$_0$ is equivalent to

$$H_0 : \gamma = T\alpha = 0[(t-1)(b-1) \times 1]. \tag{3.135}$$

Since $\alpha|\gamma \sim t[rbt - bt, E(\alpha|y), P(\alpha|y)]$, where the location and precision are given by (3.133) and (3.134), respectively, $\gamma = T\alpha$ also has a t distribution with $bt(r-1)$ degrees of freedom, location vector

$$E(\gamma|y) = TE(\alpha|y) \tag{3.136}$$

and precision matrix

$$P(\gamma|y) = [TP^{-1}(\alpha|y)T']^{-1}, \tag{3.137}$$

and a $1 - \Delta (0 < \Delta < 1)$ HPD region for γ is

$$HPD_\Delta(\gamma) = \{F(\gamma|y) \leqslant F_{\Delta,(t-1)(b-1),(r-1)bt}\} \tag{3.138}$$

where

$$F(\gamma|y) = [\gamma - E(\gamma|y)]'P(\gamma|y)[\gamma - E(\gamma|y)](t-1)^{-1}(b-1)^{-1}. \tag{3.139}$$

If no interaction is indicated, that is, if $F(\theta|\gamma) > F_{\Delta,\,(t-1)(b-1),\,bt(r-1)}$, one may analyze the additive model

$$y_{ijk} = m + a_i + b_j + e_{ijk}$$

according to the methods of the previous section.

When $r = 1$, there is no Bayesian way to test for interaction, because Jeffreys' density was used to express prior information; however, one may develop a complete Bayesian analysis if one may express prior information by a proper prior density, such as the normal-gamma density for α and τ.

The above model can be used to analyze several experimental layouts. For example, a two-factor factorial arrangement of treatments in a completely randomized design could be examined by the methods presented in this section.

THE ANALYSIS OF COVARIANCE

Analysis of covariance models combine the regression and the models for designed experiments into one model. For example, consider a completely randomized design consisting of t treatments, where the i-th treatment is assigned to n_i experimental units, and the response y_{ij} and a regressor or covariable x_{ij} are measured on the j-th unit receiving the i-th treatment. A model for such a situation is given by

$$Y_{ij} = a_i + x_{ij}\gamma + e_{ij}, \tag{3.140}$$

where $a_i \in R$, $x_{ij} \in R$, $\gamma \in R$, and the e_{ij} are n.i.d. $(0,\tau^{-1})$, where $\tau > 0$ and $i = 1, 2, \ldots, t$ and $j = 1, 2, \ldots, n_i$. The unknown parameters are the treatment effects $a_1, a_2 \ldots, a_t$, γ, and τ, and one is usually interested in the effect of treatments on the average responses $E(y_{ij}) = a_i + x_{ij}\gamma$.

The covariable or concomitant variable x is introduced to increase the precision estimating the treatment effects. The covariable should be highly related to the main response y and is not to be affected by the treatments.

Our objective is to examine the effect of the treatments on the average response after taking into account or "adjusting" for the effect of the covariable. One is referred to Cox (1958) for a discussion of experimental design considerations when one uses concomitant observations and to Graybill (1961) for the conventional analysis of variance of such designs.

From a Bayesian viewpoint, in order to study the treatment effects, adjusted for the concomitant variable, one must isolate the marginal posterior distribution of θ_1 where $\theta_1 = (a_1, a_2, \ldots, a_t)'$ is the vector of treatment effects, and it is necessary to assign a prior distribution to the parameters θ and τ where $\theta = (\theta'_1, \theta_2)$ and $\theta_2 = \gamma$. The Bayesian analysis will produce results similar to the analysis of variance if one uses

$$\xi(\theta,\tau) \propto 1/\tau, \quad \tau > 0, \quad \theta \in R^{t+1} \tag{3.141}$$

for the prior density; but it is no trouble if one uses a normal-gamma prior density.

The likelihood function for θ and τ is

$$L(\theta,\tau|y) \propto \tau^{n/2}\exp -\tfrac{\tau}{2}[(\theta - A^{-1}B)'A(\theta - A^{-1}B) + C - B'A^{-1}B], \quad \tau > 0, \quad \theta \in R^{t+1} \tag{3.142}$$

where A is a $(t + 1) \times (t + 1)$ matrix

$$A = \begin{bmatrix} n_1 & & & x_{1\cdot} \\ & \emptyset & & \\ & & n_2 & x_{2\cdot} \\ & & & \cdot \\ & & & \cdot \end{bmatrix},$$

$$A=\begin{pmatrix} & \emptyset & & . \\ & & & . \\ & & n_t & x_t. \\ x_1. & x_2. & \cdots\ x_t. & \Sigma\Sigma x^2 y \end{pmatrix},$$

B is the (t + 1) × 1 vector

$$B=\begin{pmatrix} y_1. \\ y_2. \\ . \\ . \\ . \\ y_t. \\ \Sigma\Sigma x_{ij} y_{ij} \end{pmatrix}$$

and

$$C=\sum\sum y_{ij}^2.$$

The marginal posterior density of θ is a t with n − (t + 1) degrees of freedom, location vector

$$E(\theta|y)=A^{-1}B$$

and precision matrix

$$P(\theta|y)=\frac{(n-t-1)A}{C-B'A^{-1}B}.$$

Inferences for the treatment effects are to be based on the marginal posterior distribution of θ_1, which is also a t with n − t − 1 degrees of freedom, location vector

$$E(\theta_1|y)=(I_t,\phi)A^{-1}B,$$

where ϕ is a t × 1 zero vector, and precision matrix

$$P(\theta_1|y)=[(I_t,\phi)P^{-1}(\theta|y)(I_t,\phi)\prime]^{-1}.$$

How does one test

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

for equality of treatment effects? Since is equivalent to H_0: ai − $a_{(i+1)}$ = 0, i = 1, 2, …, t − 1, consider

$$S = T\theta_1,$$

where T is a (t − 1) × t vector

$$T=\begin{pmatrix} 1 & -1 & 0 & & \dots & 0 \\ 0 & 1 & -1 & 0 & \dots & 0 \\ . & . & & & & . \\ . & . & & & & . \\ . & . & & & & . \\ 0 & 0 & \dots & 0 & 1 & -1 \end{pmatrix},$$

then S = 0[(t − 1) × 1], if and only if H_0 is true and S has a t distribution with n − t − 1 degrees of freedom, location vector TE($\theta_1|y$) and precision matrix

$$P(S|y) = [TP^{-1}(\theta_1|y)T']^{-1}.$$

It is known that

$$F(S|y) = [S - E(S|y)]\prime P(S|y)[S - E(S|y)](t-1)^{-1}$$

has an F distribution with t − 1 and n − t − 1 degrees of freedom, thus H_0 is rejected at significance level Δ (0 < Δ < 1) if $F(\theta|y) > F_{\Delta;\, t-1,\, n-t-1}$ and it can be shown that this test is equivalent to the analysis of covariance procedure.

Of course, one perhaps should first test to see if the concomitant variable x is actually needed in the model, i.e., is γ = 0? This can be done in the usual way by finding the marginal posterior distribution of γ from the marginal posterior density of θ and constructing an HPD region for γ.

The above analysis of the completely randomized design with one covariable is easily extended to other designs, for example to a randomized block design with one or more concomitant variables.

COMMENTS AND CONCLUSIONS

The Bayesian methodology for the standard statistical models of regression and for designed experiments is given in this chapter. The methodology is based on the theory which was developed in Chapter 1, and the theory gave the prior, posterior, and predictive analysis for the standard linear model.

Before analyzing the regression models, Bayesian inferences for two normal populations were examined and it was found that when the populations have a common mean but distinct variances, the marginal posterior distribution of the common mean was a poly-t, a distribution which will be again encountered in the chapters on mixed models and linear dynamic systems.

The section on regression analysis was partitioned into subsections on simple linear, multiple linear, and nonlinear regression problems. The simple and multiple linear models were studied on the basis of prior, posterior, and predictive analyses. With regard to nonlinear regression, it was shown that it is indeed difficult to specify the marginal posterior distribution of the parameters when there are several parameters, because the distribution is not well-known as it is when the regression relation is linear.

This chapter gives only a brief introduction to the Bayesian analysis of models used to analyze designed experiments. If the design is completely randomized the Bayesian analysis (with Jeffreys' prior density) is more or less equivalent to the analysis of variance procedure for testing the equality of treatment effects. On the other hand, when examining two-factor experiments, such as randomized block designs, one must reparametrize the less than full rank (design matrix) models to full rank equivalents in order to produce the standard F tests of the analysis of variance (again using Jeffreys' prior density). It was necessary to reparametrize the model to full rank because if one uses Jeffreys' prior density in conjunction with the less than full rank model, the posterior distribution of the parameters is improper.

The chapter is concluded with a Bayesian test of the equality of treatment effects, where the design is completely randomized and one concomitant variable is measured on the experimental units.

One may conclude that it is possible to develop a complete Bayesian methodology for the analysis of designed experiments and that the analysis will be more informative than the traditional analysis of variance.

EXERCISES

1. If X is a Bernoulli random variable with parameter θ, $0 \leqslant \theta \leqslant 1$ and θ has a Beta distribution with parameters α and β ($\alpha > 0$, $\beta > 0$), how would one set values for the hyperparameters α and β if one had past data $X^*_1, X^*_2, \ldots, X^*_m$, $X^*_i = 0, 1$, $i = 1, 2, \ldots, m$? Use the prior predictive distribution of X.
2. Suppose X is $n(\theta, \tau)$ and that θ and τ have a prior normal-gamma distribution with hyperparameters μ, p, α and β. If $z_1, z_2, \ldots, z_m$ represent earlier values of X, use them to set values for the hyperparameters.
3. Refer to equation (3.15) and construct: (a) a HPD region for $\theta_1 - \theta_2$, and (b) a HPD region for τ, the common precision.
4. Derive the joint prior density of τ_1 and τ_2, that is verify equation (3.18).
5. Refer to equation (3.28), the joint posterior density of θ_1 and θ_2. Are and independent? Let $\gamma = \theta_1 - \theta_2$ and develop a normal approximation to the marginal posterior density of γ.
6. In the Behrens-Fisher problem, what is the marginal posterior density of τ_1/τ_2? Refer to equation (3.21).
7. In the example of simple linear regression in this chapter (see pp. 85–94) find the mean of the marginal posterior distribution of τ^{-1}, using the data-based prior.
8. In the section on multiple linear regression, explain why equation (3.52) gives a HPD region for θ_2.
9. Refer to equation (3.103) but use Jeffreys' improper prior density (3.105), then show the analysis of variance test of H_0: $\theta_1 = \theta_2 = \ldots = \theta_t$ is equivalent to the Bayes procedure of using the HPD region of (3.103).
10. Consider the k linear models with common parameter θ

$$y_i = X_i\theta + e_i, i = 1, 2, \ldots, k$$

where the y_i are $n \times 1$, X_i is $n \times p$, θ is $p \times l$ and $e_1, e_2 \ldots, e_k$ are n.i.d. $(0, \tau_i^{-1} I_n)$, and $\theta \in R^p$ and the $\tau_i > 0$ are unknown parameters.
(a) What is the conjugate prior density for the parameters $\theta, \tau_1, \tau_2, \ldots, \tau_k$?
(b) What is the marginal posterior density of θ? Hint: See equation (3.28).
(c) Derive the marginal posterior density of $\rho = (\tau_1, \tau_2, \ldots, \tau_k)$.
To answer (a) and (b) use the conjugate prior density of the parameters.
11. Consider a linear model

$$Y_{ij} = \theta_i + \beta X_{ij} + e_{ij},$$

where $\theta_i \in R$, $\beta \in R$, the X_{ij} are known values of an independent variable X, Y_{ij} is the j-th observation of a random variable for group i, and the e_{ij} are n.i.d. $(0, \tau^{-1})$, where $i = 1, 2, \ldots, a$, $j = 1, 2, \ldots, b$, and $\tau(> 0)$ is an unknown precision. This model is that corresponding to a groups, on each of which b pairs of observations (X_{ij}, Y_{ij}) are made. The independent variable X represents a covariable which is related to the dependent variable Y. Using the improper prior density

$$\xi(\theta,\tau,\beta) \propto \frac{1}{\tau}, \quad \theta \in R^a, \quad \tau > 0, \quad \beta \in R,$$

where $\theta = (\theta_1, \theta_2, \ldots, \theta_a)'$, do the following:
(a) Find the marginal posterior density of β.

(a) Find the marginal posterior density of β.
(b) Find an HPD region for β and test H_1: $\beta = 0$.
(c) Find an HPD region for θ and test H_2:$\theta_1 = \theta_2 = \ldots = \theta_a$.

REFERENCES

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

Cochran, W. G. and G. M. Cox (1957). *Experimental Designs*, 2nd ed., John Wiley and Sons, New York.

Cox, D. R. (1958). *Planning of Experiments*, John Wiley and Sons, New York.

de Finetti, Bruno (1972). *Probability, Induction, and Statistics, the Art of Guessing*, John Wiley and Sons, London.

Draper, N. R. and H. Smith (1966). *Applied Regression Analysis*, John Wiley and Sons, Inc., New York.

Fisher, R. A. (1951). *Design of Experiments*, 6th ed., Oliver and Boyd, Edinburgh.

Graybill, F. A. (1961). *An Introduction to Linear Statistical Models*, McGraw Hill, New York.

Kadane, Joseph B., James M. Dickey, Robert L. Winkler, Wayne S. Smith and Stephen C. Peters (1980). "Interactive elicitation of opinion for a normal linear model," *Journal of American Statistical Association*, Vol. 75, pp. 845–854.

Kempthorne, O. (1952). *Design and Analysis of Experiments*, John Wiley and Sons, Inc., New York.

Lindley, D. V. (1965). *Introduction to Probability and Statistics from a Bayesian Viewpoint*, University Press, Cambridge.

Lindley, D. V. (1968). "The choice of variables in multiple regression," *Journal of the Royal Statistical Society, Series B*, Vol. 30, pp. 31–66.

Lindley, D. V. and A. F. M. Smith (1972). "Bayes estimates for the linear model," *Journal of the Royal Statistical Society, Series B*, Vol. 34, pp. 1–42.

Patil, V. H. (1964). "The Behrens-Fisher problem and its Bayesian solution," *Journal of the Indian Statistical Association*