$$A_i = (2\beta_i)^{-1} I_{m_i}, \quad i = 1, 2, \ldots, c.$$

This density is called a multiple τ or poly-t density and is discussed at length by Dreze (1977). It arises in many cases of regression analysis and Dreze discusses the various situations it will occur in in econometrics. Except in the trivial case, when c = 1, the usual t density, the poly-t distribution is difficult to work with because the normalizing constant is unknown.

For an exact Bayesian analysis of the variance components, one needs their marginal distributions. Consider the following.

Theorem 4. 4. The marginal posterior density of τ and ρ is

$$p(\tau, \rho|y) \propto A(\tau, \rho)|\beta(\tau, \rho)|^{-1/2} \exp - \tfrac{1}{2}\hat{b}'\beta_1(\rho) \times [\beta^{-1}(\tau, \rho) - \beta_1^{-1}(\rho)]\beta_1(\rho)\hat{b}, \tag{4.8}$$

where $\tau > 0$ and $\tau_i > 0$, $i = 1, 2, \ldots, c$.

$$A(\tau, \rho) = \tau^{(n-p+2\alpha)/2-1} \exp - \tfrac{\tau}{2}(2\beta + y'Ry - \hat{b}'u'Ru\hat{b}) \times \prod_1^c \tau_i^{(m_i+2\alpha_i)/2-1} \exp - \tau_i\beta_i$$

$$\beta(\tau, \rho) = \beta_1(\rho) + A(\rho),$$

where

$$\beta_1(\rho) = \tau u'Ru.$$

The above joint posterior density of τ and ρ is a very complicated expression and is analytically intractable. For instance it seems impossible to integrate it with respect to the precision components.

Another way of finding the marginal densities of the variance components is to combine the marginal density of b with the conditional densities of the variance components given b, then integrate with respect to b, but this will require an approximation to the density of b.

APPROXIMATIONS

The posterior distribution of b is a key factor in obtaining the conditional distribution of the variance components as well as determining the posterior marginal distribution of these parameters. The distribution of b is very difficult to handle, thus one way of solving the problem is to find an approximation in terms of simple expressions.

Since a multivariate t distribution may be approximated by a multivariate normal with the same first two moments as the multivariate t distribution, $[1 + (1/k)(X - \mu)'A(X - \mu)]^{-(n+k)/2}$ may approximated by $\exp - 1/2[(k - 2)/k](X - \mu)'A(X - \mu)$ for any $X \in R^n$ and any non-negative definite matrix A.

Using this approximation with each of the c factors of (4.7), one may show the density of b may be approximated by

$$p(b|y) \propto \exp - \tfrac{1}{2}(b - b^*)'A^*(b - b^*), \quad b \in R^m, \tag{4.9}$$

where

$$A^* = A_1 + A_2,$$

$$b^* = (A^*)^{-1} A_1 \hat{b},$$

$$A_1 = (n\text{-}p\text{+}2\alpha\text{-}m\text{-}2)u'Ru[2\beta\text{+}y'Ry\text{-}\hat{b}'u'Rub],$$

and

$$A_2 = \{\mathrm{DIAG}[\alpha_i - 1)/\beta_i] I_{m_i}\}$$

The matrix A_2 is m × m block diagonal with i-th diagonal matrix $[(\alpha_i - 1)/\beta_i]I_{m_i}$, where i = 1, 2, …,c.

Thus the posterior distribution of b is approximately normal with mean vector b* and precision matrix A*, however we would expect the approximation not to be good unless $\hat{b}$ is close to the zero vector. See Box and Tiao (1973), Chapter 9, where they discuss the poly-t distribution. Later, we will discuss how good the approximation is, using a one-way random model.

The mean of b, namely b*, is the approximate mean of the distribution of b and can be used as the conditioning value of b. For example, if conditional point estimates for the parameters are needed, the mean of these conditional posterior distributions of Theorem 4.2 can be taken as estimates. They are

$$\hat{\theta} = (X'X)^{-1}X'(y - ub^*)$$

for the fixed effects,

$$\hat{\sigma}^2 = (n + 2\alpha - p - 2)^{-1}(2\beta + y'Ry - \hat{b}'u'Ru\hat{b})$$

for the error variance, and

for the error variance, and

$$\hat{\sigma}_i^2 = (2\beta_i + b_i{*}'b_i{*})/(m_i + 2\alpha_i - 2),$$

for the variance components. Note, $b^* = (b_1^{*\prime}, b_2^{*\prime}, \ldots, b_c^{*\prime})'$ is the partitioned form of the approximate mean of the posterior distribution of b.

Our objective here is to determine a convenient approximation to the marginal posterior distribution of the variance components. To do this, we need the distribution of quadratic forms in normal variables.

Let Z be a random n-vector having a normal distribution with mean μ and positive definite dispersion matrix D, or symbolically let $Z \sim N(\mu, D)$. Let $Q(Z) = (Z - Z_0)'M(Z - Z_0)$, where Z_0 is a known n vector and M a known non-negative definite matrix, then it is clear the distribution of Q(Z) is the same as that of $(Z^* - Z_1)'M(Z^* - Z_1)$ where $Z_1 = Z_0 - \mu$ and $Z^* \sim N(0, D)$.

Since D is positive definite, there exists a nonsingular lower triangular matrix L such that $D = LL'$. Also, since $L'ML$ is symmetric there exists an orthogonal matrix P such that $P'L'MLP = \Lambda$, the diagonal matrix of eigenvalues of $L'ML$. Using the transformation $Z^* = LPW$, one may show the distribution of Q(Z) is that of

$$\sum_1^n \lambda_i (W_i - w_i)^2,$$

where

$$\begin{aligned}
&\Lambda = \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n),\\
&w' = (w_1, w_2, \ldots, w_n),\\
&W' = (W_1, W_2, \ldots, W_n),\\
&w = P'L^{-1}Z_1, \text{and}\\
&n = \text{number of rows of } M.
\end{aligned}$$

If n' = rank of $m < n$, then assuming the last $n - n'$ of the λ's to be zero, one may show the distribution of Q(Z) is the same as that of

$$Q(W) = \sum_1^n \lambda_i (W_i - w_i)^2.$$

Ruben (1960, 1962) has shown that for given Λ, n, and w, the distribution of Q(W) may be expressed as

$$\begin{aligned}
F_n(q/\Lambda, w) &= \Pr[Q(z) \leqslant q] \\
&= \Pr[\sum_{j=1}^n \lambda_j (w_j - w_j)^2 \leqslant q] \\
&= \sum_{j=0}^\infty e_j \Pr(\chi^2_{n+2j} < q/c),
\end{aligned} \tag{4.10}$$

where

$$\begin{aligned}
&e_0 = \exp\left(-\tfrac{1}{2}\sum_{j=1}^n w_j^2\right)[\prod_{j=1}^n (c/\lambda_j)^{1/2}],\\
&e_r = (1/2r)\sum_{j=0}^{n-1} G_{r-j} e_j \quad (r \geqslant 1),\\
&G_r = \sum_{j=1}^n (1 - c/\lambda_j)^r + rc\sum_{j=1}^n (w_j^2/\lambda_j)(1 - c/\lambda_j)^{r-1} \quad (r \geqslant 1),
\end{aligned}$$

and c is an arbitrary positive constant.

Ruben has also shown that the series on the right of (4.10) is uniformly convergent over any finite interval of q and the bound for error in the above series is the n-th term. Furthermore it is obvious that the distribution in (4.10) can be an infinite mixture of chi-square densities provided c is chosen to be less than the minimum of $(\lambda_1, \lambda_2, \ldots, \lambda_n)$, to insure each $e_i > 0$.

APPROXIMATION TO THE POSTERIOR DISTRIBUTION OF σ^2

From Theorem 4.2, the conditional distribution of τ given b has a gamma distribution with parameters $\alpha^* = (n - p + 2\alpha)/2$ and $\beta^* = (1/2)(2\beta + y'Ry - \hat{b}'u'Ru\hat{b} + Q)$, where $Q = (b - \hat{b})'u'Ru(b - \hat{b})$.

From Ruben's representation, let b^*, $\hat{b}$, A^{*-1} and $u'Ru$, replace μ, Z_0, D, and M respectively, and let s = n = rank of $u'Ru$, then the marginal density of τ is

$$P(\tau|y) = \int^\infty \frac{(\beta^*)\alpha^* \tau^{\alpha^*-1} e^{-\tau\beta^*}}{} \tag{4.11}$$

$$P(\tau|y) = \int_0 \frac{\quad}{\Gamma(\alpha^*)}$$
$$\times \sum_{j=0}^{\infty} e_j [\frac{d}{dq} Pr(\chi^2_{s+2j} < q/c)] dq$$
$$= \sum_{j=0}^{\infty} e_j \int_0^{\infty} [(d+q)/2]^k \tau^{k-1} \exp - \frac{1}{2}[q/e + (d+q)\tau]$$
$$\times (q/c)^{s/2+j-1}(1/c)dq, \quad \tau > 0$$

where $2k = n - p + 2\alpha$, and $d = 2\beta + y'Ry - \hat{b}' u'Ru\hat{b}$, and $c > 0$.

Now let k be an integer. To insure this, one needs to make a slight adjustment in the value of α, but the change in α need not be larger than 0.5. Using this assumption, the density of σ^2 is given by

$$P(\sigma^2 \big| y) = k e^{-d/2\sigma^2} (d/2)^k (1/\sigma^2)^{(k+1)} \tag{4.12}$$
$$\times \sum_{j=0}^{\infty} \frac{e_j}{\Gamma(s/2+j)(1+c/\sigma^2)^{s/2+j}}$$
$$\times \sum_{r=0}^{k} \frac{\Gamma(s/2+r+j)}{\Gamma(r+1)\Gamma(k-r+1)[(d/2)(1/c+1/\sigma^2)]^r}.$$

APPROXIMATION TO THE POSTERIOR DISTRIBUTION OF THE VARIANCE COMPONENTS

Consider the i-th precision component τ_i, then using Ruben's result, replace μ, z_0, D, and M by b_i^*, 0, A_i^{*-1}, and I_{m_i} respectively and let $s_i = m_i$ = rank of I_{m_i}, where A_i^* is the precision matrix of b_i in the normal approximation to b, given by (4.9).

The posterior marginal density of τ_i is

$$p(\tau_i|y) = \frac{1}{\Gamma(\alpha_i^*)} \int_0^{\infty} (\beta_i^*)^{\alpha_i^*} \tau_i^{\alpha_i^*-1} \exp - \beta_i^* \tau_i$$
$$\times \sum_{j=0}^{\infty} e_{ij} [\frac{d}{dq_i} Pr(\chi^2_{s_i+2j}) < q_i/c_i] dq_i,$$

where $\tau_i > 0$, $\alpha_i^* = (m_i + 2\alpha_i)/2$, $\beta_i^* = (2\beta_i + b'_i b_i)/2$, and the e_{ij} are given by (4.10), with $e_{ij} = e_j$, and c_i is an arbitrary positive constant.

One may let $k_i = (m_i + 2\alpha_i)/2$ be an integer with a slight adjustment in α_i, then letting $\sigma_i^2 = \tau_i^{-1}$ the above formula reduces to

$$p(\sigma_i^2|y) = k_i (d_i/2)^{k_i} (1/\sigma_i^2)^{(k_i+1)} \exp - d_i/2\sigma_i^2 \tag{4.13}$$
$$\times \sum_{j=0}^{\infty} \frac{e_{ij}}{\Gamma(s_i/2+j)(1+c_i/\sigma_i^2)^{(s_i/2+j)}}$$
$$\times \sum_{r=0}^{k_i} \frac{\Gamma(s_i/2+r+j)}{\Gamma(r+1)\Gamma(k_i-r+1)[(d_i/2)(1/c_i+1/\sigma_i^2)]}$$

where $\sigma_i^2 > 0$.

Using the normal approximation to b and Ruben's representation of a quadratic form in normal variables, formulas (4.12) and (4.13) are approximate densities for the marginal posterior distribution of the error variance and i-th variance component respectively. Ruben's result is exact, however, the normal distribution of b is an approximation to the true distribution of b, given by (4.9), therefore the adequacy of the approximation to the posterior distribution of the variance components and error variance depends on the adequacy of the normal approximation to b.

Thus, up to now, the exact marginal and conditional posterior distributions of the parameters of the mixed model have been derived. Since the exact marginal density (4.8) of the error variance and variance components is analytically intractable, it is necessary to represent their marginal density in a more convenient form, thus the normal approximation to b and Ruben's representation was employed. These derivations are found in Rajagopalan (1980) and Rajagopalan and Broemeling (1983).

POSTERIOR INFERENCES

Inferences for the parameters of the mixed model are to be based on the appropriate posterior distribution and two approaches will be taken here. First, in regard to estimating the error variance and variance components, a plot of the relevant marginal posterior distribution will be employed, along with the posterior mean and variance of that parameter. Secondly, joint modal estimates of all the parameters will be calculated from an iterative technique.

Marginal Posterior Inferences

Consider making inferences for the error variance σ^2 of a mixed model, then one must plot the posterior marginal distribution of σ^2, either by numerical integration, via the joint posterior density (5.8) of τ and ρ, or via the approximate density (5.12).

Of course, in a similar way, inferences for the variance components are made by plotting the marginal posterior density of that parameter, via numerical integration from (4.8), or by plotting the approximate density (4.13).

Accompanying the plot of each parameter one may compute the posterior mean and variance. Again here one has a choice. Either compute these moments from the exact joint marginal posterior density (4.8), or from the approximate densities, (4.12) and (4.13). In the latter case, one may derive formulas for the mean and variance.

Posterior Means and Variances for Variance Components

The moments of σ^2 and σ_i^2 are difficult to obtain directly from the approximate densities (4.12) and (4.13), but may be obtained relatively easily by first computing the conditional moments of these parameters given b, which are exact, then averaging conditional moments over the distribution of b, using the normal approximation to b, (4.9).

For instance, the moments of σ^2 can be computed by noting that

(i) τ given b has a gamma distribution, from Theorem 4.2, and
(ii) b has an approximate normal distribution.

Using the formulas

(i) $E(\sigma^2) = E[E(\sigma^2|b)]$, and
(ii) $\mathrm{var}(\sigma^2) = \mathrm{Var}[E(\sigma^2|b)] + E[\mathrm{Var}(\sigma^2|b)$,

where E_b denotes expectation with respect to b and Var_b means variance with respect to the distribution of b, one may show the marginal posterior mean of σ^2 is

$$\begin{aligned}E(\sigma^2|y) &= (n-p+2\alpha-2)^{-1}[2\beta + y'Ry - \hat{b}'u'Ru\hat{b}] \\ &\quad + \mathrm{Tr}(u'RuA^{*-1}) + (b^* - \hat{b})'u'Ru(b^* - \hat{b})],\end{aligned} \tag{4.14}$$

and the marginal posterior variance is

$$\begin{aligned}\mathrm{Var}(\sigma^2|y) = {} & (n-p+2\alpha-2)^{-2}(n-p+2\alpha-2)^{-1} \\ & \times [2\beta + y'Ry - \hat{b}'u'Ru\hat{b} + \mathrm{Tr}(u'RuA^{*-1}) \\ & + (b^* - \hat{b})'u'Ru(b^* - \hat{b})]^2 + (n-p+2\alpha-2)^{-1} \\ & \times (n-p+2\alpha-4)^{-1} \\ & \times [2\mathrm{Tr}(u'RuA^{*-1})^2 + 4(b^* - \hat{b})'u'RuA^{*-1} \\ & \times u'Ru(b^* - \hat{b})],\end{aligned} \tag{4.15}$$

where $n - p > 4 - 2\alpha$.

In a similar way, the first two posterior moments of σ_i^2 are

$$E(\sigma_i^2|y) = (m_i + 2\alpha_i - 2)^{-1}[2\beta_i + \mathrm{Tr}(A_i^{*-1}) + b_i^{*\prime}b_i^*] \tag{4.16}$$

and

$$\begin{aligned}\mathrm{Var}(\sigma_i^2|y) = {} & 2(m_i + 2\alpha_i - 2)^{-2}(m_i + 2\alpha_i - 4)^{-1}[2\beta_i + \mathrm{Tr}(A_i^{*-1}) \\ & + b_i^{*\prime}b_i^*]^2 + (m_i + 2\alpha_i - 2)^{-1}(m_i + 2\alpha_i - 4)^{-1} \\ & \times [2\mathrm{Tr}(A_i^{*-1})^2 + 4b_i^{*\prime}A_i^{*-1}b_i^*]\end{aligned} \tag{4.17}$$

if $m_i > 4 - 2\alpha_i$, $i = 1, 2, \ldots, c$.

All the higher moments, provided they exist, of the variance components and error variance may be obtained in a similar fashion.

An Example of Marginal Posterior Analysis

A balanced one-way random model is considered to illustrate the preceding results. The linear model describing such a design is

$$Y_{ij} = \mu$$