The Annals of Statistics
1986, Vol. 14, No. 4, 1555–1571

MAGNITUDINAL EFFECTS IN THE NORMAL MULTIVARIATE MODEL

BY IRWIN GUTTMAN,[1] U. MENZEFRICKE[2] AND DAVID TYLER[3]

University of Toronto, University of Toronto and Rutgers University

Suppose the $(k \times 1)$ vectors $\mathbf{x}$ and $\mathbf{y}$ are independent with $\mathbf{x} \sim N(\boldsymbol{\mu}, \Sigma)$ and $\mathbf{y} \sim N(\eta, \Sigma)$, Σ positive definite. If for a positive scalar c, $\boldsymbol{\eta} = c\boldsymbol{\mu}$, we find the posterior of c, using noninformative priors, given the data $\{\mathbf{x}_i\}_1^{N_1}, \{\mathbf{y}_j\}_1^{N_2}$. The $\mathbf{x}_i$ are N_1 independent observations on $\mathbf{x}$, and independent of the $\mathbf{y}_j$, which are N_2 independent observations on $\mathbf{y}$. The constant c is called the magnitudinal effect, and the posterior of c turns out to involve a truncated Student-t kernel. We also discuss the situation in which we wish to examine the truth of the statement $\boldsymbol{\eta} = c\boldsymbol{\mu}$, and proceed as follows. We first note that the matrix $\Lambda = (\lambda_{ij})$, where $\lambda_{11} = N_1\boldsymbol{\mu}'\Sigma^{-1}\boldsymbol{\mu}$, $\lambda_{12} = \lambda_{21} = \sqrt{N_1N_2}\,\boldsymbol{\mu}'\Sigma^{-1}\boldsymbol{\eta}$, and $\lambda_{22} = N_2\boldsymbol{\eta}'\Sigma^{-1}\boldsymbol{\eta}$, has a zero eigenroot if and only if $\boldsymbol{\eta} = c\boldsymbol{\mu}$, for some c. Hence, we are motivated to find the (joint) posterior distribution of ω_1, ω_2, the roots of Λ, where $\omega_1 > \omega_2 \geq 0$. Then, by integration with respect to ω_1 over the region $\omega_1 > \omega_2$, we may find the marginal of ω_2, and use it to examine the statement $\boldsymbol{\eta} = c\boldsymbol{\mu}$.

The posterior of (ω_1, ω_2) involves the multivariate hypergeometric function ${}_1F_1^{(2)}$, which in practice creates computational difficulties. Accordingly, some numerical considerations are discussed for computing of the posterior of ω_2, and an example using real data is given.

1. Introduction and summary. Kraft, Olkin and Van Eeden (1972) discuss the following interesting problem. (We refer to this paper as the KOV paper.) Suppose two methods are used and their effects can be evaluated on k characteristics, such that Method 1 produces measurement $\mathbf{x}$ and Method 2 produces measurement $\mathbf{y}$, where the k-dimensional random vectors $\mathbf{x}$ and $\mathbf{y}$ have a normal$(\boldsymbol{\mu}, \Sigma)$ and normal$(\boldsymbol{\eta}, \Sigma)$ distribution, respectively. A question of interest that may arise is whether the model

$$\boldsymbol{\eta} = c\boldsymbol{\mu}, \qquad c > 0, \tag{1.1}$$

does, or does not, hold. If the model (1.1) holds, then it may be of interest to estimate the so-called magnitudinal effect, c.

Applications of this problem abound. The KOV paper gives an interesting application to the field of medicine, where Method 1 and Method 2 are two drug treatments, measurements are on k symptoms, and if (1.1) holds, then the drugs are doing the same kind of work, up to dosage. Still another application is in chemical engineering: In the manufacture of soap, two different methods may be employed and $k = 2$ characteristics of the soap measured, that is, ability to

Received October 1983; revised September 1985.

[1] Research supported by Grant A8743 from NSERC of Canada.

[2] Research supported by Grant A5578 from NSERC of Canada.

[3] Research supported by NSF grants MCS-8301354 and DMS-8405325.

AMS 1980 *subject classifications*. Primary 62F15; secondary 62H99.

Key words and phrases. Magnitudinal effect; posterior distributions; generalized hypergeometric functions of matrix arguments.

lather, and mildness. The two methods differ in whether a catalyst of type A is used or a catalyst of type B. If model (1.1) holds, then, up to quantity, the two catalysts are equivalent.

A quite different application is in the monitoring of nursing homes and in particular whether private and government homes are spending in similar fashion with respect to four types of cost (i.e., $k = 4$), namely cost of nursing labor, cost of dietary labor, cost of plant operation and maintenance labor, and cost of housekeeping and laundry labor. Yet another application is given in Section 4. The list of applications is indeed quite extensive.

As noted earlier, the scalar c has been called the magnitudinal effect in the KOV paper, where this problem is approached from the classical sampling route. To test whether (1.1) is true, KOV employ the likelihood ratio test and supply a test procedure for large sample sizes based on the well known asymptotic distribution of the likelihood ratio test criterion. They also discuss the problem of finding the maximum likelihood estimate $\hat{c}$ of the scalar c, given that (1.1) is true and supply a confidence interval for c for large samples by finding the asymptotic distribution of $\hat{c}$ using well known limit theorems. All this is on the basis of data on $\mathbf{x}$ and $\mathbf{y}$, say N_1 independent observations $\mathbf{x}_i$ on $\mathbf{x}$, $i = 1, \ldots, N_1$, independent of N_2 observations $\mathbf{y}_j$ on $\mathbf{y}$, $j = 1, \ldots, N_2$. We let

$$X = \{\mathbf{x}_i\}_1^{N_1} \quad \text{and} \quad Y = \{\mathbf{y}_j\}_1^{N_2} \tag{1.2}$$

denote the two samples obtained from the two methods.

This paper is mainly given over to a discussion of the case that one first wishes to examine whether (1.1) is true or not. For this situation, we have a Bayesian approach to the problem, given the data (X, Y), which involves examining the smallest root of a certain matrix, and we do the examination by using the posterior distribution of this root. This is developed and explained in Section 2. Some numerical considerations are discussed in Section 3, and in Section 4 we illustrate how to use the results of this approach for an example involving real data.

We also discuss the case of estimation of the scaler c (the so-called magnitudinal effect given that (1.1) holds. This is developed, using a Bayesian approach in Section 5, and illustrated in Section 6.

2. Does the magnitudinal model hold? In this section we address the question of how to make inference about whether or not the magnitudinal model (1.1) is appropriate. We will of course, need sample information to do this, and to this end, we assume the data $X = \{\mathbf{x}_i, i = 1, \ldots, N_1\}$ and $Y = \{\mathbf{y}_j, j = 1, \ldots, N_2\}$ are available, where X and Y represent independent samples of independent observations from a normal$(\boldsymbol{\mu}, \Sigma)$ and a normal$(\boldsymbol{\eta}, \Sigma)$, respectively. We let

$$\bar{\mathbf{x}} = N_1^{-1} \sum_1^{N_1} \mathbf{x}_i, \qquad \bar{\mathbf{y}} = N_2^{-1} \sum_1^{N_2} \mathbf{y}_j, \qquad S = S_1 + S_2, \tag{2.1$'$}$$

where

$$S_1 = \sum_1^{N_1} (\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})' \quad \text{and} \quad S_2 = \sum_1^{N_2} (\mathbf{y}_j - \bar{\mathbf{y}})(\mathbf{y}_j - \bar{\mathbf{y}})'. \tag{2.1$''$}$$

We remind the reader that if the prior for μ, η and Σ^{-1} is diffuse, so that the noninformative prior for the parameters is appropriate and used, that is, if we take

$$p(\mu, \eta, \Sigma^{-1}) \propto |\Sigma^{-1}|^{-(k+1)/2}, \tag{2.2}$$

then the posterior distribution for μ, η and Σ^{-1} is characterized by

(i) the marginal posterior of Σ^{-1} is such that

$$p(\Sigma^{-1}|X, Y) \propto |\Sigma^{-1}|^{(n-k-1)/2} \operatorname{etr}\left(-\tfrac{1}{2}\Sigma^{-1}S\right), \tag{2.3}$$

with $n = N_1 + N_2 - 2 = N - 2 \geq k$, and

(ii) the conditional posterior for μ, η, given Σ^{-1}, is such that μ and η are independent with

$$N_1^{1/2}\mu \sim \operatorname{normal}\left(N_1^{1/2}\bar{\mathbf{x}}, \Sigma\right) \quad \text{and} \quad N_2^{1/2}\eta \sim \operatorname{normal}\left(N_2^{1/2}\bar{\mathbf{y}}, \Sigma\right). \tag{2.4}$$

[See, for example, Tiao and Zellner (1964).] We note that the density of (2.3) is that of a central Wishart distribution of order k, with n degrees of freedom, and positive definite parameter matrix S^{-1}.

Our objective now is to use the information in the posterior distribution to determine whether or not the magnitudinal model (1.1) holds. A somewhat obvious approach to this problem is to construct, at some level, simultaneous posterior ellipsoidal regions for μ, and for η, and examine whether any line through the origin intersects both ellipsoids. The approach, however, does not address the question directly since the construction of the ellipsoids places no special emphasis on the degree of departure of μ and η from a straight line through the origin.

Another obvious approach stems from consideration of the alternative to (1.1), namely

$$\eta = C\mu, \qquad C = \operatorname{diag}(c_1, \ldots, c_k), \qquad c_i \geq 0. \tag{2.5}$$

The approach that then suggests itself is to find the posterior of the elements $(c_1, \ldots, c_k)$. However, for reasons best explained in Guttman, Menzefricke and Tyler (1985), this approach too leads to a somewhat dead end.

Alternatively, we propose examining the posterior distribution of ω_2, the smallest eigenvalue of the 2×2 symmetric nonnegative definite matrix Λ, where

$$\Lambda = M'\Sigma^{-1}M = \begin{pmatrix} N_1\mu'\Sigma^{-1}\mu & (N_1N_2)^{1/2}\mu'\Sigma^{-1}\eta \\ (N_1N_2)^{1/2}\mu'\Sigma^{-1}\eta & N_2\eta'\Sigma^{-1}\eta \end{pmatrix} \tag{2.6}$$

and where $M = [N_1^{1/2}\mu\, N_2^{1/2}\eta]$. We note that $\omega_2 \geq 0$ with equality if and only if $\eta = c\mu$, since $\eta = c\mu$ if and only if M and hence Λ has less than full rank. It is interesting to also note that ω_2 has the following geometric interpretation. Suppose we fit a straight line through the origin that minimizes the sum of the squared distances between $N_1^{1/2}\mu$ and $N_2^{1/2}\eta$ from the line, where distance is measured by the orthogonal distance with respect to the inner product $(\mathbf{a}, \mathbf{b})_{\Sigma} = \mathbf{a}'\Sigma^{-1}\mathbf{b}$, then it can be shown that the minimum sum of squared distances so obtained is ω_2. More formally, by using fundamental results from linear algebra it can be shown that

$$\omega_2 = \min_{\mathbf{x} \in \mathbb{R}} \left\{ \min_{d_1 \in \mathbb{R}} \|N_1^{1/2}\boldsymbol{\mu} - d_1\mathbf{x}\|_\Sigma^2 + \min_{d_2 \in \mathbb{R}} \|N_2^{1/2}\boldsymbol{\eta} - d_2\mathbf{x}\|_\Sigma^2 \right\}, \tag{2.7}$$

where $\|\mathbf{a}\|_\Sigma^2 = (\mathbf{a}, \mathbf{a})_\Sigma$. In view of this, if the posterior distribution of ω_2 is "concentrated about zero," the data would be compatible with the magnitudinal model. What we mean by "concentrated about zero" is discussed at the end of this section.

We first give the joint posterior density of the roots of Λ. The proof of this result is given in Appendix 1.

THEOREM 2.1. *For $2 \le k \le n$, the posterior density of (ω_1, ω_2), where $\omega_1 > \omega_2$ are the roots of Λ defined in* (2.6), *is*

$$\begin{aligned} p(\omega_1, \omega_2 | X, Y) &= K_0(\omega_1\omega_2)^{(k-3)/2} e^{-(\omega_1+\omega_2)/2}(\omega_1 - \omega_2) \\ &\quad \times {}_1F_1^{(2)}\left\{\tfrac{1}{2}n; \tfrac{1}{2}k; \tfrac{1}{2}L_0(I + L_0)^{-1}, \Lambda_0\right\}, \qquad \omega_1 > \omega_2 \ge 0, \end{aligned} \tag{2.8}$$

where $K_0 = \pi^{1/2}\{2^k\Gamma(\frac{1}{2}k)\Gamma[\frac{1}{2}(k-1)]\}^{-1}\{(1 + l_1)(1 + l_2)\}^{-n/2}$, *with* $l_1 \ge l_2 > 0$ *being the roots of*

$$L = T'S^{-1}T, \quad \textit{with } T = \left[N_1^{1/2}\bar{\mathbf{x}}, N_2^{1/2}\bar{\mathbf{y}}\right], \tag{2.8$'$}$$

and where

$$L_0 = \operatorname{diag}\{l_1, l_2\}, \qquad \Lambda_0 = \operatorname{diag}(\omega_1, \omega_2). \tag{2.8$''$}$$

The reader is referred to the text by Muirhead (1982) for a good review of the definition and properties of the generalized hypergeometric function ${}_1F_1^{(2)}$ introduced in the theorem, in particular see his Chapter 7. It is important to note here that the function ${}_1F_1^{(2)}$ depends upon its matrix arguments only through the roots of the arguments, and is a symmetric function of these roots. Interestingly, there is a connection between the ${}_1F_1^{(2)}$ function with the classical univariate hypergeometric function of Gauss,

$${}_2F_1(a, b; c, y) = \sum_{q=0}^{\infty} \frac{(a)_q b_{(q)}}{(c)_q} \frac{y^q}{q!}, \tag{2.9}$$

where $(c)_t = c(c+1)\cdots(c+t-1)$. The relationship as given by Muirhead (1975), Lemma 1.2 [we note the omission in (1.6) of page 285 of Muirhead (1975) of a $1/k!$ (Muirhead's notation) in the summand] is

$$\begin{aligned} {}_1F_1^{(2)}\{a; c; H, U\} &= \sum_{t=0}^{\infty} \frac{(a)_t(c-a)_t}{\left(c - \frac{1}{2}\right)_t (c)_{2t}} \frac{(-h_1h_2u_1u_2)^2}{t!} \\ &\quad \times \sum_{j=0}^{\infty} \frac{(a+t)_j}{(c+2t)_j} \frac{\left\{\frac{1}{2}(h_1 + h_2)(u_1 + u_2)\right\}^j}{j!} \\ &\quad \times {}_2F_1\left(-\frac{1}{2}j, -\frac{1}{2}j + \frac{1}{2}; 1; x^2\right), \end{aligned} \tag{2.10}$$

where $H = \operatorname{diag}(h_1, h_2)$, with $h_1 > h_2$, and $U = \operatorname{diag}(u_1, u_2)$, $u_1 > u_2$, $x = \{(h_1 - h_2)(u_1 - u_2)\}/\{(h_1 + h_2)(u_1 + u_2)\}$. The ${}_2F_1$ terms arising from (2.10) have a finite expansion, since for $q > \frac{1}{2}j$, either $(-\frac{1}{2}j)_q = 0$ or $(-\frac{1}{2}j + \frac{1}{2})_q = 0$.

By inserting the expansion (2.10) into the density (2.8), making the change of indices $p = 2t + j$, $t = t$ and integrating over ω_1, we have:

THEOREM 2.2. *The marginal posterior density of ω_2, the smallest eigenvalue of the matrix Λ of* (2.6), *may be expressed as*

$$(2.11)\quad p(\omega_2|X,Y) = K_0\omega_2^{(k-3)/2}e^{-\omega_2/2}\sum_{p=0}^{\infty} c_p(\omega_2)\bar{r}^p/\left(\tfrac{1}{2}k\right)_p, \qquad \omega_2 \geq 0,$$

where K_0 is defined in Theorem 2.1, $\bar{r} = \frac{1}{2}(r_1 + r_2)$ *with* $r_i = \frac{1}{2}l_i/(1 + l_i)$, $i = 1, 2$, *and where*

$$(2.11')\quad c_p(\omega_2) = \sum_{t=0}^{[p/2]} J_t 4^t z_r^t \omega_2^t \left\{ \sum_{q=0}^{[p/2]-t} b_{p,t,q} x_r^q Q\left(\omega_2; p - 2t - 2q, 2q + 1, t + \tfrac{1}{2}k - \tfrac{3}{2}\right)\right\},$$

with $z_r = r_1r_2/(2\bar{r})^2$, $x_r = (r_1 - r_2)^2/(4\bar{r})^2$, $J_t = (-1)^t((k - n)/2)_t/\{((k - 1)/2)_t\}t!$, $b_{p,t,q} = (\frac{1}{2}n)_{p-t}/\{(q!)^2(p - 2t - 2q)!\}$, *and in general*

$$(2.11'')\quad Q(\omega_2; m_1, m_2, m_3) = \int_{\omega_2}^{\infty}(\omega_1 + \omega_2)^{m_1}(\omega_1 - \omega_2)^{m_2}\omega_1^{m_3}e^{-\omega_1/2}\,d\omega_1.$$

By direct integration, we note that an alternative form for Q in Theorem 2.2 is given by

$$(2.12)\quad Q(\omega_2; m_1, m_2, m_3) = \sum_{\nu=0}^{m_1+m_2} C_\nu 2^{\nu+m_3+1}\omega_2^{m_1+m_2-\nu}\Gamma(\nu + m_3 + 1) \times P\left\{\chi^2_{2\nu+2m_3+2} > \omega_2\right\},$$

where

$$(2.13)\qquad C_\nu = \sum_{j=\max(0,\nu-m_1)}^{\min(m_2,\nu)} \binom{m_1}{\nu - j}\binom{m_2}{j}(-1)^{m_2-j}$$

are the coefficients in the expansion $(s + 1)^{m_1}(s - 1)^{m_2} = \sum_{\nu=0}^{m_1+m_2} C_\nu s^\nu$. For the case when m_3 is also a nonnegative integer, which corresponds to k being odd, we have the form

$$(2.14)\quad Q(\omega_2; m_1, m_2, m_3) = c^{-\omega_2/2}\sum_{\nu=0}^{m_1+m_3} B_\nu 2^{m_2+\nu+1}\Gamma(m_2 + \nu + 1)\omega_2^{m_1+m_3-\nu},$$

where

$$(2.14')\qquad B_\nu = \sum_{j=\max(0,\nu-m_3)}^{\max(m_1,\nu)} \binom{m_1}{j}\binom{m_3}{\nu - j}2^{m_1-j}$$

are the coefficients in the expansion $(s + 2)^{m_1}(s + 1)^{m_3} = \sum_{\nu=0}^{m_1+m_3} B_\nu s^\nu$.

The posterior moments of ω_2 can be found directly from equations (2.11), (2.11′) and (2.11″). Using the notation of Theorem 2.2, we have

$$E(\omega_2^m|X, Y) = K_0 \sum_{p=0}^{\infty} \mu_{m,p}\Gamma(p + k + m)2^{p+1}\bar{r}^p/(\tfrac{1}{2}k)_p, \tag{2.15}$$

where

$$\mu_{m,p} = \sum_{t=0}^{[p/2]} J_t z_r^t \left\{ \sum_{q=0}^{[p/2]-t} b_{p,t,q} x_r^q a_{m,t,q} \right\} \tag{2.16}$$

and

$$a_{m,t,q} = \sum_{\nu=0}^{m} (-1)^\nu \binom{m}{\nu} \beta\{\tfrac{1}{2}(\nu + 2q + 2), \tfrac{1}{2}(2t + k - 1)\},$$

with $\beta(\cdot, \cdot)$ denoting the beta function.

We now address the question of how to decide whether the posterior distribution of ω_2 is sufficiently "concentrated about zero" and hence that the magnitudinal model is feasible. One approach, which may seem intuitively appealing, would be to see if zero is, for some level, in the highest posterior density (HPD) interval. The drawback to this approach is that for $k = 2$, the posterior for ω_2 has an asymptote at zero, regardless of the data, and hence the HPD interval will always contain zero. For $k \geq 4$, the posterior density of ω_2 at zero is always zero, regardless of the data, and hence the HPD interval will never contain zero for these cases.

In order to understand the information given in the posterior distribution of ω_2, some reference point is needed. This necessity of a reference point becomes apparent after noting that when $l_1 = l_2 = 0$, which occurs when $\bar{\mathbf{x}} = \bar{\mathbf{y}} = \mathbf{0}$, the posterior distribution of ω_2 corresponds to the distribution of the smallest root of a central Wishart distribution of order 2, on k degrees of freedom, and with positive definite matrix parameter I. This implies that for large k, the posterior distribution of ω_2/k would be highly concentrated about 1 whenever $\bar{\mathbf{x}} = \bar{\mathbf{y}} = \mathbf{0}$. However, $\bar{\mathbf{x}} = \bar{\mathbf{y}} = \mathbf{0}$ would be an extreme case in support of the magnitudinal model, and the "upward" bias present in this extreme case in the posterior of ω_2 can be attributed to the fact that ω_2 is a distance measurement.

But because we feel a reference point is needed, we proceed as follows. We first note that if we observe $\bar{\mathbf{x}}$ and $\bar{\mathbf{y}}$ to be proportional (that is, $\bar{\mathbf{y}}_i = \tilde{c}\bar{\mathbf{x}}_i$, $i = 1, \dots, k$), then $l_2 = 0$. Such an event gives extreme sample evidence support that the magnitudinal model (1.1) holds. In order to assess whether the actual observed l_2 is too large for the magnitudinal model to be feasible, we propose comparing the posterior distribution of ω_2 when we observe l_1, l_2, to the posterior distribution of ω_2 under the same l_1, but with $l_2 = 0$. We call the posterior $p(\omega_2|l_1, l_2 = 0)$ a *reference posterior*. The proposed use of such a posterior is, we believe, new, but of course, the use of reference priors has been employed for various situations before (e.g., see Chapter 1 of Box and Tiao (1973)). An application of this approach is illustrated in Section 4.

It is interesting to draw some parallels between the frequentist approach and the Bayesian approach. The likelihood ratio test for testing whether the magnitudinal model holds is based on l_2, whose exact sampling distribution depends on both ω_1 and ω_2. If $\omega_2 = 0$, the distribution of l_2 still depends on the nuisance parameter ω_1, although its asymptotic distribution does not. The KOV paper gives details. In our Bayesian approach, the posterior distribution of ω_2 is a function of l_1 and l_2, which of course are both known for any given sample.

3. Numerical considerations and a mixture representation. In this section, some difficulties that may arise in the computation of the posterior density of ω_2 are discussed. We first note that the expansion for the hypergeometric function ${}_1F_1^{(2)}$ given by (2.10) can converge slowly if the matrix arguments are large, or if a is large. This convergence problem of the power series representation is common to most hypergeometric functions of matrix arguments. Quoting Muirhead ((1978), page 5): "These series... tend to converge extremely slowly for cases of particular interest and it is very difficult to obtain from them any feeling for the behavior of the density... ." We now discuss in more detail how this convergence problem specifically relates to the expansion for the posterior density of ω_2 given in Theorem 2.2.

By inserting the expansion (2.10) into the density (2.8) and making the change of indices $p = 2t + j$ and $t = t$, we have

$$(3.1)\quad p(\omega_1, \omega_2|X, Y) = 2^{-(k-3)}K_0 \sum_{p=0}^{\infty} H_p(V)G_p(\theta)\bar{r}^p/\left(\tfrac{1}{2}k\right)_p, \qquad \omega_1 > \omega_2 > 0,$$

where $V = \omega_1 + \omega_2$, $\theta = (\omega_1 - \omega_2)/(\omega_1 + \omega_2)$, $H_p(V) = V^{p+k-2}e^{-V/2}$, and

$$(3.2)\qquad G_p(\theta) = \sum_{t=0}^{[p/2]} J_t z_r^t \left\{ \sum_{q=0}^{[p/2]-t} b_{p,t,q} x_r^q (1-\theta^2)^{t+(k-3)/2}\theta^{2q+1} \right\}.$$

The constants K_0, z_r, x_r, J_t and $b_{p,t,q}$ are as in Theorem 2.2. We note that integrating (3.1) term by term over ω_1 yields the expansion for the posterior density of ω_2 given by (2.11). Although the constant J_t can be negative, by using the zonal polynomial expansion for the hypergeometric function ${}_1F_1^{(2)}$ in (3.3) and comparing it to the polynomial expansion (3.1), it can be shown that $G_p(\theta)$ is a nonnegative function. The reader is again referred to Muirhead ((1982), Chapter 7) for a discussion on the zonal polynomial expansions for hypergeometric functions of matrix arguments. If the value of $n - k$ is even, then we note that $J_t > 0$ for $t \le \frac{1}{2}(n-k)$ and $J_t = 0$ for $t > \frac{1}{2}(n-k)$.

Since the Jacobian of the transformation (made in (3.1)), $(\omega_1, \omega_2) \to (V, \theta)$, is $\frac{1}{2}V$, the joint posterior density of (V, θ) is thus

$$(3.3)\qquad p(V, \theta|X, Y) = \sum_{p=0}^{\infty} a_p \chi^2_{2p+2k}(V) g_p(\theta), \qquad 0 < \theta < 1,\ V > 0,$$

where $\chi^2_\nu(\cdot)$ represents a chi-square density on ν degrees of freedom, $g_p(\theta)$ is the density obtained from normalizing $G_p(\theta)$, that is,

$$g_p(\theta) = G_p(\theta) \Big/ \int_0^1 G_p(\alpha)\, d\alpha, \tag{3.4}$$

and the weights a_p are given by

$$a_p = 2^{p+2} K_0 \bar{r}^p \Gamma(p+k) \int_0^1 G_p(\alpha)\, d\alpha \big/ \left(\tfrac{1}{2}k\right)_p. \tag{3.5}$$

The posterior distribution of (ω_1, ω_2) can thus be viewed as a compound distribution where, for a given p, the quantities $V = \omega_1 + \omega_2$ and $\theta = (\omega_1 - \omega_2)/(\omega_1 + \omega_2)$ are independent with densities $\chi^2_{2p+2k}(V)$ and $g_p(\theta)$, respectively, and where the probability mass function for p is given by a_p.

By inspecting the weights a_p, we can determine which values of p are important in calculating (3.3) and hence in calculating (2.11). Thus, a more detailed analysis of these weights is warranted. We proceed by first noting that the posterior marginal density of $V = \omega_1 + \omega_2$ can be obtained by integrating (3.3) over θ. This gives

$$p(V|X, Y) = \sum_{p=0}^{\infty} a_p \chi^2_{2p+2k}(V), \qquad V > 0. \tag{3.6}$$

A more direct method for obtaining the posterior density of V is as follows. Since $V = \operatorname{tr}(\Lambda)$, where Λ is defined by (2.6), we note from (3.3) that the posterior distribution of V given Σ^{-1} is a noncentral chi-square on $2k$ degrees of freedom with noncentrality parameter $\delta = N_1 \bar{\mathbf{x}}' \Sigma^{-1} \bar{\mathbf{x}} + N_2 \bar{\mathbf{y}}' \Sigma^{-1} \bar{\mathbf{y}}$. By using the familiar expansion for the noncentral chi-square density as a weighted infinite sum of central chi-square densities, and then taking the expectation with respect to Σ^{-1} we obtain (3.6) with

$$a_p = \{(1 + l_1)(1 + l_2)\}^{-n/2} E\left\{\left(r_1 \chi^2_{1,n} + r_2 \chi^2_{2,n}\right)^p\right\} / p!, \tag{3.7}$$

with $r_j = (l_j/2)/(1 + l_j)$ for $j = 1, 2$, and where $\chi^2_{1,n}$ and $\chi^2_{2,n}$ have independent chi-square distributions, both on n degrees of freedom. Evaluation of the expectation in (3.7) gives

$$a_p = \{(1 + l_1)(1 + l_2)\}^{-n/2} 2^p \sum_{\nu=0}^{p} \left\{\left(\tfrac{1}{2}n\right)_\nu \left(\tfrac{1}{2}n\right)_{p-\nu} r_1^\nu r_2^{p-\nu}\right\} / \{\nu!(p-\nu)!\}. \tag{3.8}$$

Now using a_p as mass function for p (see (3.5)), we may calculate the mean and variance of this mass function, obtaining

$$E(p) = \tfrac{1}{2}n(l_1 + l_2) \quad \text{and} \quad \operatorname{Var}(p) = \tfrac{1}{2}n\left(l_1^2 + l_2^2\right). \tag{3.9}$$

If the mean and variance of p are not too large, calculation of the posterior density of ω_2 may proceed directly via (2.11). The function $Q(\omega_2; m_1, m_2, m_3)$ can be calculated by using either (2.12), (2.14) or recursively for fixed m_3 and ω_2 since in cases of interest m_1 and m_2 are nonnegative integers.

If the mean and variance of p are large, direct computation of the posterior density of ω_2 by (2.11) is not feasible, primarily due to the computation of the function $Q(\omega_2; m_1, m_2, m_3)$ within the summations. In such cases, computations can be saved by first computing the posterior joint density of (ω_1, ω_2) given by (3.1) over a nonrectangular grid

$$\omega_2 = \omega_{2,i} \quad \text{and} \quad (\omega_1 - \omega_2)/(\omega_1 + \omega_2) = \theta_i \tag{3.10}$$

for a range of values of $\omega_{2,i}$ and θ_i, and then numerically integrating over ω_1. This procedure, used in the example discussed in the next section, saves computations since the values for $G_p(\theta_i)$ defined by (3.2) can be stored.

Alternatively, the ${}_1F_1^{(2)}$ function appearing in the posterior joint density of (ω_1, ω_2) given by (2.10) can be calculated by applying numerical integration procedures to an integral representation of the ${}_1F_1^{(2)}$ function. This function can be represented as an integral over the group of orthogonal matrices of order 2 where the integrand is a ${}_1F_1$ hypergeometric function. The ${}_1F_1$ function can in turn be represented as an integral over the set of symmetric positive definite matrices of order 2 where the integrand is a ${}_1F_0$ hypergeometric function, which has the closed form ${}_1F_0(c, M) = |I - M|^{-c}$. For more detail, we again refer the reader to Muirhead ((1982), Chapter 7). Tiao and Fienberg (1969) have used numerical integration for integrating over the group of orthogonal matrices of order 2 in calculating the ${}_0F_0^{(2)}$ hypergeometric function. Finally, asymptotic approximations for the ${}_1F_1^{(2)}$ function can be used when appropriate; see Muirhead (1978) and Glynn (1980).

4. Does the magnitudinal model hold?—An example. Johnson and Wichern ((1982), page 243) give an interesting set of data on electrical usage. Samples of sizes $N_1 = 45$ and $N_2 = 55$ were taken of Wisconsin homeowners with and without air conditioning, respectively. The observations were on electrical usage (kilowatt hours) during

$$(\text{on-peak hours, off-peak hours})'$$

for both groups, and if the magnitudinal model (1.1) holds, the utility can plan accordingly, aim advertising campaigns accordingly, etc.

The resulting statistics (with notation of Section 2) are

$$\begin{aligned} \bar{\mathbf{x}} &= (204.4, 556.6)', \qquad N_1 = 45, \\ \bar{\mathbf{y}} &= (130.0, 355.0)'; \qquad N_2 = 55. \end{aligned} \tag{4.1}$$

(Johnson and Wichern state that the off-peak consumption is higher than the on-peak consumption because there are more off-peak hours in a month.) Also

$$S = \begin{pmatrix} 1{,}074{,}443 & 2{,}107{,}539 \\ 2{,}107{,}539 & 6{,}238{,}807 \end{pmatrix}. \tag{4.2}$$

A quick calculation shows (using the notation in Theorem 2.1) that the matrix

$$L = \begin{pmatrix} 2.26787 & 1.59829 \\ 1.59829 & 1.12642 \end{pmatrix}, \tag{4.3}$$

with roots

$$l_1 = 3.39428, \qquad l_2 = 0.0000134. \tag{4.3'}$$

The next step is to compute $p(\omega_2|X, Y) = p(\omega_2|l_1 = 3.39428,\ l_2 = 0.0000134)$, the posterior of ω_2, for the above data, from (2.11). In accordance to the dictates of Section 3, we first compute the mean and standard deviation of the index p, and find, using (3.9) that

$$E(p) = 16.632, \qquad \text{S.D.}(p) = 23.7599. \tag{4.4}$$

Because of the above results, we use (2.11) with the index of summation ranging over $0 \le p \le 300$, and a calculation of $\sum_{p=0}^{300} a_p$ shows that this range amounts for 99.999% of the probability for p. The results are listed in Table 1 and graphed as

TABLE 1

Tabulation of the posterior $p(\omega_2|X, Y) = p(\omega_2|l_2 = 3.39428,\ l_2 = 0.0000134) = p(\omega_2)$ *and the reference posterior* $p^*(\omega_2|l_1 = 3.39428,\ l_2 = 0) = p^*(\omega_2)$. *The actual output contains* 34 *other values of* $p(\omega_2)$ *and* $p(\omega_2^*)$, *at selected values of* ω_2 *ranging from* 8.1206 *to* 27. *For all these cases,* $|p(\omega_2) - p^*(\omega_2)| < 2 \times 10^{-5}$. *In fact, as* ω_2 *increases,* $|p(\omega_2) - p^*(\omega_2)|$ *decreases rapidly, so that, for example,* $|p(27) - p^*(27)| = 1.75 \times 10^{-9}$. *The entire output is available from the authors. Since the distributions of* ω_2 *are peaked near zero and long-tailed to the right, we have used nonequally spaced points to obtain a clear picture of the distributions.*

ω_2	$p(\omega_2)$	$p^*(\omega_2)$	ω_2	$p(\omega_2)$	$p^*(\omega_2)$
0.000027	60.7312	60.7715	1.06121	0.227792	0.227783
0.000216	28.4825	28.5013	1.15762	0.207818	0.207796
0.000729	14.7067	14.7164	1.25971	0.189290	0.189257
0.001728	9.6441	9.65047	1.36763	0.172109	0.172067
0.003375	6.87252	6.88706	1.48154	0.156190	0.156140
0.005832	5.21769	5.22113	1.60161	0.141454	0.141397
0.009261	4.13050	4.13321	1.72800	0.127830	0.127768
0.013824	3.37184	3.37405	1.86087	0.115251	0.115135
0.019683	2.81818	2.82001	2.00038	0.103658	0.103589
0.027000	2.39723	2.39877	2.14669	0.092992	0.092921
0.035937	2.06858	2.06990	2.29997	0.083201	0.083129
0.046656	1.80579	1.80693	2.46037	0.074233	0.074161
0.059319	1.59136	1.59235	2.62807	0.066040	0.065968
0.074088	1.41344	1.41431	2.80322	0.058573	0.058503
0.091125	1.26365	1.26441	2.98598	0.051788	0.051719
0.110592	1.13592	1.13659	3.17652	0.045640	0.045574
0.132651	1.02579	1.02638	3.37500	0.040088	0.040025
0.157464	0.929879	0.930397	3.58158	0.035090	0.035030
0.185193	0.845615	0.846074	3.79642	0.020606	0.030549
0.216000	0.771007	0.771407	4.01968	0.026597	0.026543
0.250047	0.704476	0.704826	4.25153	0.023026	0.022976
0.287496	0.644784	0.645088	4.49212	0.019858	0.019812
0.328509	0.590932	0.591194	4.74163	0.017058	0.017015
0.373248	0.542104	0.542329	5.00021	0.014593	0.014554
0.421875	0.497635	0.497825	5.26802	0.012432	0.012397
0.474552	0.456985	0.457143	5.54523	0.010547	0.010515
0.531441	0.419705	0.419835	5.83200	0.008908	0.008880
0.592704	0.385414	0.385517	6.12849	0.007491	0.007465
0.658503	0.353796	0.353876	6.43486	0.005270	0.006248
0.729000	0.324591	0.324648	6.75127	0.005225	0.005205
0.804357	0.297567	0.297605	7.07789	0.004332	0.004315
0.884736	0.272532	0.272552	7.41487	0.003575	0.003560
0.970299	0.249323	0.249327	7.76239	0.002936	0.002923

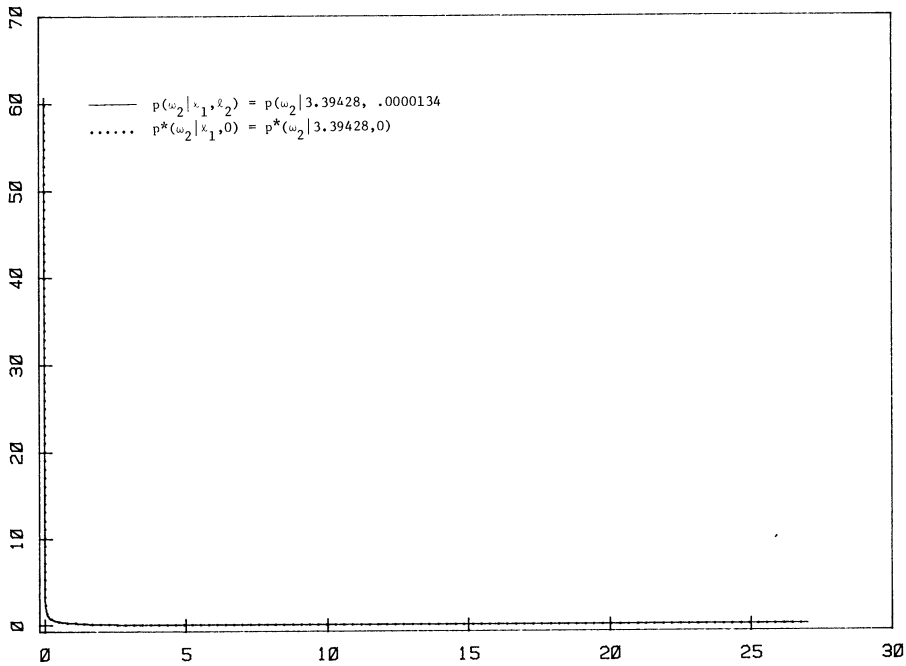

FIG. 1. *The posterior and reference posteriors for the electricity data.*

the solid line in Figure 1. We also (in Table 1) give values of $p^*(\omega_2|l_1 = 3.39428, l_2 = 0)$, the reference posterior that would give extreme sample corroboration that the magnitudinal model holds, and graph $p^*(\omega_2|l_1 = 3.39428, l_2 = 0)$ as the dotted curve in Figure 1. As we can see, there is little to choose between $p(\omega_2|3.39428, 0.0000134)$ and $p^*(\omega_2|3.394282, 0)$. The latter was computed, incidentally, for the same range of the index p using (2.11), since the mean and standard deviation for p for the case $l_1 = 3.339428$, $l_2 = 0$ have values, to two decimal places which are the same as given in (4.4).

The mean and variance of the posterior $p(\omega_2|l_1 = 3.339428, l_2 = 0.0000134)$ turn out to be:

$$(4.5)\qquad \begin{aligned} E(\omega_2|X, Y) &= E(\omega_2|l_1 = 3.39428, l_2 = 0.0000134) = 0.9983, \\ \operatorname{Var}(\omega_2|X, Y) &= \operatorname{Var}(\omega_2|l_1 = 3.39428, l_2 = 0.0000134) = 1.9930. \end{aligned}$$

In contrast, the mean and variance of the reference posterior $p^*(\omega_2|l_1 = 3.39428, l_2 = 0)$ are

$$(4.6)\qquad \begin{aligned} E^*(\omega_2|l_1 = 3.39428, l_2 = 0) &= 0.9969, \\ \operatorname{Var}^*(\omega_2|l_1 = 3.39428, l_2 = 0) &= 1.9878. \end{aligned}$$

Inspection of the differences between the actual posterior and the reference posterior as indicated in Table 1, Figure 1 and (4.5)–(4.6) leads to the conclusion, as indicated, that the data supports the statement

$$(4.7)\qquad \omega_2 = 0, \qquad \eta = c\mu.$$

Hence, we may wish to adopt (4.7) for this set of data and turn now to the question of estimating c. This is discussed and developed in the ensuing section, and the development there is illustrated with the set of data of this section.

5. The posterior distribution of the magnitudinal effect. In this section, we assume that the magnitudinal effect model (1.1) holds, and that interest is in making inference about the magnitudinal parameter c. Specifically, we assume that the k-dimensional random vectors $\mathbf{x}$ and $\mathbf{y}$ are independent and distributed as

$$\mathbf{x} \sim \text{normal}(\boldsymbol{\mu}, \Sigma) \quad \text{and} \quad \mathbf{y} \sim \text{normal}(\boldsymbol{\eta}, \Sigma), \tag{5.1}$$

where

$$\boldsymbol{\eta} = c\boldsymbol{\mu}, \qquad \boldsymbol{\mu} \neq \mathbf{0}, \qquad c > 0. \tag{5.2}$$

As an example of how this model may arise, suppose we are dealing with two drugs so that c represents a change needed in the dosage of the first drug whose effects on k symptoms are measured by $\mathbf{x}$ to make it equivalent to that dosage of the second drug whose effects on the same k symptoms are measured by $\mathbf{y}$. In such a setting, it is not uncommon to have $\text{Var}(\mathbf{x}) = \text{Var}(\mathbf{y}) = \Sigma$ and we operate under this assumption (as does the work in the KOV paper).

Suppose, then that N_1 independent observations on $\mathbf{x}$ are generated, say $X = \{\mathbf{x}_1, \dots, \mathbf{x}_{N_1}\}$ and that X is independent of the N_2 independent observations $\mathbf{y}_j$ taken on $\mathbf{y}$, say $Y = \{\mathbf{y}_1, \dots, y_{N_2}\}$. The likelihood based on X and Y of the parameters $\boldsymbol{\mu}$, c, Σ^{-1} is then

$$\begin{aligned} &l(\mu, c, \Sigma^{-1}|X, Y) \\ &\quad\propto |\Sigma^{-1}|^{N/2} \operatorname{etr}\left[-\tfrac{1}{2}\Sigma^{-1}\left\{\sum_1^{N_1}(\mathbf{x}_i - \boldsymbol{\mu})(\mathbf{x}_i - \boldsymbol{\mu})' \right.\right. \\ &\qquad\qquad \left.\left. + \sum_1^{N_2}(\mathbf{y}_j - c\boldsymbol{\mu})(\mathbf{y}_j - c\boldsymbol{\mu})'\right\}\right], \end{aligned} \tag{5.3}$$

where $N = N_1 + N_2$ and $\operatorname{etr}(\cdot)$ stands for $\{\exp \operatorname{trace}(\cdot)\}$. We assume that we are in a situation where, a priori, the noninformative prior of Section 2 is appropriate, so that

$$p(\boldsymbol{\mu}, c, \Sigma^{-1}) = p(\boldsymbol{\mu}, c)p(\Sigma^{-1}) \tag{5.4}$$

is such that

$$p(\boldsymbol{\mu}, c, \Sigma^{-1}) \propto |\Sigma^{-1}|^{-(k+1)/2}. \tag{5.4$'$}$$

We show below that combining (5.4′) with (5.3) and then integrating out $\boldsymbol{\mu}$, Σ^{-1} leads to the posterior of c, given the data X, Y, which is such that

$$\begin{aligned} p(c|X, Y) &= K\left(N_1 + N_2 c^2\right)^{(N-k-1)/2} \\ &\qquad \left\{(b_1 + 1/N_1)c^2 - b_2 c + (b_3 + 1/N_2)\right\}^{-(N-1)/2}, \qquad c > 0, \end{aligned} \tag{5.5}$$

where

(5.5′) $b_1 = \bar{\mathbf{x}}'S^{-1}\bar{\mathbf{x}}, \qquad b_2 = \bar{\mathbf{x}}'S^{-1}\bar{\mathbf{y}}, \qquad b_3 = \bar{\mathbf{y}}'S^{-1}\bar{\mathbf{y}}$ and $S = S_1 + S_2$,

with $\bar{\mathbf{x}}$, $\bar{\mathbf{y}}$, S_1 and S_2 as defined in (2.1a)–(2.1b). The last factor of (5.5) could be rewritten by completing the square in c, and we would then find the equivalent form

$$(5.6)\quad p(c|X,Y) = K\left(N_1 + N_2c^2\right)^{(N-k-1)/2}\left\{1 + a_0\left(c - \tilde{c}\right)^2\right\}^{-(N-1)/2}, \qquad c > 0,$$

where $a_0 = N_2(1 + N_1b_1)/t$, $\tilde{c} = N_1b_2/(1 + N_1b_1)$, $t = N_1\{1 + (N_1b_3 + N_1N_2b_1q)/(1 + N_1b_1)\}$, $q = (\bar{\mathbf{y}} - b\bar{\mathbf{x}})'S^{-1}(\bar{\mathbf{y}} - b\bar{\mathbf{x}})$ and $b = b_2/b_1$. In (5.5) or (5.6), K is a normalizing constant and may be determined numerically, and similarly, we note here that it is an easy matter to tabulate (5.6) and to determine the mode graphically. An example is given in Section 6.

To derive the result (5.5)–(5.5′), we first note that, after some algebra and application of Bayes's theorem, we may write the posterior of μ, c, Σ^{-1}, given X and Y, in the form

$$(5.7)\qquad p(\mu, c, \Sigma^{-1}|X,Y) \propto |\Sigma^{-1}|^{(N-k-1)/2}\exp\left\{-\tfrac{1}{2}\left[\mathrm{tr}^{-1}S + Q\right]\right\},$$

where

$$Q = N_1(\mu - \bar{\mathbf{x}})'\Sigma^{-1}(\mu - \bar{\mathbf{x}}) + N_2(c\mu - \bar{\mathbf{y}})'\Sigma^{-1}(c\mu - \bar{\mathbf{y}}),$$

with $\bar{\mathbf{x}}$, $\bar{\mathbf{y}}$ and S as defined in (2.1a)–(2.1b). A routine completion of the square in μ, with some tedious simplification leads to

$$(5.8)\qquad Q = (\mu - \tilde{\mu})'A(\mu - \tilde{\mu}) + \frac{N_1N_2}{N_1 + N_2c^2}(\bar{\mathbf{y}} - c\bar{\mathbf{x}})'\Sigma^{-1}(\bar{\mathbf{y}} - c\bar{\mathbf{x}}),$$

where

$$(5.8')\qquad A = \left(N_1 + c^2N_2\right)\Sigma^{-1}, \qquad \tilde{\mu} = \left(N_1 + N_2c^2\right)^{-1}\left[N_1\bar{\mathbf{x}} + N_2c\bar{\mathbf{y}}\right].$$

Inserting Q in (5.7) and integrating with respect to μ, we find, using properties of the k-variate normal, that

$$(5.9)\qquad \begin{aligned} p(c, \Sigma^{-1}|X, y) &\propto |A|^{-1/2}|\Sigma^{-1}|^{(N-k-1)/2} \\ &\quad\times\exp\left\{-\tfrac{1}{2}\mathrm{tr}\Sigma^{-1}\left[S + \frac{N_1N_2}{N_1 + N_2c^2}(\bar{\mathbf{y}} - c\bar{\mathbf{x}})(\bar{\mathbf{y}} - c\bar{\mathbf{x}})'\right]\right\}. \end{aligned}$$

Since Σ^{-1} is $(k \times k)$, $|A| = (N_1 + c^2N_2)^k|\Sigma^{-1}|$, and using this in (5.9), we have

$$(5.9')\qquad \begin{aligned} p(c, \Sigma^{-1}|X,Y) &\propto \left(N_1 + c^2N_2\right)^{-k/2}|\Sigma^{-1}|^{[(N-1)-k-1]/2} \\ &\quad\times\exp\left\{-\tfrac{1}{2}\mathrm{tr}\Sigma^{-1}\left[S + \frac{N_1N_2}{N_1 + N_2c^2}(\bar{\mathbf{y}} - (\bar{\mathbf{x}})(\bar{\mathbf{y}} - (\bar{\mathbf{x}})'))\right]\right\}. \end{aligned}$$

Using properties of the k-order Wishart distribution, we now integrate (5.9′)

with respect to Σ^{-1} and find (since S is given once X and Y observed)

$$p(c|X,Y) \propto \left(N_1 + N_2c^2\right)^{-k/2} \tag{5.10}$$
$$\times \left| I + \frac{N_1N_2}{N_1 + N_2c^2} S^{-1}(\bar{\mathbf{y}} - c\bar{\mathbf{x}})(\bar{\mathbf{y}} - c\bar{\mathbf{x}})'| ^{-(N-1)/2}. \right|$$

But $|I + AB| = |I + BA|$, so that, after simplification,

$$p(c|X,Y) \propto \left(N_1 + N_2c^2\right)^{-k/2} \tag{5.10$'$}$$
$$\times \left(1 + \frac{N_1N_2}{N_1 + N_2c^2}(\bar{\mathbf{y}} - c\bar{\mathbf{x}})'S^{-1}(\bar{\mathbf{y}} - c\bar{\mathbf{x}})\right)^{-(N-1)/2}$$

or

$$p(c|X,Y) \propto \left(N_1 + N_2c^2\right)^{(N-1-k)/2} u^{-(N-1)2} \tag{5.11}$$

TABLE 2

The posterior $p(c|X,Y) = p(c)$ for the electricity data. Taken from output that tables $p(c)$ for values of $c = 0.300(0.005)$ to 1.095 (available from the authors).

c	$p(c)$	c	$p(c)$
0.300	0.000029	0.705	3.411992
0.315	0.000082	0.720	2.806116
0.330	0.000228	0.735	2.239178
0.345	0.000605	0.750	1.737182
0.360	0.001539	0.765	1.312955
0.375	0.003739	0.780	0.968628
0.390	0.008665	0.795	0.698872
0.405	0.019118	0.810	0.494057
0.420	0.040105	0.825	0.342823
0.435	0.079900	0.840	0.233895
0.450	0.151035	0.855	0.157160
0.465	0.270717	0.870	0.104162
0.480	0.459943	0.885	0.068197
0.495	0.740631	0.900	0.044169
0.510	1.130533	0.915	0.028336
0.525	1.636566	0.930	0.018029
0.540	2.248232	0.945	0.011390
0.555	2.933566	0.960	0.007153
0.570	3.639796	0.975	0.004695
0.585	4.299798	0.990	0.002781
0.600	4.843352	1.005	0.001726
0.615	5.210431	1.020	0.001068
0.630	5.362763	1.035	0.000660
0.645	5.290479	1.050	0.000407
0.660	5.012257	1.065	0.000251
0.675	4.569532	1.080	0.000155
0.690	4.016958	1.095	0.000096

with

$$u = (N_1 + N_2c^2) + N_1N_2(\bar{\mathbf{y}}-c\bar{\mathbf{x}})'S^{-1}(\bar{\mathbf{y}}-c\bar{\mathbf{x}}). \tag{5.12}$$

Some straightforward algebra yields

$$u = N_1N_2\left\{\left(\frac{1}{N_1} + b_1\right)c^2 - 2b_2c + \left(\frac{1}{N_2} + b_3\right)\right\} \tag{5.12$'$}$$

and insertion of (5.12′) in (5.11) is the result (5.5), which of course after completion of the square in c and introducing the notation of (5.6), leads to the result (5.6). We turn now to an illustrative example.

6. Estimation of *c* for the electricity data. Suppose we assume that the data of Section 4 were generated under the conditions specified by (5.1)–(5.2). The posterior of the *magnitudinal effect c* is then given by (5.5), where for this set of data

$$b_1 = 5.0397111 \times 10^{-2}, \qquad b_2 = 3.2126838 \times 10^{-2}, \qquad b_3 = 2.0480364 \times 10^{-2}, \tag{6.1}$$

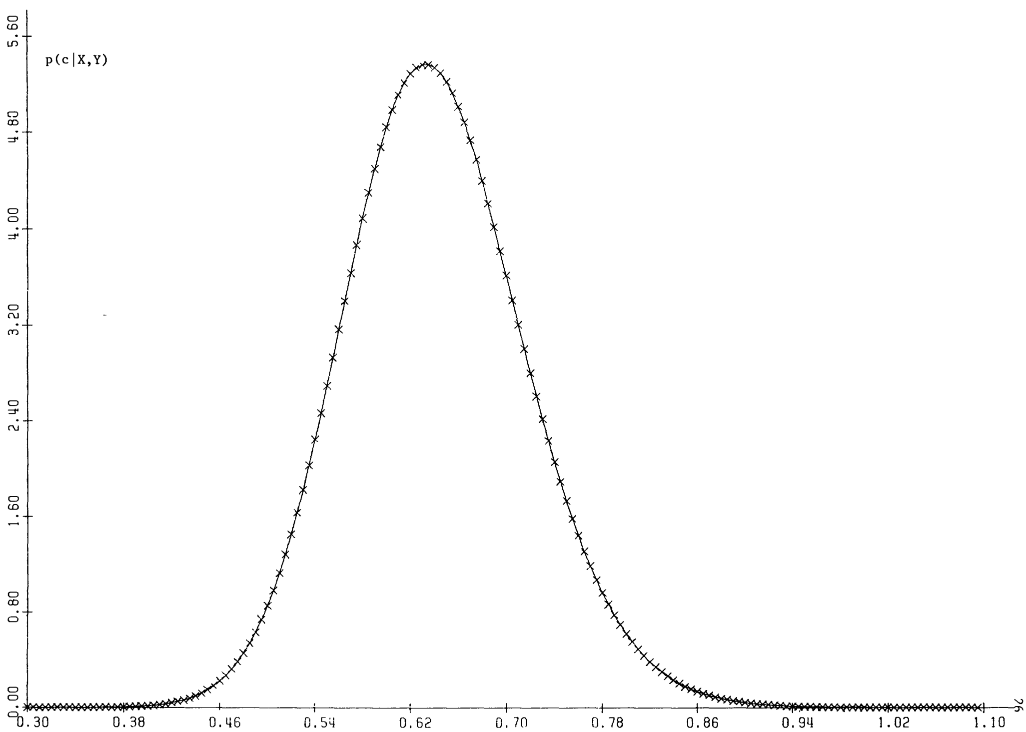

FIG. 2. *The posterior $p(c|X, Y)$ for the electricity data. The mode is at $c = c_m = 0.635$.* [*See Table 3 for the* $1 - \alpha$ *intervals.*]

TABLE 3
Posterior intervals $[c_1, c_2]$ *for c; see* (6.5).

$1-\alpha$	0.90	0.95	0.99
c_1	0.5217	0.5006	0.4600
c_2	0.7709	0.8006	0.8600

and S is as quoted in (4.2). We also have

$$N_1 = 45, \qquad N_2 = 55, \qquad N = 100. \tag{6.2}$$

Using the above data, we tabulate $p(c|X, Y)$ in Table 2, and graph $p(c|X, Y)$ for this set of data in Figure 2. Inspection of Figure 2 shows a slight skewness to the right. We note that the mode, say c_m, of this posterior is at

$$c_m = 0.635000. \tag{6.3}$$

Numerical integration yields

$$E(c|X, Y) = 0.641198, \qquad \operatorname{Var}(c|X, Y) = 0.005776, \tag{6.4}$$

so that S.D. $(c|X, Y) = 0.076$.

The question of $1 - \alpha$ (posterior) confidence intervals is quickly determined by numerical integration. For this purpose we let c_1 and c_2 be such that

$$\int_0^{c_1} p(c|X, Y)\, dc = \alpha/2 = \int_{c_2}^{\infty} p(c|X, Y)\, dc. \tag{6.5}$$

The intervals $[c_1, c_2]$ are then of content $1 - \alpha$. For this set of data, and for $\alpha = 0.90, 0.95, 0.99$, the resulting intervals, as found using numerical integration, are as given in Table 3. [To avoid the infinite limit, we actually found c_2 using $\int_0^{c_2} p(c|X, Y)\, dc = 1 - \alpha/2$].

APPENDIX

PROOF OF THEOREM 2.1. Using (2.4), we obtain as a special case of (68) in James (1964), that the conditional posterior density of (ω_1, ω_2), given Σ^{-1}, is

$$\begin{aligned} p(\omega_1, \omega_2|X, Y, \Sigma^{-1}) = {} & K_1 e^{-(\omega_1+\omega_2)/2}(\omega_1\omega_2)^{(k-3)/2}(\omega_1 - \omega_2) \\ & \times \operatorname{etr}\{-\tfrac{1}{2}T'\Sigma^{-1}T\}\,{}_0F_1^{(2)}(\tfrac{1}{2}k; \tfrac{1}{2}T'\Sigma^{-1}T, \Lambda), \\ & \omega_1 > \omega_2 \geq 0, \end{aligned} \tag{A.1}$$

where $K_1 = \pi^{1/2}\{2^k\Gamma(\frac{1}{2}k)\Gamma(\frac{1}{2}(k-1))\}^{-1}$, and T is defined in the statement of the theorem. Thus, to find $p(\omega_1, \omega_2|X, Y)$, we need to compute

$$E\{\operatorname{etr}(-\tfrac{1}{2}W)\,{}_0F_1^{(2)}(\tfrac{1}{2}k; \tfrac{1}{2}W, \Lambda), \tag{A.2}$$

where the expectation is with respect to the distribution of Σ^{-1} given in (2.3), and $W = T'\Sigma^{-1}T$. Now using well known properties of the Wishart distribution, $W = T'\Sigma^{-1}T$ has a central Wishart distribution of order 2 and n degrees of freedom, with positive definite matrix parameter $L = T'S^{-1}T$. By making the transformation $Z = Q^{-1}WQ^{-1}$, where Q is a 2×2 symmetric matrix with $Q^2 = 2L(1 + L)^{-1}$, (A.2) can be expressed as

$$(\text{A.3}) \qquad K'|I + L|^{-n/2} \int_{Z>0} \operatorname{etr}(-Z)|Z|^{(n-3)/2} {}_0F_1^{(2)}\{\tfrac{1}{2}k; \tfrac{1}{2}QZQ, \Lambda\}\, dZ,$$

where $K' = \{\pi^{1/2}\Gamma(\frac{1}{2}n)\Gamma[\frac{1}{2}(n-1)]\}^{-1}$. As a special case of equation (31) in James (1964), (A.3) is ${}_1F_1^{(2)}\{\frac{1}{2}n; \frac{1}{2}k; \frac{1}{2}L(I+L)^{-1}, \Lambda\}$. The theorem then follows by replacing the matrix arguments with the diagonal matrices consisting of their eigenvalues. □

Acknowledgments. This work was started at Old Dominion University, continued at the University of Toronto, and completed at Stanford University. The authors are grateful for the comments of R. Dahiya and M. Srivastava, and gratefully acknowledge the computer assistance of Gary Wong of Old Dominion University.

REFERENCES

Box, G. E. P. and Tiao, G. C. (1973). *Bayesian Analysis in Statistical Inference*. Addison-Wesley, Reading, Mass.

Glynn, W. J. (1980). Asymptotic representations of the densities of canonical correlations and latent roots in MANOVA when the population parameters have arbitrary multiplicity. *Ann. Statist.* **8** 958–976.

Guttman, I., Menzefricke, U. and Tyler, D. (1985). Magnitudinal effects in the normal multivariate model. Technical Report No. 10, Department of Statistics, Univ. of Toronto.

James, A. T. (1964). Distribution of matrix variates and latent roots derived from normal samples. *Ann. Math. Statist.* **35** 475–501.

Johnson, R. A. and Wichern, D. W. (1982). *Applied Multivariate Statistical Analysis*. Prentice-Hall, Englewood Cliffs, N.J.

Kraft, C. H., Olkin, I. and van Eeden, C. V. (1972). Estimation and testing for differences in magnitude or displacement in the mean vectors of two multivariate normal populations. *Ann. Math. Statist.* **43** 455–467.

Muirhead, R. J. (1975). Expressions for some hypergeometric functions of matrix arguments with applications. *J. Multivariate Anal.* **5** 283–293.

Muirhead, R. J. (1978). Latent roots and matrix variates: a review of some asymptotic results. *Ann. Statist.* **6** 5–33.

Muirhead, R. J. (1982). *Aspects of Multivariate Statistical Theory*. Wiley, New York.

Tiao, G. C. and Fienberg, S. (1969). Bayesian estimation of latent roots and vectors with special reference to the bivariate normal distribution. *Biometrika* **56** 97–108.

Tiao, G. C. and Zellner, A. (1964). On the Bayesian estimation of multivariate regression. *J. Roy. Statist. Soc. Ser. B* **26** 277–285.

I. Guttman
Department of Statistics
University of Toronto
Toronto, Ontario M5S 1A1
Canada

U. Menzefricke
Faculty of Management Studies
University of Toronto
Toronto, Ontario M5S 1A1
Canada

D. Tyler
Department of Statistics
Rutgers University
New Brunswick, New Jersey 08903