac{\Sigma}{N_2} \left(\frac{\partial g_2}{\partial \bar{y}}\right)' \\ &\quad\quad + 2n \operatorname{tr}\left[\left(\frac{\partial g_1}{\partial S}\right) \Sigma \left(\frac{\partial g_2}{\partial S}\right) \Sigma\right], \end{aligned} \tag{3.9}$$

where, for $l = 1, 2$,

$$\left(\frac{\partial g_l}{\partial \bar{x}}\right) = \left(\frac{\partial g_l}{\partial \bar{x}_1}, \cdots, \frac{\partial g_l}{\partial \bar{x}_p}\right), \qquad \left(\frac{\partial g_l}{\partial \bar{y}}\right) = \left(\frac{\partial g_l}{\partial \bar{y}_1}, \cdots, \frac{\partial g_l}{\partial \bar{y}_p}\right),$$

$$\left(\frac{\partial g_l}{\partial S}\right) = \left(\frac{\partial g_l^*}{\partial s_{ij}}\right), \qquad \frac{\partial g_l^*}{\partial s_{ii}} = \frac{\partial g_l}{\partial s_{ii}}, \qquad \frac{\partial g_l^*}{\partial s_{ij}} = \frac{1}{2} \frac{\partial g_l}{\partial s_{ij}}, \qquad (i \neq j),$$

and where all derivatives are evaluated at the means μ, ν and $n\Sigma$ for $\bar{x}$, $\bar{y}$ and S. In particular, $t_{ij} \equiv t_{ij}(\bar{x}, \bar{y}, S)$, so that we need to evaluate the following derivatives (evaluated at the means):

$$\frac{\partial t_{11}}{\partial \bar{x}} = \frac{2\mu\Sigma^{-1}}{n}, \qquad \frac{\partial t_{22}}{\partial \bar{x}} = 0, \qquad \frac{\partial t_{12}}{\partial \bar{x}} = \frac{\nu\Sigma^{-1}}{n}, \qquad \frac{\partial t_{11}}{\partial \bar{y}} = 0,$$

$$(3.10) \qquad \frac{\partial t_{22}}{\partial \bar{y}} = \frac{2\nu\Sigma^{-1}}{n}, \qquad \frac{\partial t_{12}}{\partial \bar{y}} = \frac{\mu\Sigma^{-1}}{n}, \qquad \frac{\partial t_{11}}{\partial S} = -\frac{\Sigma^{-1}\mu'\mu\Sigma^{-1}}{n^2},$$

$$\frac{\partial t_{22}}{\partial S} = -\frac{\Sigma^{-1}\nu'\nu\Sigma^{-1}}{n^2}, \qquad \frac{\partial t_{12}}{\partial S} = -\frac{\Sigma^{-1}(\mu'\nu + \nu'\mu)\Sigma^{-1}}{2n^2}.$$

Using (3.10) with (3.9) yields the covariances of the asymptotic distribution:

$$\begin{aligned}
\operatorname{Var}(t_{11}) &= 2\tau_{11}(2n + N_1\tau_{11})/n^3N_1, \\
\operatorname{Var}(t_{22}) &= 2\tau_{22}(2n + N_2\tau_{22})/n^3N_2, \\
(3.11) \qquad \operatorname{Var}(t_{12}) &= [n(N_1\tau_{11} + N_2\tau_{22}) + N_1N_2(\tau_{12}^2 + \tau_{11}\tau_{22})]/n^3N_1N_2, \\
\operatorname{Cov}(t_{11}, t_{22}) &= 2\tau_{12}^2/n^3, \\
\operatorname{Cov}(t_{11}, t_{12}) &= 2\tau_{12}(n + N_1\tau_{11})/n^3N_1, \\
\operatorname{Cov}(t_{22}, t_{12}) &= 2\tau_{12}(n + N_2\tau_{22})/n^3N_2.
\end{aligned}$$

We recapitulate by stating the following theorem.

THEOREM 1. *If* $\mathscr{L}(\bar{x}) = \mathscr{N}(\mu, \Sigma/N_1)$, $\mathscr{L}(\bar{y}) = \mathscr{N}(\nu, \Sigma/N_2)$, $\mathscr{L}(S) = \mathscr{W}(\Sigma; p, n)$, $\bar{x}, \bar{y}, S$ *independent, and*

$$t_{11} = \bar{x}S^{-1}\bar{x}', \qquad t_{22} = \bar{y}S^{-1}\bar{y}', \qquad t_{12} = \bar{x}S^{-1}\bar{y}',$$

then as $N_1 \to \infty$, $N_2 \to \infty$ *with* N_1/N_2 *fixed,* (t_{11}, t_{22}, t_{12}) *is asymptotically normally distributed, with mean* $(\tau_{11}, \tau_{22}, \tau_{12})$ *given by* (3.8), *and covariances given by* (3.11).

From (3.9) we find that the variance of the asymptotic distribution of $\hat{c} \equiv h(t_{11}, t_{22}, t_{12})$, is given by

$$V_\infty(\hat{c}) \equiv \left(\frac{\partial h}{\partial t_{11}}, \frac{\partial h}{\partial t_{22}}, \frac{\partial h}{\partial t_{12}}\right) \psi \left(\frac{\partial h}{\partial t_{11}}, \frac{\partial h}{\partial t_{22}}, \frac{\partial h}{\partial t_{12}}\right)',$$

where the derivatives are evaluated at the means μ, ν and $n\Sigma$ and where ψ is the covariance matrix of the asymptotic distribution of (t_{11}, t_{22}, t_{12}). A tedious but direct computation yields

$$\left(\frac{\partial h}{\partial t_{11}}, \frac{\partial h}{\partial t_{22}}, \frac{\partial h}{\partial t_{12}}\right) = \frac{n}{\tau_{22}(N_1c^2 + N_2)} (N_1c, -N_2c, N_2 - N_1c^2).$$

Consequently, we obtain the

COROLLARY 2. *The asymptotic distribution of* $\hat{c}$ *is normal with mean* c *and variance* $(N_1c^2 + N_2)/N_1N_2\tau_{22}$.

Approximate confidence bounds for c may be obtained from the Corollary by solving a quadratic equation. Alternatively, we may use the variance stabilizing transformation

$$A(c) = N_1^{-\frac{1}{2}} \operatorname{arc\,sinh} N_2^{-\frac{1}{2}} N_1^{\frac{1}{2}} c \,,$$

in which case

$$A(\hat{c}) \to_d N\left(A(c), \frac{1}{N_1 N_2 \tau_{22}}\right).$$

Since $p \lim t_{22}/n = \tau_{22}$, we have that

$$(N_1 N_2/n)^{\frac{1}{2}}[A(\hat{c}) - A(c)]t_{22} \to_d N(0, 1) \,.$$

Because $A(c)$ is a monotonically increasing function of c, confidence bounds for $A(c)$ may easily be converted to confidence bounds for c.

3.3. MLE *of displacement and magnitude parameters*. From (3.4) we need to minimize $g(c, d)$ with respect to c and d. Since $g(c, d)$ is quadratic in d, it is straightforward to find the minimizer, $\hat{d}$, with respect to d, namely,

$$\hat{d} = \frac{eS^{-1}\bar{x}' - ceS^{-1}\bar{y}'}{eS^{-1}e'} = \frac{eS^{-1}(\bar{x} - c\bar{y})'}{eS^{-1}e'} \,, \tag{3.12}$$

from which, after some simplification,

$$g(c, \hat{d}) = \frac{(\bar{x} - c\bar{y})A(\bar{x} - c\bar{y})'}{N_2 + N_1 c^2} \,, \tag{3.13}$$

where

$$A = S^{-1} - \frac{S^{-1}e'eS^{-1}}{eS^{-1}e'} \,. \tag{3.14}$$

The minimizer of $g(c, \hat{d})$ with respect to c is now obtained as in Section 3.1 to be

$$\hat{c} = \frac{N_1 \bar{x}A\bar{x}' - N_2 \bar{y}A\bar{y}' + [(N_1 \bar{x}A\bar{x}' - N_2 \bar{y}A\bar{y}')^2 + 4N_1 N_2 (\bar{x}A\bar{y}')^2]^{\frac{1}{2}}}{2\bar{x}A\bar{y}'N_1} \,. \tag{3.15}$$

After considerable algebra, we find that

$$2N_1 N_2 g(\hat{c}, \hat{d}) = N_1 t_{11}^* + N_2 t_{22}^* - [(N_1 t_{11}^* - N_2 t_{22}^*)^2 + 4N_1 N_2 t_{12}^{*2}]^{\frac{1}{2}} \,, \tag{3.16}$$

where

$$t_{11}^* = \bar{x}A^{-1}\bar{x}' \,, \qquad t_{12}^* = \bar{x}A^{-1}\bar{y}' \,, \qquad t_{22}^* = \bar{y}A^{-1}\bar{y}' \,.$$

Following the method of Section 3.2, we now prove

THEOREM 3. *If* $\bar{x}$, $\bar{y}$, *and* S *are independently distributed with* $\mathscr{L}(\bar{x}) = \mathscr{N}(c\nu + de, \Sigma/N_1)$, $\mathscr{L}(\bar{y}) = \mathscr{N}(\nu, \Sigma/N_2)$, $\mathscr{L}(S) = \mathscr{W}(\Sigma; p, n)$, $n = N_1 + N_2 - 2$, *then as* $N_1 \to \infty$, $N_2 \to \infty$ *with* N_1/N_2 *fixed*, $(\hat{c}, \hat{d})$ *is asymptotically normally distributed* *with mean vector* (c, d) *and covariance matrix*

$$\begin{pmatrix} V_\infty(\hat{c}) & \mathrm{Cov}_\infty(\hat{c}, \hat{d}) \\ \mathrm{Cov}_\infty(\hat{c}, \hat{d}) & V_\infty(\hat{d}) \end{pmatrix} = \frac{N_2 + N_1 c^2}{N_1 N_2 \Delta} \begin{pmatrix} e\Sigma^{-1}e' & -e\Sigma^{-1}\nu' \\ -\nu\Sigma^{-1}e' & \nu\Sigma^{-1}\nu' \end{pmatrix},$$

where $\Delta = (e\Sigma^{-1}e')(\nu\Sigma^{-1}\nu') - (e\Sigma^{-1}\nu')^2$.

PROOF. For simplicity of notation, define

$$z_{11} = (\bar{x}A\bar{x}')(eS^{-1}e') = (\bar{x}S^{-1}\bar{x}')(eS^{-1}e') - (\bar{x}S^{-1}e')^2,$$
$$z_{22} = (\bar{y}A\bar{y}')(eS^{-1}e') = (\bar{y}S^{-1}\bar{y}')(eS^{-1}e') - (\bar{y}S^{-1}e')^2,$$
$$z_{12} = (\bar{x}A\bar{y}')(eS^{-1}e') = (\bar{x}S^{-1}\bar{y}')(eS^{-1}e') - (\bar{x}S^{-1}e')(\bar{y}S^{-1}e'),$$

so that

$$\hat{c} = \frac{N_1 z_{11} - N_2 z_{22} + [(N_1 z_{11} - N_2 z_{22})^2 + 4N_1 N_2 z_{12}^2]^{\frac{1}{2}}}{2N_1 z_{12}} \equiv \frac{Q}{2N_1 z_{12}}.$$

The z_{ij} evaluated at the means $E\bar{x} = c\nu + de$, $Ey = \nu$, $ES = n\Sigma$, and denoted by a superscript 0, are:

$$z_{11}^0 = c^2\Delta/n^2, \qquad z_{22}^0 = \Delta/n^2, \qquad z_{12}^0 = c\Delta/n^2.$$

Thus it is easily checked that $\hat{c}^0 = c$ and $\hat{d}^0 = d$. Note also that $Q^0 = 2N_1 c^2\Delta/n^2$.

For the covariances, we need

$$V_\infty(\hat{c}) = \left(\frac{\partial \hat{c}}{\partial \bar{x}}\right)\frac{\Sigma}{N_1}\left(\frac{\partial \hat{c}}{\partial \bar{x}}\right)' + \left(\frac{\partial \hat{c}}{\partial \bar{y}}\right)\frac{\Sigma}{N_2}\left(\frac{\partial \hat{c}}{\partial \bar{y}}\right)' + 2n \operatorname{tr}\left[\left(\frac{\partial \hat{c}}{\partial S}\right)\Sigma\left(\frac{\partial \hat{c}}{\partial S}\right)\Sigma\right],$$
$$V_\infty(\hat{d}) = \left(\frac{\partial \hat{d}}{\partial \bar{x}}\right)\frac{\Sigma}{N_1}\left(\frac{\partial \hat{d}}{\partial \bar{x}}\right)' + \left(\frac{\partial \hat{d}}{\partial \bar{y}}\right)\frac{\Sigma}{N_2}\left(\frac{\partial \hat{d}}{\partial \bar{y}}\right)' + 2n \operatorname{tr}\left[\left(\frac{\partial \hat{d}}{\partial S}\right)\Sigma\left(\frac{\partial \hat{d}}{\partial S}\right)\Sigma\right],$$
$$\mathrm{Cov}_\infty(\hat{c}, \hat{d}) = \left(\frac{\partial \hat{c}}{\partial \bar{x}}\right)\frac{\Sigma}{N_1}\left(\frac{\partial \hat{d}}{\partial \bar{x}}\right)' + \left(\frac{\partial \hat{c}}{\partial \bar{y}}\right)\frac{\Sigma}{N_2}\left(\frac{\partial \hat{d}}{\partial \bar{y}}\right)' + 2n \operatorname{tr}\left[\left(\frac{\partial \hat{c}}{\partial S}\right)\Sigma\left(\frac{\partial \hat{d}}{\partial S}\right)\Sigma\right],$$

where the derivatives are evaluated at the means. We now show that

$$\left(\frac{\partial \hat{c}}{\partial \bar{x}}\right) = \frac{q_1}{\Delta}, \qquad \left(\frac{\partial \hat{c}}{\partial \bar{y}}\right) = -\frac{cq_1}{\Delta}, \qquad \left(\frac{\partial \hat{c}}{\partial S}\right) = 0, \tag{3.17}$$

$$\left(\frac{\partial \hat{d}}{\partial \bar{x}}\right) = \frac{q_2}{\Delta}, \qquad \left(\frac{\partial \hat{d}}{\partial \bar{y}}\right) = -\frac{cq_2}{\Delta}, \qquad \left(\frac{\partial \hat{d}}{\partial S}\right) = 0, \tag{3.18}$$

where

$$q_1 = (e\Sigma^{-1}e')\nu\Sigma^{-1} - (e\Sigma^{-1}\nu')e\Sigma^{-1}, \qquad q_2 = (\nu\Sigma^{-1}\nu')e\Sigma^{-1} - (e\Sigma^{-1}\nu')\nu\Sigma^{-1},$$

from which the variances and covariances follow easily.

To evaluate the terms in (3.17), first note

$$\frac{\partial z_{11}}{\partial \bar{x}} = \frac{2cq_1}{n^2}, \qquad \frac{\partial z_{12}}{\partial \bar{x}} = \frac{q_1}{n^2}, \qquad \frac{\partial z_{22}}{\partial \bar{x}} = 0,$$
$$\frac{\partial z_{11}}{\partial \bar{y}} = 0, \qquad \frac{\partial z_{12}}{\partial \bar{y}} = \frac{cq_1}{n^2}, \qquad \frac{\partial z_{22}}{\partial \bar{y}} = \frac{2q_1}{n^2}.$$

After some simplification in each step,

$$\frac{\Delta}{n^2}\left(\frac{\partial \hat{c}}{\partial \bar{x}}\right) = \frac{1}{2N_1 c}\frac{\partial Q}{\partial \bar{x}} - \frac{\partial z_{12}}{\partial \bar{x}}$$

$$= \frac{1}{N_1 c^2 + N_2}\left[N_1 c \frac{\partial z_{11}}{\partial \bar{x}} + 2N_2 \frac{\partial z_{12}}{\partial \bar{x}}\right] - \frac{\partial z_{12}}{\partial \bar{x}} = \frac{q_1}{n^2},$$

$$\frac{\Delta}{n^2}\left(\frac{\partial \hat{c}}{\partial \bar{y}}\right) = \frac{1}{2N_2 c}\frac{\partial Q}{\partial \bar{y}} - \frac{\partial z_{12}}{\partial \bar{y}}$$

$$= \frac{1}{N_1 c^2 + N_2}\left[-N_2 c \frac{\partial z_{22}}{\partial \bar{y}} + 2N_2 \frac{\partial z_{12}}{\partial \bar{y}}\right] - \frac{\partial z_{12}}{\partial \bar{y}} = \frac{-cq_1}{n^2}.$$

Define

$$W = -(\nu\Sigma^{-1}\nu')\Sigma^{-1}e'e\Sigma^{-1} - (e\Sigma^{-1}e')\Sigma^{-1}\nu'\nu\Sigma^{-1} + (\nu\Sigma^{-1}e')\Sigma^{-1}(e'\nu + \nu'e)\Sigma^{-1}.$$

A direct but cumbersome computation then yields

$$\left(\frac{\partial z_{11}}{\partial S}\right) = \frac{c^2 W}{n^2}, \qquad \left(\frac{\partial z_{22}}{\partial S}\right) = \frac{W}{n^2}, \qquad \left(\frac{\partial z_{12}}{\partial S}\right) = \frac{cW}{n^2},$$

and

$$n^2\frac{\partial Q}{\partial S} = N_1\frac{\partial z_{11}}{\partial S} - N_2\frac{\partial z_{22}}{\partial S} + \frac{(N_1 c^2 - N_2)\left(N_1\frac{\partial z_{11}}{\partial S} - N_2\frac{\partial z_{22}}{\partial S}\right) + 4N_1 N_2 c\frac{\partial z_{12}}{\partial S}}{N_1 c^2 + N_2}$$

$$= 2N_1 c^2 W.$$

Hence

$$\frac{\Delta}{n^2}\left(\frac{\partial \hat{c}}{\partial S}\right) = \frac{1}{2N_1 c}\frac{\partial Q}{\partial S} - \frac{\partial z_{12}}{\partial S} = 0,$$

which completes the derivation of (3.17).

The derivation of (3.18) may now be carried out in a fairly direct form:

$$\frac{\partial \hat{d}}{\partial \bar{x}} = \frac{1}{e\Sigma^{-1}e'}\left[e\Sigma^{-1} - (e\Sigma^{-1}\nu')\frac{\partial \hat{c}}{\partial \bar{x}}\right] = \frac{1}{e\Sigma^{-1}e'}\left[e\Sigma^{-1} - \frac{(e\Sigma^{-1}\nu')q_1}{\Delta}\right]$$

$$= \frac{1}{\Delta}[(\nu\Sigma^{-1}\nu')e\Sigma^{-1} - (\nu\Sigma^{-1}e')\nu\Sigma^{-1}] = \frac{q_2}{\Delta},$$

$$\frac{\partial \hat{d}}{\partial \bar{y}} = -\frac{1}{e\Sigma^{-1}e'}\left[ce\Sigma^{-1} + (e\Sigma^{-1}\nu')\frac{\partial \hat{c}}{\partial \bar{y}}\right] = -\frac{1}{e\Sigma^{-1}e'}\left[ce\Sigma^{-1} - (e\Sigma^{-1}\nu')\frac{cq_1}{\Delta}\right]$$

$$= \frac{c}{\Delta}q_2,$$

$$\frac{\partial \hat{d}}{\partial S} = \frac{1}{e\Sigma^{-1}e'}\left[\frac{\partial Q_1}{\partial S} + d\Sigma^{-1}e'e\Sigma^{-1}\right],$$

where $Q_1 = eS^{-1}\bar{x}' - ceS^{-1}\bar{y}'$. After some algebra, one obtains $\partial Q_1/\partial S = -d\Sigma^{-1}e'e\Sigma^{-1}$, which completes the proof. □

In order to obtain a confidence region for (c, d), we may use Theorem 3 as follows:

$$(N_1N_2)^{\frac{1}{2}}(\hat{c} - c, \hat{d} - d) \sim N(0, G)\,,$$

where

$$G = \frac{N^{\frac{3}{2}}\hat{\Delta}}{N_1\hat{c}^2 + N_2}\begin{pmatrix} eS^{-1}e' & -eS^{-1}\bar{y}' \\ -\bar{y}S^{-1}e' & \bar{y}S^{-1}\bar{y}' \end{pmatrix}^{-\frac{1}{2}};$$

$$\hat{\Delta} = (eS^{-1}e')(\bar{y}S^{-1}\bar{y}') - (eS^{-1}\bar{y}')^2\,.$$

4. Testing for magnitude and direction. The main hypotheses of interest are $H_1: \mu = c\nu$, $H_2: \mu = c\nu + de$, where c and d are unknown. The MLE for unrestricted μ, ν, and Λ are

$$\hat{\mu} = \bar{x}\,, \qquad \hat{\nu} = \bar{y}\,, \qquad \hat{\Lambda} = NS^{-1}\,,$$

so that

$$P_\Omega \equiv \max_{\mu,\nu,\Lambda} p(\bar{x}, \bar{y}, S; \mu, \nu, \Lambda) = K(S)|NS^{-1}|^{N/2}e^{-pN/2}\,. \tag{4.1}$$

From (3.16) combined with (3.3),

$$\begin{aligned} P_{\omega_1} &\equiv \max_{\mu=c\nu;\Lambda>0} p(\bar{x}, \bar{y}, S; \mu, \nu, \Lambda) \\ &= \frac{K(S)|NS^{-1}|^{N/2}e^{-\frac{1}{2}pN}}{\{1 + \frac{1}{2}[N_1t_{11} + N_2t_{22} - [(N_1t_{11} - N_2t_{22})^2 + 4N_1N_2t_{12}^2]^{\frac{1}{2}}]\}^{N/2}}\,; \end{aligned} \tag{4.2}$$

from (3.3), (3.13) and (3.15), we have

$$\begin{aligned} P_{\omega_2} &\equiv \max_{\mu=c\nu+de;\Lambda>0} p(\bar{x}, \bar{y}, S; \mu, \nu, \Lambda) \\ &= \frac{K(S)|NS^{-1}|^{N/2}e^{-\frac{1}{2}pN}}{[1+\frac{1}{2}\{N_1\bar{x}A\bar{x}'+N_2\bar{y}A\bar{y}'-[(N_1\bar{x}A\bar{x}'-N_2\bar{y}A\bar{y}')^2+4N_1N_2(\bar{x}A\bar{y}')^2]^{\frac{1}{2}}\}]^{N/2}}\,, \end{aligned} \tag{4.3}$$

where A is defined by (3.14).

From (4.1) and (4.2) the LRT for testing $H_1: \mu = c\nu$, $-\infty < c < \infty$, $-\infty < \nu_j < \infty$ versus $H: -\infty < \mu_j < \infty$, $-\infty < \nu_j < \infty$, $j = 1, \cdots, p$ is then given by: reject H_1 if λ_1 is small, where

$$\begin{aligned} \lambda_1^{-2/N} = 1 + \tfrac{1}{2}\{&N_1\bar{x}S^{-1}\bar{x}' + N_2\bar{y}S^{-1}\bar{y}' \\ &- [(N_1\bar{x}S^{-1}\bar{x}' - N_2\bar{y}S^{-1}\bar{y}')^2 + 4N_1N_2(\bar{x}S^{-1}\bar{y}')^2]^{\frac{1}{2}}\}\,. \end{aligned} \tag{4.4}$$

This test is obtained by Anderson (1951, Section 7) in a more general context, and the right-hand side of (4.4) is $1 + \varphi$ where φ is the smallest root of the matrix

$$\begin{pmatrix} N_1t_{11} & (N_1N_2)^{\frac{1}{2}}t_{12} \\ (N_1N_2)^{\frac{1}{2}}t_{12} & N_2t_{22} \end{pmatrix}.$$

(Also see Cochran (1943).) As an approximate test we have that when H_1 is true, $-2\log\lambda_1$ has a χ^2_{p-1} distribution. Under the alternative hypothesis $-2\log\lambda_1$ has an asymptotic distribution which is non-central chi-square with $p - 1$ degrees of freedom and non-centrality parameter as given in Section 5.1.

From (4.1) and (4.2), the LRT for testing $H_2: \mu = c\nu + de$, $-\infty < c, d < \infty$, $-\infty < \nu_j < \infty$ versus H is given by: reject H_2 if λ_2 is small, where

$$\lambda_2^{-2/N} = 1 + \tfrac{1}{2}\{N_1\bar{x}A\bar{x}' + N_2\bar{y}A\bar{y}' - [(N_1\bar{x}A\bar{x}' - N_2\bar{y}A\bar{y}')^2 + 4N_1N_2(\bar{x}A\bar{y}')^2]^{\frac{1}{2}}\} .$$

To carry out the test we use the fact that when hypothesis H_2 is true, $-2\log\lambda_2$ has an asymptotic χ^2-distribution with $p-2$ degrees of freedom. Under the alternative hypothesis, $-2\log\lambda_2$ has an asymptotic distribution which is non-central chi-square with $p-2$ degrees of freedom and non-centrality parameter as given in Section 5.2.

Finally, suppose $\mu_i = c\nu_i + d$, $1 \leqq i \leqq p$, and we wish to test that $d = 0$. The LRT is then given by the ratio of (4.2) to (4.3), namely

$$\lambda_{12}^{2/N} = \frac{1+\frac{1}{2}\{N_1\bar{x}S^{-1}\bar{x}'+N_2\bar{y}S^{-1}\bar{y}'+[(N_1\bar{x}S^{-1}\bar{x}'-N_2\bar{y}S^{-1}\bar{y}')^2+4N_1N_2(\bar{x}S^{-1}\bar{y}')^2]^{\frac{1}{2}}\}}{1+\frac{1}{2}\{N_1\bar{x}A\bar{x}'+N_2\bar{y}A\bar{y}'+[(N_1\bar{x}A\bar{x}'-N_2\bar{y}A\bar{y}')^2+4N_1N_2(\bar{x}A\bar{y}')^2]^{\frac{1}{2}}\}} .$$

When the hypothesis $\mu = c\nu$ is true, the asymptotic distribution of $-2\log\lambda_{12}$ is χ^2 with 1 degree of freedom, and under the alternative the asymptotic distribution of $-2\log\lambda_{12}$ is a non-central χ^2 distribution with 1 degree of freedom and non-centrality parameter as given in Section 5.3.

5. Asymptotic distribution for the non-central case. Given the density function $p(x; \theta_1, \cdots, \theta_l)$ we wish to test $H: \xi_1(\theta) = \cdots = \xi_r(\theta) = 0$ against general alternatives. If certain regularity conditions are satisfied, Wald (1943) showed that under the alternative hypothesis $-2\log\lambda$, where λ is the likelihood ratio statistic, approaches a non-central chi-squared distribution with r degrees of freedom and non-centrality parameter given as follows. Define $c_{ij} = -E\partial^2 \log p/\partial\theta_i\partial\theta_j$, $i, j = 1, \cdots, l$ and $C = (c_{ij}): l \times l$; $\xi_{ij} = \partial\xi_i/\partial\theta_j$, $i = 1, \cdots, r$; $j = 1, \cdots, l$, $\Xi = (\xi_{ij}): r \times l$; $\xi = (\xi_1, \cdots, \xi_r)$, $\xi_j \equiv \xi_j(\theta)$. The non-centrality parameter is the quadratic form $\delta = \xi(\Xi C^{-1}\Xi')^{-1}\xi'$.

As our starting point we have the joint density

$$p(\bar{x}, \bar{y}, S) = K(S)|\Lambda|^{N/2} \exp -\tfrac{1}{2}[(\bar{x} - c\nu - de)\Lambda(\bar{x} - c\nu - de)' + (\bar{y} - \nu)\Lambda(\bar{y} - \nu)' + \operatorname{tr}\Lambda S] . \tag{5.1}$$

5.1. *Test for magnitude.* Given the joint density (5.1) with $d = 0$, we wish to test

$$H: \mu_1 - \frac{\mu_p}{\nu_p}\nu_1 = \cdots = \mu_{p-1} - \frac{\mu_p}{\nu_p}\nu_{p-1} = 0$$

against general alternatives. Here we use the parametrization $\xi_i = \mu_i - (\mu_p/\nu_p)\nu_i$, $i = 1, \cdots, p-1$ $(r = p - 1)$. If we partition

$$\Xi = (\Xi_1, \Xi_2, \Xi_3) ,$$

where

$$\Xi_1 = \left(\frac{\partial\xi_i}{\partial\mu_j}\right), \qquad \Xi_2 = \left(\frac{\partial\xi_i}{\partial\nu_j}\right), \qquad \Xi_3 = \left(\frac{\partial\xi}{\partial\lambda_{\alpha\beta}}\right), \tag{5.2}$$

with $\Lambda = (\lambda_{\alpha\beta})$, we see that $\Xi_3 = 0$. In order to determine C, we have

$$\frac{\partial^2 \log p}{\partial\mu_i \partial\mu_j} = \frac{\partial^2 \log p}{\partial\nu_i \partial\nu_j} = -\lambda_{ij}\,, \qquad \frac{\partial^2 \log p}{\partial\mu_i \partial\nu_j} = 0\,,$$

$$E\frac{\partial^2 \log p}{\partial\mu_i \partial\lambda_{jk}} = 0\,, \qquad E\frac{\partial^2 \log p}{\partial\nu_i \partial\lambda_{jk}} = 0\,,$$

so that $C = \operatorname{diag}(\Lambda, \Lambda, C_{33})$. Because $\Xi_3 = 0$, the value of C_{33} is not required. Let $\dot\nu = (\nu_1, \cdots, \nu_{p-1})$, then

$$\Xi_1 = \left(I, -\frac{\dot\nu'}{\nu_p}\right), \qquad \Xi_2 = -\frac{\mu_p}{\nu_p}\left(I, -\frac{\dot\nu'}{\nu_p}\right) = -\frac{\mu_p}{\nu_p}\,\Xi_1\,,$$

and hence

$$\Xi C^{-1}\Xi' = \Xi_1\Sigma\Xi_1' + \Xi_2\Sigma\Xi_2' = \left(I, -\frac{\dot\nu'}{\nu_p}\right)\Sigma\left(I, -\frac{\dot\nu'}{\nu_p}\right)'\left(1 + \frac{\mu_p^2}{\nu_p^2}\right).$$

If we write $\dot\mu = (\mu_1, \cdots, \mu_{p-1})$, then the non-centrality parameter becomes

$$\delta = (\dot\mu - \dot\nu\mu_p/\nu_p)\,\frac{\left[\left(I, -\frac{\dot\nu'}{\nu_p}\right)\Sigma\left(I, -\frac{\dot\nu'}{\nu_p}\right)'\right]^{-1}}{1 + (\mu_p^2/\nu_p^2)}\,(\dot\mu - \dot\nu\mu_p/\nu_p)'\,.$$

5.2. *Test for magnitude and displacement.* Given the joint density (5.1) with $d = 0$, we wish to test

$$H: \xi_j \equiv \mu_j - \mu_p - \frac{\mu_{p-1} - \mu_p}{\nu_{p-1} - \nu_p}(\nu_j - \nu_p) = 0\,, \qquad j = 1, \cdots, p-2\,.$$

For convenience, we write

$$\eta_j = \mu_j - \mu_p\,, \qquad \tau_j = \nu_j - \nu_p\,, \qquad j = 1, \cdots, p\,.$$

Partition $\Xi = (\Xi_1, \Xi_2, \Xi_3)$ as in (5.2), and further

$$\Xi_1 = (\Xi_{11}, \Xi_{12}, \Xi_{13})\,, \qquad \Xi_2 = (\Xi_{21}, \Xi_{22}, \Xi_{23})\,,$$

$$\Xi_{11} = \frac{\partial\xi_i}{\partial\mu_j}\,, \quad j = 1, \cdots, p-2\,, \qquad \Xi_{12} = \frac{\partial\xi_i}{\partial\mu_{p-1}}\,, \qquad \Xi_{13} = \frac{\partial\xi_i}{\partial\mu_p}\,,$$

$$\Xi_{21} = \frac{\partial\xi_i}{\partial\nu_j}\,, \quad j = 1, \cdots, p-2\,, \qquad \Xi_{22} = \frac{\partial\xi_i}{\partial\nu_{p-1}}\,, \qquad \Xi_{23} = \frac{\partial\xi_i}{\partial\nu_p}\,.$$

A direct computation yields

$$\Xi_1 = \left(I, -\frac{\tau'}{\tau_{p-1}}, -e' + \frac{\tau'}{\tau_{p-1}}\right)$$

$$\Xi_2 = -\frac{\eta_{p-1}}{\tau_{p-1}}\left(I, -\frac{\tau}{\tau_{p-1}}, -e' + \frac{\tau'}{\tau_{p-1}}\right) = -\frac{\eta_{p-1}}{\tau_{p-1}}\,\Xi_1\,,$$

where $\tau = (\tau_1, \cdots, \tau_{p-2})$. With C as in Section 5.1, the non-centrality parameter is given by

$$(\eta - \tau\eta_{p-1}/\tau_{p-1}) \frac{\left[\left(I, -\frac{\tau'}{\tau_{p-1}}, -e' + \frac{\tau'}{\tau_{p-1}}\right)\Sigma\left(I, -\frac{\tau'}{\tau_{p-1}}, -e' + \frac{\tau'}{\tau_{p-1}}\right)'\right]^{-1}}{1 + (\eta_{p-1}^2/\tau_{p-1}^2)} \times (\eta - \tau\eta_{p-1}/\tau_{p-1})'$$

where $\eta = (\eta_1, \cdots, \eta_{p-2})$.

5.3. *Test for direction in the presence of a magnitude change.* We are now given the joint density (5.1), and we wish to test $H: d = 0$. It is straightforward to verify that

$$C = \begin{pmatrix} e\Lambda e' & e\Lambda\nu' & ce\Lambda & 0 \\ \nu\Lambda e' & \nu\Lambda\nu' & c\nu\Lambda & 0 \\ c\Lambda e' & c\Lambda\nu' & (1 + c^2)\Lambda & 0 \\ 0 & 0 & 0 & L \end{pmatrix},$$

where the diagonal elements are the expectations of the derivatives of $\log p$ with respect to d, c, ν, and Λ, respectively. Since ξ is a single element equal to d, $\Xi = (1, 0)$, so that $\Xi C^{-1}\Xi' = c^{11}$, where c^{11} is found from

$$1/c^{11} = e\Lambda e' - (e\Lambda\nu', ce\Lambda)\begin{pmatrix} \nu\Lambda\nu' & c\nu\Lambda \\ c\Lambda\nu' & (1 + c^2)\Lambda \end{pmatrix}^{-1}\begin{pmatrix} e\Lambda\nu' \\ c\Lambda e' \end{pmatrix}$$
$$= \frac{e\Lambda e'\nu\Lambda\nu' - (e\Lambda\nu')^2}{(1 + c^2)\nu\Lambda\nu'};$$

consequently, the non-centrality parameter is d^2/c^{11}.

REFERENCES

[1] Anderson, T. W. (1951). Estimating linear restrictions on regression coefficients for multivariate normal distributions. *Ann. Math. Statist.* **22** 327–351.

[2] Anderson, T. W. (1958). *An Introduction to Multivariate Statistical Analysis.* Wiley, New York.

[3] Cochran, W. G. (1943). The comparison of different scales of measurement for experimental results. *Ann. Math. Statist.* **14** 205–216.

[4] Paulson, E. (1942). A note on the estimation of some mean values for a bivariate distribution. *Ann. Math. Statist.* **13** 440–445.

[5] Sitgreaves, R. (1952). On the distribution of two random matrices used in classification procedures. *Ann. Math. Statist.* **23** 263–270.

[6] Wald, A. (1943). Tests of statistical hypotheses concerning several parameters when the number of observations is large. *Trans. Amer. Math. Soc.* **54** 426–482.