## Additive Models

Let us first consider an additive model,

$$y_{ijk} = m + a_i + b_j + e_{ijk} \tag{3.109}$$

where i = 1, 2, …, t, j = 1, 2, …, b, k = 1, 2, …, r, and $e_{ijk} \sim n(0, \tau^{-1})$, $\tau > 0$, $a_i \in R$, $b_j \in R$, and $m \in R$. If $\theta = (\theta_0, \theta'_1, \theta'_2)'$ where $^{\theta}0 = m$,

$$\theta_1 = \begin{matrix} a_1 \\ a_2 \\ \vdots \\ a_t \end{matrix},$$

and

$$\theta_2 = \begin{matrix} b_1 \\ b_2 \\ \vdots \\ b_b \end{matrix},$$

then the model is written as

$$y = x\theta + e, \tag{3.110}$$

where x is a n × p (p = 1 + b + t) design matrix, $y = (y_{111}, y_{112}, \ldots, y_{11n_{11}}, \ldots, y_{tb1}, \ldots, y_{tbn_{tb}})'$, and e is the n × 1 vectors of error terms $e_{ijk}$, thus $\theta_1$ is the t × 1 vector of the effects of the first factor, and $\theta_2$ the b × 1 vector of effects of the second factor. The design matrix x is such that the first $n_{11}$ rows consist of the (1 + t + b) vector (1; 1, 0, …, 0; 1, 0, …, 0), the next $n_{12}$ rows consist of the vector (1; 1, 0, …, 0; 0, 1, 0, …, 0), and so on, hence x is of less than full rank p, namely of rank p − 2.

Since x is less than full rank, if one combines the likelihood function of θ and τ induced by (3.110) with Jeffreys' improper prior density

$$\xi(\theta, \tau) \propto 1/\tau, \quad \tau > 0, \quad \theta \in R^p \tag{3.111}$$

the posterior density of the parameters will not be proper. Of course this difficulty can be avoided by employing a normal-gamma prior density in lieu of (3.111); however, some may object to this alternative because it may be very difficult to assign a proper prior distribution.

Fortunately, one may transform a less than full-rank model (3.110) to full rank

$$y = z\alpha + e, \tag{3.112}$$

where z is n × k (k < p) of full rank, α is k × 1, namely,

$$\alpha = u\theta$$

where u is a k × p known matrix, and as before, $e \sim n(0, \tau^{-1} I_n)$

According to Graybill (1961, pages 235–236), z and u are constructed as follows. Since x'x is a p × p and symmetric positive semi-definite matrix, there exists a non-singular w* (p × p) matrix such that

$$(w*)'x'xw* = \begin{pmatrix} B & 0 \\ 0 & 0 \end{pmatrix}$$

where B is k × k of rank k. If w* = (w, $w_1$), where w is p × k and $(w*)^{-1} = u* = (u', u'_1)'$, where u is k × p, then

$$z = xw$$

and

$$\alpha = u\theta.$$

Consider the two-way additive model (3.109) with p = b + t + 1 parameters where x is of rank k = p − 2 = b + t − 1, then obviously, one may use the above reparametrization to arrive at a full-rank representation. How should one choose α, u, and w? Since the primary purpose of the analysis is to investigate the effect of the two factors on the average response $m + a_i + b_j$ one way to choose α is to let

$$\begin{matrix} m + a_1 + b_1 \\ a_1 - a_2 \\ a_2 - a_3 \end{matrix} \tag{3.113}$$

$$\alpha = \begin{pmatrix} a_2 - a_3 \\ \vdots \\ a_t - a_{t-1} \\ b_1 - b_2 \\ b_2 - b_3 \\ \vdots \\ {}^b b - b_{t-1} \end{pmatrix} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix},$$

where $\alpha_1 = m + a_1 + b_1$. This means u must be

$$u = \begin{pmatrix} 1; & 1, & 0, & & \ldots, & 0; & 1, & 0, & \ldots, & 0 \\ 0; & 1, & -1, & 0, & \ldots, & 0; & 0, & 0, & \ldots, & 0 \\ 0; & 0, & 1, & -1, & \ldots, & 0; & 0, & 0, & \ldots, & 0 \\ & & & & \vdots & & & & & \\ 0; & 0, & \ldots, & & 1; & -1; & 0, & 0, & \ldots, & 0 \\ 0; & 0, & \ldots, & & 0; & 1, & -1, & 0, & \ldots, & 0 \\ & & & & \vdots & & & & & \\ 0; & 0, & \ldots, & & 0; & 0, & 0, & \ldots, & 1, & -1 \end{pmatrix}, \tag{3.114}$$

which is $k \times p$. Thus let

$$u^* = \begin{pmatrix} u \\ u_1 \end{pmatrix}$$

where $u_1$ is $(p - k) \times p$ so that $u^*$ is of full rank. Now let $w^* = (u^*)^{-1}$, where $w^* = (w, w_1)$, then $z = xw$, and the reparametrization is complete. The likelihood function for $\alpha$ and $\tau$ is

$$L(\alpha, \tau|y) \propto \tau^{rbt/2} \exp - \tfrac{\tau}{2}(y - z\alpha)'(y - z\alpha), \quad \alpha \in R^k, \quad \tau > 0,$$

where $\alpha$ is of order $b + t - 1$ and when combined with the Jeffreys' improper prior density

$$\xi(\alpha, \tau) \propto 1/\tau, \quad \tau > 0, \quad \alpha \in R^k,$$

gives

$$\xi(\alpha,\tau|y) \propto \tau^{rbt/2-1} \exp - \tfrac{\tau}{2}[(\alpha - \widehat{\alpha})'z'z(\alpha - \alpha - \widehat{\alpha}) + (y - 2\widehat{\alpha})'(y - z\widehat{\alpha})], \tag{3.115}$$

for the joint posterior distribution of $\alpha$ and $\tau$, where $\widehat{\alpha} = (z'z)^{-1} z'y$, $\alpha \in R^k$, and $\tau > 0$, and the usual Bayesian analysis follows. It can be shown that $\widehat{\alpha}$ is such that its first component is $\overline{y}_{\ldots} + (\overline{y}_{1..} - \overline{y}_{\ldots}) + (\overline{y}_{.2.} - \overline{y}_{\ldots})$, the next $a - 1$ components are $\overline{y}_{i..} - \overline{y}_{i+1..}$, $i=1,2,\ldots, a - 1$, and the last $b - 1$ components are $\overline{y}_{.j.} - \overline{y}_{.j+1.}$, where $j = 1, 2, \ldots, b - 1$, where

$$\overline{y}_{\ldots} = \sum_i \sum_j \sum_k y_{ijk}/btr$$

$$\overline{y}_{i..} = \sum_i \sum_k y_{ijk}/rb$$

and

$$\overline{y}_{.j.} = \sum_i \sum_k y_{ijk}/rt.$$

Suppose one wants to make inferences about the first factor $(a_1, a_2, \ldots, a_t)$, then one would want to know the marginal posterior distribution of $\alpha_2 = (a_1 - a_2, a_2 - a_3, \ldots, a_{t-1} - a_t)$, which is known once one knows the posterior distribution of $\alpha$, which, from (3.115), is a t-distribution with $rbt - (t + b + 1)$ degrees of freedom, location

$$E(\alpha|y) = \widehat{\alpha}, \tag{3.116}$$

and precision

$$P(a|y) = \frac{[rbt-(t+b+1)]z'z}{(y-z\widehat{\alpha})'(y-z\widehat{\alpha})} \tag{3.117}$$

$(y-z\hat{\alpha})'(y-z\hat{\alpha})$

thus $\alpha_2$ also has a t-distribution with rbt – (t + b+1) degrees of freedom, location

$$E(\alpha_2|y) = (\phi_1, I_{t-1}, \phi_2)\hat{\alpha}, \tag{3.118}$$

where $\phi_1$ is $(t-1) \times 1$ matrix of zeros, $1_{t-1}$ is the identity matrix of order $t-1$, and $\phi_2$ is a zero matrix of order $(t-1) \times (b-1)$. The precision matrix of $\alpha_2$ is

$$P(\alpha_2|y) = [(\phi_1, I_{t-1}, \phi_2)P^{-1}(\alpha|y)(\phi_1, I_{t-1}, \phi_2)']^{-1}. \tag{3.119}$$

As for $\alpha_3 = (b_1 - b_2, b_2 - b_3, \ldots, b_b - b_{b-1})'$, the other factor in the model, its posterior distribution is also a t with rbt – (t + b + 1) degrees of freedom, location vector

$$E(\alpha_3|y) = (\phi_1^*, \phi_2^*, I_{b-1})E(\alpha_2|y), \tag{3.120}$$

where $\phi^*_1$ is a $(b-1) \times 1$ zero vector, and $\phi^*_2$ is a $(b-1) \times (t-1)$ matrix of zeros, thus the precision matrix of the posterior distribution of $\alpha_3$ is

$$P(\alpha_3|y) = [(\phi_1^*, \phi_2^*, I_{b-1})P^{-1}(\alpha|y)(\phi_1^*, \phi_2^*, I_{b-1})']^{-1}. \tag{3.121}$$

Of course, the marginal posterior distribution of $\tau$ is gamma with parameters $[rbt - (t + b + 1)]/2$ and $(y - z\hat{\alpha})'(y - z\hat{\alpha})/2$.

In problems like this, one is usually interested in the effect of the levels of the two factors on the average response and in particular if all the levels of one factor have the same effect. For example, do all the t levels of the first factor have the same effect, that is, is $H_0$: $a_1 = a_2 = \ldots = a_t$, which is true whenever $\alpha_2 = 0$? The usual approach is to do an analysis of variance and the Bayesian approach is to find an HPD region for $\alpha_2$.

Since the random variable

$$F(\alpha_2|y) = (t-1)^{-1}(\alpha_2 - E(\alpha_2|y)'P(\alpha_2|y)\left(\alpha_2 - \hat{\alpha}\right) \tag{3.122}$$

has an F distribution with $(t-1)$ and rbt – (t + b + 1) degrees of freedom, a $1 - \Delta$ $(0 < \Delta < 1)$ HPD region for $\alpha_2$ is

$$HPD_\Delta(\alpha_2) = \{\alpha_2 : F(\alpha_2|y) \leqslant F_{\Delta;t-1,rdt-t-b-1}\} \tag{3.123}$$

and

$$H_0 : a_1 = a_2 = \ldots = a_t$$

is rejected when $\alpha_2 = 0[(t-1) \times 1]$ is not contained in the region or equivalently whenever $F(\theta\,|\,y) > F_{\Delta t-1,\, rbt-t-b-1}$

One may test the hypothesis

$$H_0^* : b_1 = b_2 = \ldots = b_b(\alpha_3 = 0)$$

in a similar way by constructing an HPD region for $\alpha_3$.

## The Randomized Block Design

The two-way model (3.109) may be used to analyze the experimental results of a randomized block design. In such a layout, the n experimental units are arranged into b blocks such that each block has t units, where t treatments are assigned at random to the t units of each block assuming no interaction between the treatments and experimental material, a randomized block design is modeled by (3.109), where r = 1, and the levels of the first factor $a_1$, $a_2$, …, $a_t$ are called treatment effects, while the block effects are $b_1$, $b_2$, …, $b_b$ the levels of the second factor. The model is

$$Y_{ij} = m + a_i + b_j + e_{ij} \tag{3.124}$$

where the $e_{ij}$'s are n.i.d. $(0, \tau^{-1})$, $m \in R$, $a_i \in R$, and $b_j \in R$.

Obviously, one may use the results of the preceding section to analyze a randomized block design. First, one would reparameterize the model to full rank and test for equality of treatment effects by using the HPD region for $\alpha_2$ given by (3.123).

## Experiments with Interaction

The two-way classification model with interaction is given by

$$y_{ijk} = m + a_i + b_j + (ab)_{ij} + e_{ijk}, \tag{3.125}$$

where i = 1, 2, …, t, j = 1, 2, …, b, and k = 1, 2, …, r. Also, $y_{ijk}$ is the k-th observation when the i-th level of the first factor and the j-th level of the second are present. The $a_i$, i = 1, 2, …, t, are the effects of the first factor while the effects of the second are the $b_j$, j = 1, 2, …, b, and $(ab)_{ij}$ is called the second are present. The $a_i$, $i = 1, 2, \ldots, t$, are the effects of the first factor while the effects of the second are the $b_j$, $j = 1, 2, \ldots, b$, and $(ab)_{ij}$ is called the interaction between the i-th and the j-th levels of the two factors. As before, assume the $e_{ijk}$,'s are n.i.d. $(0,\tau^{-1})$, $\tau > 0$, then one may express (3.125) as $y = x\theta + e$, where y is the n(=rbt) × 1 vector of observations, x is the $n \times p(= bt + b + t + 1)$ design matrix, θ is the p × 1 vector

$$\theta = \begin{pmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \\ \theta_4 \end{pmatrix}, \tag{3.126}$$

where = m, $\theta_2$ is t × 1, namely

$$\theta_2 = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_t \end{pmatrix}, \tag{3.127}$$

$\theta_3 = (b_1, b_2, \ldots, b_b)'$ and $\theta_4$ is tb × 1, namely

$$\theta_4 = \begin{pmatrix} (ab)_{11} \\ \vdots \\ (ab)_{tb} \end{pmatrix}. \tag{3.128}$$

Finally, e is a n × 1 normal random vector distributed as $n(0, \tau^{-1} I_n)$ and $e = (e_{111}, e_{112}, \ldots, e_{tbr})'$

Knowing θ allows one to construct the design matrix x which is n × p and of rank k* = bt. Thus since x is of less than full rank, the model y = x θ + e cannot be analyzed with Jeffreys' improper prior density. But, one may transform to a full rank model

$$y = z\alpha + e, \tag{3.129}$$

where α is k* × 1 and z is n × k* and is of full rank.

One way to choose α is to let α be (bt) × 1, where

$$\alpha = \begin{pmatrix} m + a_1 + b_1 + (ab)_{11} \\ m + a_1 + b_2 + (ab)_{12} \\ \vdots \\ m + a_t + b_b + (ab)_{tb} \end{pmatrix} = u\theta, \tag{3.130}$$

thus u is bt × (1 + t + b + bt) and known, since one knows θ from (3.126). Now let $u^* = (u', u'_1)'$, where $u_1$ is any matrix such that u* is non-singular. Now let $w^* = (u^*)^{-1}$, where $w^* = (w, w_1)$ and w is p × k*, then z = xw and the full rank representation (3.129) of the original model is complete.

One now analyzes this model in much the same way the additive model was analyzed in the previous section. We see from (3.129) that y is normally distributed with mean of zα and precision matrix $\tau^{-1} I_n$, thus if α and τ have a Jeffreys' prior density

$$\xi(\alpha, \tau) \propto 1/\tau, \quad \tau > 0, \quad \alpha \in R^{k^*}$$

the joint posterior density of α and τ is

$$\xi(\alpha, \tau | y) \propto \tau^{rbt/2-1} \exp - \tfrac{\tau}{2}[(\alpha - \hat{a})' z' z \left(\alpha - \hat{\alpha}\right) \\ +(y - z\alpha')' \left(y - z\hat{\alpha}\right)], \tag{3.131}$$

where $\alpha \in r^{k^*}$, τ > 0, and $\hat{\alpha} = (z'z)^{-1} z' y$. The marginal posterior density of α is

$$\xi(\alpha|y) \propto [\left(y - z\hat{\alpha}\right)' \left(y - z\hat{\alpha}\right) + \left(\alpha - \hat{\alpha}\right)' z' z \left(\alpha - \hat{\alpha}\right)]^{rbt/2}, \\ \alpha \in R^{bt} \tag{3.132}$$

hence α has a t distribution with rbt – bt degrees of freedom, location vector

$$E(\alpha|y) = \hat{\alpha}, \tag{3.133}$$

and precision matrix

$$P(\alpha|y) = \frac{bt(r-1)z'z}{\left(y - z\hat{\alpha}\right)' \left(y - z\hat{\alpha}\right)}. \tag{3.134}$$

$(y-z\alpha)\ (y-z\alpha)$

It is interesting to observe that if $r = 1$, there are zero degrees of freedom and the marginal posterior distribution is improper, which is also a problem when one analyzes the experiment the conventional way, because then the error sum of squares is zero and one is unable to estimate $\tau^{-1}$, the error variance. One way for the Bayesian to avoid this is to use a normal-gamma prior density for α and τ, but it perhaps would be a difficult task to assign values to the hyperparameters. Let us, for the time being, assume $r > 1$ and avoid the difficulty.

How does one analyze such a model? Remember, one is interested in the way the levels of the two factors effect the average response

$$S_{ij} = m + a_i + b_j + (ab)_{ij},$$

for $i = 1, 2, \ldots, a$ and $j = 1, 2, \ldots, b$, and the model is said to have no interaction if $(S_{ij} - S_{i'j}) - (S_{ij'} - S_{i'j'}) = 0$ for all i, i’, j and j’, which is equivalent to saying $(ab)_{ij} - (ab)_{i'j\cdot} - (ab)_{ij'} + (ab)_{i'j'} = 0$ for all i, i’, j, and j’, otherwise the model is said to be nonadditive or is a model with interaction. In the analysis of such models, one first checks for the presence of interaction and if there is none, one examines the levels $a_1, a_2, \ldots, a_t$ of the first factor and then the b levels of the second factor.

Thus, the first task is to develop a test for no interaction, namely, test the hypothesis

$$H_0 : (ab)^*_{ij} = (ab)_{ij} - (\overline{ab})_{i\cdot} - (\overline{ab})_{\cdot j} + (\overline{ab})_{\cdot\cdot}, \quad \text{and}$$

$$(\overline{ab})_{i\cdot} = \sum_{j=1}^{b} (ab)_{ij}/b, \quad (\overline{ab})_{\cdot j} = \sum_{i=1}^{t} (ab)_{ij}/t, \quad \text{and}$$

$$(\overline{ab})_{\cdot\cdot} = \sum_{1}^{t}\sum_{1}^{b} (ab)_{ij}/tb.$$

One may show there is no interaction if and only if $(ab)^*_{ij} = 0$ for all i and j.

For example, if $t = b = 2$, $H_0$ implies $T\alpha = 0$ $(1 \times 1)$, where T is the matrix $T = (1, -1, -1, 1)'$, thus to test for no interaction, one must find an HPD region for $T\alpha$.

In general, to test for no interaction, one may find a $(t - 1).\ (b - 1) \times bt$ matrix T, such that no interaction $H_0$