MAGNITUINAL EFFECTS IN THE NORMAL MULTIVARIATE MODEL

BY

IRWIN GUTTMAN, and DAVID TYLER

TECHNICAL REPORT NO. 190
JUNE 1983

PREPARED UNDER THE AUSPICES OF
NATIONAL SCIENCE FOUNDATION GRANT
MCS 81-04262
Ingram Olkin, Project Director

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA
MAGNITUINAL EFFECTS IN THE NORMAL MULTIVARIATE MODEL

by

Irwin Guttman and David Tyler

TECHNICAL REPORT NO. 190
JUNE 1983

PREPARED UNDER THE AUSPICES
OF
NATIONAL SCIENCE FOUNDATION GRANT
MCS 81-04262
Ingram Olkin, Project Director

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA
0. ABSTRACT

Suppose the \((k \times 1)\) vectors \(x\) and \(y\) are independent with \(x \sim N(\mu, \Sigma)\) and \(y \sim N(\eta, \Sigma)\), \(\Sigma\) positive definite. If for a positive scalar \(c\), \(\eta = c\mu\), we find the posterior of \(c\), using noninformative priors, given the data \(\{x_i\}_1^N, \{y_j\}_1^N\). The \(x_i\) are \(N_1\) independent observations on \(x\), and independent of the \(y_j\), which are \(N_2\) independent observations on \(y\). 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 \(\eta = c\mu\), and proceed as follows. We first note that the matrix \(\Lambda = (\lambda_{ij})\), where \(\lambda_{11} = N_1\mu'\Sigma^{-1}\mu\), \(\lambda_{12} = \lambda_{21} = \sqrt{N_1N_2}\mu'\nu_1^{-1}\eta\), \(\lambda_{22} = N_2\nu_2^{-1}\eta\) has a zero eigen root if and only if \(\eta = c\mu\). Hence, we are motivated to find the posterior (joint) distribution of \(\omega_1, \omega_2\), the roots of \(\Lambda\), where \(\omega_1 > \omega_2\). Then, by integration with respect to \(\omega_1\) over the region \(\omega_1 > \omega_2\), we may find the marginal of \(\omega_2\), and use it to examine the statement \(\omega_2 = 0 \Leftrightarrow \eta = c\mu\).

The posterior of \((\omega_1, \omega_2)\) involves the multivariate hypergeometric function \(\text{E}_1^{(2)}\), which in practice creates computational difficulties. Accordingly, some numerical considerations are discussed for the computing of the posterior of \(\omega_2\), and an example using real data is given.
1. INTRODUCTION AND SUMMARY

Kraft, Olkin and Van Eeden (1972) discuss the following interesting problem. (We will refer to this paper as the KOVE paper.) Suppose two methods are used and their effects can be evaluated on \( k \) characteristics, such that Method 1 produces measurement \( \tilde{x} \) and Method 2 produces measurement \( \tilde{y} \), where the \( k \)-dimensional random vectors \( \tilde{X} \) and \( \tilde{Y} \) have a Normal(\( \mu,\Sigma \)) and Normal(\( \eta,\Psi \)) distribution respectively. Applications abound in Chemical Engineering and in Medicine, notably drug evaluation studies, in which the question of interest lies in examining whether the model

\[
\tilde{y} = c\mu, \quad c > 0
\]

holds, and if so, what to do about estimating the scalar \( c \).

The scalar \( c \) has been called the magnitudinal effect in the KOVE paper, who approach this problem from the classical sampling route. To test whether (1.1) is true, they 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. Furthermore, given that the model (1.1) is true, they find the maximum likelihood estimate \( \hat{c} \) of the scalar \( c \) and supply a confidence interval for large samples by finding the asymptotic distribution for \( \hat{c} \) using well known limit theorems.

Our approach is different. Given that (1.1) is true and using a non-informative prior, we find the posterior distribution of \( c \), valid for any sample size, and use it to determine confidence intervals and point estimates. This is done in Section 2. If we first wish to examine whether
(1.1) is true or not, we have a Bayesian approach to the problem which involves examining the smallest root of a certain matrix. This is further explained and developed in Section 3. Some numerical considerations are discussed in Section 4, and in Section 5 we illustrate how to use the results of this paper on a typical data set by using data found in Timm (1975).

2. THE POSTERIOR DISTRIBUTION OF THE MAGNITUDEAL EFFECT

In this section, we assume that it is known that the magnitudeal effect model holds, and that interest is in making inference about the magnitudeal parameter. Specifically, we assume that the k-dimensional random vectors $\tilde{x}$ and $\tilde{y}$ are independent and distributed as

$$
\tilde{x} \sim \text{Normal}(\mu, \tau) \quad \text{and} \quad \tilde{y} \sim \text{Normal}(\eta, \tau)
$$

(2.1)

where

$$
\eta = c\mu, \quad \mu \neq 0, \quad c > 0.
$$

(2.2)

As mentioned in Section 1, we may be dealing with two drugs and $c$ represents a change needed in the dosage of that drug where effects on k symptoms are measured by $\tilde{x}$ to make it equivalent to that dosage of the drug whose effect on the k symptoms are measured by $\tilde{y}$. In such a setting, it is not uncommon to have $\text{Var}(\tilde{x}) = \text{Var}(\tilde{y}) = \frac{\tau}{c}$ and we are operating under this assumption.

Suppose that $N_1$ independent observations on $\tilde{x}$ are generated, say $X = \{\tilde{x}_1, \ldots, \tilde{x}_{N_1}\}$ and that $X$ is independent of $N_2$ independent observations $\tilde{y}_j$ taken on $\tilde{y}$, say $Y = \{\tilde{y}_1, \ldots, \tilde{y}_{N_2}\}$. The likelihood based on these observations of the parameters $\mu, c, \frac{\tau}{c}$ is then
\[ \ell(\mu, c, \psi^{-1} | X, Y) \]  
\[ |\psi^{-1}|^{N/2} \exp\left\{-\frac{1}{2} |\psi^{-1}| \left\{ \sum_{i=1}^{N_1} (x_i - \mu)^T (x_i - \mu) + \sum_{j=1}^{N_2} (y_j - \psi \zeta_j)^T (y_j - \psi \zeta_j) \right\} \right\} \]

where \( N = N_1 + N_2 \) and \( \exp\{\text{trace}(\cdot)\} \) stands for \( \exp\{\text{trace}(\cdot)\} \). We assume that we are in a situation where, a priori, the noninformative prior is appropriate, so that we use as prior

\[ p(\mu, c, \psi^{-1}) \propto |\psi^{-1}|^{-\frac{1}{2}(k+1)} . \]

In Appendix I of this paper, we show that combining (2.4) with (2.3), and then integrating out \( \mu, \psi^{-1} \), leads to the posterior of \( c \), given the data \( X, Y \), which is such that

\[ p(c|X,Y) = K(N_1 + N_2) \cdot c^{\frac{k}{2}(N-k-1)} \cdot (b_1 + 1/N_1) c^{2} - 2b_2 c \]

\[ + (b_3 + 1/N_2) \cdot c^{-\frac{k}{2}(N-1)} \]

where \( b_1 = \bar{\xi} S^{-1} \bar{\xi} \), \( b_2 = \bar{\xi} S^{-1} \bar{\gamma} \), \( b_3 = \bar{\gamma} S^{-1} \bar{\gamma} \) and \( S = S_1 + S_2 \) with

\[ S_1 = \sum_{i=1}^{N_1} (x_i - \bar{x}) (x_i - \bar{x})', S_2 = \sum_{j=1}^{N_2} (y_j - \bar{y}) (y_j - \bar{y})', \bar{\xi} = N_1^{-1} \sum_{i=1}^{N_1} x_i \]

and

\[ \bar{\gamma} = N_2^{-1} \sum_{j=1}^{N_2} y_j. \]

The last factor of (2.5) could be rewritten by completing the square in \( c \) and we would then find the equivalent form

\[ p(c|X,Y) = K(N_1 + N_2) \cdot c^{\frac{k}{2}(N-k-1)} \cdot (1 + a_0 (c - \bar{c})^2)^{-\frac{k}{2}(N-1)} \]

where \( a_0 = N_2 (1 + N_1 b_1) / T, \bar{c} = N_1 b_2 / (1 + N_1 b_1), T = N_1 \left\{ 1 + (N_2 b_3 + N_1 N_2 b_1 q) / (1 + N_1 b_1) \right\}, q = (\bar{y} - b_2 \bar{x})' S^{-1} (\bar{y} - b_2 \bar{x}) \) and \( b = b_2 / b_1 \). In (2.5) or (2.6), \( K \) is a normalizing constant and may be determined numerically, and similarly, we note here that it is an easy matter to tabulate (2.6) and to determine the mode graphically. An example is given in Section 4.
3. WIDENING THE MODEL TO EXAMINE THE MAGNITUDEAL EFFECT MODEL

The aim of this section is to postulate a model which is more general than that of (2.1) - (2.2), but contains (2.1) - (2.2) as a special case. An equivalent objective is to ask how to go about, on the basis of data described in Section 2, the question of whether (2.1) - (2.2) holds.

The authors first tried the following formulation. Suppose (2.1) holds, and that we replace (2.2) with

$$\eta = \zeta \mu, \quad \mu \neq 0$$  \hspace{1cm} (3.1)

where $C = \text{diag}(c_1, \ldots, c_k)$ with $c_j > 0$. In principle we could find the posterior of $c_j$, and from that the posterior of $d_j$, where $d_j$ are a set of $k-1$ contrasts given by

$$d_1 = (c_1 - c_2)/\sqrt{2}$$  \hspace{1cm} (3.2)

$$d_2 = (c_1 + c_2 - 2c_3)/\sqrt{(3)(2)}$$

$$\vdots$$

$$d_{k-1} = (c_1 + \ldots + c_{k-1} - (k-1)c_k)/\sqrt{k(k-1)}$$

After obtaining the joint posterior of $\hat{d} = (d_1, \ldots, d_{k-1})'$, a posterior region for $\hat{d}$ could then be found, say at level $1-\alpha$, and if it contains $0$, we would have a basis for saying that the data supports the assumption that $c_1 = \ldots = c_k = c$, say at level $1-\alpha$.

However, the posterior of $(c_1, \ldots, c_k)$ is extremely hard to work with, that is not tractable in $(c_1, \ldots, c_k)$. The problem lies in the fact that integration of $p(\hat{\eta}^{-1}, C|X,Y)$ with respect to $\hat{\eta}^{-1}$ cannot be accomplished in closed form. Indeed, it turns out (Appendix II gives the details)
\[
p(\hat{\beta}^{-1}, C | X, Y) \propto |\hat{\beta}^{-1}|^{1/2(N_2-k-1)} \left| (N_1 \hat{\beta}^{-1} + N_2 C \hat{\beta}^{-1} C)^{-1} \right|^{-1/2} \\
\times \text{etr}(-\frac{1}{2} \{\hat{\beta}^{-1}S + (\hat{y} - CX)' A(\hat{y} - CX)\})
\]

where \( A = N_2 \hat{\beta}^{-1} - N_2 \hat{\beta}^{-1} C (N_1 \hat{\beta}^{-1} + N_2 C \hat{\beta}^{-1} C)^{-1} C \hat{\beta}^{-1} \). The intractability of the above is now apparent.

Because of the above, we proceed as follows. Consider the 2x2 symmetric positive definite matrix

\[
\Omega = M' \hat{\beta}^{-1} M = \begin{pmatrix}
N_1 \mu'^{-1} \mu & (N_1 N_2)^{1/2} \mu'^{-1} \eta \\
(N_1 N_2)^{1/2} \mu'^{-1} \eta & N_2 \eta'^{-1} \eta
\end{pmatrix}
\]

where \( M = [N_1^{3/2} \mu \ N_2^{3/2} \eta] \). Under the model (3.1) the rank of \( M \) is one, if and only if \( \eta = c \mu \) for \( c > 0 \), and thus the smallest eigenvalue of \( \Omega \), say \( \omega_2 \), is zero if and only if the magnitudinal model holds. In view of this, our aim is to find the posterior distribution of \( \omega_2 \), given the data. If the distribution is concentrated about zero, the data would be compatible with the magnitudinal model.

We remind the reader that if (2.1) holds, and if the noninformative prior for \( \mu, \eta \) and \( \hat{\beta}^{-1} \) is used, that is

\[
p(\mu, \eta, \hat{\beta}^{-1}) \propto |\hat{\beta}^{-1}|^{-1/2(k+1)}
\]

then the posterior distribution for \( \mu, \eta \) and \( \hat{\beta}^{-1} \) is characterized by

(i) the posterior, marginally, of \( \hat{\beta}^{-1} \) is such that

\[
p(\hat{\beta}^{-1} | X, Y) \propto |\hat{\beta}^{-1}|^{1/2(n-k-1)} \text{etr}(-\frac{1}{2} \hat{\beta}^{-1} S)
\]
where $S$ is given after (2.5), with $n = N_1 + N_2 - 2 \geq k$, and (ii) the posterior for $\mu, \eta$ given $\Sigma^{-1}$ is such that $\mu$ and $\eta$ are independent with

$$\begin{align*}
N_{1\mu}^{\frac{1}{2}} & \sim \text{Normal}(N_{1\mu}^{\frac{1}{2}}x, \Sigma) \quad \text{and} \quad N_{2\eta}^{\frac{1}{2}} & \sim \text{Normal}(N_{2\eta}^{\frac{1}{2}}y, \Sigma) 
\end{align*}$$

(3.7)

[see Tiao and Zellner (1964)]. We note that the density (3.6) is that of a central Wishart distribution of order $k$, $n$ degrees of freedom, with positive definite parameter matrix $S^{-1}$.

The posterior density for the roots of $\Omega$ may now be derived. This density and its derivation requires the introduction of the generalized hypergeometric functions $\text{ _0F_1 (2) } (c; H, U)$ and $\text{ _1F_1 (2) } (a; c; H, U)$, where the arguments $H$ and $U$ are $2 \times 2$ symmetric matrices. The reader is referred to either James (1964) or Muirhead (1978) for an excellent review of the hypergeometric functions with matrix arguments. In particular, these functions depend on their matrix arguments only through the roots of the arguments, and are symmetric functions of these roots.

**THEOREM 3.1.** For $2 \leq k \leq n$, the posterior density of $(\omega_1, \omega_2)$, where $\omega_1 \geq \omega_2$ are the roots of $\Omega$ defined in (3.4), is

$$
p(\omega_1, \omega_2 | X, Y) = K_0(\omega_1, \omega_2)^{\frac{1}{2}(k-3)} e^{-\frac{1}{2}(\omega_1 + \omega_2)} (\omega_1 - \omega_2)$$

$$\times \text{ _1F_1 (2) } \left\{ \frac{\omega_1}{2}; \frac{\omega_2}{2}; \frac{1}{2}L_0(1 + L_0)^{-1}, \Omega_0 \right\}, \quad \omega_1 > \omega_2 \geq 0$$

(3.8)

where $K_0 = \frac{\pi^\frac{1}{2}}{2^\frac{k}{2}} (\frac{k}{2})^\frac{(k-1)}{2} \Gamma[\frac{k}{2}(k-1)]^{-1} (1 + \frac{1}{2} \omega_1) (1 + \frac{1}{2} \omega_2) \Gamma[\frac{1}{2}(k-1)]^{-2\mu}$ with $\frac{1}{2} \mu_1 > \frac{1}{2} \mu_2 > 0$

being the roots of $L = T'S^{-1}T$ and $T = [N_{1\mu}^{\frac{1}{2}} \ N_{2\eta}^{\frac{1}{2}} ]$, $L_0 = \text{ diag}(\mu_1, \mu_2)$ and $\Omega_0 = \text{ diag}(\omega_1, \omega_2)$. 

6
PROOF. From (3.7), we obtain as a special case of (68) in James (1964) the conditional posterior density of \((\omega_1, \omega_2)\) given \(X, Y, \frac{q}{n}\), which is

\[
p(\omega_1, \omega_2 | X, Y, \frac{q}{n}) = K_1 \exp \left(-\frac{k}{2}(\omega_1^2 + \omega_2^2) \right) \omega_1^{\frac{k}{2}(k-3)} (\omega_1 - \omega_2) \prod_{j=1}^{k} F_1^{(2)} \left(\frac{1}{2}; \frac{1}{2}; \frac{q}{n+1}, \Omega, \omega_1 \right) \prod_{j=1}^{k} F_1^{(2)} \left(\frac{1}{2}; \frac{1}{2}; \frac{q}{n+1}, \Omega, \omega_2 \right)
\]

(3.9)

where \(K_1 = \pi^{\frac{k}{4}} (2^k \Gamma(\frac{k}{2}) \Gamma(\frac{k}{2}(k-1)))^{-1}\), and \(T\) is defined in the statement of the theorem. Thus, to find \(p(\omega_1, \omega_2 | X, Y)\), we need to compute

\[
E\left\{\exp\left(-\frac{q}{n} W\right) \prod_{j=1}^{k} F_1^{(2)} \left(\frac{1}{2}; \frac{1}{2}; \frac{q}{n+1}, \Omega\right)\right\}
\]

(3.10)

where, from (3.6) and using well known properties of the Wishart distribution, \(W = T' T^{-1} T\) has a central Wishart distribution of order 2, \(n\) degrees of freedom, with positive definite matrix parameter \(L = T'ST\). By making the transformation \(Z = Q^{-1} W Q^{-1}\), where \(Q\) is a \(2 \times 2\) symmetric matrix with \(Q^2 = \frac{1}{2} L(I+L)^{-1}\), (3.10) can be expressed as

\[
K' |I+L|^{-\frac{3}{2}n} \int_{Z>0} \exp(-Z) |Z|^{-\frac{3}{2}(n-3)} \prod_{j=1}^{k} F_1^{(2)} \left(\frac{1}{2}; \frac{1}{2}; \frac{1}{2} QZQ, \Omega\right) dZ
\]

(3.11)

where \(K' = \left\{ \pi^{\frac{k}{4}} (2^k \Gamma(\frac{k}{2}) \Gamma(\frac{k}{2}(n-1)))^{-1}\right\}\). As a special case of equation (31) in James (1964), (3.11) is \(1 \prod_{j=1}^{k} F_1^{(2)} \left(\frac{1}{2}; \frac{1}{2}; \frac{1}{2} L(I+L)^{-1}, W\right)\). The theorem then follows by replacing the matrix arguments with the diagonal matrices consisting of their eigenvalues. □

There is a connection between the \(1 \prod_{j=1}^{k} F_1^{(2)}\) function with the classical univariate hypergeometric function

\[
\gamma_1(a, b; c; z) = \sum_{q=0}^{\infty} \frac{(a)_{q} (b)_{q} z^q}{(c)_{q} q!}
\]

(3.12)
where \((c)_t = c(c+1) \ldots (c+t-1)\). The relationship as given by Muirhead (1975), Lemma 1.2, is:

\[
_1F_1(2)\{a; c; H, U\} = \sum_{t=0}^{\infty} \frac{(a)_t (c-a)_t}{(c-k)_t (c)_t^{2t}} \frac{(-h_1 h_2 u_1 u_2)^t}{t!} 
\]

\[
x \sum_{j=0}^{\infty} \frac{(a+t)_j}{(c+2t)_j} \frac{[\frac{1}{2}(h_1 + h_2) (u_1 + u_2)]^j}{j!} 2F_1(-\frac{j}{2}, -\frac{j+1}{2}; 1; x^2)
\]

where \(H = \text{diag}(h_1, h_2)\), with \(h_1 > h_2\), and \(U = \text{diag}(u_1, u_2), x = \{(h_1 - h_2)(u_1 - u_2)\}/\{(r_1 + r_2)(u_1 + u_2)\}\). The \(_2F_1\) terms arising from (3.13) have a finite expansion since for \(q > \frac{1}{2}\), either \((-\frac{1}{2} + \frac{1}{2})_q = 0\) or \((-\frac{1}{2} + \frac{1}{2})_q = 0\).

By inserting the expansion (3.13) into the density (3.8), making the change of indices \(p = 2t+j, t = t\), and integrating over \(\omega_1\), we have

**THEOREM 3.2.** The marginal posterior density of \(\omega_2\), the smallest eigenvalue of the matrix \(\Omega\) of (3.4), with (3.5) holding, may be expressed as

\[
p(\omega_2 | X, Y) = K_0 \omega_2^{\frac{1}{2}(k-3)} e^{-\frac{1}{2} \omega_2} \sum_{p=0}^{\infty} c_p(\omega_2) \frac{\omega^p}{(\frac{k}{2})_p} \quad \omega_2 \geq 0
\]

where \(K_0\) is defined in Theorem 3.1, \(\bar{r} = \frac{1}{2}(r_1 + r_2)\) with \(r_i = \frac{k}{2} \omega_i / (1 + \omega_i)\)

i=1, 2, and where

\[
c_p(\omega_2) = \sum_{t=0}^{\left[\frac{k}{2}\right]} \frac{\omega_2^t}{t!} \sum_{q=0}^{\left[\frac{k}{2}\right]-t} b_p, t, q \times x_q F(\omega_2; p-2t-2q, 2q+1, t+\frac{3}{2}) (3.15)
\]

with \(z_r = r_1 r_2 / (2r)^2, x_r = (r_1 - r_2)^2 / (4r)^2, j_t = (-1)^t \frac{k-n}{2} \sqrt{(\frac{k}{2} - t)} t!\), \(b_p, t, q = \frac{(\omega_2)^p}{p!} \times (q!)^2 (p-2t-2q)\), and in general

\[
F(\omega_2; m_1, m_2, m_3) = \int_{\omega_2}^{\infty} (\omega_1 + \omega_2) \omega_1^{-m_1} (\omega_1 - \omega_2)^{m_2} \omega_1^{-\frac{k}{2}} e^{-\frac{1}{2} \omega_2} d\omega_1.
\]

*We note the omission in (1.6), of page 285 of Muirhead (1975) of a 1/k! (Muirhead's notation) in the summand.*
For computational purposes, the function $F$ given by (3.16) can be calculated recursively for fixed $m_3, \omega_2$ since in cases of interest $m_1$ and $m_2$ are nonnegative integers. By direct integration, we note that an alternative form for $F$ is given by

$$F(\omega_2; m_1, m_2, m_3) = \sum_{\nu=0}^{m_1+m_2} C_{\nu} 2^{\nu+m_3+1} \omega_2^{\nu} \Gamma(\nu+m_3+1) \frac{\chi^2_{2\nu+2m_3+2}}{\.1 \text{ after some rearrangements}}$$

(3.17)

where

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

(3.18)

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

$$F(\omega_2; m_1, m_2, m_3) = e^{-\frac{1}{2} \omega_2} \sum_{\nu=0}^{m_1+m_3} B_{\nu} 2^{\nu} \Gamma(m_2+\nu+1) \omega_2^{m_1+m_3-\nu}$$

(3.19)

where

$$B_{\nu} = \sum_{j=max(0,\nu-m_3)}^{\min(m_1,\nu)} \binom{m_1}{j} \binom{m_3}{\nu-j} (-1)^{m_1-j}$$

(3.20)

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 $\omega_2$ can be found directly from equations (3.14), (3.15) and (3.16). Using the notation of Theorem 3.2, we have

$$E(\omega_2^m | X, Y) = K_0 \sum_{p=0}^{\infty} \mu_{m,p} \Gamma(p+k+m) \frac{\tau^p}{(\tau k)^p}$$

(3.21)

where

$$\mu_{m,p} = \sum_{t=0}^{[\frac{3p}{2}]} 2^{p-2t+1} \sum_{t=0}^{[\frac{3p}{2}]-t} b_{p,t,q} x_1^2 \sum_{q=0}^{[\frac{3q}{2}]-t} a_{m,t,q}$$

(3.22)
\[ a_{m,t,q} = \sum_{v=0}^{m} (-1)^v \binom{m}{v} \beta\left(\frac{1}{2}(v+2q+2), \frac{1}{2}(2t+k-1)\right), \quad (3.23) \]

with \( \beta(\cdot, \cdot) \) denoting the Beta function. The application of the posterior density of \( \omega_2 \) and the first two moments of \( \omega_2 \) is illustrated in Section 5.

4. NUMERICAL CONSIDERATIONS

In this section, some difficulties which may arise in the computation of the posterior density of \( \omega_2 \) are discussed. We first note that the expansion for the hypergeometric function \( _1F_1^{(2)} \) given by (3.13) can converge slowly if either \( a \) is large, the argument \( h_1, h_2, u_1 \) or \( u_2 \) is large, or if the absolute difference \( |h_1-h_2| \) or \( |u_1-u_2| \) is large. This convergence problem is common to all hypergeometric functions of matrix arguments. For further discussion, we again refer the reader to James (1964) or Muirhead (1978). We now discuss in more detail how this convergence problem specifically relates to the expansion for the posterior density of \( \omega_2 \) given in Theorem 3.2.

By inserting the expansion (3.13) into the density (3.8) and making the change of indices \( p = 2t+j \) and \( t = t \), we have

\[ p(\omega_1,\omega_2|X,Y) = 2^{-(k-3)} K_0 \sum_{p=0}^{\infty} H_p(V) G_p(\theta) \frac{\Gamma_p}{(\frac{1}{2}k)_p}, \quad \omega_1 > \omega_2 > 0 \quad (4.1) \]

where \( V = \omega_1 + \omega_2, \theta = (\omega_1-\omega_2)/(\omega_1+\omega_2), H_p(V) = \nu^{p+k-2} e^{-\frac{1}{2}V}, \) and

\[ G_p(\theta) = \sum_{t=0}^{[\nu p]} \sum_{q=0}^{[\nu p]-t} b_{p,t,q} x^q (1-\theta^2)^{t+\frac{1}{2}(k-3)} \theta^{2q+1}. \quad (4.2) \]
The constants $K_0$, $z_r$, $x_r$, $J_t$ and $b_{p,t,q}$ are as in Theorem 3.1. We note integrating (4.1) term by term over $\omega_1$ yields the expansion for the posterior density of $\omega_2$ given by (3.14). Although the constant $J_t$ can be negative, by using the zonal polynomial expansion for the hypergeometric function $\frac{1}{p} F_1^{(2)}$ and comparing it to (4.1), it can be shown that $G_p(\theta)$ is a nonnegative function. The reader is referred to James (1964) and Muirhead (1978) 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 $t \leq \frac{1}{2}(n-k)$ and $J_t = 0$ for $t > \frac{1}{2}(n-k)$.

Since the Jacobian of the transformation $(\omega_1, \omega_2) \rightarrow (V, \theta)$ is $\frac{1}{2}V$, (see (4.1)), the joint posterior density of $(V, \theta)$ is thus

$$p(V, \theta | X, Y) = \sum_{p=0}^{\infty} a_p \chi^2_{2p+2k}(V) \cdot g_p(\theta) \quad 0 < \theta < 1, \quad V > 0 \quad (4.3)$$

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

$$g_p(\theta) = G_p(\theta)/\int_0^1 G_p(\alpha) \, d\alpha \quad (4.4)$$

and the weights $a_p$ are given by

$$a_p = 2^{p+2} c_0 \cdot \Gamma(p+k) \int_0^1 G_p(\alpha) \, d\alpha / (\pi k)_p \quad (4.5)$$

The posterior distribution of $(\omega_1, \omega_2)$ 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 (4.3) and hence in calculating (3.14). 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 (4.3) over $\theta$. This gives

$$p(V|X,Y) = \sum_{p=0}^{\infty} a_p \gamma_{2p+2k}^2(V), \quad V > 0. \quad (4.5)$$

A more direct method for obtaining the posterior density of $V$ is as follows. Since $V = \text{trace}(\Omega)$, where $\Omega$ is defined by (3.4), we note from (3.7) that the posterior distribution of $V$ given $\sigma^{-1}$ is a noncentral chi-square on $2k$ degrees of freedom with noncentrality parameter

$$\delta = N_1 \bar{x}^{\sigma^{-1}}x + N_2 \bar{x}^{\sigma^{-1}}x.$$

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 $\sigma^{-1}$ we obtain

$$a_p = ((1+\lambda_1)(1+\lambda_2))^{-\frac{p}{2}} \frac{1}{p!} \left( r_1^2 + r_2^2 \right)^p / \Gamma(p+1). \quad (4.6)$$

with $r_j = (\ell_j/2)/(1+\ell_j)$ for $j = 1, 2$, and where $X_{1,n}^2$ and $X_{2,n}^2$ have independent chi-square distributions, both on $n$ degrees of freedom.

Evaluation of the expectation in (4.6) gives

$$a_p = ((1+\lambda_1)(1+\lambda_2))^{-\frac{p}{2}} \frac{1}{p!} \left\{ \sum_{\nu=0}^{p} \binom{p}{\nu} \binom{n}{\nu} \frac{r_1^\nu}{\nu!} \frac{r_2^{p-\nu}}{(p-\nu)!} \right\}. \quad (4.7)$$

Furthermore, we can calculate the mean and variance of the weights by using (4.6). This gives

$$\text{E}(p) = \frac{1}{2} \lambda (\lambda_1 + \lambda_2) \quad \text{and} \quad \text{var}(p) = \frac{1}{2} \lambda (\lambda_1^2 + \lambda_2^2). \quad (4.8)$$
If the mean and variance of $p$ is large, direct computation of the posterior density of $\omega_2$ by (3.14) is not feasible. In such cases, computations can be saved by first computing the posterior joint density of $(\omega_1,\omega_2)$ given by (4.1) over a non-rectangular grid

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

(4.9)

for a range of values of $\omega_{2,i}$ and $\theta_i$, and then numerically integrating over $\omega_1$. This procedure, used in the example discussed in the next section, saves computations since the values for $G_p(\theta_i)$ defined by (4.2) can be stored.

Alternatively, the $\mathbf{1}_{F_1}^{(2)}$ function appearing in the posterior joint density of $(\omega_1,\omega_2)$ given by (3.8) can be calculated by applying numerical integration procedures to an integral representation of the $\mathbf{1}_{F_1}^{(2)}$ function. This function can be represented as an integral over the group of orthogonal matrices of order 2 where the integrand is a $\mathbf{1}_{F_1}$ hypergeometric function. The $\mathbf{1}_{F_1}$ 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 $\mathbf{1}_{F_0}$ hypergeometric function, which has the closed form $\mathbf{1}_{F_0}(c,M) = |I-M|^{-c}$. For more detail, see either James (1964) or Muirhead (1978). Tiao and Fienberg (1969) have used numerical integration for integrating over the group of orthogonal matrices of order 2 in calculating the $\mathbf{0}_{F_0}^{(2)}$ hypergeometric function. Finally, asymptotic approximations for the $\mathbf{1}_{F_1}^{(2)}$ function can be used when appropriate, see Muirhead (1978).

5. ANALYSIS OF A TYPICAL DATA SET

Timm (1975, p. 167) gives a set of 28 observations on a bivariate (k=2) variable, which we have arbitrarily split into two groups. The
first \( N_1 = 14 \) observations are regarded as the first sample, and the second \( N_2 = 14 \) observations are regarded as the second sample. The observations for sample 1 are

\[
X = \begin{bmatrix}
(72), & (60), & (56), & (41), & (32), & (30), & (39) \\
(66), & (53), & (57), & (29), & (32), & (35), & (35) \\
(42), & (37), & (33), & (32), & (63), & (54), & (47) \\
(43), & (40), & (29), & (30), & (45), & (46), & (51)
\end{bmatrix}
\]

(5.1)

and the observations for sample 2 are

\[
Y = \begin{bmatrix}
(91), & (56), & (79), & (81), & (78), & (46), & (39) \\
(79), & (60), & (65), & (80), & (55), & (38), & (35) \\
(32), & (60), & (35), & (39), & (50), & (43), & (48) \\
(30), & (50), & (37), & (36), & (34), & (37), & (54)
\end{bmatrix}
\]

(5.2)

Here, both \( X \) and \( Y \) are \((2 \times 14)\) matrices. We find

\[
\bar{x}_1 = \begin{pmatrix} \bar{x}_1 \\ \bar{x}_2 \end{pmatrix} = \begin{pmatrix} 45.5714 \\ 42.5000 \end{pmatrix} \quad \text{and} \quad \bar{y}_1 = \begin{pmatrix} \bar{y}_1 \\ \bar{y}_2 \end{pmatrix} = \begin{pmatrix} 55.000 \\ 49.2851 \end{pmatrix}
\]

(5.3)

while the matrix \( S \), as defined after (2.5) is

\[
S = 10^4 \begin{pmatrix} .715093 & .552600 \\ .552600 & .532860 \end{pmatrix}
\]

(5.4)

Hence, the matrix \( L \) defined in Theorem 3.1 is found to be

\[
L = \begin{pmatrix} 4.76787 & 5.56822 \\ 5.56882 & 6.57197 \end{pmatrix}
\]

(5.5)
and its eigenvalues are $\lambda_1 = 11.3107$ and $\lambda_2 = .029105$. Inserting the above information in Theorems 3.1 and 3.3 yields the joint posterior of $(\omega_1, \omega_2)$, and the marginal posterior of $\omega_2$ respectively.

Before graphing $p(\omega_2 | X, Y)$, we note from (4.8) that the expected value of $p$ is 147.4 with a standard deviation of 40.8. The posterior joint density of $(\omega_1, \omega_2)$ is computed from (4.1) using the values $0 \leq p \leq 399$ and for a grid of the form (4.9). Further computations based on (4.7) show that this range accounts for at least .99999 of the probability for $p$. Numerical integration over $\omega_1$ gives the resulting Table 5.1 and Figure 5.1 for the posterior density of $\omega_2$.

**TABLE 5.1**

The Marginal Posterior $p(\omega_2 | X, Y) = p(\omega_2)$

<table>
<thead>
<tr>
<th>$\omega_2$</th>
<th>.01</th>
<th>.05</th>
<th>.1</th>
<th>.5</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p(\omega_2)$</td>
<td>1.946</td>
<td>.879</td>
<td>.620</td>
<td>.268</td>
<td>.182</td>
<td>.117</td>
<td>.085</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$\omega_2$</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>8</th>
<th>10</th>
<th>12</th>
<th>14</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p(\omega_2)$</td>
<td>.065</td>
<td>.051</td>
<td>.041</td>
<td>.027</td>
<td>.017</td>
<td>.014</td>
<td>.008</td>
</tr>
</tbody>
</table>

It is easy to see from Table 5.1 and Figure 5.1 that indeed the posterior of $\omega_2$ is concentrated near zero. In fact defining

$$H(m) = P(\omega_2 > m \sqrt{\text{Var}(\omega_2 | X, Y)} | X, Y)$$  \hspace{1cm} (5.6)

and tabling $H$ for selected values of $m$ in Table 5.2, we see also from examination of $H$ that the posterior of $\omega_2$ is concentrated near zero.

Numerical integration is used to produce values of (5.6). In addition, by applying (3.21) we obtain

$$E(\omega_2 | X, Y) = 1.682, \quad \sqrt{\text{Var}(\omega_2 | X, Y)} = 2.208.$$  \hspace{1cm} (5.7)
TABLE 5.2

Selected values of $H(m)$

<table>
<thead>
<tr>
<th>$m$</th>
<th>.5</th>
<th>1.5</th>
<th>2.5</th>
<th>3.5</th>
<th>4.5</th>
<th>5.5</th>
<th>6.5</th>
<th>7.5</th>
<th>8.5</th>
<th>9.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H(m)$</td>
<td>.4474</td>
<td>.1693</td>
<td>.0676</td>
<td>.0271</td>
<td>.0108</td>
<td>.0043</td>
<td>.0017</td>
<td>.0007</td>
<td>.0003</td>
<td>.0001</td>
</tr>
</tbody>
</table>

With all the above as evidence, we accept the assertion $\frac{\eta}{c} = c \mu$ and look at the posterior (2.6) of $c$. For the data (5.1) and (5.2) we arrive at the posterior graphed in Figure 5.2 and tabulated in Table 5.3. The posterior of $c$ was tabulated in increments of .0231 for $c$ from 0 to 5. Using numerical integration, we obtain

$$E(c|X,Y) = 1.1186 \text{ and } \text{Var}(c|X,Y) = .023,$$

(5.8)

and "central" posterior intervals for $c$, which are given in Table 5.4 below. The Bayes estimate of $c$ is 1.1186, the modal posterior estimate is 1.1568.

TABLE 5.3

The posterior of $c$, $p(c) = p(c|X,Y)$

<table>
<thead>
<tr>
<th>$c$</th>
<th>0</th>
<th>.51</th>
<th>1.02</th>
<th>1.1568</th>
<th>1.53</th>
<th>2.04</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p(c)$</td>
<td>0</td>
<td>$118 \times 10^{-7}$</td>
<td>1.5711</td>
<td>2.8225</td>
<td>.2367</td>
<td>.001345</td>
</tr>
</tbody>
</table>

TABLE 5.4

Central $\beta$-level intervals for $c$ based on $p(c|X,Y)$

$$\int_0^{\beta} p(c|X,Y)dc = \int_{u_B}^{\infty} p(c|X,Y)dc = \frac{1}{2}(1-\beta)$$

<table>
<thead>
<tr>
<th>$\beta$</th>
<th>$(\ell_B, u_B)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>.8</td>
<td>(0.00436, 1.38038)</td>
</tr>
<tr>
<td>.9</td>
<td>(.95864, 1.45022)</td>
</tr>
<tr>
<td>.95</td>
<td>(.91928, 1.51716)</td>
</tr>
<tr>
<td>.99</td>
<td>(.84218, 1.67078)</td>
</tr>
</tbody>
</table>
Acknowledgements  The research of I. Guttman and U. Menzefricke was supported by grants A8743 and A5578, respectively, from NSERC of Canada. The research of D. Tyler was supported by a Summer Research Grant (81-968) from Old Dominion University, where this work was started. The work was 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


APPENDIX I

As mentioned in Section 2, this appendix is devoted to establishing the results (2.5) and (2.6). Using (2.3) and (2.4), we have that the posterior of $\mu$, $c$ and $\frac{1}{\psi}$, given the data $X,Y$ is such that

$$p(\mu, c, \frac{1}{\psi} | X, Y) \propto \left| \frac{1}{\psi} \right|^{-\frac{N-k-1}{2}} \exp(-\frac{1}{2} \frac{1}{\psi} S)$$

$$\times \exp -\frac{1}{2} \left[ N_1 (\bar{X} - \mu)' \frac{1}{\psi} (\bar{X} - \mu) + N_2 (\bar{Y} - c \xi)' \frac{1}{\psi} (\bar{Y} - c \xi) \right].$$

By straightforward and tedious algebra involving "completing the square" in $\mu$, it is easy to see that, apart from the $(-\frac{1}{2})$, the exponent of the last factor of (AI.1) may be written as

$$(N_1 + N_2 c^2)(\bar{X} - \xi)' \frac{1}{\psi} (\bar{X} - \xi) + a$$

where

$$\xi = (N_1 + N_2 c^2)^{-1} [N_1 \bar{X} + c N_2 \bar{Y}]$$

and

$$a = N_1 \bar{X} \frac{1}{\psi} \frac{1}{\xi} + N_2 \bar{Y} \frac{1}{\psi} \frac{1}{\xi} - (N_1 + N_2 c^2) \frac{1}{\xi} \frac{1}{\psi} \frac{1}{\xi}.$$

Substituting (AI.2) in (AI.1), and integrating out $\mu$, yields the posterior of $c$, $\frac{1}{\psi}$, and is easily seen to be such that

$$p(c, \frac{1}{\psi} | X, Y) \propto \left| (N_1 + N_2 c^2) \frac{1}{\psi} - \frac{1}{2} \frac{N-k-1}{2} \right| \exp(-\frac{1}{2} \frac{1}{\psi} S)$$

$$\times \exp -\frac{1}{2} a$$

or, equivalently,
\[ p(c, \frac{1}{x}, X, Y) \propto (N_1 + N_2 c^2)^{-k/2} \left| \frac{1}{x} \right|^2 \text{etr} \left( -\frac{1}{2} \frac{1}{x} \right) S \times \exp \left( -\frac{1}{2} a \right) \quad (AI.6) \]

Using the definition of \( \frac{1}{x} \) (given in (AI.3)), it is easily seen that we may write (AI.4) as

\[ a = \frac{N_1 N_2}{N_1 + N_2 c^2} \left( (\tilde{\gamma} - c\tilde{x})' \frac{1}{x} (\tilde{\gamma} - c\tilde{x}) \right) \quad (AI.7) \]

so that

\[ p(c, \frac{1}{x}, X, Y) \propto \left( N_1 + N_2 c^2 \right)^{-k/2} \left| \frac{1}{x} \right|^2 \text{etr} \left( -\frac{1}{2} \frac{1}{x} \right) \left[ S + \frac{N_1 N_2}{N_1 + N_2 c^2} (\tilde{\gamma} - c\tilde{x})(\tilde{\gamma} - c\tilde{x})' \right] \quad (AI.8) \]

Using the identity based on the Wishart distribution, we find, on integrating (AI.8) with respect to \( \frac{1}{x} \), that

\[ p(c|X, Y) \propto (N_1 + N_2 c^2)^{-k/2} \left| S + \frac{N_1 N_2}{N_1 + N_2 c^2} (\tilde{\gamma} - c\tilde{x})(\tilde{\gamma} - c\tilde{x})' \right|^2 \quad (AI.9) \]

But the determinant mentioned in (AI.9) may be written as

\[ \left| S \right| \left( 1 + \frac{N_1 N_2}{N_1 + N_2 c^2} (\tilde{\gamma} - c\tilde{x})(\tilde{\gamma} - c\tilde{x})' \right) \quad (AI.10) \]

and using \( |I + AB| = |I + BA| \), we find that (AI.10) may be written as

\[ \left| S \right| (1 + \frac{N_1 N_2}{N_1 + N_2 c^2} (\tilde{\gamma} - c\tilde{x})(\tilde{\gamma} - c\tilde{x})') \quad (AI.11) \]

or

\[ \frac{\left| S \right| N_1 N_2}{N_1 + N_2 c^2} \left\{ c^2 \left( \frac{1}{N_1} + b_1 \right) - 2 b_2 c + \left( b_3 + \frac{1}{N_2} \right) \right\} \quad (AI.12) \]
Substituting (AI.12) in (AI.9), and absorbing factors involving $|S|$ and $N_1N_2$ into the constant of proportionality, we find $p(c|X,Y)$ is as stated in (2.5).

We wish now to establish (2.6). This is straightforward, and involves, of course, completing the square of the factor in the braces \{ \} of (AI.12), and this is easily seen to be

$$
\left( \frac{N_1b_1+1}{N_1} \right) \left( c - \frac{N_1b_2}{N_1b_1+1} \right)^2 + \frac{(N_2b_3+1)(N_1b_1+1) - N_2N_1b_2^2}{N_2(N_1b_1+1)}.
$$

(AI.13)

Hence (AI.13) is proportional to

$$
\left\{ \frac{N_2(N_1b_1+1)^2}{N_1[(N_2b_3+1)(N_1b_1+1) - N_2N_1b_2^2]} \right\} \left( c - \frac{N_1b_1}{1+N_1b_1} \right)^2 + 1,
$$

(AI.14)

which if written as in (2.6), is

$$
a_0(c-\bar{c})^2 + 1
$$

(AI.15)

with

$$
\bar{c} = \frac{N_1b_2}{(1+N_1b_1)}
$$

(AI.15a)

which agrees with the result advertised in (2.6), and

$$
a_0 = \frac{N_2(N_1b_1+1)^2}{N_1[(N_2b_3+1)(N_1b_1+1) - N_2N_1b_2^2]}.
$$

(AI.15b)

Hence it remains to show that $a_0$ above may be expressed as in (2.6). But this is trivial, for we may write

$$
a_0 = \frac{N_2(N_1b_1+1)}{N_1} \left[ 1 + \frac{N_2b_3 + N_2N_1b_1(b_3^2 - b_2^2/b_1)}{1 + N_1b_1} \right].
$$

(AI.16)
Now

\[ b_2 - b_2^2/b_1 = \tilde{y}'S^{-1}\tilde{y} - (\tilde{x}'S^{-1}\tilde{y})^2/\tilde{x}'S^{-1}\tilde{x} \]  

\[ = (\tilde{y} - b\tilde{x})' S^{-1}(\tilde{y} - b\tilde{x}) \]

\[ = q \]

where \( q \) is notation used in (2.6), and

\[ b = b_2/b_1 = \frac{\tilde{x}'S^{-1}\tilde{y}}{\tilde{x}'S^{-1}\tilde{x}}. \]  

(AI.18)

Hence

\[ a_0 = \frac{N_2 (N_1 b_1 + 1)/N_1}{1 + \frac{N_2 b_3 + N_2 N_1 b_1 q}{1 + N_1 b}} \]  

\[ = \frac{N_2 (N_1 b_1 + 1)/T}{1 + \frac{N_2 b_3 + N_2 N_1 b_1 q}{1 + N_1 b}} \]  

(AI.19)

where \( T \) is as given in (2.6). This then, establishes (2.6).
APPENDIX II

In this appendix, we establish the result (3.3). Using the model (2.1), write the relation (3.1) holding, gives the likelihood

\[ \ell(\mu, C, \tau^{-1} | X, Y) \propto |\tau^{-1}|^{N/2} \text{etr}[-\frac{1}{2} \tau^{-1} \sum_{i=1}^{N_1} (x_i - \mu) (x_i - \mu)^\prime] \]

\[ + \sum_{j=1}^{N_2} (y_j - C\mu) (y_j - C\mu)^\prime] \]

which may be written as

\[ \ell(\mu, C, \tau^{-1} | X, Y) \propto |\tau^{-1}|^{N/2} \text{etr}[-\frac{1}{2} \tau^{-1} S] \]

\[ \times \exp -\frac{1}{2} \left\{ N_1 (\mu - \tilde{x})^\prime \tau^{-1} (\mu - \tilde{x}) + N_2 (C\mu - \tilde{y})^\prime \tau^{-1} (C\mu - \tilde{y}) \right\} . \]

Using the prior (2.4) yields the posterior of \( \mu, C, \tau^{-1} \), given the data, and is such that

\[ p(\mu, C, \tau^{-1} | X, Y) \propto |\tau^{-1}|^{\frac{N-k-1}{2}} \text{etr}[-\frac{1}{2} \tau^{-1} S] \times \exp -\frac{1}{2} d \]

where \( d \) is the expression in the braces \( \{ \} \) of (AII.2). But "completing the square" in \( \tau \) for \( d \) yields, as is easily seen,

\[ d = (\mu - \tilde{A}^{-1} m)^\prime \tilde{A}(\mu - \tilde{A}^{-1} m) + \left[ N_1 \tilde{x}^\prime \tilde{x}^{-1} \tilde{x} + N_2 \tilde{y}^\prime \tau^{-1} \tilde{y} - m^\prime \tilde{A}^{-1} m \right] , \]

where

\[ m = N_1 \tilde{x}^{-1} \tilde{x} + N_2 C\tilde{y}^{-1} \tilde{y} . \]

and

\[ \tilde{A} = N_1 \tilde{x}^{-1} + N_2 C\tilde{y}^{-1} C . \]
Substitution of (AII.4) into (AII.3), and integrating with respect to \( \mu \) yields

\[
p(C, \xi^{-1} | X, Y) = |A|^{-\frac{1}{2}} |\xi|^{-\frac{1}{2}} \times \text{etr}^{-\frac{1}{2}} \xi^{-1} S \times \exp^{-\frac{1}{2}} \bar{d}
\]  

(AII.7)

where \( \bar{d} \) is the expression in the braces \{ \} of (AII.4). We note that the first two factors of (AII.7) match the first two factors of (3.3), so that inspection of (3.3) and (AII.7) implies that we need only show

\[
\bar{d} = (\bar{\gamma} - C\tilde{x})' A(\bar{\gamma} - C\tilde{x})
\]

(AII.8)

where \( A \) is defined in (3.3). Now from the definition of \( \bar{d} \) given after (AII.7) above, and using (AII.5), we find

\[
\bar{d} = \bar{\gamma}' A_1 \bar{\gamma} + \bar{\xi}' A_2 \bar{\xi} - 2\bar{\gamma}' A_3 \bar{\xi},
\]

(AII.9)

with the \( A_j \) given by

\[
A_1 = N_2 \xi^{-1} - N_2 \xi^{-1} C A^{-1} C \xi^{-1},
\]

(AII.9a)

\[
A_2 = N_1 \xi^{-1} - N_1 \xi^{-1} A^{-1} \xi^{-1},
\]

(AII.9b)

and

\[
A_3 = N_1 N_2 \xi^{-1} C A^{-1} \xi^{-1}.
\]

(AII.9c)

We note that \( A_1 = A \), where \( A \) is defined in (5.3), since \( \bar{A} \) is as given in (AII.6). We also note that if we can establish that

\[
A_2 = CA_1 C = CAC
\]

(AII.10)

and

\[
A_3 = A_1 C = AC
\]

(AII.11)
then we would have, from (AII.9), that

\[
\tilde{d} = \tilde{\gamma}'A\tilde{\gamma} + \tilde{\chi}'CAC\tilde{\chi} - 2\tilde{\gamma}'AC\tilde{\chi} \\
= (\tilde{\gamma} - C\tilde{\chi})' A(\tilde{\gamma} - C\tilde{\chi}),
\]

(AII.12)

establishing (AII.8), and hence (3.3) would be proved. We now prove (AII.10) and (AII.11).

Firstly, we note that \( AC \) may be written as

\[
AC = N_2 \tilde{\gamma}^{-1}C - N_2 \tilde{\chi}^{-1}C\tilde{\chi}^{-1}C\tilde{\chi}^{-1}C
\]

(AII.13)

so that

\[
AC = N_2 \tilde{\gamma}^{-1}C[1 - N_2 \tilde{\gamma}^{-1}C\tilde{\chi}^{-1}C]
\]

so that

\[
AC = N_2 \tilde{\gamma}^{-1}C\tilde{\chi}^{-1}(\tilde{\alpha} - N_2 C\tilde{\chi}^{-1}C)
\]

(AII.13a)

Using the definition of \( \tilde{\alpha} \) given in (AII.6), we find

\[
AC = N_2 \tilde{\gamma}^{-1}C\tilde{\chi}^{-1}(N_1 \tilde{\gamma}^{-1})
\]

(AII.13b)

so that from (AII.9c), we have now proved (AII.11).

Similarly, we note that

\[
CA_1C = CAC = N_2 C\tilde{\chi}^{-1}C(I - N_2 \tilde{\gamma}^{-1}C\tilde{\chi}^{-1}C)
\]

(AII.14)

so that again using (AII.6) we find

\[
CAC = N_2 C\tilde{\chi}^{-1}C\tilde{\chi}^{-1}(N_1 \tilde{\gamma}^{-1})
\]

(AII.14a)

\[
= N_1 N_2 C\tilde{\chi}^{-1}C\tilde{\chi}^{-1} \tilde{\gamma}^{-1}.
\]
But we also have, from (AII.9b), that

\[ A_2 = N_1^{\dagger -1} - N_1^2 \tilde{A}^{-1} \tilde{A}^{-1} \tilde{A}^{\dagger -1} \]  
\[ = (I - N_1^{\dagger -1} \tilde{A}^{-1}) N_1^{\dagger -1} \]
\[ = (\tilde{A} - N_1^{\dagger -1}) \tilde{A}^{-1} N_1^{\dagger -1} \]
\[ = (N_2 \tilde{C}^{\dagger -1} C) \tilde{A}^{-1} N_1^{\dagger -1} \]
\[ = N_1 N_2 \tilde{C}^{\dagger -1} C \tilde{A}^{-1} \tilde{A}^{\dagger -1} \]

that is, \( A_2 = \text{CAC} = N_1 N_2 \tilde{C}^{\dagger -1} C \tilde{A}^{-1} \tilde{A}^{\dagger -1} \). Hence we have proved (AII.10), and (3.3) is now established.