CHAPTER II

STUDENT $t$-VARIABLES IN BAYESIAN INFERENCE

PROBLEMS

1. Introduction

In many applied decision problems, e.g., the multiple comparisons problem in analysis of variance setting, we are concerned with comparing several normal populations $\mathcal{N}(\mu_i, \sigma^2), i = 1, 2, \ldots$. Here we assume that, a priori, the means $\mu_i$ have independent normal distributions with equal means $\bar{\mu}$ and equal variances $\sigma^2_{\mu}$, where $\bar{\mu}$ and $\sigma^2_{\mu}$ are unknown but the ratio $\sigma^2_{\mu}/\sigma^2$ is known. Using this prior knowledge, Duncan (1961, 1965) obtained a Bayes solution of the multiple comparisons problem concerning the means $\mu_i$ by restricting to a class of invariant rules. By using the prior knowledge of the parameters $\mu_i$, we construct a new prior distribution and adopt it for making posterior inferences. We obtain posterior distributions of some functions of parameters useful for comparing the normal populations. These results point out the use of standard statistics—the $t$-ratio, the $F$-ratio in the Bayesian inference problems and incidently provide a full Bayesian justification of Duncan's (1961) solution of the multiple comparisons problem.
2. Comparison of two normal populations

2.1 Consider a problem where we wish to compare two normal populations \( N(\mu_1, \sigma^2) \), \( i = 1, 2 \). Let \( x_{11}, x_{12}, \ldots, x_{1n} \) and \( x_{21}, x_{22}, \ldots, x_{2n} \) denote samples of size \( n \) each from the two normal populations. Suppose it is known that, apriori, \( \mu_1, \mu_2 \) have independent normal distributions with equal means \( \mu \) and equal variances \( \sigma^2 \), where \( \mu \) and \( \sigma^2 \) are unknown but the ratio \( \sigma^2_1 / \sigma^2 = 1/n' \) is known. See Duncan (1961). For purposes of Bayesian comparison of the two populations, we may confine to posterior inferences about \( (\mu_1 - \mu_2) / \sigma \); this quantity measures the difference in means \( \mu_1, \mu_2 \) in standard deviation units.

The likelihood \( l(x|\theta) \) of the samples

\[
y = (x_{11}, x_{12}, \ldots, x_{1n}, x_{21}, x_{22}, \ldots, x_{2n}) \text{ and parameters } \theta, i.e., \text{the joint probability density of the sample observations } x \text{ is}
\]

\[
l(x|\theta) = (\sqrt{2\pi} \sigma)^{-2n} \exp \left[ -\frac{1}{2\sigma^2} \left( \frac{1}{n} \sum_{i=1}^{n} (x_{1i} - \mu)^2 + \frac{1}{n} \sum_{i=1}^{n} (x_{2i} - \mu)^2 \right) \right]
\]

\[
= (\sqrt{2\pi} \sigma)^{-2n} \exp \left[ -\frac{1}{2\sigma^2} \sum_{i,j=1}^{n} (d_{ij} - \bar{d}_{ij})^2 \right]
\]

\[
= \frac{1}{(\sqrt{2\pi} \sigma)^{-2n}} \exp \left[ -\frac{1}{2\sigma^2} \sum_{i=1}^{n} \frac{1}{2} \left( \frac{x_{1i} + x_{2i} - \mu_1 - \mu_2}{\sigma} \right)^2 \right]
\]

(2.1)
where $d_{i|j}$ are linear functions in the observations $x_{i|j}$, 
$
\delta_{i|j}$ are linear functions in $\mu_1$ and $\mu_2$, $x_{i|1} = \frac{\sum x_{i|1}}{n}$
and $x_{i|2} = \frac{\sum x_{i|2}}{n}$.

The parameters involved in the likelihood of the sample are $\mu_1, \mu_2$ and $\sigma^2$. Since prior density of only $\mu_1, \mu_2$ (some parameters in this density being not known) is given, the posterior inferences cannot be made directly. For this reason, we obtain new prior distribution for $(\mu_1, \mu_2, \sigma^2)$ by using the method suggested in Section 6 of Chapter 1, and adopt it for making posterior inferences.

It is given that the prior density of $\mu_1$ and $\mu_2$ is

$$g_0(\mu_1, \mu_2) = \left( \frac{n'}{2\pi \sigma^2} \right)^{\frac{1}{2}} \exp\left\{ -\frac{n'}{2\sigma^2} \left( \frac{(\mu_1 - m)^2 + (\mu_2 - m)^2}{2} \right) \right\}. $$

... (3.2)

Let us define $s_1 = (\mu_1 - \mu_2)/\sqrt{2}$, $s_2 = (\mu_1 + \mu_2)/\sqrt{2}$. The prior density of $s_1, s_2$ is

$$g_0(s_1, s_2) = \left( \frac{\sqrt{n'}}{\sqrt{2\pi} \sigma} \right)^{\frac{1}{2}} \exp\left\{ -\frac{n'}{2\sigma^2} s_1^2 \right\} \left[ \frac{\sqrt{n'}}{\sqrt{2\pi} \sigma} \exp\left\{ -\frac{n'}{2\sigma^2} s_2^2 \right\} \right]. $$

... (3.3)
where \( g_1(\theta_1) \) and \( g_2(\theta_2) \) are the marginal densities of \( \theta_1 \) and \( \theta_2 \) respectively.

Since \( n \) is not known, we obtain a reduced prior distribution \( \tilde{f}_0(\theta_1, \theta_2) \):

\[
\tilde{f}_0(\theta_1, \theta_2) = g_1(\theta_1) \cdot p(\theta_2) \quad \ldots \tag{2.4}
\]

where

\[
p(\theta_2) \propto c \\ \ldots \tag{2.5}
\]

That is, the function \( \tilde{f}_0(\theta_1, \theta_2) \) is obtained by adopting an uniform prior distribution (See Jeffreys(1931)) for the location parameter \( \theta_2 \) instead of the distribution \( g_2(\theta_2) \).

Now since the scale parameter \( \sigma \) is not known, the final prior distribution for \( (\theta_1, \theta_2, \sigma) \) is taken as

\[
\tilde{f}(\theta_1, \theta_2, \sigma) \propto \sigma^{-1} \tilde{f}_0(\theta_1, \theta_2) \\
\propto \sigma^{-2} \exp \left\{ -\frac{n}{2\sigma^2} \theta_1 \right\}. \quad \ldots \tag{2.6}
\]
2.2 The posterior density of $\tau = \sqrt{n\delta_1/\sigma}$

Using $g(\delta_1, \delta_2, \sigma)$ for the prior distribution, the posterior distribution of $\delta_1, \delta_2, \sigma$ is obtained as

$$p(\delta_1, \delta_2, \sigma | x) \propto l(x | \delta_1, \delta_2, \sigma) g(\delta_1, \delta_2, \sigma)$$  \hspace{1cm} (2.7)

where $l(x | \delta_1, \delta_2, \sigma)$ is given by (2.1).

Let us transform $\delta_1, \delta_2, \sigma$ to $\tau, \delta_2, \sigma$

where

$$\tau = \sqrt{n\delta_1/\sigma}.$$  \hspace{1cm} (2.8)

The posterior joint density of $\tau, \delta_2, \sigma$ is expressable as

$$p(\tau, \delta_2, \sigma | x) \propto l(x | \tau, \delta_2, \sigma) \sigma^{-1} q(\tau)$$  \hspace{1cm} (2.9)

where

$$q(\tau) = \frac{1}{\sqrt{(2\pi)^\gamma}} \exp \left( - \frac{\tau^2}{2\gamma^2} \right)$$  \hspace{1cm} (2.10)

defines the density function of a normal variable with
mean zero and variance \( \tau^2 \) and \( \nu^2 = \frac{n}{n'} \).

Let us define the usual \( t \)-ratio

\[
t = \sqrt{n} \frac{\bar{x}_1 - \bar{x}_2}{s}
\]

where

\[
s^2 = \sum (x_i - \bar{x}_i)^2 / v, \quad v = 2(n-1)
\]

The posterior density \( p(\tau | \bar{x}) \) of \( \tau \) is obtained after integrating (2.9) over \( \sigma_2 \) and \( \sigma_0 \). Thus

\[
p(\tau | \bar{x}) \propto \gamma(\nu) \int \sigma^{-1} l(\tau | \bar{x}, \sigma_2, \sigma_0) d\sigma_2 d\sigma
\]

It can be easily verified that

\[
\int \sigma^{-1} l(\tau | \bar{x}, \sigma_2, \sigma) d\sigma_2 d\sigma \propto \int u^v \exp \left[ -\frac{1}{2} \left\{ u^2 v + (ut - \tau)^2 \right\} \right] du
\]

where \( t \) is defined in (2.11).

The right hand side of (2.14) is a function of \( t \) and \( \tau \). We thus have
\[ p(v|z) = p(v|t) \]

\[ \propto \frac{1}{2} f_v(t, z) \] (2.15)

where

\[ f_v(t, z) \propto \int_0^\infty u^v \exp \left[ -\frac{1}{2} \left( u^2 + (ut-z)^2 \right) \right] du \] (2.16)

Thus the statistic \( t \) is marginally sufficient for \( z \) w.r.t. the prior distribution \( \mathcal{N}(\mu, \sigma^2) \). We also note that the function \( f_v(t, z) \) defines the density function of a noncentral t-variable with noncentrality parameter \( z \) and \( v \) degrees of freedom.

On simplification of (2.15), we obtain

\[
p(v|t) = \frac{2^v v!}{(v-1)/2 \Gamma(v-1/2)} \left\{ \frac{v+t^2/(1+z^2)^2}{v+t^2} \right\}^{(v+1)/2} \exp \left\{ -\frac{1}{2} z^2 \left( \frac{t^2}{v+t^2} \right) \right\} I_v(-z) \] (2.17)

where \( I_v(x) \) is defined as in Fisher (1931) -
\[ I_v(x) = \int_0^\infty (\sqrt{2\pi})^{-1} w^v \exp \left\{ -\frac{1}{2} (w + x)^2 \right\} dw \]  

.. (2.18)

and

\[ x = \frac{\alpha \gamma}{(\nu + \tau^2)} , \quad \beta^2 = 1 + 1/\gamma^2 \quad .. \quad (2.19) \]

The function \( I_v(x) \) is tabulated in British Assn. Moth. Tables Volume 1 by Airey (1937).

We now obtain the moment generating function of \( p(\tau|\nu) \). Let \( h \) be a real number. We have

\[ E_q \{ \exp (zh) \} = \frac{\int_0^\infty \exp (zh) f(t, \tau) \hat{\psi}(\tau) d\tau}{\int_0^\infty f(t, \tau) \hat{\psi}(\tau) d\tau} \quad .. \quad (2.20) \]

(We write the expectations in the posterior distribution with a suffix 1).

On solving the integrals on the r.h.s. of (2.20), we obtain
\[ E_1 \left\{ \exp(\tau h) \right\} = \left\{ (v/2) \right\} \cdot 2^{(v+2)/2} \exp \left\{ \frac{t^2}{2 \beta^2} \left( 1 + \frac{t^2}{v+t^2/(1+\tau^2)} \right) \right\} \]

\[ I_v(- \theta h / \beta(v+t^2/(1+\tau^2))^{1/2}) \quad (2.21) \]

By using Fisher's (1931) expansion of \( I_v(x) = \)

\[ I_v(x) = \frac{2^{-(v+2)/2} (v-1)!}{(v-2)!} y_2 - \frac{2^{-(v+1)/2} (v-1)!}{(v-2)!} y_1 \quad \ldots \quad (2.22) \]

where \( y_1 \) and \( y_2 \) are given by

\[ y_1 = x + \frac{(v-1)}{3!} x^3 + \frac{(v-1)(v-3)}{5!} x^5 + \ldots \quad \ldots \quad (3.23) \]

\[ y_2 = 1 + \frac{\sqrt{v}}{2!} x^2 + \frac{v(v-2)}{4!} x^4 + \ldots \quad \ldots \quad \]

we can express the r.h.s of (2.21) as a power series in \( h \)

and obtain the moments of the distribution \( p(\tau | t) \).

The first four moments of the distribution

are given by
\[ E_1(\gamma) = (t/\beta^2) L_1 \left\{ (v/2)! / ((v-1)/2)! \right\} \]

\[ E_1(\gamma^2) = 1/\beta^2 + (t/\beta^2)^2 L_1 \left\{ ((v+1)/2)! / ((v-1)/2)! \right\} \]

\[ E_1(\gamma^3) = E_1(\gamma) \left\{ 3/\beta^2 + (t/\beta^2)^2 L_1 \left\{ ((v+2)/2)! / (v/2)! \right\} \right\} \]

\[ E_1(\gamma^4) = \frac{1}{\beta^2} \left\{ \frac{3}{\beta^2} + 6(t/\beta^2)^2 L_1 \left\{ \frac{(v+1)!}{(y-1)/2)!} + \frac{t^4}{\beta^6} L_1 \left\{ \frac{(v+3)}{(y-2)\/2)!} \right\} \right\} \]

where

\[ L_1 = \sqrt{2} / (v + t^2 / (1 + t^2)) \]

The second, third and fourth moments of about the mean are given by

\[ \text{Var}_1(\gamma) = \frac{1}{\beta^2} + (t/\beta)^2 L_1 \left\{ (v+1)/2! - ((v/2)! / ((v-1)/2)!) \right\} \]

\[ E_1(\gamma - E_1(\gamma))^3 = \frac{1}{2} (t/\beta^2)^3 L_1 \left\{ 2 \left( \frac{(v+1)}{(v-1)/2)! \right) - (v+1) \right\} \]

\[ \ldots \]

\[ (2.24) \]

\[ (2.25) \]

\[ (2.26) \]
When $v$ is large, we can use the approximation for gamma functions $\frac{n(p+b)}{p} \approx p^h$ (where $p$ is large and $h$ small). Using this approximation

$$\text{Var}_1(\tau) = \frac{1}{\beta^2}$$

$$E_1(\tau - E_1(\tau)) = 0$$

$$E_1(\tau - E_1(\tau))^4 = 3/\beta^4.$$  \hspace{1cm} (2.27)

Thus, for moderate values of $t$, when $v$ is large, the approximate posterior density of $\tau$ is normal with mean and variance given in (2.24) and (2.26) respectively.

3. Comparison of several normal populations.

3.1. We consider a problem of comparing $(p+1)$ normal populations $N(\mu_i, \sigma^2)$, $i = 1, 2, \ldots, p+1$. Let:
Let \( \mathbf{Z} = \{ z_{ij}, j = 1, 2, \ldots, n, i = 1, 2, \ldots, p+1 \} \) denote the samples of size \( n \) from each of the \((p+1)\) populations. Suppose it is known that, apriori, \( \mu_i \)'s have independent normal distributions with equal means \( \mu \) and equal variances \( \sigma^2 \), where \( \mu, \sigma^2 \) are unknown, but the ratio \( \sigma^2 / \sigma^2 \) is known. For purposes of Bayesian comparisons of the \((p+1)\) populations, we may confine to posterior inferences about a set of orthogonal contrasts in \( \mu_i / \sigma, i = 1, 2, \ldots, p+1 \).

Let \( \delta_1, \delta_2, \ldots, \delta_p \) denote the set of \( p \) orthogonal linear functions of \( \mu \)'s: \( \delta_k = \sum_{j=1}^{p} \gamma_j \mu_j \)

where \( \sum_{j=1}^{p} \gamma_j^2 = 1, \sum_{j=1}^{p} \gamma_j \mu_j = 0, i \neq i' \).

As in section 2, the prior distribution of the parameters involved in the likelihood is not completely specified. We therefore adopt a reduced prior measure to obtain the posterior inferences.

It is given that the prior density of 

\[
(\delta_1, \delta_2, \ldots, \delta_p, \Sigma_{\mu_i})
\]

is

\[
g_0(\delta_1, \delta_2, \ldots, \delta_p, \Sigma_{\mu_i}) = \prod_{i=1}^{p} \left\{ \frac{\sqrt{n_i}}{\sqrt{(2\pi)\sigma}} \exp \left( - \frac{n_i \delta_i^2}{2 \sigma^2} \right) \right\}
\]

\[
\frac{\sqrt{n_i}}{\sqrt{2\pi(p+1)\sigma}} \exp \left\{ - \frac{n_i}{2(p+1)\sigma^2} \left( \frac{\mu_1 - (p+1)\mu}{\Sigma_{\mu_i}} \right) \right\}
\]

\[
= g_0(\delta_1) \cdots g_0(\delta_p) g_0(\Sigma_{\mu_i}) \quad (3.1)
\]
Since \( \mu \) is not known, we obtain the function

\[
\bar{g}_d(\theta_1, \ldots, \theta_p, \Sigma_{\mu_1}) = g_1(\theta_1) \cdots g_1(\theta_p)
\]

by replacing the uniform prior distribution for \( \Sigma_{\mu_1} \) instead of \( g_1(\Sigma_{\mu_1}) \). The 'reduced prior' for \((\theta_1, \ldots, \theta_p, \Sigma_{\mu_1}, \sigma)\) is given by

\[
\bar{g}(\theta_1, \ldots, \theta_p, \Sigma_{\mu_1}, \sigma) = \sigma^{-1} g_1(\theta_1) \cdots g_1(\theta_p)
\]

\[
\times \sigma^{-(p+1)} \exp \left\{ -\frac{n}{2\sigma^2} \sum \delta^2 \right\}. \tag{3.3}
\]

3.2. The posterior distribution of \( \xi = \sqrt{n} \bar{\theta}/\sigma \).

The posterior joint density of

\[
g = (\theta_1, \ldots, \theta_p), \quad \Sigma_{\mu_1}, \sigma \text{ w.r.t. the prior distribution}
\]

\[
\bar{g}(\xi, \Sigma_{\mu_1}, \sigma) \text{ is}
\]

\[
p(\xi, \Sigma_{\mu_1}, \sigma | x) = \bar{g}(\xi, \Sigma_{\mu_1}, \sigma) l(\xi | \bar{\theta}, \Sigma_{\mu_1}, \sigma)
\]

\[
\tag{3.4}
\]

where \( l(\xi | \bar{\theta}, \Sigma_{\mu_1}, \sigma) \) is the likelihood of the sample \( x \).
Let \( r_i = \sqrt{n} \frac{\delta_i}{\sigma}, i = 1, 2, \ldots, p \). The posterior joint density of \( \mathbf{\xi}^T = (\xi_1, \ldots, \xi_p)^T \), \( \Sigma/\mu_1\), \( \sigma \) is given by

\[
p(\xi, \Sigma/\mu_1, \sigma \mid x) \propto \frac{1}{\sqrt{(2\pi)^p \sigma}} \exp \left\{ \frac{1}{2} \frac{1}{\sigma} \sum_{i=1}^{p} \Sigma \right\}
\]

(3.5)

where

\[
\frac{1}{\sqrt{(2\pi)^{p-1} \sigma}} \exp \left\{ \frac{1}{2} \frac{1}{\sigma} \sum_{i=1}^{p} \Sigma \right\}
\]

(3.6)

and \( r^2 = n/p' \).

Let us define the vector \( \mathbf{t} = (t_1, t_2, \ldots, t_p)^T \)

where

\[
t_i = \frac{\sqrt{n} \sum_{j=1}^{n} \xi_j}{s} \quad \ldots \quad (3.7)
\]

and

\[
s^2 = \frac{\sum_{i=1}^{p} \sum_{j=1}^{n} (\xi_{ij} - \bar{x}_i)^2}{v}, \quad v = (p+1)(n-1)
\]

(3.8)

The posterior density of \( \mathbf{\xi} \) is given by

\[
p(\mathbf{\xi} \mid x) \propto \frac{1}{\sqrt{(2\pi)^p \sigma}} \int \sigma^{-1} \prod_{i=1}^{p} \left( \xi_i, \Sigma/\mu_1, \sigma \right) d(\Sigma/\mu_1) d\sigma
\]

(3.9)

It can be verified that

\[
\int \sigma^{-1} \prod_{i=1}^{p} \left( \xi_i, \Sigma/\mu_1, \sigma \right) d(\Sigma/\mu_1) d\sigma
\]

\[
\propto \xi_v (\mathbf{t}, \mathbf{\xi}) \quad \ldots \quad (3.10)
\]
where \( f_v(t; \xi) \) defines the density function of the multivariate \( t \)-distribution (see Dunnek and Sobel (1983)) for the \( t_i \)-variables having noncentrality parameters \( \xi_i \) and \( v \) degrees of freedom -

\[
f_v(t; \xi) \propto \int_0^{\infty} u^{(v+p-1)} \exp \left[ -\frac{1}{2} \left( u^2 v + \sum \left( u t_i - \xi_i \right)^2 \right) \right] du. \tag{3.11}
\]

The posterior density of \( \xi \) is therefore given by

\[
p(\xi | x) = p(\xi | t) \propto f_v(t; \xi) \frac{1}{f(t)} \tag{3.12}
\]

On simplification we obtain

\[
p(\xi | x) = \frac{\beta^p v^p (p-1)! (v+\sum t_i^2)/(1+\sum_i) \exp \left[ -\frac{1}{2} \left( \beta^p \sum r_i^2 - x^2 \right) \right] \Gamma_{v+p-1}(-x)}{(2v)^{(p-1)/2} v^v (v+p-2)! (v+p-2)/(2v^2)!} \tag{3.13}
\]

where
\[ x = \frac{\sum t_i x_i}{(v + \sum t_i^2)^{1/2}} \quad \ldots \quad (3.14) \]

and \( I_v(x), \beta^2 \) are as defined in (2.18) and (2.19).

The posterior marginal density of \( x_1 \) is

\[
p(x_1 | \mathbf{t}) = \frac{\beta (v + p - 1)}{2} \frac{v + \sum t_j^2 / (1 + r^2)}{(v + p - 2) / 2} \frac{v + \sum t_j^2 / (1 + r^2) + t_1^2}{v + \sum t_j^2 / (1 + r^2) + t_1^2}
\]

\[
\exp \left\{ -\frac{1}{2} (\beta t_1^2 - x_1^2) \right\} I_{v+p-1}(-x_1^2) \quad \ldots \quad (3.15)
\]

where

\[
x_1' = \frac{t_1 x_1}{(v + \sum t_j^2 / (1 + r^2) + t_1^2)^{1/2}} \quad \ldots \quad (3.16)
\]

The mean vector and the variance-covariance matrix of the distribution \( p(x_1 | \mathbf{t}) \) are

\[
E_1(x) = \frac{k_1}{\beta^2} \mathbf{t} \quad \ldots \quad (3.17)
\]

\[
E(x_1 x_1') - E(x_1) E(x_1') = (1/\beta^2) I + \frac{k_2}{\beta^2} \mathbf{tt}' \quad \ldots \quad (3.18)
\]

where
\[ k_1 = \left( \frac{\sqrt{2}}{(v + z_1^2/(1 + r^2))} \right) \left( \frac{(v + p - 1)/2}{(v + p - 2)/2} \right) \]

and

\[ k_2 = \frac{2}{(v + \Sigma t_1^2/(1 + r^2))} \cdot \left\{ \frac{v + p - 1}{2} \right\} \left( \frac{v + p - 2}{2} \right)^2 \]

When \( t_1 \)'s are moderate and \( v \) is large, the approximate posterior distribution of \( \bar{r} \) is \( p \)-variate normal with the mean vector and the variance-covariance matrix given by (3.17) and (3.18) respectively.

3.3. Posterior distribution of \( \bar{z} \left( \mu_1 - \bar{\mu} \right)^2 / \sigma^2 \)

A common problem in comparisons of several normal populations having the same variance arising for example in the comparison of treatments in analysis of variance setting is the inference about the null hypothesis, namely the hypothesis which states that the several means are equal. A measure of distance of the null from the true hypothesis is provided by \( \bar{z} \left( \mu_1 - \bar{\mu} \right)^2 / \sigma^2 \), \( \bar{\mu} = \mu_1/(p + 1) \). (See Bulmer (1958).)
We may wish to conclude that the treatments are homogeneous if $\Sigma (\mu_1 - \bar{\mu})^2/\sigma^2$ is less than a certain value say $k_c$. Thus given sample observations, we may wish to obtain the posterior probability of the statement $\Sigma (\mu_1 - \bar{\mu})^2/\sigma^2 < k_c$.

We here derive the posterior distribution of $\Sigma (\mu_1 - \bar{\mu})^2/\sigma^2$.

From (3.12), we have

$$f_v(t_2^2) = \int \text{const.} \exp \left\{ -\frac{1}{2} v^2 \left( t_2^2 / (1 + t_2^2) \right) \right\} \exp \left\{ -\frac{1}{2} \Sigma (\mu_1^2 - \bar{\mu}^2)^2 \right\} \frac{v^{v-1}}{2^{v/2}} dv$$

Let us define $\phi = \beta^2 \Sigma t_2^2 = \beta^2 \Sigma (\mu_1 - \bar{\mu})^2 / \sigma^2$.

We have

$$\int (\beta^2 \Sigma t_2^2) \exp \left\{ -\frac{1}{2} \Sigma (\mu_1^2 - \bar{\mu}^2)^2 \right\} d\Sigma$$

$$= \left\{ 2^{v/2} \frac{\Gamma(v/2)}{\sqrt{v}} \right\} \exp \left\{ -\frac{1}{2} \left( \phi + \frac{\beta^2}{\sigma^2} \right) \right\} \frac{1}{\beta^2} \exp \left\{ -\frac{1}{2} \frac{\beta^2}{\sigma^2} \right\} \int_0^\infty \exp \left\{ -\frac{v^2 \phi}{2 \beta^2} \right\} \frac{\Gamma(v-1)}{2} d\phi$$

where

$$\phi = \Sigma t_2^2 / \rho$$

Therefore we obtain
\[ \int f_Y(z) f(z \mid \phi) dz = \text{const.} \exp \left\{ -\phi / 2 \right\} \phi^{(p-2)/2} \]
\[ \int_0^{1/2} \phi^{-1} d\phi = 0^\infty \int_0^{1/2} \phi^{-1} d\phi + \frac{1}{2} d\phi \]
\[ \sum_{r=0}^{\infty} \frac{(pF \phi)}{(2r)!} \frac{1}{(v+p)_{(v+p)/2 + r}} \frac{(v+p+r-1)!}{(v+pF)_{(v+p)/2 + r}} \]
\[ B\left(\frac{v-1}{2}, \frac{r+1}{2}\right) d\phi \]

On dividing by the normalising factor and simplifying, we obtain the posterior density of \( \phi \)

\[ p(\phi \mid x) = p(\phi \mid F) \]
\[ = \left\{ \frac{p/2}{(v+pF(1+\gamma^2))} \right\}^{(-1)} \frac{(v+pF(1+\gamma^2))}{v+pF} \exp \left\{ -\phi / 2 \right\} \]
\[ \phi^{(p-2)/2} \sum_{r=0}^{\infty} \frac{pF \phi}{2^r (v+pF)} \frac{1}{(v+pF)_{(v+p)/2 + r}} \frac{(v+p+r-1)!}{(v+pF)_{(v+p)/2 + r}} \]

This density of \( \phi \) can be expressed as the weighted sum of the densities of chi-square variables -

\[ p(\phi \mid F) = \sum_{r=0}^{\infty} \frac{pF \phi}{2^r (v+pF)} \frac{1}{(v+pF)_{(v+p)/2 + r}} \frac{(v+p+r-1)!}{(v+pF)_{(v+p)/2 + r}} \]

where \( f(x^2_{p+2r}) \) is the density function of a chi-square variable with \( (p+2r) \) degrees of freedom and

\[ t = \alpha \]
\[ \nu_p = \left\{ \frac{v + pF}{v + pF} \right\} \]
\[ \rho = \frac{(v+p+r-1)\rho}{\rho^{(v+p)/2}} \]
\[ \frac{\rho^{pF}}{\rho^{2(v+pF)}} \]

The moments of \( \phi \) can be obtained by using (3.26).

\[ E_1(\phi) = \sum \nu_p E \left( \chi^2_{\nu + 2r} \right) = \sum \nu_p (p + 2r) \]
and
\[ E_1(\phi^2) = \sum \nu_p E \left( \chi^2_{\nu + 2r} \right)^2 = \sum \nu_p \left\{ 2(p + 2r) + (p + 2r)^2 \right\} \quad (3.29) \]

On substituting \( \nu_p \) in (3.28) and (3.29), we obtain on simplification

\[ E_1(\phi) = p + k_3 pF/ \beta^2 \]
and
\[ \text{Var}_1(\phi) = 2p + 4k_3 pF/ \beta^2 + (k_4 - k_3^2)(pF/\beta^2)^2 \]
where
\[ k_3 = (v + p)/(v + pF/(1 + r^2)) \]
and
\[ k_4 = (v + p)(v + p + 2)/(v + pF/(1 + r^2))^2 \]

For large \( v \), the approximate distribution of \( \phi \) is that of a \( \chi^2 \)-variable, where \( c \) and \( p' \)

\[ c = \{\text{Var}_1(\phi)\} / \{2E_1(\phi)\} \quad \text{and} \quad p' = \{2E_1^2(\phi)\} / \{\text{Var}_1(\phi)\} \]

are obtained by using (3.30) to (3.33). (See Patnaik (1944)).
4. Posterior Distributions Concerning Multivariate

Normal Samples.

4.1 Suppose that we are given a sample $x \in \mathbb{R}^p$
size $n = x_1, x_2, \ldots, x_n, X_j = (x_{1j}, \ldots, x_{pj})^\prime$, from a $p$-variate
normal population $N(\mu, \Sigma)$ where $\Sigma = B \sigma^2$, $B$ is a $p \times p$ positive
definite matrix of known constants and $\mu$ and $\sigma^2$ are unknown.
Suppose it is known that, apriori, $\mu$ has a normal distribution
with mean vector $\bar{\mu}^* \sigma$ ($\bar{\mu}^*$ is known) and variance-covariance
matrix $\bar{\Sigma} \sigma^2$ ($\bar{\Sigma}$ is a $p \times p$ positive definite matrix of known
constants). We obtain below the posterior distribution of
$\sigma^{-1} \mu$.

The prior density of $\mu$ is
\[ g_0(\mu) \propto 1 \sigma^* \sigma^2 \sigma^{-1/2} \exp \left\{ -\frac{1}{2} (\mu - \bar{\mu} \sigma)^\prime (\bar{\Sigma} \sigma^2)^{-1} (\mu - \bar{\mu} \sigma) \right\} \]  \hspace{1cm} (4.1)

The prior measure for $(\mu, \sigma)$ is taken as
\[ \bar{e}(\mu, \sigma) \propto \sigma^{-1} g_0(\mu). \quad \quad \hspace{1cm} (4.2) \]

The posterior joint density of $(\mu, \sigma)$
\text{v.r.t.} \bar{e}(\mu, \sigma) is
\[ p(\mu, \sigma \mid x) \propto 1(x \mid \mu, \sigma) \bar{e}(\mu, \sigma) \quad \quad \hspace{1cm} (4.3) \]
where $1(x \mid \mu, \sigma)$ is the likelihood of the sample.
Let $\mathbf{y} = (y_1, y_2, \ldots, y_p)'$ where $y_i = \sqrt{n} \mu_i / \sigma$. The posterior density of $\mathbf{y}$ is given by

$$p(\mathbf{y} | \mathbf{x}) = \begin{array}{l} \frac{1}{(2\pi)^{-p/2}} \int \sigma^{-1} \exp \left[ - \frac{1}{2} \mathbf{y}' \mathbf{y} + (\mathbf{y} - \mathbf{\mu})' \mathbf{R}^{-1} (\mathbf{y} - \mathbf{\mu}) \right] d\sigma \end{array} (4.4)$$

where $\gamma(\mathbf{y} | \mathbf{\mu}, \mathbf{R})$ is the function defining the $\mathbf{p}$-variate normal density for $\mathbf{y}$ with mean vector $\mathbf{\mu} = \sqrt{n} \mathbf{\mu}^*$ and variance-covariance matrix $\mathbf{R} = \mathbf{n} \mathbf{R}^*$.

Let us define the vector $\mathbf{t} = (t_1, t_2, \ldots, t_p)'$ where

$$t_1 = \sqrt{n} \bar{x}_1 / s$$

and

$$\bar{x}_1 = \frac{1}{n} \sum x_{1j} / n \quad \bar{x} = (\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_p)'$$

$$s^2 = \frac{1}{n} \sum (x_{1j} - \bar{x})^2, \quad \mathbf{R}^{-1}(x_{1j} - \bar{x}) / v, \quad v = p(n-1).$$

It is easily verified that

$$\int \sigma^{-1} \exp \left[ - \frac{1}{2} \mathbf{y}' \mathbf{y} + (\mathbf{y} - \mathbf{\mu})' \mathbf{R}^{-1} (\mathbf{y} - \mathbf{\mu}) \right] d\sigma = f'_v(\mathbf{t}, \mathbf{z}) \quad \cdots \quad (4.7)$$

where

$$f'_v(\mathbf{t}, \mathbf{z}) = \int u^{p-1} \exp \left[ - \frac{1}{2} \mathbf{t}' \mathbf{R}^{-1} \mathbf{t} + (\mathbf{t} - \mathbf{z})' \mathbf{R}^{-1} (\mathbf{t} - \mathbf{z}) \right] du \quad \cdots \quad (4.8)$$

We can therefore write
\[ p(\tau|t) = p(\tau|\xi) \]
\[ = \frac{\xi}{\nu} (\xi|\xi) \frac{\xi}{\nu} (\xi|\xi, \xi) \]
\[ = \int_{0}^{\infty} u^{v+p-1} \exp \left[ -\frac{1}{2} \left( u\nu + (u\xi - \tau) R^{-1}(u\xi - \tau) \right) \right. \]
\[ \left. + (\xi - \xi') (\xi - \xi') \right] du \quad \ldots \quad (4.9) \]
\[ = \exp \left[ -\frac{1}{2} \left( \xi (R^{-1} + C^{-1}) \xi - 2 \xi' C^{-1} \xi' - x' x' \right) \right] \]
\[ I_{v+p-1}(-x') \]

where
\[ x' = (t'R^{-1} t') / (v + t' R^{-1} t')^{1/2} \quad \ldots \quad (4.10) \]

On dividing by the normalising constant we obtain the posterior density of \( \xi \),

\[ p(\tau|t) (2\pi)^{-p/2} | (R^{-1} + C^{-1})|^{1/2} \left( \frac{v'}{v + t' R^{-1} t'} \right)^{(v+p)/2} \]
\[ \exp \left[ -\frac{1}{2} \left( \xi (R^{-1} + C^{-1}) \xi - 2 \xi' C^{-1} \xi' \right. \right. \]
\[ \left. + (\xi - \xi') (\xi - \xi') \right] \]
\[ \{ I_{v+p-1}(-x') \} / \{ I_{v+p-1}(-z') \} \quad \ldots \quad (4.11) \]

where
\[ z' = \left( t'R^{-1} (R^{-1} + C^{-1})^{-1} C^{-1} \xi \right) / (v')^{1/2} \quad (4.12) \]

and
\[ v' = \left( v + t' (R^{-1} - R^{-1} (R^{-1} + C^{-1})^{-1} R^{-1}) \right) t \quad (4.13) \]
The mean vector and the variance-covariance matrix of the distribution are given by

\[ E_1(\xi) = (R^{-1} + \zeta^{-1})^{-1} \left( k_5 R^{-1} \xi + \zeta^{-1} \zeta \right) \]  

(4.14)

and

\[ E(\xi' \xi') - E(\xi) E(\xi') = (R^{-1} + \zeta^{-1})^{-1} \left( \frac{\partial^2}{\partial \xi \partial \xi'} \right) \]

\[ + (k_5 - k_5^2) (R^{-1} \xi R^{-1} \xi) (R^{-1} + \zeta^{-1})^{-1} \]  

(4.15)

where

\[ k_5 = (v')^{-1} (v+p) I_{v+p+1}(-z') / I_{v+p-1}(-z') \]  

(4.16)

and

\[ k_6 = (v')^{-1} (v+p)(v+p+1) I_{v+p+1}(-z') / I_{v+p-1}(-z') \]  

(4.17)

The posterior marginal distribution of \( \xi \)
is given by

\[ p(\xi_1|t) = \frac{1}{\sqrt{2\pi c_{11}1}} \left\{ \frac{v'}{v_1} \right\}^{(v+p)/2} \exp \left[ -\frac{1}{2} c_{11} (\xi_1 - l_{(1)})^2 \right] \]

\[ \frac{I_{v+p-1}(-x_1)}{I_{v+p-1}(-z')} \]  

(4.18)

where

\[ l_{(1)} \] is the first row in \( (R^{-1} + \zeta^{-1})^{-1} \)

\[ c_{11} \] is the \((1,1)\) element in \( (R^{-1} + \zeta^{-1})^{-1} \)
\[
x_1 = \frac{\left\{ v + t' \left( R^{-1} - R^{-1} g R^{-1} \right) t + \frac{(1(1) R^{-1} t)^2}{c_{11}} \right\}^{1/2}}{c_{11}}
\]
\[
v_1 = \left\{ v + t' \left( R^{-1} - R^{-1} g R^{-1} \right) t + \frac{(1(1) R^{-1} t)^2}{c_{11}} \right\}
\]

4.2 A singular prior distribution

Given the sample \( x \) of size \( n \) from \( N(\mu, R \sigma^2) \) as above, we shall suppose it is known that, a priori, \( \mu \) has a singular normal distribution and that

\[
\mu = A \lambda
\]

where \( A \) is a \( p \times r \) matrix of rank \( r \) (\( r \leq p \)) and \( \lambda \) has a nonsingular \( r \)-variate normal distribution with mean \( \bar{\sigma}^* \sigma \) and variance-covariance matrix \( \Sigma^{* \sigma^2} \) (\( A, \bar{\sigma}^*, \Sigma^{* \sigma^2} \) are known).

A prior measure for \((\lambda, \sigma)\) is taken as

\[
g(\lambda, \sigma) \propto \sigma^{-1} \left[ I R^{* \sigma^2} \right]^{-1/2} \exp \left\{ -\frac{1}{2} \left( \lambda - \bar{\sigma}^* \sigma \right)' \left( \Sigma^{* \sigma^2} \right)^{-1} \left( \lambda - \bar{\sigma}^* \sigma \right) \right\}
\]

\]
Let \( \tilde{\nu}^{(1)} = (\bar{X}/\sigma)_{\lambda} \). We shall put
\( \theta = \sqrt{n} \theta^* \) and \( \bar{T} = n \bar{T}^* \). The posterior density of \( \tilde{\nu}^{(1)} \) is obtained as

\[
p(\tilde{\nu}^{(1)} | \hat{t}) = \frac{(2\pi)^{-p/2}}{(\lambda_{1} R^{-1}_{1} + T^{-1})^{1/2}} \left( \frac{v''}{v + \hat{t}'_{1} R^{-1}_{1} \hat{t}} \right) \exp \left\{ -\frac{1}{2} \left[ (\lambda_{1} R^{-1}_{1} + T^{-1}) \tilde{\nu}^{(1)} - 2 \tilde{\nu}^{(1)} \theta^{-1} \tilde{\nu}^{(1)} \theta' \right. \right.
\]
\[
\left. + (T^{-1} \theta)' (A_{1} R^{-1}_{1} + \theta^{-1})^{-1} (T^{-1} \theta) - x''_{p} + z''_{p} \right\} \]
\[
I_{V_{p-1}} (-x'') / I_{V_{p-1}} (-z'') \quad \ldots \quad (4.22)
\]

where \( t = (t_{1}, \ldots, t_{p})' \) is as defined in (4.5)

\[
x'' = (\hat{t}'_{1} R^{-1}_{1} \tilde{\nu}^{(1)}) / (v + \hat{t}'_{1} R^{-1}_{1} \hat{t})^{1/2} \quad \ldots \quad (4.23)
\]
\[
z'' = (A_{1} R^{-1}_{1})' (A_{1} R^{-1}_{1} + T^{-1})^{-1} T^{-1} \theta / (v'')^{1/2} \quad \ldots \quad (4.24)
\]
\[
v'' = v + \hat{t}'_{1} (R^{-1}_{1} - R^{-1}_{1} A_{1} R^{-1}_{1} + \theta^{-1})^{-1} A_{1} R^{-1}_{1} \hat{t} \quad \ldots \quad (4.25)
\]

The mean vector and the variance-covariance matrix of \( p(\tilde{\nu}^{(1)} | \hat{t}) \) are

\[
E_{1}(\tilde{\nu}^{(1)}) = (A_{1} R^{-1}_{1} + T^{-1})^{-1} (R_{1} R^{-1}_{1} \hat{t} + T^{-1} \theta) \quad \ldots \quad (4.26)
\]
and

\[ \Sigma_1(\xi_1, \xi_1') = \Sigma_1(\xi_1) K_1(\xi_1') = \left( A_1 R^{-1}_1 + T^{-1}_1 \right)^{-1} \]

\[ + \left( k_8 - k_7 \right) (A_2 R^{-1}_2 + t_2 R^{-1}_2) \left( A_2 R^{-1}_2 + T^{-1}_2 \right)^{-1} \] (3.27)

The quantities \( k_7 \) and \( k_8 \) are obtained from \( k_5 \) and \( k_6 \) respectively by replacing \( z' \) by \( z'' \) and \( \sqrt{ } \) by \( \sqrt{ } \).

5. **Remarks**

It is shown in sections 2 and 3 that with respect to the prior distribution considered, the student t-variables are marginally sufficient for the differences in means expressed in standard deviation units. It may be noted that due to the adopting of the uniform prior distribution for the location parameter \( \xi_{1g} \), integration over this parameter gives the posterior marginal distribution independent of \( \xi_{1g} \). Further the uniform prior for log \( \sigma \) effects a reduction to only the t-statistics, these statistics being invariant to transformations \( x_{1j} \sim \sigma x_{1j} + b \) \( e \neq 0 \).
References


2 N.G. Bulmer, "Confidence intervals for distance in analysis of variance", Biometrika 41 (1953), 360-370.


