A BAYESIAN APPROACH TO PARALLEL LINE BIOASSAY

by

SARAH CAROLINE DARBY

Thesis submitted for the degree of Doctor of Philosophy of the University of London.

London School of Hygiene & Tropical Medicine.

October, 1976

Se non è vero, è molto ben trovato.

16th Century Anonymous.

Abstract

This thesis considers parallel line bioassay from a Bayesian point of view along the lines laid out by Lindley (1972) and de Finetti (1975). The mathematical model used for the analysis is a non-linear one in which the log potency ratio appears explicitly as a parameter. This enables prior knowledge about the log potency ratio to be incorporated straightforwardly in the analysis. The method of analysis follows closely the ideas of Lindley and Smith (1972) for the linear model. Extended models in which experimental design features such as randomized blocks and Latin squares are accounted for are also considered, and a method for the use of prior information to design an assay is given.

In addition to the analysis of a single assay the problem of combining information from several assays is considered and two different models which combine such information are discussed.

Acknowledgements

The author would like to thank Professor P. Armitage for his guidance and supervision during the research. Thanks are also due to Miss M.V. Mussett and Mr. T. Kirkwood of the National Institute for Biological Standards and Control, and Mrs. M.J. Ellis of St. Thomas's Hospital Medical School for supplying the data. Finally thanks are due to the Medical Research Council for their financial support during the course of this research.

Contents

Figures

Figures / cont.

Tables

Tables / cont.

Transparencies (inside back cover)

1. Approximate large sample density of μ:N(0.243, 0.0298)

2. N(0.0, 0.0291) density

3. N(0.0, 0.0149) density

4. N(0.0, 0.00993) density

5. Approximate marginal posterior density of μ, neglecting prior information about α and β, for data from first tobramycin assay.

6. Approximate marginal posterior density of μ, assuming σ^2 to be known and equal to its value at the mode of the joint density of μ and σ^2, for data from the first tobramycin assay.

7. Approximate marginal posterior density of μ for tobramycin data, assuming $\underset{\sim}{S}$ to be known and equal to its value at the mode of the joint density of μ and S^{-1}. Prior parameters are $v=0, \underset{\sim}{\phi}^{-1}=0$, $p=2$, $\underset{\sim}{R} = \begin{pmatrix} .4\times10^4 & .1\times10^4 \\ .1\times10^4 & .4\times10^3 \end{pmatrix}$

8 Approximate marginal posterior density of μ for tobramycin data, assuming $\underset{\sim}{S}$ to be known and equal to its value at the mode of the joint density of μ and $\underset{\sim}{S}^{-1}$. Prior parameters are $v=0$, $\underset{\sim}{\phi}^{-1}=0$, $p=2$, $\underset{\sim}{R} = \begin{pmatrix} .4\times10^6 & .1\times10^6 \\ .1\times10^6 & .4\times10^5 \end{pmatrix}$

Chapter 1. Introduction

Many drugs in use at the present time are of such a complex nature that it is impossible to predict at all accurately the strength of a particular preparation by considering the ingredients and processes involved in producing it. In such cases the strength of every preparation of the drug has to be determined experimentally after the manufacturing process is complete. Experiments of this nature involving biological material are called biological assays or, more commonly, bioassays.

In its most general form the experiment consists of measuring the activity of a preparation of a drug, which we shall call the test preparation, in a biological system. This information alone is of little practical use since the activity of the test preparation will depend very heavily on the particular biological material used, and it is likely to vary considerably from experiment to experiment. What is required is a measure of the activity of the test preparation that is independent of the biological system used to determine it. Such a measure is obtained by carrying out simultaneously a similar experiment using a standard preparation. A measure of the activity of the test preparation relative to the standard preparation is then available and this should be independent of the biological medium involved in the experimentation. Standard preparations of drugs are normally of an arbitrarily defined strength. For many drugs national or international standards have been adopted, and samples of these are available from an agreed issuing laboratory.

Bioassay experiments take several different forms depending on the substances and the assay medium concerned. One possibility is that specified doses of both test and standard preparations are administered to experimental units and the resulting quantitative responses recorded. Dose-response relationships are of various types, but for a wide class of drugs the log-dose response curve is roughly linear for a range of doses, and flattens out for doses above or below this range giving a sigmoid curve altogether. In the ideal bioassay the test and standard preparations behave as if they contain different concentrations of the same active ingredient, and so the two log-dose response curves will have identical shapes but will be displaced horizontally. In practice the active ingrodiont of the two preparations is usually similar but not identical so this is only approximately true. In these assays the linear sections of the log-dose response curves for the two substances will be approximately parallel, and consequently they are known as parallel-line assays. The feature of interest in the assay is the horizontal distance between the linear sections of the two log-dose response curves, which is called the log potency ratio. Commonly occuring pharmaceutical substances calibrated in this way are insulin, vitamin D, and many antibiotics.

The results of parallel line bioassays have been analysed for many years using sampling theory techniques. Parallel regression lines are fitted to the linear sections of the two log-dose response curves using the method of least squares, and normal residuals are assumed. The equations of the fitted lines are

$$Y_S = \bar{y}_S + b(x_S - \bar{x}_S),$$

and

$$Y_T = \bar{y}_T + b(x_T - \bar{x}_T),$$

where b is the common slope of the lines, $\bar{x}_S$ and $\bar{y}_S$ are the means of the log-doses and responses for the standard preparation and Y_S is the fitted response for a log-dose x_S of the standard preparation. The suffix T refers to the test preparation in a similar way. The estimated log potency ratio M is then the difference in the log-doses of the two substances required to give the same fitted response, that is

$$M = \bar{x}_S - \bar{x}_T - \frac{(\bar{y}_S - \bar{y}_T)}{b}$$

The sampling distributions of $(\bar{y}_S - \bar{y}_T)$ and b are both normal distributions and are mutually independent so confidence limits for the log-potency ratio can be calculated using Fieller's theorem. Frequently information from several assays needs to be combined, and if one takes the above approach this proves a difficult problem which has remained unsolved for many years. Several empirical methods, in the form of weighted averages, were suggested by Finney (1964), and more recently a procedure has been described by Armitage et al (1975) which is equivalent both to generalized least squares and to maximum likelihood estimation.

In this thesis we have considered the problem outlined above from a Bayesian point of view, along the lines laid out by Lindley (1971a) and de Finetti (1975).

We begin by taking a critical look at the parametrization of the standard approach. An unusual feature is that the parameter of central interest, the log potency ratio, does not appear in the basic model. In the Bayesian framework information about the likely value of a parameter is expressed, both before and after an experiment, in the form of a distribution. This seems very difficult to do unless those parameters in which one is primarily interested occur explicitly in the model. Hence our first decision about the model we should use is that the log potency ratio should occur explicitly in our basic formulation. There now remains the task of deciding on the remaining parametrization of the model. Mathematically a model for two parallel linear regressions set at a certain distance apart can be described using three parameters. Physically one can associate four simple meaningful quantities with the situation: the horizontal distance between the lines, the joint slope of the lines and the two intercepts of the lines. The decision before us is which two of the last three quantities to include as parameters in our model. We have come to the conclusion that the correct model will depend on the precise experimental situation under consideration. The problem we are primarily concerned to study is that of calibrating a relatively unknown test substance with a relatively wellknown standard. In this case we believe that the experimenter would be most happy about quantifying his prior beliefs about the regression line for the standard preparation completely, and then quantifying, possibly independently, his prior beliefs about the likely log potency ratio of the test preparation when compared with the standard. If normally distributed errors are assumed then we have the following model for observations on the standard preparation:

$$y \sim N(\alpha+\beta x,\sigma^2),$$

where y is the response, x is the log-dose, β is the slope of the regression line, α its intercept and σ^2 the residual variance. Also we have the following model for observations on the test preparation:

$$y \sim N(\alpha+\beta(\mu+x),\sigma^2),$$

where μ is the log potency ratio. Combining these two into a single equation the basic model is

$$y \sim N(\alpha+\beta\mu z+\beta x,\sigma^2),$$

where z is a dummy variable taking the value 0 when a dose of the standard preparation is used and 1 when a dose of the test preparation is used.

This model has an obvious disadvantage in that it is nonlinear; however we believe that our parameterization is a more natural one than the one used in the standard sampling theory analysis, and in particular we believe that the problem of combining information from several different assays on the same pair of substances is made logically simpler by this approach.

In the following chapters we explore the consequences of adopting this model and we follow closely the ideas set out by Lindley & Smith [1972] for the linear model, adapting them where necessary to this non-linear case.

Chapter 2. Analysis of a Single Assay With Known Residual Variance

2.1 The Model

The first analysis we shall attempt is that of a single assay. For initial simplicity we shall assume that the residual variance is known, and then in a later chapter we shall remove this restriction. To carry out our first analysis we shall use the following two stage model:

$$\text{1st stage:} \quad y \sim N\{(\alpha+\beta\mu z+\beta x),\sigma^2\}$$

$$\text{2nd stage:} \quad \begin{pmatrix}\alpha\\ \beta\\ \mu\end{pmatrix} \sim N\left\{\begin{pmatrix}\alpha_o\\ \beta_o\\ \mu_o\end{pmatrix}, \underset{\sim}{\Sigma}\right\} \qquad (2.1)$$

where y is the response, x is the log-dose, and z is a dummy variable taking value 0 when a dose of the standard preparation is used and 1 when a dose of the test preparation is used. The second stage of the model describes prior knowledge about the parameters in the first stage; α_o, β_o, μ_o and the elements of $\underset{\sim}{\Sigma}$ are assumed known. We have considered a general case where all the elements of $\underset{\sim}{\Sigma}$ can be non-zero, but in many cases some of the off-diagonal elements will be zero. The appropriate form in any particular case will depend on the precise nature of the prior information available.

As an example of a case where some of the elements of $\underset{\sim}{\Sigma}$ are zero, let us consider the following situation. Suppose we want to determine the activity of a test preparation of vitamin D by comparison with a well known standard, and suppose we are going to carry out this particular assay on chickens. It so happens that we have carried out many assays on this medium using our current standard and other test preparations, but the only assays we have done with our current pair of substances have used rats instead of chickens.

By considering the results we have obtained in the past for the standard preparation in assays on chickens, we should be able to form an idea of what to expect next time. Let the intercept with the x-axis, and the slope of the linear part of the log-dose response curve be α & β respectively. We construct values α_o, β_o, Σ_{11}, Σ_{12}, Σ_{22} such that to a reasonable approximation

$$\begin{pmatrix}\alpha\\ \beta\end{pmatrix} \sim N\left\{\begin{pmatrix}\alpha_o\\ \beta_o\end{pmatrix}, \begin{bmatrix}\Sigma_{11} & \Sigma_{12}\\ \Sigma_{12} & \Sigma_{22}\end{bmatrix}\right\}.$$

Also, by considering the extent of the linear part of the log-dose response curve in past assays we should be able to decide on the range of doses to be used for the standard preparation.

Quite independently of the above we now consider the results of the rat assays. Let the log potency ratio of the two substances concerned be μ. We construct values μ_o and Σ_{33} such that approximately

$$\mu \sim N(\mu_o, \Sigma_{33})$$

We can now decide on the range of doses to be used for the test preparation and then on the final design. A method for designing assays is discussed in Chapter 3.

Amalgamating the prior information from the two separate sources the second stage of the model becomes

$$\begin{pmatrix}\alpha\\ \beta\\ \mu\end{pmatrix} \sim N\left\{\begin{pmatrix}\alpha_o\\ \beta_o\\ \mu_o\end{pmatrix}, \begin{bmatrix}\Sigma_{11} & \Sigma_{12} & 0\\ \Sigma_{12} & \Sigma_{22} & 0\\ 0 & 0 & \Sigma_{33}\end{bmatrix}\right\}.$$

The situation described above will occur rather infrequently. However, the implied structure for Σ will hold approximately in many cases where prior information about the log potency ratio of the two substances concerned is assessed separately from prior information about the behaviour of the standard preparation using the current assay medium.

2.2 Posterior Distributions

After the assay results have been obtained we can multiply together the likelihood and the prior density, as given by 2.1, to form the posterior density of the three parameters α, β and μ up to a multiplicative constant.
This gives:

$$\pi(\alpha,\beta,\mu|\underset{\sim}{y}) \propto \exp -\tfrac{1}{2}\left\{\alpha^2\left(\frac{n}{\sigma^2}+\Sigma^{11}\right)+2\alpha\beta\left(\frac{\Sigma x_i}{\sigma^2}+\frac{\mu\Sigma z_i}{\sigma^2}+\Sigma^{12}\right)+\beta^2\left(\frac{\Sigma x_i^2}{\sigma^2}+\frac{2\mu\Sigma x_i z_i}{\sigma^2}+\frac{\mu^2\Sigma z_i^2}{\sigma^2}+\Sigma^{22}\right)\right.$$

$$-2\alpha\left(\frac{\Sigma y_i}{\sigma^2}+\alpha_0\Sigma^{11}+\beta_0\Sigma^{12}-(\mu-\mu_0)\Sigma^{13}\right)$$

$$-2\beta\left(\frac{\Sigma x_i y_i}{\sigma^2}+\frac{\mu\Sigma y_i z_i}{\sigma^2}+\beta_0\Sigma^{22}+\alpha_0\Sigma^{12}-(\mu-\mu_0)\Sigma^{23}\right)$$

$$\left.+\mu^2\Sigma^{33}-2\mu(\mu_0\Sigma^{33}+\beta_0\Sigma^{23}+\alpha_0\Sigma^{13})\right\} \qquad (2.2)$$

where n is the number of subjects in the assay. z_i and x_i refer to the ith subject, Σ^{ij} is the $(ij)^{th}$ element of $\underset{\sim}{\Sigma}^{-1}$, and summations are from i=1 to i = n unless otherwise indicated.

As might be expected, this does not correspond to any standard distribution, and consequently its properties are difficult to examine. For example, we have been unable to find either the mean or the variance analytically. We can, however, find the mode. This occurs at

$$\alpha=\frac{\dfrac{\Sigma y_i}{\sigma^2}-\dfrac{\beta\mu\Sigma z_i}{\sigma^2}-\dfrac{\beta\Sigma x_i}{\sigma^2}+\alpha_0\Sigma^{11}-(\beta-\beta_0)\Sigma^{12}-(\mu-\mu_0)\Sigma^{13}}{\dfrac{n}{\sigma^2}+\Sigma^{11}}$$

$$\beta=\frac{\dfrac{\Sigma x_i y_i}{\sigma^2}+\dfrac{\mu\Sigma y_i z_i}{\sigma^2}-\dfrac{\alpha\Sigma x_i}{\sigma^2}-\dfrac{\alpha\mu\Sigma z_i}{\sigma^2}+\beta_0\Sigma^{22}-(\alpha-\alpha_0)\Sigma^{12}-(\mu-\mu_0)\Sigma^{23}}{\dfrac{\Sigma x_i^2}{\sigma^2}+\dfrac{2\mu\Sigma x_i z_i}{\sigma^2}+\dfrac{\mu^2\Sigma z_i^2}{\sigma^2}+\Sigma^{22}}, \qquad (2.3)$$

$$\mu = \frac{\frac{\beta\Sigma y_i z_i}{\sigma^2} - \frac{\beta^2\Sigma x_i z_i}{\sigma^2} - \frac{\alpha\beta\Sigma z_i}{\sigma^2} + \mu_0\Sigma^{33} - (\alpha-\alpha_0)\Sigma^{13} - (\beta-\beta_0)\Sigma^{23}}{\frac{\beta^2\Sigma z_i^{\,2}}{\sigma^2} + \Sigma^{23}}$$

If one has very little prior knowledge about α, β & μ, the elements of Σ will become extremely large, and consequently the elements of Σ^{-1} will become very small. In the limiting case of no prior knowledge they will all be zero and the mode will occur at

$$\alpha = \bar{y} - \beta\mu\bar{z} - \beta\bar{x} ,$$

$$\beta = \frac{\Sigma x_i y_i + \mu\Sigma y_i z_i - \alpha n\bar{x} - \alpha\mu n\bar{z}}{\Sigma x_i^{\,2} + 2\mu\Sigma x_i z_i + \mu^2\Sigma z_i^{\,2}} , \qquad (2.4)$$

$$\mu = \frac{\Sigma y_i z_i - \beta\Sigma x_i z_i - \alpha n\bar{z}}{\beta\Sigma z_i^{\,2}} ,$$

where $\bar{y}$ is the average of $y_1, y_2, \ldots y_n$, $\bar{z}$ is the average of $z_1, z_2, \ldots, z_n$ and $\bar{x}$ is the average of $x_1, x_2, \ldots, x_n$. Substituting for α in the expression for μ, and for α and μ in the expression for β gives

$$\beta = \frac{S_{xy} - \frac{S_{xz}S_{yz}}{S_{zz}}}{S_{xx} - \frac{(S_{xz})^2}{S_{zz}}} , \qquad (2.5)$$

$$\mu = \frac{S_{yz} - \beta S_{xz}}{\beta S_{zz}} ,$$

where $S_{xy} = \Sigma(x_i - \bar{x})(y_i - \bar{y})$ and similarly for S_{xx}, S_{xz}, S_{yz}, S_{zz}

The expressions for β and μ, although disguised by the use of the dummy variable z, are exactly the estimates of slope of regression line and log potency ratio obtained by the standard sampling theory analysis. This can easily be seen as follows. If we dispense with the dummy variable z we have the following relationships:

$$S_{xy}=\sum_{S}(x_i-\bar{x}_S)(y_i-\bar{y}_S)+\sum_{T}(x_i-\bar{x}_T)(y_i-\bar{y}_T)+\frac{n_S n_T}{n}(\bar{x}_T-\bar{x}_S)(\bar{y}_T-\bar{y}_S)\ ,$$

$$S_{xz}=\frac{n_S n_T}{n}(x_T-x_S)\ ,$$

$$S_{yz}=\frac{n_S n_T}{n}(y_T-y_S)\ ,$$

$$S_{zz}=\frac{n_S n_T}{n}\ ,$$

$$S_{xx}=\sum_{S}(x_i-\bar{x}_S)^2+\sum_{T}(x_i-\bar{x}_T)^2+\frac{n_S n_T}{n}(\bar{x}_S-\bar{x}_T)^2\ ,$$

where suffices S and T refer to standard and test preparations respectively. On substituting these relationships into the modal values for β and μ we get

$$\beta=\frac{\sum_{S}(x_i-\bar{x}_S)(y_i-\bar{y}_S)+\sum_{T}(x_i-\bar{x}_T)(y_i-\bar{y}_T)}{\sum_{S}(x_i-\bar{x}_S)^2+\sum_{T}(x_i-\bar{x}_T)^2}\ ,$$

$$\mu=\bar{x}_S-\bar{x}_T+\frac{(\bar{y}_T-\bar{y}_S)}{\beta}\ .$$

By examining the form of the joint posterior density given in 2.2, it can be seen that the joint distribution of α and β for a fixed value of μ is in the form of a bivariate normal distribution. We can therefore integrate over α and β

to obtain the marginal posterior density of μ up to a multiplicative constant. This calculation gives

$$\pi(\mu|\underset{\sim}{y}) \propto |\underset{\sim}{V}|^{\frac{1}{2}} \exp -\tfrac{1}{2}\{\mu^2\Sigma^{33} - 2\mu(\alpha_o\Sigma^{13}+\beta_o\Sigma^{23}+\mu_o\Sigma^{33}) - \begin{bmatrix} a \\ b \end{bmatrix}^T \underset{\sim}{V} \begin{bmatrix} a \\ b \end{bmatrix}\},$$

where
$$\underset{\sim}{V} = \begin{bmatrix} \left(\frac{n}{\sigma^2}+\Sigma^{11}\right) & \left(\frac{\Sigma x_i}{\sigma^2}+\frac{\mu\Sigma z_i}{\sigma^2}+\Sigma^{12}\right) \\ \left(\frac{\Sigma x_i}{\sigma^2}+\frac{\mu\Sigma z_i}{\sigma^2}+\Sigma^{12}\right) & \left(\frac{\Sigma x_i^2}{\sigma^2}+\frac{2\mu\Sigma x_i z_i}{\sigma^2}+\frac{\mu^2\Sigma z_i^2}{\sigma^2}+\Sigma^{22}\right) \end{bmatrix}^{-1},$$

$$a = \frac{n\bar{y}}{\sigma^2}+\alpha_o\Sigma^{11}+\beta_o\Sigma^{12}-(\mu-\mu_o)\Sigma^{13},$$

$$b = \frac{\Sigma x_i y_i}{\sigma^2}+\frac{\mu\Sigma y_i z_i}{\sigma^2}+\beta_o\Sigma^{22}+\alpha_o\Sigma^{12}-(\mu-\mu_o)\Sigma^{23}$$

Again, this density does not correspond to any standard distribution, and it is even more intractable than the joint posterior density in the sense that the mode cannot be found analytically. For a closer investigation of its behaviour we have resorted to numerical techniques in special cases; see section 2.5.

The posterior marginal density of β can be found in a similar fashion and appears no less complicated.

In our subsequent discussion, either for theoretical simplicity, or as an approximation to a real situation, we may wish to consider the case where we have little or no prior information about one or more of the parameters in our model. For example, reduction of prior information about β would cause Σ_{22} to get bigger, and eventually to tend to infinity. Before allowing the limiting situation of no prior knowledge to occur we should examine carefully the consequences for the posterior distributions involved.

In the following argument we show that prior ignorance about μ causes the joint posterior density to be unnormed. This does not happen when there is no prior knowledge about α or β. We assume throughout that for at least one of the preparations at least two different doses are administered.

The cases we wish to consider are to let one or more of Σ_{11}, Σ_{22}, Σ_{33} tend to infinity in $\underset{\sim}{\Sigma}$. If $\Sigma_{ii} \to \infty$, $\Sigma^{ij} = \Sigma^{ji} = 0$ for $j = 1,2,3$. Let the expression on the right hand side of the $\propto$ sign in 2.2 be $f(\alpha,\beta,\mu)$, then $\pi(\alpha,\beta,\mu|\underset{\sim}{y})$ will be a normed density function only when $\iiint f(\alpha,\beta,\mu)\,d\alpha\, d\beta\, d\mu$ is finite.

From 2.6 $\iint f(\alpha,\beta,\mu)\,d\alpha\, d\beta = \{A(\mu)\}^{-\frac{1}{2}} B(\mu) \exp\tfrac{1}{2}C(\mu)$

where
$$A(\mu) = \left\{\left(\frac{n}{\sigma^2} + \Sigma^{11}\right)\left(\frac{\Sigma x_i^{\,2}}{\sigma^2} + 2\mu\frac{\Sigma x_i z_i}{\sigma^2} + \mu^2\frac{\Sigma z_i^{\,2}}{\sigma^2} + \Sigma^{22}\right) - \left(\frac{\Sigma x_i}{\sigma^2} + \mu\frac{\Sigma z_i}{\sigma^2} + \Sigma^{12}\right)^2\right\}$$

$$B(\mu) = \exp -\tfrac{1}{2}\{\mu^2\Sigma^{33} - 2\mu(\alpha_o\Sigma^{13} + \beta_o\Sigma^{23} + \mu_o\Sigma^{33})\}$$

$$C(\mu) = \frac{1}{A(\mu)} \times \left\{\left(\frac{n\bar{y}^2}{\sigma^2} + \alpha_o\Sigma^{11} + \beta_o\Sigma^{12} - (\mu-\mu_o)\Sigma^{13}\right)^2\left(\frac{\Sigma x_i^{\,2}}{\sigma^2} + 2\mu\frac{\Sigma x_i z_i}{\sigma^2} + \mu^2\frac{\Sigma z_i^{\,2}}{\sigma^2} + \Sigma^{22}\right)\right.$$

$$-2\left(\frac{\Sigma x_i}{\sigma^2} + \mu\frac{\Sigma z_i}{\sigma^2} + \Sigma^{12}\right)\left(\frac{n\bar{y}}{\sigma^2} + \alpha_o\Sigma^{11} + \beta_o\Sigma^{12} - (\mu-\mu_o)\Sigma^{13}\right)\left(\frac{\Sigma x_i y_i}{\sigma^2} + \mu\frac{\Sigma y_i z_i}{\sigma^2} + \beta_o\Sigma^{22} + \alpha_o\Sigma^{12} - (\mu-\mu_o)\Sigma^{23}\right)$$

$$\left. + \left(\frac{\Sigma x_i y_i}{\sigma^2} + \mu\frac{\Sigma y_i z_i}{\sigma^2} + \beta_o\Sigma^{22} + \alpha_o\Sigma^{12} - (\mu-\mu_o)\Sigma^{23}\right)^2\left(\frac{n}{\sigma^2} + \Sigma^{11}\right)\right\}.$$

This result is true for all the cases we wish to consider, although various terms in $A(\mu)$, $B(\mu)$ and $C(\mu)$ will be zero when one or more of Σ_{11}, Σ_{22}, $\Sigma_{33} \to \infty$.

We can rewrite $A(\mu)$ in the form

$$A(\mu) = \Sigma^{11}\Sigma^{22} - (\Sigma^{12})^2 + \frac{1}{\sigma^2}\left(\Sigma^{11}\Sigma(x_i+\mu z_i)^2 - 2\Sigma^{12}\Sigma(x_i+\mu z_i) + n\Sigma^{22}\right)$$
$$+ \frac{n}{\sigma^4}\Sigma\{x_i - \bar{x} + \mu(z_i - \bar{z})\}^2 \qquad (2.7)$$

For all the cases we wish to consider the matrix $\begin{bmatrix}\Sigma^{11} & \Sigma^{12}\\ \Sigma^{12} & \Sigma^{22}\end{bmatrix}$ is positive semidefinite and so its determinant will be non-negative, that is $\Sigma^{11}\Sigma^{22}-(\Sigma^{12})^2 \geq 0$.

Also

$\{\Sigma^{11}\Sigma(x_i+\mu z_i)^2-2\Sigma^{12}\Sigma(x_i+\mu z_i)+n\Sigma^{22}\} \geq 0$, since it is the sum of n quadratic forms in $\begin{bmatrix}\Sigma^{11} & \Sigma^{12}\\ \Sigma^{12} & \Sigma^{22}\end{bmatrix}$. Lastly $\Sigma\{x_i-\bar{x}+\mu(z_i-\bar{z})\}^2 > 0$,

since we have assumed that at least two different doses are used for at least one of the preparations. Hence we have that $A(\mu) > 0, \forall\mu$.

Firstly let us consider the case when the coefficient of μ^2 in $A(\mu)$ is strictly positive, that is

$\left\{\frac{n}{\sigma^4} S_{zz} + \frac{\Sigma^{11}\Sigma z_i^2}{\sigma^2}\right\} > 0$. We can rewrite $C(\mu)$ in the form

$$C(\mu) = \left[\frac{\mu^2(\Sigma^{13})^2\Sigma z_i^2}{\frac{n}{\sigma^2} S_{zz}+\Sigma z_i^2\Sigma^{11}}\right] + 2\Sigma^{13}a\mu + \left[\frac{b\mu^2+c\mu+d}{A(\mu)}\right]$$

where a,b,c & d are constants which do not depend on μ

Let

$$B^*(\mu) = B(\mu)\exp\tfrac{1}{2}\left\{\frac{\mu^2(\Sigma^{13})^2\Sigma z_i^2}{\frac{n}{\sigma^2} S_{zz} + \Sigma^{11}\Sigma z_i^2} + 2\Sigma^{13}a\mu\right\},$$

and

$$C^*(\mu) = \frac{b\mu^2+c\mu+d}{A(\mu)} .$$

Since $A(\mu)$ has no real roots $C^*(\mu)$ will be bounded above and below, and $\{A(\mu)\}^{-\frac{1}{2}}$ will be bounded above. It follows that there exist ξ_L, ξ_u and η_u all strictly positive such that

$$\xi_L \leq \exp\tfrac{1}{2}C^*(\mu) \quad \xi_u$$

and $(A(\mu))^{-\frac{1}{2}} \leqslant \eta_u$

for all μ.

Suppose $\Sigma_{33} < \infty$, then

$$\iiint f(\alpha,\beta,\mu)\,d\alpha\, d\beta\, d\mu = \int \{A(\mu)\}^{-\frac{1}{2}} B^*(\mu) \exp\tfrac{1}{2} C^*(\mu)\,d\mu$$

$$\leqslant \xi_u \eta_u \int B^*(\mu)\,d\mu$$

$$= \xi_u \eta_u \int_{-\infty}^{\infty} \exp -\tfrac{1}{2}\left[\mu^2 \left\{\Sigma^{33} - \frac{(\Sigma^{13})^2 \Sigma z_i^2}{\frac{n}{\sigma^2} S_{zz} + \Sigma^{11} \Sigma z_i^2}\right\} - 2\mu\{(\alpha_o - a)\Sigma^{13} + \beta_o \Sigma^{23} + \mu_o \Sigma^{33}\right] d\mu$$

$< \infty$, since $\left\{\Sigma^{33} - \dfrac{(\Sigma^{13})^2 \Sigma z_i^2}{\frac{n}{\sigma^2} S_{zz} + \Sigma^{11} \Sigma z_i^2}\right\} > 0$ in all the

cases under consideration.

Now suppose $\Sigma_{33} \to \infty$. $B^*(\mu) = 1$, and

$$\iiint f(\alpha,\beta,\mu)\,d\alpha\, d\beta\, d\mu = \int \{A(\mu)\}^{-\frac{1}{2}} B^*(\mu) \exp\tfrac{1}{2} C^*(\mu)\,d\mu$$

$$\geqslant \xi_L \int_{-\infty}^{\infty} \{A(\mu)\}^{-\frac{1}{2}}\,d\mu$$

$= \infty$ as is shown below.

From 2.7 we can write $A(\mu)$ in the form $A(\mu) = e(\mu+g)^2 + h$ where

$$e = \frac{n}{\sigma^4} S_{zz} + \Sigma^{11} \frac{\Sigma z_i^2}{\sigma^2},$$

and

$$h = \Sigma^{11}\Sigma^{22} - (\Sigma^{12})^2 + \frac{1}{\sigma^2}\{\Sigma x_i^2 \Sigma^{11} - 2\Sigma^{12}\Sigma x_i + n\Sigma^{22}\} + \frac{n}{\sigma^4} Sxx .$$

In the present case both e and h are strictly positive. Let us transform from μ to t where $\tan t = \left(\frac{e}{h}\right)^{\frac{1}{2}}(\mu+g)$, then

$$\int_{-\infty}^{\infty}\{A(\mu)\}^{-\frac{1}{2}}d\mu=\int_{-\infty}^{\infty}\frac{1}{\{e(\mu+g)^2+h\}^{\frac{1}{2}}}\,d\mu$$

$$=2e^{-\frac{1}{2}}\int_0^{\pi/2}\sec t\,dt$$

$$=2e^{-\frac{1}{2}}\lim_{\delta\to o}\Big[\log(\sec t+\tan t)\Big]_o^{\pi/2\;-\delta}$$

$$=2e^{-\frac{1}{2}}\lim_{\delta\to o}\log\left[\sec\left(\frac{\pi}{2}-\delta\right)+\tan\left(\frac{\pi}{2}-d\right)\right]$$

$$=\infty$$

This completes the argument when the coefficient of μ^2 in $A(\mu)$ is strictly positive. This coefficient cannot be negative, but it can be zero, and we now consider this case.

We are considering the case $\frac{n}{\sigma^4}S_{zz}+\Sigma^{11}\frac{\Sigma z_1{}^2}{\sigma^2}=0$. This can happen in two different ways; either we can have $S_{zz}=0$ and $\Sigma^{11}=0$ or we can have $S_{zz}=0$ and $\Sigma z_1{}^2=0$. If the first of these possibilities is true then

$$A(\mu)=\frac{n}{\sigma^2}\Sigma^{22}+\frac{n}{\sigma^4}Sxx\ ,$$

and

$$C(\mu)=\frac{(\Sigma^{23})^2\mu^2}{\Sigma^{22}+\frac{Sxx}{\sigma^2}}+2j\mu+k\ .$$

where j & k are constants independent of μ.

Suppose $\Sigma_{33} < \infty$, $\iiint f(\alpha,\beta,\mu)\,d\alpha\,d\beta\,d\mu = \int \{A(\mu)\}^{-\frac{1}{2}} B(\mu) \exp \frac{1}{2} C(\mu)\,d\mu$

$$= \left\{ \frac{n}{\sigma^2} \left(\Sigma^{22} + \frac{Sxx}{\sigma^2} \right) \right\}^{-\frac{1}{2}} \int_{-\infty}^{\infty} \exp -\frac{1}{2} \left[\mu^2 \left\{ \Sigma^{33} - \frac{(\Sigma^{23})^2}{\Sigma^{22} + \frac{Sxx}{\sigma^2}} \right\} - 2\mu \{ \alpha_o \Sigma^{13} + \beta_o \Sigma^{23} + \mu_o \Sigma^{33} + j \} - k \right] d\mu$$

$< \infty$ since $\left\{ \Sigma^{33} - \dfrac{(\Sigma^{23})^2}{\Sigma^{22} + \frac{Sxx}{\sigma^2}} \right\} > 0$ in all the cases under consideration.

Now suppose $\Sigma_{33} \to \infty$. $B(\mu)=1$ and both the terms in $C(\mu)$ involving μ dissappear, hence $C(\mu)=k$.

$$\iiint f(\alpha,\beta,\mu)\,d\alpha\,d\beta\,d\mu = \left\{ \frac{n}{\sigma^2} \left(\Sigma^{22} + \frac{Sxx}{\sigma^2} \right) \right\}^{-\frac{1}{2}} \exp \frac{k}{2} \int_{-\infty}^{\infty} 1 . d\mu$$
$$= \infty$$

We now consider the final case. Here we have $S_{zz}=0$ and $\Sigma z_i^{\,2}=0$.
In this case

$$A(\mu) = \Sigma^{11}\Sigma^{22} - (\Sigma^{12})^2 + \frac{1}{\sigma^2} \{ \Sigma^{11} \Sigma x_i^{\,2} - 2\Sigma^{12} \Sigma x_i + n\Sigma^{22} \} + \frac{n}{\sigma^4} Sxx = \ell \text{ say.}$$

$$C(\mu) = \left[\Sigma^{11}(\Sigma^{23})^2 + \Sigma^{22}(\Sigma^{13})^2 - 2\Sigma^{12}\Sigma^{13}\Sigma^{23} + \frac{1}{\sigma^2} \{ (\Sigma^{13})^2 \Sigma x_i^{\,2} - 2\Sigma^{13}\Sigma^{23}\Sigma x_i + n(\Sigma^{23})^2 \} \right] \mu^2$$
$$+ 2m\mu + n,$$

where ℓ, m & n are constants independent of μ.
Suppose $\Sigma_{33} < \infty$, $\iiint f(\alpha,\beta,\mu)\,d\alpha\,d\beta\,d\mu = \ell^{-\frac{1}{2}}$

$$\times \int \exp -\frac{1}{2} \left\{ \mu^2 p - 2\mu(\alpha_o \Sigma^{13} + \beta_o \Sigma^{23} + \mu_o \Sigma^{33} + m) - n \right\} d\mu$$

where

$$p=\Sigma^{33}-\left[\frac{\Sigma^{11}(\Sigma^{23})^2+\Sigma^{22}(\Sigma^{13})^2-2\Sigma^{12}\Sigma^{13}\Sigma^{23}+\frac{1}{\sigma^2}\left\{(\Sigma^{13})^2\Sigma x_i^{\,2}-2\Sigma^{13}\Sigma^{23}\Sigma x_i+n(\Sigma^{23})^2\right\}}{\Sigma^{11}\Sigma^{22}-(\Sigma^{12})^2+\frac{1}{\sigma^2}\left(\Sigma^{11}\Sigma x_i^{\,2}-2\Sigma^{12}\Sigma x_i+n\Sigma^{22}\right)+\frac{n}{\sigma^4}S_{xx}}\right]$$

It can easily be shown that $p > 0$ for all the cases under consideration, and hence

$$\iiint f(\alpha,\beta,\mu)\,d\alpha\, d\beta\, d\mu < \infty .$$

Now suppose $\Sigma_{33} \to \infty$. $B(\mu)=1$, and as in the previous case $C(\mu)$ becomes a constant. Hence

$$\iiint f(\alpha,\beta,\mu)\,d\alpha\, d\beta\, d\mu = k^{-1}\exp\frac{n}{2}\int_{-\infty}^{\infty} 1.d\mu$$

$$= \infty$$

This completes the argument.

If we had not satisfied the initial assumption of at least two doses being used on one preparation, our argument would still have held provided $A(\mu) > 0$ for all μ. From 2.7 this will be true if

$$\Sigma^{11}\Sigma(x_i+\mu z_i)^2-2\Sigma^{12}\Sigma(x_i+\mu z_i)+n\Sigma^{22} > 0,$$

that is if either $\Sigma^{22} > 0$, or $\Sigma^{11} > 0$ and a non-zero dose of the standard is used. This will happen when either we have some prior information about the slope of the log-dose response line of the standard, or we have some prior information about the intercept of this line with the y-axis and experimental knowledge about some other point on it, thus enabling the slope to be estimated.

In the light of the preceding result we shall in our subsequent discussion consider using uniform priors for α and β. but not for ω. The result perhaps indicates that in a parallel line bioassay one obtains information about log potency ratio in a rather indirect way and consequently the resulting information is imprecise. The result compares with the fact that in the standard sampling theory analysis the log potency ratio is estimated by the ratio of two statistics whose sampling distributions are normal and mutually independent. Consequently the sampling distribution of the estimate of log potency ratio has no finite moments.

2.3 Large Sample Distributions

Lindley (1961) has shown that given n independent observations $\underset{\sim}{y}=(y_1,y_2,\ldots y_n)^T$ each with probability density $p(y|\underset{\sim}{\theta})$, where $\underset{\sim}{\theta}=(\theta_1,\theta_2,\ldots,\theta_p)$ is a vector of parameters, then provided $p(y|\underset{\sim}{\theta})$ is sufficiently regular, the asymptotic distribution of $\underset{\sim}{\theta}$ is

$$\pi(\underset{\sim}{\theta}|\underset{\sim}{y}) \sim (2\pi)^{-p/2}\,|\underset{\sim}{W}|^{-\frac{1}{2}}\exp-\tfrac{1}{2}\{(\underset{\sim}{\theta}-\underset{\sim}{\hat{\theta}})^T\underset{\sim}{W}^{-1}(\underset{\sim}{\theta}-\underset{\sim}{\hat{\theta}})\}$$

where the $(i,j)^{th}$ element of $\underset{\sim}{W}^{-1}$ is

$$\frac{-\partial^2}{\partial\theta_i\partial\theta_j}\left\{\sum_{k=1}^{n}\log p(y_k|\underset{\sim}{\theta}=\underset{\sim}{\hat{\theta}})\right\},$$

and $\underset{\sim}{\hat{\theta}}$ is the usual maximum likelihood value of $\underset{\sim}{\theta}$.

Considering the current model, the regularity conditions are satisfied, and the maximum likelihood values are

$$\hat{\alpha}=\bar{y}-\hat{\beta}\hat{\mu}\bar{z}-\hat{\beta}\bar{x},$$

$$\hat{\beta}=\left\{Sxy-\frac{SyzSxz}{Szz}\right\}\Big/\left\{Sxx-\frac{(S_{xz})^2}{S_{zz}}\right\},$$

$$\hat{\mu}=\frac{-S_{xz}}{Szz}+\frac{Syz}{\hat{\beta}Szz}\,. \qquad (2.8)$$

and

$$\underset{\sim}{W}=\frac{\sigma^2}{\{SxxS_{zz}-(S_{xz})^2\}}\begin{bmatrix}\frac{1}{n}(\Sigma z_i^2\Sigma x_i^2-(\Sigma x_iz_i)^2) & \{\bar{z}Sxz-\bar{x}Szz\} & \frac{1}{\hat{\beta}^2}\{\bar{x}Syz-\bar{z}Sxy\} \\ \{\bar{z}Sxz-\bar{x}Szz\} & S_{zz} & \frac{-Syz}{\hat{\beta}} \\ \frac{1}{\hat{\beta}^2}(\bar{x}Syz-\bar{z}Sxy) & \frac{-Syz}{\hat{\beta}} & \frac{1}{\hat{\beta}^2}(\hat{\mu}^2Szz+2\hat{\mu}S_{xz}+Sxx)\end{bmatrix}$$

Hence we have that for assays with an infinite number of responses the three parameters are normally distributed with means equal to the mode of the joint posterior density for finite samples when the terms involving the prior knowledge are neglected, see 2.4 & 2.5. As is indicated in section 3.2, $\bar{\beta}$ and $\overset{+}{\mu}$ are the estimates of β & μ given by the standard sampling theory analysis. It can easily be shown that the variance of β is equal to the sampling variance of the standard estimate of slope, and that the variance of μ is equal to the approximate formula frequently used as the sampling variance of the standard estimate of log potency ratio.

2.4 Estimation of Log Potency Ratio

Following de Finetti (1975) we feel that, within the Bayesian framework, the natural way to present the solution of a statistical problem is to give the relevant posterior distribution. In the present case this is the marginal posterior distribution of μ. In the context of bioassay. However, drugs need to be labelled with particular strengths and so there is a need for a more concise representation of the available information in the form of a point estimate of μ and also possibly a confidence interval.

We shall approach the problem of point estimation from a decision theoretic point of view, and we shall assume for the sake of definiteness that a quadratic loss function is appropriate. In this case the best estimate of log potency ratio will be the marginal posterior mean of μ, calculation of which will involve two one-dimensional numerical integrations. At the present time there are fast and reliable computer packages which perform one-dimensional numerical integrations of the type required and so this calculation should not present too great a problem. If necessary, however, one could approximate the marginal posterior mean by the marginal posterior mode, the calculation of which is a much simpler problem numerically.

A further possible estimate of the log potency ratio is the value of μ at the mode of the joint posterior distribution of α, β, and μ as given by 2.3. If large quantities of data were available the joint posterior distribution of α, β and μ would be approximately multivariate normal, and the joint mode would be approximately equal to the marginal posterior means. However, data from a single assay are unlikely to be sufficiently extensive for this to be the case.

	Test Preparation		Standard Preparation	
Log dose	1	2	1.5	2.5
	0.419	0.959	0.391	1.551
	1.193	1.757	0.083	1.537
	0.937	1.415	0.411	0.833
	0.233	1.135	0.388	1.409
	0.303	1.619	0.980	2.330
	0.698	1.401	1.179	1.799
	-0.574	1.305	0.918	1.557
	0.639	1.496	1.108	2.340

Table 2.1 Generated data set

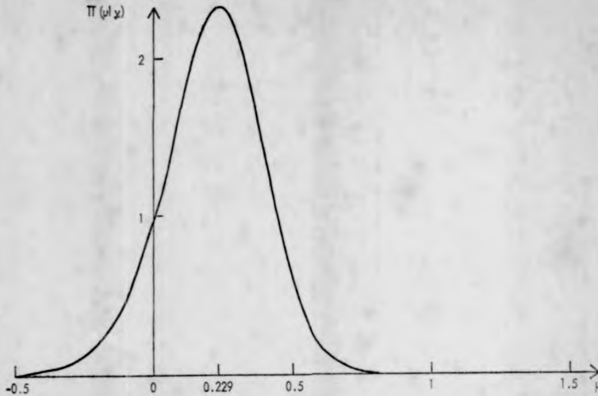


Figure 2.1 Marginal posterior density of μ for the generated data set when the prior distribution for μ is $\mu \sim N(0.000, 0.500)$

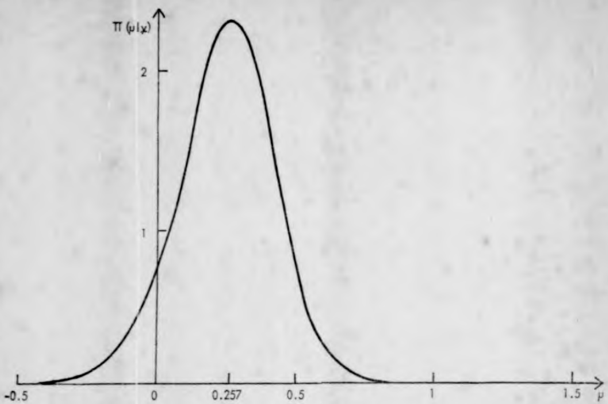


Figure 2.2 Marginal posterior density of μ for the generated data set when the prior distribution for μ is μ ~ N(0.500, 0.500)

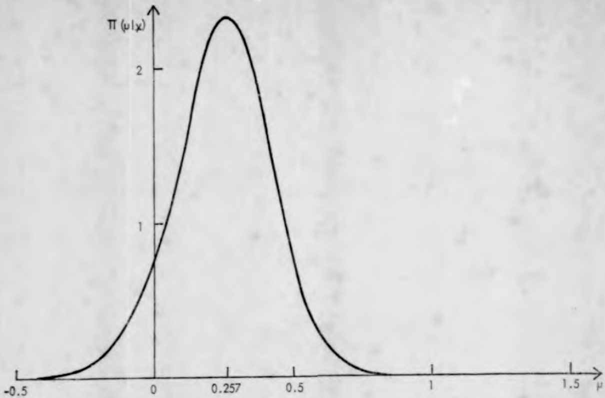


Figure 2.2 Marginal posterior density of μ for the generated data set when the prior distribution for μ is $\mu \sim N(0.500, 0.500)$

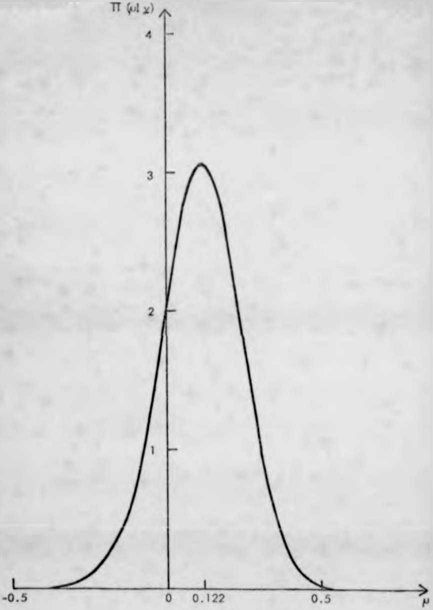


Figure 2.3 Marginal posterior density of μ for the generated data set when the prior distribution for μ is N(0.000, 0.0298)

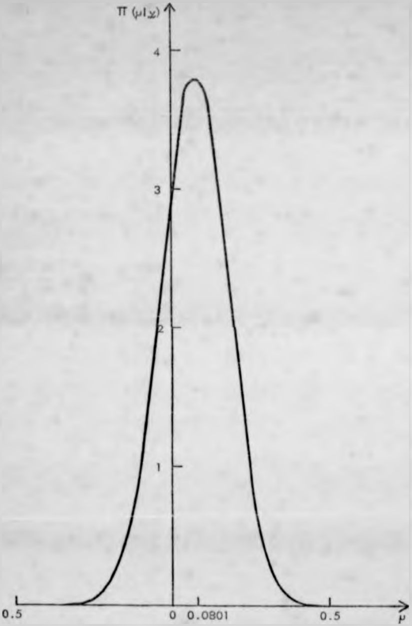


Figure 2.4 Marginal posterior density of μ for the generated data set when the prior distribution for μ is $N(0.000, 0.0148)$

Prior mean of μ	Prior variance of μ	Marginal posterior mean of μ	Variance of marginal posterior distribution of μ	Value of μ at marginal posterior mode	Value of at mode of joint posterior density $\pi(\alpha,\beta,\mu \mid y)$
0.000	0.500	0.207	0.0338	0.236	0.229
0.500	0.500	0.253	0.0313	0.264	0.257
1.000	0.500	0.270	0.0313	0.280	0.265
0.000	0.0298	0.110	0.0172	0.120	0.114
0.500	0.0298	0.369	0.0146	0.369	0.368
1.000	0.0298	0.632	0.0182	0.619	0.621
0.000	0.0149	0.0719	0.0111	0.0763	0.0723
0.500	0.0148	0.412	0.00974	0.411	0.410
1.000	0.0149	0.777	0.0131	0.769	0.772

Table 2.2 Features of some posterior distributions using the generated data set for varying prior distributions.

Parameters of the prior distribution		Parameters of the approximate normal posterior distribution	
Mean μ_0	Variance Σ_{33}	Mean μ_2	Variance σ_2^2
0.000	0.500	0.229	0.0281
0.500	0.500	0.257	0.0281
1.000	0.500	0.288	0.0281
0.000	0.0298	0.122	0.0149
0.500	0.0298	0.372	0.0149
1.000	0.0298	0.622	0.0149
C.000	0.0149	0.0801	0.00983
0.500	0.0149	0.414	0.00993
1.000	0.0149	0.748	0.00983

Table 2.3 Parameters of the approximate normal posterior distribution using the generated data set for varying prior distributions.

2.5 A Generated Data Set

In this section we shall illustrate the ideas laid out in the previous sections with the aid of an artificially generated data set. Data for a 4 - point assay with 8 measurements at each point were constructed with the following parameter values:

$$\alpha = -1.0,$$
$$\beta = 1.0,$$
$$\mu = 0.5,$$
$$\sigma^2 = 0.2.$$

The log doses were 1.0 and 2.0 for the test preparation and 1.5 and 2.5 for the standard preparation. The data are given in Table 2.1.

Taking the prior distributions to be uniform for α and β and $N(\mu_o, \Sigma_{33})$ for μ, the posterior density of μ is

$$\pi(\mu|y) = (Sxx+2\mu Sxz+\mu^2 Szz)^{-\frac{1}{2}} \exp\left\{\frac{(Sxy+\mu Syz)^2}{2\sigma^2(Sxx+2\mu Sxz+\mu^2 Szz)}\right\}$$

$$\times \exp -\frac{(\mu-\mu_o)^2}{2\Sigma_{33}}$$

Using large sample theory the approximate posterior distribution of μ is $N(0.243, 0.0298)$. For various prior distributions of μ the constant of integration was found numerically using Gauss-Hermite quadrature as described by Froberg (1985). Some examples of the resulting posterior densities are illustrated in Figures 2.1 - 2.4. The values of 0.0288 and 0.0149 for the prior variance of μ are intended to represent situations where the prior information carries approximately the same amount of information as the data, and approximately twice as much information as the data. For each of the prior distributions considered the value of μ at the mode of the joint posterior density of α, β and μ, the value of μ at the mode of the marginal posterior density of μ, and the mean and variance of the marginal posterior distribution of μ were calculated. The results are given in Table 2.2. The marginal posterior mean of μ is theoretically the best point estimate of μ, but we can see that in this case both the value of μ at the mode of the marginal density of μ and its value at the mode of the joint posterior density of α, β, and μ are good approximations to the marginal posterior mean. Of these two modal approximations the one based on the marginal posterior distribution should on theoretical grounds be the better one, although for this data set the estimate based on the joint distribution is closer to the marginal posterior mean for almost all the prior distributions considered.

On inspection the densities illustrated in Figures 2.1 - 2.4 look as if they may not be very different from normal densities. This raises the question as to whether they can be reasonably approximated by normal densities. If satisfactory approximations could be found it might be possible to apply them without access to a computer. The density corresponding to the large sample approximate distribution is illustrated in Transparency 1 inside the back cover. Comparison of the transparency with Figures 2.1 - 2.4 shows this density to be a reasonable approximation to the small sample density only when there is little prior information available. A more useful approximation might be obtained by combining the prior information with the approximate large sample distribution in some way. Suppose the approximate large sample distribution of μ for a data set is $N(M, S^2)$, and suppose we treat the experimental data as if it were a single observation M from a normal distribution with variance S^2. The posterior distribution of μ would then be $\mu \sim N(\mu_2, \sigma_2^2)$ where

$$\mu_2 = \frac{\dfrac{M}{S^2} + \dfrac{\mu_0}{\Sigma_{33}}}{1/S^2 + 1/\Sigma_{33}}$$

and

$$\sigma_2^2 = \frac{1}{{}^1/_{S^2} + {}^1/_{\Sigma_{33}}}$$

The posterior means and variances which this approximation gives for our data set with various prior distributions are given in Table 2.3. Also normal densities with variances corresponding to the situations illustrated in Figures 2.1 - 2.4 are illustrated in Transparencies 2 - 4. For this data set the approximate procedure outlined above seems to give reasonably good results. We regret to say, however, that we have been unable to justify it theoretically.

Chapter 3. Use of the Prior Distribution in Designing the Experiment.

3.1 Introduction

When we have available prior information about the parameters in an assay, it seems reasonable that this information should influence the doses used.

The use of prior distributions in designing experiments for parameter estimation in non-linear models has been discussed by Draper & Hunter (1967). We shall now give a short summary of the relevant parts of this paper. Suppose we wish to make n observations of the form

$$y_i = f(\underline{x}_i, \underline{\theta}) + \varepsilon_i, \ (i = 1, 2 \ldots n)$$

where the ε_i's are independently normally distributed with zero mean and variance σ^2. $\underline{x} = [x_1, x_2, \ldots x_k]^T$ is a vector of k variables, $\underline{\theta} = (\theta_1, \theta_2, \ldots \theta_p]^T$ is a vector of p parameters to be estimated, and $f(\underline{x}, \underline{\theta})$ is a non-linear function of $\underline{x}$ & $\underline{\theta}$. Suppose we also have available prior information about the $\underline{\theta}$'s in the form of a multivariate normal distribution with mean $\underline{\theta}$ and covariance matrix $\underline{\Sigma}$.

We should like to choose the n points $\underline{x}_i (i=1, 2, \ldots n)$ to obtain the best posterior distribution. The criterion for best is taken to be to maximize the final posterior density both with respect to $\underline{\theta}$ and $\underline{x}_i (i=1, 2, \ldots n)$. By approximating $f(\underline{x}_i, \underline{\theta})$ by the first two terms of its Taylor expansion about $\underline{\theta}_n$, the maximum likelihood estimator of $\underline{\theta}$ after the experiment has been carried out, the best design is found to be that which maximizes

$$|\underline{X}^T\underline{X} + \sigma^2\underline{\Sigma}^{-1}|$$

with respect to $\underline{x}_i (i=1, 2, \ldots n)$, where the $(i,j)^{th}$ element of $\underline{X}$ is

$$\left.\frac{\partial f(\underline{x}_i, \underline{\theta})}{\partial \theta_j}\right|_{\underline{\theta}=\underline{\theta}_n}$$

so the $(j,k)^{th}$ element of $\underset{\sim}{X}^T \underset{\sim}{X}$ is

$$\sum_{i=1}^{n} \left[\frac{\partial f(\underset{\sim}{x}_i, \underset{\sim}{\theta})}{\partial \theta_j} \cdot \frac{\partial f(\underset{\sim}{x}_i, \underset{\sim}{\theta})}{\partial \theta_k} \right] \Bigg|_{\underset{\sim}{\theta} = \underset{\sim}{\theta}_n}$$

Since $\hat{\underset{\sim}{\theta}}_n$ is not available before the experiment is performed, we have to approximate $\hat{\underset{\sim}{\theta}}_n$ by $\underset{\sim}{\theta}_o$, thus obtaining a practically applicable criterion.

3.2. Application to Parallel Line Bioassay

In using this procedure to design a parallel line bioassay we shall use the model as stated at the beginning of chapter 2.

In this particular application a further constraint will be imposed by the biological system on which the assay is performed, because the assay is restricted to lie in the linear part of the log-dose response curve. We shall assume that the log-dose response response curve is linear for both test and standard preparations for responses lying between two particular values which we estimate to be y_1 & y_2. We must try and restrict the doses used so that the responses will lie between these two values. We have to decide on the doses before carrying out the assay, and so we must rely on our prior information in doing this. Consequently we shall choose the points (x_i, z_i) such that

$$y_1 \leq \alpha_0 + \beta_0 \mu_0 z_i + \beta_0 x_i \leq y_2 \ , \quad i=1, 2, \ldots n \ . \qquad (3.1)$$

The region which satisfies these constraints is a convex hull and we shall call it the feasible region.

To return to the optimizing criterion of Draper & Hunter, in this application $f(x, \underset{\sim}{\theta}) = \alpha + \beta\mu z + \beta x$ and $\underset{\sim}{\theta}^T_o = (\alpha_o, \beta_o, \mu_o)$.

Hence

$$\underset{\sim}{X}^T\underset{\sim}{X}=\begin{vmatrix} n & (\mu_o\Sigma z_i+\Sigma x_i) & \beta_o\Sigma z_i \\ (\mu_o\Sigma z_i+\Sigma x_i) & \Sigma(\mu_o z_i+x_i)^2 & \beta_o(\mu_o\Sigma z_i^2+\Sigma z_i x_i) \\ \beta_o\Sigma z_i & \beta_o(\mu_o\Sigma z_i^2+\Sigma z_i x_i) & \beta_o^2\Sigma z_i^2 \end{vmatrix}.$$

Suppose n_T of the doses are on the test preparation, and the average of these log-doses is $\bar{x}_T$. Similarly n_S of the doses are on the standard preparation and the average of the log-doses is $\bar{x}_S$. Also let $\sigma^{-2}\underset{\sim}{S}=\underset{\sim}{\Sigma}$, and $\underset{\sim}{\Sigma}^{-1}=\begin{vmatrix} S^{11} & S^{12} & S^{13} \\ S^{12} & S^{22} & S^{23} \\ S^{13} & S^{23} & S^{33} \end{vmatrix}$.

In practice σ^2 would usually be unknown and so $\underset{\sim}{S}$ would have to be estimated rather than Σ. This notation gives

$$|\underset{\sim}{X}^T\underset{\sim}{X}+\sigma^2\underset{\sim}{\Sigma}^{-1}|=\begin{vmatrix} (n+S^{11}) & (n_T\bar{x}_T+n_S\bar{x}_S+n_T\mu_o+S^{12}) & (n_T\beta_o+S^{13}) \\ (n_T\bar{x}_T+n_S\bar{x}_S+n_T\mu_o+S^{12}) & (\Sigma x_i^2+2n_T\mu_o\bar{x}_T+n_T\mu_o^2+S^{22}) & (n_T\beta_o\bar{x}_T+n_T\beta_o\mu_o+S^{23}) \\ (n_T\beta_o+S^{13}) & (n_T\beta_o\bar{x}_T+n_T\beta_o\mu_o+S^{23}) & (n_T\beta_o^2+S^{33}) \end{vmatrix}$$

$$=(n_Tn_S\beta_o^2+nS^{33}-2n_T\beta_oS^{13}+n_T\beta_o^2S^{11}+S^{11}S^{33}-(S^{13})^2)\,\Sigma x_i^2$$

$$+(-n_S\beta_o^2-\beta_o^2S^{11}+2\beta_oS^{13}-S^{33})n_T^2\bar{x}_T^2$$

$$+(-n_T\beta_o^2-S^{33})n_S^2\bar{x}_S^2$$

$$+2(\beta_oS^{13}-S^{33})\,n_Tn_S\bar{x}_T\bar{x}_S$$

$$+2\{\mu_o S^{11}S^{33}-\mu_o(S^{13})^2+n_S\mu_o S^{33}-n_S\beta_o S^{23}-\beta_o S^{11}S^{23}+\beta_o S^{12}S^{13}-S^{12}S^{33}+S^{13}S^{23}\}$$

$$\times n_T\bar{x}_T$$

$$+2\{n_T\beta_o S^{23}+n_T\beta_o\mu_o S^{13}-n_T\beta_o{}^2S^{12}-n_T\mu_o S^{33}-S^{12}S^{33}+S^{13}S^{23}\}n_S\bar{x}_S \qquad (3.2)$$

+ terms not involving the x_i .

If we fix n_T at a particular positive integer no bigger than n, the above expression will be a convex function of the x_i if the matrix $\underset{\sim}{C}$ is positive definite, where

$$\underset{\sim}{C} = \left[\begin{array}{c:c} p\underset{\sim}{I}_{n_T}+q\underset{\sim}{J}_{n_T} & s\underset{\sim}{J}_{n_T\times n_S} \\ \hdashline s\underset{\sim}{J}_{n_S\times n_T} & p\underset{\sim}{I}_{n_S}+r\underset{\sim}{J}_{n_S} \end{array}\right]$$

and

$$p=n_Tn_S\beta_o{}^2+nS^{33}-2n_T\beta_o{}^2S^{11}+n_T\beta_o{}^2S^{11}+S^{11}S^{33}-(S^{13})^2$$

$$q=-n_S\beta_o{}^2-\beta_o{}^2S^{11}+2\beta_o S^{13}-S^{33},$$

$$r=-n_T\beta_o{}^2-S^{33},$$

$$s=\beta_o S^{13}-S^{33},$$

$\underset{\sim}{I}_j$ is the $j\times j$ identity matrix ,
$\underset{\sim}{J}_{j\times k}$ is the $j\times k$ matrix whose elements are all 1.

$\underset{\sim}{C}$ will be positive definite if and only if all its principal minors are positive. This implies two sets of conditions:

1. $p+mq>0,\qquad 0\leqslant m\leqslant n_T$
2. $\{(p+lr)(p+n_Tq)-n_Tls^2\}>0\ ,\quad 1\leqslant l\leqslant n_S$.

Considering the first set of conditions,

$$p+mq=(n_T-m)(n_S\beta_o{}^2+\beta_o{}^2S^{11}-2\beta_o S^{13}+S^{33})+n_S S^{33}+S^{11}S^{33}-(S^{13})^2 .$$

From its definition, $\underset{\sim}{S}=\sigma^{-2}\underset{\sim}{\Sigma}$ where $\underset{\sim}{\Sigma}$ is the covariance matrix of a multivariate normal distribution and σ^2 is a variance. Hence $\underset{\sim}{S}$ will be positive definite and consequently $\underset{\sim}{S}^*=\begin{bmatrix} S^{11} S^{13} \\ S^{13} S^{33} \end{bmatrix}$ will also be positive definite. This implies that

$$\beta_o{}^2S^{11}-2\beta_oS^{13}+S^{33}=(\beta_o-1)\underset{\sim}{S}^*(\beta_o-1)^T,\ S^{33},\ \text{and}\ \{S^{11}S^{33}-(S^{13})^2\}=|\underset{\sim}{S}^*|$$

are all strictly positive; so it follows that $p+mq$ will be strictly positive for $m=0, 1, \ldots n_T$ and the first set of conditions is always satisfied.

Considering the second set of conditions, on substitution $(p+lr)\times(p+n_Tq) - n_Tls^2=(n_S-l)S^{33}+S^{11}S^{33}-(S^{13})^2$. This will be strictly positive for $l=1, 2, \ldots n_S$ from the positive definiteness of $\underset{\sim}{S}^*$.

Hence we have the result that for fixed n_T $|\underset{\sim}{X}^T\underset{\sim}{X}+\sigma^2\underset{\sim}{\Sigma}^{-1}|$ is a convex function of the x_i .

3.3 Maximization of $|\underset{\sim}{X}^T\underset{\sim}{X}+\sigma^2\underset{\sim}{\Sigma}^{-1}|$.

We can now apply the criterion of Draper & Hunter by first fixing the number of doses on each of the test and standard preparations and maximising the resulting expression for $|\underset{\sim}{X}^T\underset{\sim}{X}+\sigma^2\underset{\sim}{\Sigma}^{-1}|$. We can then consider the resulting maximum and maximize it with respect to n_T.

First let us fix the number of doses on the test preparation at n_T, leaving $(n-n_T)$ doses on the standard preparation. Maximization of $|\underset{\sim}{X}^T\underset{\sim}{X}+\sigma^2\underset{\sim}{\Sigma}^{-1}|$ over the feasible region then amounts to maximizing a convex function over a convex hull. The maximum will therefore lie in a vertex of the feasible region. This means that for each of the two preparations the doses will lie at the ends of the permitted range. Suppose k_T doses of the test preparation and k_S doses of the standard preparation are at the highest permitted levels. Then from the constraints, 3.1, k_T of the x_i will take value $\left(\dfrac{y_2-\alpha_o-\mu_o}{\beta_o}\right)$.

$(n_T - k_T)$ of them will take value $\left(\dfrac{y_1-\alpha_o}{\beta_o} - \mu_o\right)$, k_S of them will take value $\left(\dfrac{y_2-\alpha_o}{\beta_o}\right)$ and $(n_S - k_S)$ of them will take value $\left(\dfrac{y_1-\alpha_o}{\beta_o}\right)$. In preparation for writing $|\underset{\sim}{X}^T\underset{\sim}{X}+\sigma^2\underset{\sim}{\Sigma}^{-1}|$ as given by 3.2 as a function of k_T and k_S, if we let $y_2 - y_1 = r$, we have

$$\Sigma x_i^2 = k_T\frac{r}{\beta_o}\left\{2\left(\frac{y_1-\alpha_o}{\beta_o}-\mu_o\right)+\frac{r}{\beta_o}\right\}+k_S\frac{r}{\beta_o}\left\{\frac{2(y_1-\alpha_o)}{\beta_o}+\frac{r}{\beta_o}\right\}+n\left\{\frac{y_1-\alpha_o}{\beta_o}\right\}^2$$

$$-2\mu_o n_T\left\{\frac{y_1-\alpha_o}{\beta_o}\right\}+n_T\mu_o^2 .$$

$$n_T^2\bar{x}_T^2 = k_T^2\frac{r^2}{\beta_o^2}+2k_Tn_T\frac{r}{\beta_o}\left\{\frac{y_1-\alpha_o}{\beta_o}-\mu_o\right\}.$$

$$n_S^2\bar{x}_S^2 = \frac{k_S^2r^2}{\beta_o^2}+2k_Sn_S\frac{r}{\beta_o}\left\{\frac{y_1-\alpha_o}{\beta_o}\right\}+n_S^2\left\{\frac{y_1-\alpha_o}{\beta_o}\right\}^2 .$$

$$n_Tn_S\bar{x}_T\bar{x}_S = k_Tk_S\frac{r^2}{\beta_o^2}+k_Tn_S\frac{r}{\beta_o}\left\{\frac{y_1-\alpha_o}{\beta_o}\right\}+k_Sn_T\frac{r}{\beta_o}\left\{\frac{y_1-\alpha_o}{\beta_o}-\mu_o\right\}$$

$$+n_Sn_T\left\{\frac{y_1-\alpha_o}{\beta_o}\right\}\left\{\frac{y_1-\alpha_o}{\beta_o}-\mu_o\right\} .$$

$$n_T\bar{x}_T = k_T\frac{r}{\beta_o}+n_T\left\{\frac{y_1-\alpha_o}{\beta_o}-\mu_o\right\} .$$

$$n_S\bar{x}_S = k_S\frac{r}{\beta_o}+n_S\left\{\frac{y_1-\alpha_o}{\beta_o}\right\} .$$

Inserting these expressions into 3.2, we have

$$|\underset{\sim}{X}^T\underset{\sim}{X}+\sigma^2\underset{\sim}{\Sigma}^{-1}| = k_T^2\frac{r^2}{\beta_o^2}\{-n_S\beta_o^2-\beta_o^2S^{11}+2\beta_o S^{13}-S^{33}\}$$

$$+k_S^2\frac{r^2}{\beta_o^2}\{-n_T\beta_o^2-S^{33}\}+2k_Tk_S\frac{r^2}{\beta_o^2}\{-S^{23}+\beta_o S^{12}\}$$

$$+k_T\left[\frac{r^2}{\beta_o^2}\{n_Tn_S\beta_o^2+nS^{33}-2n_T\beta_oS^{13}+n_T\beta_o^2S^{11}+S^{11}S^{33}-(S^{13})^2\}\right.$$

$$+\frac{2r}{\beta_o}\left\{\frac{y_1-\alpha_o}{\beta_o}\right\}\{S^{11}S^{33}-(S^{13})^2+n_S\beta_o S^{13}\}$$

$$\left.+\frac{2r}{\beta_o}\{-n_S\beta_oS^{23}-\beta_oS^{11}S^{23}+S^{13}S^{23}+\beta_oS^{12}S^{13}-S^{13}S^{33}\}\right]$$

$$+k_S\left[\frac{r^2}{\beta_o^2}\{n_Tn_S\beta_o^2+nS^{33}-2n_T\beta_oS^{13}+n_T\beta_o^2S^{11}+S^{11}S^{33}-(S^{13})^2\}\right.$$

$$+\frac{2r}{\beta_o}\left\{\frac{y_1-\alpha_o}{\beta_o}\right\}\{n_T\beta_o^2S^{11}-n_T\beta_oS^{13}+S^{11}S^{33}-(S^{13})^2\}$$

$$\left.+\frac{2r}{\beta_o}\{-n_T\beta_o^2S^{12}+n_T\beta_oS^{23}-S^{12}S^{33}+S^{13}S^{23}\}\right]$$

+ terms not involving k_T or k_S. (3.3)

Considering $|\underset{\sim}{X}^T\underset{\sim}{X}+\sigma^2\underset{\sim}{\Sigma}^{-1}|$ as a quadratic form in $(k_T,k_S)^T$, $|\underset{\sim}{X}^T\underset{\sim}{X}+\sigma^2\underset{\sim}{\Sigma}^{-1}|$ will be con:ave if the matrix $\underset{\sim}{H}$ is positive definite, Inserting these expressions into 3.2, we have

$$|\underset{\sim}{X}^T\underset{\sim}{X}+\sigma^2\underset{\sim}{\Sigma}^{-1}| = k_T^2\frac{r^2}{\beta_o^2}\{-n_S\beta_o^2-\beta_o^2S^{11}+2\beta_oS^{13}-S^{33}\}$$

$$+k_S^2\frac{r^2}{\beta_o^2}\{-n_T\beta_o^2-S^{33}\}+2k_Tk_S\frac{r^2}{\beta_o^2}\{-S^{33}+\beta_oS^{13}\}$$

$$+k_T\left[\frac{r^2}{\beta_o^2}\{n_Tn_S\beta_o^2+nS^{33}-2n_T\beta_oS^{13}+n_T\beta_o^2S^{11}+S^{11}S^{33}-(S^{13})^2\}\right.$$

$$+\frac{2r}{\beta_o}\left\{\frac{y_1-\alpha_o}{\beta_o}\right\}\{S^{11}S^{33}-(S^{13})^2+n_S\beta_oS^{13}\}$$

$$\left.+\frac{2r}{\beta_o}\{-n_S\beta_oS^{23}-\beta_oS^{11}S^{23}+S^{13}S^{23}+\beta_oS^{12}S^{13}-S^{13}S^{33}\}\right]$$

$$+k_S\left[\frac{r^2}{\beta_o^2}\{n_Tn_S\beta_o^2+nS^{33}-2n_T\beta_oS^{13}+n_T\beta_o^2S^{11}+S^{11}S^{33}-(S^{13})^2\}\right.$$

$$+\frac{2r}{\beta_o}\left\{\frac{y_1-\alpha_o}{\beta_o}\right\}\{n_T\beta_o^2S^{11}-n_T\beta_oS^{13}+S^{11}S^{33}-(S^{13})^2\}$$

$$\left.+\frac{2r}{\beta_o}\{-n_T\beta_o^2S^{12}+n_T\beta_oS^{23}-S^{12}S^{33}+S^{13}S^{23}\}\right]$$

$$+\text{ terms not involving } k_T \text{ or } k_S . \qquad (3.3)$$

Considering $|\underset{\sim}{X}^T\underset{\sim}{X}+\sigma^2\underset{\sim}{\Sigma}^{-1}|$ as a quadratic form in $(k_T, k_S)^T$, $|\underset{\sim}{X}^T\underset{\sim}{X}+\sigma^2\underset{\sim}{\Sigma}^{-1}|$ will be concave if the matrix $\underset{\sim}{H}$ is positive definite.

where

$$\underset{\sim}{H} = \frac{r^2}{\beta_o^{\,2}} \begin{bmatrix} (n_S\beta_o^{\,2}+\beta_o^{\,2}S^{11}-2\beta_o S^{13}+S^{33}) & (S^{33}-\beta_o S^{13}) \\ (S^{33}-\beta_o S^{13}) & (n_T\beta_o^{\,2}+S^{33}) \end{bmatrix} .$$

For $\underset{\sim}{H}$ to be positive definite we need

1. $n_S\beta_o^{\,2}+\beta_o^{\,2}S^{11}-2\beta_o S^{13}+S^{33} > 0,$

2. $\beta_o^{\,2}\{n_T n_S\beta_o^{\,2}+n_T\beta_o^{\,2}S^{11}-2n_T\beta_o S^{13}+nS^{33}+S^{11}S^{33}-(S^{13})^2\} > 0.$

These conditions are both satisfied due to the positive definiteness of $\underset{\sim}{S}^*$.

It follows that $|\underset{\sim}{X}^T\underset{\sim}{X}+\sigma^2\underset{\sim}{\Sigma}^{-1}|$ will achieve its maximum at the solution of the two simultaneous linear equations

$$\frac{\partial}{\partial k_T}\,|\underset{\sim}{X}^T\underset{\sim}{X}+\sigma^2\underset{\sim}{\Sigma}^{-1}| = 0 ,$$

&

$$\frac{\partial}{\partial k_S}\,|\underset{\sim}{X}^T\underset{\sim}{X}+\sigma^2\underset{\sim}{\Sigma}^{-1}| = 0 .$$

From 3.3 this is the point

$$k_T = \frac{n_T}{2} - \frac{S^{23}}{r} + \frac{(y_1-\alpha_o+{}^r/_2)\,S^{13}}{r\beta_o} , \tag{3.4}$$

$$k_S = \frac{n_S}{2} + \frac{S^{23}}{r} - \frac{(y_1-\alpha_o+{}^r/_2)\,(S^{13}-\beta_o S^{11})}{r\beta_o} - \frac{\beta_o S^{12}}{r} .$$

Assuming the values obtained for k_T and k_S are such that k_T lies in the interval $[0,n_T]$ and k_S lies in the interval $[0,n_S]$ we can now substitute these values back into $|\underset{\sim}{X}^T\underset{\sim}{X}+\sigma^2\underset{\sim}{\Sigma}^{-1}|$

and we get

$$|\underset{\sim}{X}^T\underset{\sim}{X}+\sigma^2\underset{\sim}{\Sigma}^{-1}| = \left\{\frac{nr^2}{4}+(y_1-\alpha_0+{}^r/_2)^2S^{11}-2(y_1-\alpha_0+{}^r/_2)\beta_0S^{12}+\beta_0^{\,2}S^{22}\right\}\times$$

$$\left\{-n_T^{\,2}+n_T\left(n+S^{11}-\frac{2S^{13}}{\beta_0}\right)\right\} + \text{terms not involving } n_T\ .$$

This will have a turning point at

$$n_T=\frac{n}{2}+\frac{S_{11}}{2}-\frac{S^{13}}{\beta_0}\ . \tag{3.5}$$

Since $\underset{\sim}{S}$ is positive definite $\begin{bmatrix} S^{11} & S^{12} \\ S^{12} & S^{22}\end{bmatrix}$ will be positive definite also, and so $(y_1-\alpha_0+{}^r/_2)^2S^{11}-2(y_1-\alpha_0+{}^r/_2)\beta_0S^{12}+\beta_0^{\,2}S^{22}$

$$=\begin{bmatrix}(y_1-\alpha_0+{}^r/_2)\\ -\beta_0\end{bmatrix}^T\begin{bmatrix} S^{11} & S^{12} \\ S^{12} & S^{22}\end{bmatrix}\begin{bmatrix}(y_1-\alpha_0+{}^r/_2)\\ -\beta_0\end{bmatrix}$$

will be positive. Consequently the coefficient of $n_T^{\,2}$ in the above expression is negative, and the turning point is a maximum. Assuming the value of n_T at the turning point lies in the interval $[0,n]$ we can substitute it into the expressions for k_T and k_S to get

$$k_T=\frac{n}{4}+\frac{S^{11}}{4}+\left(\frac{y_1-\alpha_0}{r\beta_0}\right)S^{13}-\frac{S^{23}}{r}\ .$$

$$(n_T-k_T)=\frac{n}{4}+\frac{S^{11}}{4}-\left(\frac{y_1-\alpha_0}{r\beta_0}\right)S^{13}-\frac{S^{13}}{\beta_0}+\frac{S^{23}}{r}\ .$$

$$k_S = \frac{n}{4} + \frac{S^{11}}{4} + \left(\frac{y_1-\alpha_0}{r}\right) S^{11} - \frac{\beta_0 S^{12}}{r} - \left(\frac{y_1-\alpha_0}{r}\right)\frac{S^{13}+S^{23}}{r} \; , \tag{3.6}$$

$$(n_S-k_S) = \frac{n}{4} - \frac{3S^{11}}{4} - \left(\frac{y_1-\alpha_0}{r}\right) S^{11} + \frac{\beta_0 S^{12}}{r} + \left(\frac{y_1-\alpha_0}{r\beta_0}\right) S^{13} - \frac{S^{23}}{r} + \frac{S^{13}}{\beta_0}$$

Hence we have the result that the optimal design is to place k_T and k_S doses at the highest possible dose for the test and standard preparations respectively, and (n_T-k_T) and (n_S-k_S) doses at the lowest possible dose for the test and standard preparations, where k_T, k_S (n_T-k_T) & (n_S-k_S) are as given above.

This procedure does not guarantee to place an integral number of doses at each point in the design. To overcome this difficulty we suggest the pragmatic approach of setting n_T equal to that integer nearest to the value given by 3.5, and then using this integral value of n_T, finding k_T and k_S from 3.4 by the same method.

In order for the solution 3.6 to be meaningful, n_T must lie in the interval $[o, n]$, k_T in the interval $[o, n_T]$, and k_S in the interval $[o, n_S]$. This implies the following inequalities:

$$-n \leq S^{11} - \frac{2S^{13}}{\beta_0} \leq n \; ,$$

$$0 \leq \frac{n}{4} + \frac{S^{11}}{4} + \left(\frac{y_1-\alpha_0}{r\beta_0}\right) S^{13} - \frac{S^{23}}{r} \leq \frac{n}{2} + \frac{S^{11}}{2} - \frac{S^{13}}{\beta_0} \; , \tag{3.7}$$

$$0 \leq \frac{n}{4} + \frac{S^{11}}{4} + \left(\frac{y_1-\alpha_0}{r}\right) S^{11} - \frac{\beta_0 S^{12}}{r} - \left(\frac{y_1-\alpha_0}{r\beta_0}\right) S^{13} + \frac{S^{23}}{r} \leq \frac{n}{2} - \frac{S^{11}}{2} + \frac{S^{13}}{\beta_0}$$

It does not seem possible to interpret these inequalities in any detail for the general experiment. One case when they will all hold is when the elements of $\underset{\sim}{S}^{-1}$ are small compared with n, that is the elements of $\underset{\sim}{\Sigma}^{-1}$ are small compared with n/σ^2. This will occur when the prior information is rather diffuse when compared with the amount of information one hopes to gain from the experiment. It is quite possible to find examples where not all the inequalities hold. Suppose the optimal value for n_T given by 3.5 is greater than n. Intuitively this means that there is so much more prior information available about the standard preparation that even if we devoted the whole experiment to the test preparation we would still know less about it than about the standard preparation. A first suggestion would be to set n_T equal to n and then use 3.4 to find k_T. However, even in the case where a great deal is already known about the standard preparation it will rarely be desirable to carry out an assay where the standard preparation is not used at all. A possible compromise might be to use just two doses of the standard, one at each of the extreme dosage levels. Cases where n_T lies in the interval $[0,n]$ but k_T lies outside its permitted range might be more happily solved by setting k_T equal to 0 or n_T, whichever was appropriate. The same applies to k_S .

3.4 Two Examples

Suppose we wish to calibrate a relatively new test preparation with a well-known standard. Typically our prior knowledge about the test preparation will be vague compared with our prior knowledge about the standard preparation. However, just considering one preparation, our prior opinions about the response for different doses will be equally precise, or in other words the variance of our prior predictions of responses at different doses will be equal. Suppose we consider the following model:

1st stage: $y \sim N\{(\alpha^1+\beta\mu z+\beta(x-x_{MS})),\sigma^2\}$

2nd stage: $\begin{pmatrix}\alpha^1\\ \beta\\ \mu\end{pmatrix} \sim N\left\{\begin{pmatrix}\alpha_o^1\\ \beta_o\\ \mu_o\end{pmatrix}, \begin{bmatrix}\Sigma_1 & 0 & 0\\ 0 & \Sigma_2 & 0\\ 0 & 0 & \Sigma_3\end{bmatrix}\right\}.$

where x_{MS} is the mid-point of the permitted range of log-doses for the standard preparation. We need only consider the four extreme doses which figure in the optimal design. If we estimate α^1, β & μ by a_o^1, β_o & μ_o, and if we let x_{US} be the highest log-dose and x_{LS} the lowest log-dose in the permitted range for the standard, then our predicted response for the highest possible dose on the standard is $y = a_o^1+\beta_o(x_{US}-x_{MS})$ with variance $V(y) = \Sigma_1+(x_{US}-x_{MS})^2\,\Sigma_2$. The predicted response for the lowest possible dose on the standard is $y = a_o^1+\beta_o(x_{LS}-x_{MS})$ with variance $V(y) = \Sigma_1+(x_{LS}-x_{MS})^2\,\Sigma_2$. The two variances are equal. The predicted responses for the highest and lowest possible doses on the test preparation are the same as those for the standard. The variances are again equal, this time with value $\Sigma_1+(x_{US}-x_{MS})^2\Sigma_2+(\Sigma_2+\beta_o^2)\Sigma_3$. This is greater than the corresponding variance for the standard preparation by the quantity $(\Sigma_2+\beta_o^2)\Sigma_3$. It follows that this model describes the required situation.

This model is a special case of the more general model described in the previous sections of this chapter. To illustrate this we need to set $\alpha=\alpha^1-\beta x_{MS}$ and $a_o=\alpha_o^1-\beta_o x_{MS}$ in the general model. It follows from the first of these relations, and from the diagonal covariance matrix in this example, that we need to set

$$\begin{bmatrix} \Sigma_{11}\Sigma_{12}\Sigma_{13} \\ \Sigma_{12}\Sigma_{22}\Sigma_{23} \\ \Sigma_{13}\Sigma_{23}\Sigma_{33} \end{bmatrix} = \begin{bmatrix} (\Sigma_1+x_{MS}^2\,\Sigma_2) & -x_{MS}\Sigma_2 & 0 \\ -x_{MS}\Sigma_2 & \Sigma_2 & 0 \\ 0 & 0 & \Sigma_3 \end{bmatrix}.$$

From this the elements of $\underset{\sim}{S}^{-1} = \sigma^2\underset{\sim}{\Sigma}^{-1}$ are σ^2 $$\begin{bmatrix} \dfrac{1}{\Sigma_1} & \dfrac{x_{MS}}{\Sigma_1} & 0 \\ \dfrac{x_{MS}}{\Sigma_1} & \left(\dfrac{1}{\Sigma_2}+\dfrac{x_{MS}^2}{\Sigma_1}\right) & 0 \\ 0 & 0 & \dfrac{1}{\Sigma_3} \end{bmatrix}.$$

In terms of the constraints given by 3.1 $x_{MS} = \dfrac{y_1 + y_2 - 2\alpha_o}{2\beta_o}$

Substituting these values of the elements of S^{-1} in the general optimal design given by 3.6 we have $k_T = n_T - k_T = \dfrac{n}{4} + \dfrac{\sigma^2}{4\Sigma_1}$.

and $k_S = (n_S - k_S) = \dfrac{n}{4} - \dfrac{\sigma^2}{4\Sigma_1}$. Hence the optimal design in this example is to place $\dfrac{n}{4} + \dfrac{\sigma^2}{4\Sigma_1}$ doses at each of the extremities of the possible range for the test preparation and $\dfrac{n}{4} - \dfrac{\sigma^2}{4\Sigma_1}$ doses at each of the extremities of the possible range for the standard preparation. The inequalities given by 3.7 reduce to the single inequality $\Sigma_1 \geq \sigma^2/n$. If this is not satisfied it indicates that a priori a great deal is known about the standard preparation and one should devote all the available resources to exploring the test preparation.

A second commonly occurring situation is that the prior knowledge about test and standard preparations is symmetric in the sense that we know as much about one substance as we do about the other. We can model this situation as follows:

$$\text{1st stage: } y \sim N\left\{\left(\alpha^1 + \beta\mu(z - \tfrac{1}{2}) + \beta(x - x_{MST})\right), \sigma^2\right\}$$

$$\text{2nd stage: } \begin{pmatrix} \alpha^1 \\ \beta \\ \mu \end{pmatrix} \sim N\left\{\begin{pmatrix} \alpha_o^1 \\ \beta_o \\ \mu_o \end{pmatrix}, \begin{bmatrix} \Sigma_1 & 0 & 0 \\ 0 & \Sigma_2 & 0 \\ 0 & 0 & \Sigma_3 \end{bmatrix}\right\}$$

where x_{MST} is the average of the mid-points of the permitted range of log-doses for the two substances. The predicted responses for doses occurring in the optimal design are for the highest doses on both preparations $y = \alpha_o^1 - \tfrac{1}{2}\beta_o\mu_o + \beta_o(x_{US} - x_{MST})$

and for the lowest doses on both preparations $y=\alpha_o^1-\frac{1}{2}\beta_o\mu_o+\beta_o\left(x_{LS}-x_{MST}\right)$

All these four predictions have variance $\Sigma_1+\Sigma_2(x_{US}-x_{MST})^2+\frac{1}{4}(\Sigma_2+\beta_o^2)\Sigma_3$

We can relate the general model to this example by setting $\alpha=\alpha^1-\frac{\beta\mu}{2}-\beta x_{MST}$ and $\alpha_o=\alpha_o^1-\frac{\beta_o\mu_o}{4}-\beta_o x_{MST}$ in the general model.

From the first of these relations and from the diagonal form of the covariance matrix, it follows that we need in the general model

$$\begin{bmatrix} \Sigma_{11} & \Sigma_{12} & \Sigma_{13} \\ \Sigma_{12} & \Sigma_{22} & \Sigma_{23} \\ \Sigma_{13} & \Sigma_{23} & \Sigma_{33} \end{bmatrix} = \begin{bmatrix} \{\Sigma_1+(x_{MST}+\frac{1}{2}\mu_o)^2\Sigma_2-\frac{1}{4}\Sigma_2\Sigma_3+\frac{1}{4}\beta_o^2\Sigma_3\} & \{-(x_{MST}+\frac{1}{2}\mu_o)\Sigma_2\} & -\frac{1}{2}\beta_o\Sigma_3\} \\ -(x_{MST}+\frac{1}{2}\mu_o)\Sigma_2 & \Sigma_2 & 0 \\ -\frac{1}{2}\beta_o\Sigma_3 & 0 & \Sigma_3 \end{bmatrix}.$$

Hence the elements of $\underset{\sim}{S}^{-1}$ are

$$\frac{\sigma^2}{\Sigma_1-\frac{1}{4}\Sigma_2\Sigma_3} \begin{vmatrix} 1 & (x_{MST}+\frac{1}{2}\mu_o) & \frac{1}{2}\beta_o \\ (x_{MST}+\frac{1}{2}\mu_o) & \left(\frac{\Sigma_1-\frac{1}{4}\beta_o^2\Sigma_3}{\Sigma_2}\right) & \frac{1}{2}\beta_o(x_{MST}+\frac{1}{2}\mu_o) \\ \frac{1}{2}\beta_o & \frac{1}{2}\beta_o(x_{MST}+\frac{1}{2}\mu_o) & \left(\frac{\Sigma_1-\frac{1}{4}(x_{MST}+\frac{1}{2}\mu_o)^2\Sigma_2}{\Sigma_3}\right) \end{vmatrix}.$$

Substituting these values of the elements of $\underset{\sim}{S}^{-1}$ into the general optimal design given by 3.6 we have

$k_T=n_T=k_T=k_S=k_S=n_S=\frac{n}{4}$. Hence the optimal design in this case is to place one quarter of the available doses at each of the four extreme dose points. As one might expect from the general symmetry of the situation the inequalities given by 3.7 are always satisfied in this case.

Chapter 4. Analysis of a Single Assay With Unknown Residual Variance.

4.1 Model and Posterior Distributions

In chapter 2 we made the assumption that the residual variance was known. In practice this will rarely be the case so we now remove this unrealistic assumption and obtain a model which is suitable for the analysis of data. If the residual variance is unknown it will be a parameter in the model and consequently we shall need to specify a prior distribution for it. We shall use the relevant conjugate prior distribution which is that $\frac{\nu\lambda}{\sigma^2}$ has a χ^2-distribution on ν degrees of freedom

where ν and λ are known constants whose values depend on our prior knowledge about σ^2. The prior density of σ^2 will therefore be

$$\pi(\sigma^2|\nu,\lambda) \propto (\sigma^2)^{-\frac{\nu+2}{2}} \exp-\left\{\frac{\nu\lambda}{2\sigma^2}\right\}, \sigma^2 > 0 .$$

We shall assume that our prior knowledge about σ^2 is independent of our prior knowledge about the other parameters.

For a given set of assay results we can obtain the joint posterior density of the four parameters α, β, μ and σ^2 up to a multiplicative constant. We get

$$\pi(\alpha,\beta,\mu,\sigma^2|\underset{\sim}{y}) \propto (\sigma^2)^{-\frac{n+\nu+2}{2}} \exp-\tfrac{1}{2}\left[\frac{\Sigma y_i^2}{\sigma^2} + \frac{\nu\lambda}{\sigma^2} + \alpha^2\left(\frac{n}{\sigma^2} + \Sigma^{11}\right) + 2\alpha\beta\left(\frac{\mu\Sigma z_i}{\sigma^2} + \frac{\Sigma x_i}{\sigma^2} + \Sigma^{12}\right)\right.$$

$$+ \beta^2\left(\frac{\mu^2\Sigma z_i^2}{\sigma^2} + \frac{2\mu\Sigma x_i z_i}{\sigma^2} + \frac{\Sigma x_i^2}{\sigma^2} + \Sigma^{22}\right)$$

$$-2\alpha\left(\frac{\Sigma y_i}{\sigma^2} + \alpha_0\Sigma^{11} + \beta_0\Sigma^{12} - (\mu-\mu_0)\Sigma^{13}\right)$$

$$-2\beta\left(\frac{\mu\Sigma y_i z_i}{\sigma^2}+\frac{\Sigma x_i y_i}{\sigma^2}+\alpha_o\Sigma^{12}+\beta_o\Sigma^{22}-(\mu-\mu_o)\Sigma^{23}\right.$$

$$\left.\left.+\mu^2\Sigma^{33}-2\mu(\alpha_o\Sigma^{13}+\beta_o\Sigma^{23}+\mu_o\Sigma^{33})\right)\right] \qquad (4.1)$$

This is an obvious extension of the joint posterior density of α,β and μ for known σ^2 given by 2.2 . Its mode is at the point given by 2.3 where σ^2 is now given by

$$\sigma^2=\frac{\Sigma(y_i-\alpha-\beta\mu z_i-\beta x_i)^2+\nu\lambda}{n+\nu+2}$$

As in the case where σ^2 is known, we can integrate over α and β in 4.1 to obtain the posterior density of μ and σ^2 up to a multiplicative constant. We get

$$\pi(\mu,\sigma^2|\underset{\sim}{y})\propto(\sigma^2)^{-\frac{(n+\nu+2)}{2}}|\underset{\sim}{V}|^{\frac{1}{2}}\exp-\frac{1}{2}\left\{\frac{\Sigma y_i^2+\nu\lambda}{\sigma^2}+\frac{\mu^2\Sigma^{33}}{\sigma^2}-2\mu(\alpha_o\Sigma^{13}+\beta_o\Sigma^{23}+\mu_o\Sigma^{33})-\begin{bmatrix}a\\b\end{bmatrix}^T\underset{\sim}{V}\begin{bmatrix}a\\b\end{bmatrix}\right\} \qquad (4.2)$$

where a,b and $\underset{\sim}{V}$ are as given by 2.6 . We can also integrate over σ^2 in 4.1 to obtain the joint posterior density of α,β and μ :

$$\pi(\alpha,\beta,\mu|\underset{\sim}{y})\propto(\nu\lambda+\Sigma(y_i-\alpha-\beta\mu z_i-\beta x_i)^2)^{-\frac{n+\nu}{2}}\exp-\frac{1}{2}\begin{bmatrix}\alpha-\alpha_o\\\beta-\beta_o\\\mu-\mu_o\end{bmatrix}^T\underset{\sim}{\Sigma}^{-1}\begin{bmatrix}\alpha-\alpha_o\\\beta-\beta_o\\\mu-\mu_o\end{bmatrix}. \qquad (4.3)$$

We cannot in general perform analytically the integrations necessary to obtain the marginal posterior distribution of μ.

The large sample results obtained in section 2.3 carry over to the unknown residual variance case, except that now σ^2 is normally distributed with mean

$\hat{\sigma}^2 = \Sigma(y_i - \hat{\alpha} - \hat{\beta}\hat{\mu}z_i - \hat{\beta}x_i)^2/n$ and variance $2\hat{\sigma}^4/n$. Also in 3.1, the expression for the large sample variance of α, β and μ, $\hat{\sigma}^2$ replaces σ^2.

4.2 A Special Case.

If we consider the case where we have uniform prior distributions for α and β , the joint posterior distribution of μ and σ^2 as given by 4.2 becomes

$$\pi(\mu,\sigma^2|\underset{\sim}{y})\propto(\sigma^2)^{-\frac{(n+\nu)}{2}}(Sxx+2\mu Sxz+\mu^2Szz)^{-\frac{1}{2}}$$

$$\times\exp-\frac{1}{2\sigma^2}\left\{\nu\lambda+Syy-\frac{(Sxy+Syz)^2}{(Sxx+2\mu Sxz+\mu^2Szz)}\right\}\exp-\tfrac{1}{2}(\mu^2-2\mu\mu_0)\Sigma^{33}.$$

In this special case we can perform the necessary integration over σ^2 to obtain the marginal posterior distribution of μ up to a multiplicative constant. We get

$$\pi(\mu|\underset{\sim}{y})\propto(Sxx+2\mu Sxz+\mu^2Szz)^{-\frac{1}{2}}\left(\nu\lambda+Syy-\frac{(Sxy+\mu Syz)^2}{Sxx+2\mu Sxz+\mu^2Szz}\right)^{-\frac{(n+\nu-2)}{2}}$$

$$\times\exp-\tfrac{1}{2}(\mu^2-2\mu\mu_0)\Sigma^{33} \qquad (4.4)$$

Before proceding any further we can now show that provided $\Sigma_{33}<\infty$, and we have more than two observations, then a vague prior for σ^2 , that is one where $\nu=0$, does not cause the joint posterior density of α,β ,μ and σ^2 to be unnormed, whether or not we have uniform priors for α and β. We use the notation $\pi^*(\ )$ to indicate unnormalized density functions as calculated. The joint posterior density of α,β ,μ and σ^2 will be normed provided $\iiiint\pi^*(\alpha,\beta,\mu,\sigma^2|\underset{\sim}{y})d\alpha d\beta d\sigma^2d\mu<\infty$

We know that $\pi^*(\alpha,\beta,\mu|\sigma^2,y)\leqslant k\pi^*(\alpha,\beta,\mu|\underset{\sim}{y},\sigma^2;\Sigma_{11},\Sigma_{22}\to\infty)$ for some positive constant k, so

$$\iiiint\pi^*(\alpha,\beta,\mu,\sigma^2|\underset{\sim}{y},\nu=0,\Sigma_{33}<\infty)d\alpha d\beta d\sigma^2d\mu$$

$$\leqslant k\iiiint\pi^*(\alpha,\beta,\mu,\sigma^2|\underset{\sim}{y},\nu=0,\Sigma_{33}<0;\Sigma_{11},\Sigma_{22}\to\infty)d\alpha d\beta d\sigma^2d\mu \quad \Sigma_{33}<\infty$$

$$=\int (Sxx+2\mu Sxz+\mu^2Szz)^{-\frac{1}{2}}\left(Syy-\frac{(Sxy+\mu Syz)^2}{Sxx+2\mu Sxz+\mu^2Szz}\right)^{-\frac{(n-2)}{2}} \exp-\tfrac{1}{2}(\mu^2-2\mu\mu_o)\Sigma^{33}d\mu$$

$$\leqslant \left(Sxx-\frac{S^2xz}{Szz}\right)^{-\frac{1}{2}}\left(Syy-\frac{Sxy+\hat{\mu}Syz}{\hat{\beta}}\right)^{-\frac{(n-2)}{2}} \int \exp-\tfrac{1}{2}(\mu^2-2\mu\mu_o)\Sigma^{33}d\mu$$

where $\hat{\mu}$ and $\hat{\beta}$ are the large sample estimates of μ and β
$< \infty$ since $\Sigma^{33} > 0$.

This result is not surprising since one would expect that the data contain, in some sense, quite a lot of information about the residual variance.

Let us return to the marginal posterior distribution of μ when $\Sigma_{11},\Sigma_{22} \to \infty$, given up to a constant by 4.4 . If our prior distribution for μ had been a t-distribution of a particular form instead of a normal distribution then we would be able to write down the posterior distribution of μ exactly rather than just up to a multiplicative constant. Using the notation $x \sim t_v(a,b)$ to indicate that $(x-a)/\sqrt{b}$ follows a t-distribution with v degrees of freedom, let the prior distribution for μ be

$$\mu \sim t_{n+v-4}\left\{\frac{-Sxz}{Szz}, \frac{(SxxSzz-S^2xz)}{(n+v-4)\ S^2zz}\right\} .$$

that is

$$\pi(\mu)\propto(Sxx+2\mu Sxz+\mu^2Szz)^{-\frac{(n+v-3)}{2}}$$

This is a nonsensical prior distribution in that the mean depends on the design to be used and the variance on the number of observations to be taken, however multiplying the above density with the likelihood and integrating over α and β we get

$$\pi(\mu|\underset{\sim}{y})\propto\left\{(v\lambda+Syy)(Sxx+2\mu Sxz+\mu^2Szz)-(Sxy+\mu Syz)^2\right\}^{-\frac{(n+v+2)}{2}} .$$

that is the posterior distribution of μ is $t_{n+v-3}(a,b)$, where $a = -\dfrac{\{(\nu\lambda+Syy)Sxz-SyzSxy\}}{\{(\nu\lambda+Syy)Szz-S^2yz\}}$.

and $b = \dfrac{1}{n+\nu-3}\left[\dfrac{\{(\nu\lambda+Syy)Sxx-S^2xy\}}{\{(\nu\lambda+Syy)Szz-S^2yz\}} - \dfrac{\{(\nu\lambda+Syy)Sxz-SxySyz\}^2}{\{(\nu\lambda+Syy)Szz-S^2yz\}^2}\right]$.

Hence the posterior mean of μ is a, and its posterior variance is $(n+\nu-3)b/(n+\nu-5)$.

In the case of vague prior knowledge for σ^2 these simplify to

$$\frac{SxySyz-SxzSyy}{SyySzz-S^2yz}$$

for the mean, and

$$\frac{Syy(SxxSyySzz-SxxS^2yz-SyyS^2xz-SzzS^2xy+2SxySxzSyz)}{(n-5)\ (SyySzz-S^2yz)^2}$$

for the variance. These results do not seem to correspond in any simple way to the large sample results, and the result appears to be of no practical value.

4.3 Estimation of Log Potency Ratio

Suppose we are in the position of uniform prior knowledge for α and β . The way to proceed is then clear. We can obtain the marginal posterior distribution of μ up to a multiplicative constant, as given by 4.4 , and with the help of one-dimensional numerical integrations we can obtain the posterior mean of μ and a confidence interval for it.

Unfortunately, the above will rarely be the case, and we shall have to resort either to more complex numerical techniques or to approximations. An exact numerical treatment would find the marginal posterior density of μ numerically from the joint posterior density of μ and σ^2 , as given by 4.2 , and then base inferences and decisions concerning μ on this numerical density.

This procedure requires a two-dimensional numerical integration. Such integrations are quite possible as will be demonstrated in section 4.5, however the computing power required is considerable, possibly more than might be available to a laboratory carrying out bioassays. In addition we have not found any satisfactory computer packages that will carry out numerical integrations in more than one-dimension. As a result of this we feel that approximations which require fewer computing facilities are worth considering.

Suppose we have available a certain amount of prior knowledge about α and β , but not a great deal. One possibility would be to disregard this information and proceed as in the first paragraph of this section. We shall demonstrate in section 4.4 that the posterior density for μ converges uniformly to the posterior density for μ given uniform prior distributions for α and β, as prior knowledge about α and β becomes more and more vague.

If there is substantial prior knowledge about α and β then the above approximation will not be satisfactory since it neglects a substantial amount of information. In this case there are two possible types of approach.

The first is to estimate μ by its value at the mode of a joint density. There are several joint densities to choose from, for example $\pi(\alpha,\beta,\mu,\sigma^2|y)$, $\pi(\alpha,\beta,\mu|y)$ and $\pi(\mu,\sigma^2|y)$. Of these one would expect the mode of $\pi(\mu,\sigma^2|y)$ to be the best approximation to the marginal posterior mean of μ since it is based on the joint distribution of two parameters rather than three or four. All these modal estimators suffer from the defect that there is no obvious confidence interval that can be associated with them, unless the assays are large enough for the joint densities to be approximately normal.

The second type of approach is based on a suggestion by Box & Tiao (1973). The data should contain quite a lot of information about σ^2 , and consequently the density $\pi(\sigma^2|y)$ should be reasonably sharp, with most of its probability mass concentrated over a small region about its marginal mode, $\hat{\sigma}^2$ say. Consequently, integrating over σ^2 in $\pi(\mu,\sigma^2|y)$ will be nearly equivalent to assigning the modal value to σ^2 in $\pi(\mu|\sigma^2,y)$. Unfortunately we cannot obtain $\hat{\sigma}^2$ analytically. We can, however, obtain it numerically by carrying out a series of one-dimensional numerical integrations. If this is not possible, due to restrictions on the use of computing time, one could approximate $\hat{\sigma}^2$ by the value of σ^2 at the mode of $\pi(\mu,\sigma^2|y)$. This type of approach leads to an approximate numerical posterior density for μ from which the posterior mean and a confidence interval could be estimated.

4.4 An Argument Supporting an Approximation Suggested in Section 4.3.

In this section we shall show that, as prior knowledge about α and β becomes more and more vague, the posterior density of μ converges uniformly to the posterior density of μ assuming uniform prior distributions for α and β as given by 4.4.

We shall assume throughout that S_{xx}, S_{yy} and S_{zz} are strictly positive, and that the number of observations made is greater than two.

Let

$$f=\{nX(\mu)\}^{-\frac{1}{2}}(\sigma^2)^{-\frac{(n+\nu)}{2}}\exp-\frac{1}{2\sigma^2}\left\{\nu\lambda+Syy-\frac{Y^2(\mu)}{X(\mu)}\right\}\exp-\tfrac{1}{2}(\mu^2-2\mu\mu_o)\Sigma^{33}, \quad (4.5)$$

and let

$$f_m=\{nX(\mu)\}^{-\frac{1}{2}}(\sigma^2)^{-\frac{(n+\nu)}{2}}\exp-\frac{1}{2\sigma^2}\{\nu\lambda+\Sigma y_i^{\,2}\}\exp-\tfrac{1}{2}\left\{\mu^2\Sigma^{33}-2\mu\left(\mu_o\Sigma^{33}+\frac{[illegible]}{m}+\frac{[illegible]}{m}\right)\right\}$$

$$\times\left\{1+\frac{\sigma^2}{m}\frac{W(\mu)}{X(\mu)}+\frac{\sigma^4 Z}{m^2 nX(\mu)}\right\}^{-\frac{1}{2}}\exp\tfrac{1}{2}\begin{vmatrix}a_m\\ b_m\end{vmatrix}^T V_m\begin{vmatrix}a_m\\ b_m\end{vmatrix}, \quad m=1,2,3,\ldots, \qquad (4.6)$$

where $W(\mu)=\Sigma^{11}(\Sigma x_i^{\,2}+2\mu\Sigma x_iz_i+\mu^2\Sigma z_i^{\,2})-2\Sigma^{12}\Sigma(x_i+\mu z_i)+n\Sigma^{22}$,

$X(\mu)=Sxx+2\mu Sxz+\mu^2 Szz$,

$Y(\mu)=Sxy+\mu Syz$,

$Z \quad =\Sigma^{11}\Sigma^{22}-(\Sigma^{12})^2$,

$$a_m \quad =\frac{n\bar{y}}{\sigma^2}+\frac{\alpha_o\Sigma^{11}}{m}+\frac{\beta_o\Sigma^{12}}{m}-\frac{(\mu-\mu_o)\Sigma^{13}}{m}$$

$$b_m \quad =\frac{\Sigma x_iy_i}{\sigma^2}+\frac{\mu\Sigma y_iz_i}{\sigma^2}+\frac{\beta_o\Sigma^{22}}{m}+\frac{\alpha_o\Sigma^{12}}{m}-\frac{(\mu-\mu_o)\Sigma^{23}}{m} \quad ,$$

$$V_m=\begin{bmatrix}\left(\frac{n}{\sigma^2}+\frac{\Sigma^{11}}{m}\right) & \left(\frac{\Sigma x_i}{\sigma^2}+\frac{\mu\Sigma z_i}{\sigma^2}+\frac{\Sigma^{12}}{m}\right)\\ \left(\frac{\Sigma x_i}{\sigma^2}+\frac{\mu\Sigma z_i}{\sigma^2}+\frac{\Sigma^{12}}{m}\right) & \left(\frac{\Sigma x_i^{\,2}}{\sigma^2}+\frac{2\mu\Sigma x_iz_i}{\sigma^2}+\frac{\mu^2\Sigma z_i^{\,2}}{\sigma^2}+\frac{\Sigma^{22}}{m}\right)\end{bmatrix}^{-1} .$$

This is equivalent to considering a sequence of prior distributions for α, β and μ whose variance matrices have inverses

$$\Sigma_m^{-1} = \begin{bmatrix} \frac{\Sigma^{11}}{m} & \frac{\Sigma^{12}}{m} & \frac{\Sigma^{13}}{m} \\ \frac{\Sigma^{12}}{m} & \frac{\Sigma^{22}}{m} & \frac{\Sigma^{23}}{m} \\ \frac{\Sigma^{13}}{m} & \frac{\Sigma^{23}}{m} & \frac{\Sigma^{33}}{m} \end{bmatrix} , \quad m = 1,2,3,\ldots$$

Every matrix in this sequence is positive definite if the first member Σ_1^{-1} is positive definite.

We wish to show that for all $\epsilon > 0$, there exists an M such that for all $m > M$,

$$\left| \int_{\sigma^2=0}^{\infty} f d\sigma^2 - \int_{\sigma^2=0}^{\infty} f_m d\sigma^2 \right| < \epsilon$$

for all μ.

It will be enought to show that there exist M and $\delta > 0$ such that

$$\text{(i)} \quad \left| \int_{\delta}^{\infty} f d\sigma^2 \right| < \epsilon/3 ,$$

$$\text{(ii)} \quad \left| \int_{\delta}^{\infty} f_m d\sigma^2 \right| < \epsilon/3 ,$$

$$\text{(iii)} \quad \left| \int_{0}^{\delta} f d\sigma^2 - \int_{0}^{\delta} f_m d\sigma^2 \right| < \epsilon/3 ,$$

for all $m > M$ and all μ.

We shall consider these three points in turn.

(i) Since $X(\mu) = Sxx+2\mu Sxz+\mu^2 Szz \geqslant Sxx-S^2xz/Szz > 0$, we know that

$$\{X(\mu)\}^{-\frac{1}{2}} \leqslant \left\{Sxx-\frac{S^2xz}{Szz}\right\}^{-\frac{1}{2}}$$

for all μ. It can be shown that $\{v\lambda+Syy-Y^2(\mu)/X(\mu)\} > 0$, so

$$\left|\exp-\frac{1}{2\sigma^2}\left\{v\lambda+Syy-\frac{Y^2(\mu)}{X(\mu)}\right\}\right| < 1$$

for all μ and all $\sigma^2 \in [\delta,\infty)$ Since $(\mu-\mu_o)^2\Sigma^{33} \geqslant 0$,

$$\left|\exp-\tfrac{1}{2}(\mu^2-2\mu\mu_o)\Sigma^{33}\right| \leqslant \left|\exp\tfrac{1}{2}\mu^2{}_o\Sigma^{33}\right|$$

for all μ. Hence, from 4.5 ,

$$\left|\int_\delta^\infty f d\sigma^2\right| \leqslant n^{-\frac{1}{2}}\left\{Sxx-\frac{S^2xz}{Szz}\right\}^{-\frac{1}{2}}\exp\tfrac{1}{2}\mu_o{}^2\Sigma^{33}\left|\int_\delta^\infty (\sigma^2)^{-\frac{(n+v)}{2}} d\sigma^2\right|$$

$$\leqslant n^{-\frac{1}{2}}\left\{Sxx-\frac{S^2xz}{Szz}\right\}^{-\frac{1}{2}}\exp\tfrac{1}{2}\mu_o{}^2\Sigma^{33}\cdot\frac{2}{(n+v-2)}\cdot\frac{1}{(\delta)^{\left(\frac{n+v-2}{2}\right)}}$$

$$< \varepsilon/3$$

for all sufficiently large δ .

(ii) Since n, $W(\mu)$, $X(\mu)$ and Z are all strictly positive,

$$1 > \left\{1+\frac{\sigma^2}{mn}\cdot\frac{W(\mu)}{X(\mu)}+\frac{\sigma^4}{m^2 n}\cdot\frac{Z}{X(\mu)}\right\}^{-\frac{1}{2}} > 0$$

for all m, all μ and all $\sigma^2 \in [\delta,\infty)$. Let

$$\xi(m,\mu,\sigma^2) = -\frac{1}{2}\mu^2\Sigma^{33}+\mu\left(\mu_o\Sigma^{33}+\alpha_o\frac{\Sigma^{13}}{m}+\beta_o\frac{\Sigma^{23}}{m}\right)+\tfrac{1}{2}\begin{bmatrix}a_m\\b_m\end{bmatrix}^T \underline{V}_m\begin{bmatrix}a_m\\b_m\end{bmatrix}.$$

It can easily be shown that for positive μ

$$\xi(m,\mu,\sigma^2) \leqslant \frac{\xi_4\mu^4+\xi_3\mu^3+\xi_2\mu^2+\xi_1\mu+\xi_0}{\frac{n}{\delta^2}\left\{ Sxx-\frac{S^2xz}{Szz} \right\}}$$

and for negative μ

$$\xi(m,\mu,\sigma^2) \leqslant \frac{\xi_4\mu^4-\xi_3\mu^3+\xi_2\mu^2-\xi_1\mu+\xi_0}{\frac{n}{\delta^2}\left\{ Sxx-\frac{S^2xz}{Szz} \right\}}$$

where ξ_0, ξ_1, ξ_2, and ξ_3 are constant independent of m, μ or σ^2 and

$$\xi_4 = -\Sigma^{23}\frac{n}{\sigma^4}S_{zz}-\frac{\Sigma z_i^2}{\sigma^2}\left\{ \frac{\Sigma^{11}\Sigma^{33}}{m}-\frac{(\Sigma^{12})^2}{m^2} \right\}$$

ξ_4 will be strictly negative for all m and all $\sigma^2\epsilon[\delta,\infty)$. hence $\xi(m,\mu,\sigma^2)$ will be bounded above, that is $\xi(m,\mu,\sigma^2) \leqslant \xi_{max}$ for all m, all μ and all $\sigma^2\epsilon[\delta,\infty)$. Lastly $\nu\lambda+\Sigma y_i^2 > 0$, so

$$\left| \exp -\frac{1}{2\sigma^2}(\nu\lambda+\Sigma y_i^2) \right| < 1$$

for all $\sigma^2\epsilon[\delta,\infty)$. Relating these inequalities to 4.8 we have

$$\left| \int_\delta^\infty f_m d\sigma^2 \right| \leqslant n^{-\frac{1}{2}}\left\{ Sxx-\frac{S^2xz}{Szz} \right\}^{-\frac{1}{2}} \exp \xi_{max} \int_\delta^\infty (\sigma^2)^{\frac{n+\nu}{2}} d\sigma^2$$

$$\leqslant n^{-\frac{1}{2}}\left\{ Sxx-\frac{S^2xz}{Szz} \right\}^{-\frac{1}{2}} \exp \xi_{max}\cdot\left(\frac{2}{n+\nu-2}\right)\frac{1}{(\delta)^{\frac{n+\nu-2}{2}}}$$

$$< \epsilon/_3$$

for all m and all μ for sufficiently large δ .

(iii) We would like to show that for any large δ there exists an M such that for all $m > M$

$$\left| \int_0^{\delta} f d\sigma^2 - \int_0^{\delta} f_m d\sigma^2 \right| < \varepsilon/_3$$

for all μ . This will be true if there exists an M such that for all $m > M$, $|f - f_m|$ $\frac{\delta\varepsilon}{3}$ for all μ and all $\sigma^2 \epsilon [o, \delta)$.

We shall need the result that

$$(\sigma^2)^{-\frac{(n+\nu)}{2}} \exp -\frac{1}{2\sigma^2}\left\{\nu\lambda + Syy - \frac{Y^2(\mu)}{X(\mu)}\right\} \leq \left\{\frac{\nu\lambda + Syy - \frac{Y^2(\hat{\mu})}{X(\hat{\mu})}}{n+\nu}\right\}^{-\frac{(n+\nu)}{2}} \exp -\frac{(n+\nu)}{2}$$

for all μ and all $\sigma^2 \epsilon [0, \delta]$, where $\bar{\mu}$ is the large sample mean of

Let us first consider the case where μ is either very large and positive or very large and negative. Applying identities already obtained to 4.5 and 4.6 we have that

$$|f - f_m| \leq A(\exp -\tfrac{1}{2}(\mu^2 - 2\mu\mu_o)\Sigma^{33} + \exp\zeta(m, \mu, \sigma^2)) ,$$

where

$$A = n^{-\frac{1}{2}}\left\{Sxx - \frac{S^2xz}{Szz}\right\}^{-\frac{1}{2}} \left\{\nu\lambda + Syy - \frac{Y^2(\hat{\mu})}{X(\hat{\mu})}\right\}^{-\frac{(n+\nu)}{2}} \exp -\frac{(n+\nu)}{2} ,$$

and

$$\zeta(m, \mu, \sigma^2) = -\frac{1}{2}\mu^2\Sigma^{33} + \mu(\mu_o\Sigma^{33} + \alpha_o\Sigma^{13} + \beta_o\Sigma^{23}) - \tfrac{1}{2}\begin{bmatrix} a_m \\ b_m \end{bmatrix}^T \underset{\sim}{V}_m \begin{bmatrix} a_m \\ b_m \end{bmatrix} + \frac{n\bar{y}^2}{2\sigma^2} + \frac{Y^2(\mu)}{2\sigma^2 X(\mu)}$$

For any δ and ε, $\exp -\tfrac{1}{2}(\mu^2 - 2\mu\mu_o)\Sigma^{33} < \delta\varepsilon/6A$ for all μ such that $|\mu| > K$, for sufficiently large K.

It can easily be shown that for positive μ

$$\zeta(m,\mu,\sigma^2) \leq \frac{\zeta_4\mu^4+\zeta_3\mu^3+\zeta_2\mu^2+\zeta_1\mu+\zeta_0}{n\left\{Sxx-\frac{S^2xz}{Szz}\right\}} ,$$

and for negative μ

$$\zeta(m,\mu,\sigma^2) \leq \frac{\zeta_4\mu^4-\zeta_3\mu^3+\zeta_2\mu^2-\zeta_1\mu+\zeta_0}{n\left\{Sxx-\frac{S^2xz}{Szz}\right\}} ,$$

where ζ_0, ζ_1, ζ_2 and ζ_3 are constants independent of m μ or σ^2 , and

$$\zeta_4=-nSzz-\Sigma z_1{}^2\left\{\frac{\Sigma^{11}\Sigma^{33}}{m}-\frac{(\Sigma^{13})^2}{m^2}\right\} .$$

ζ_4 will be strictly negative for all m, so $\zeta(m,\mu,\sigma^2) \to -\infty$ as $\mu \to \pm\infty$. Consequently for any δ and ϵ, $\left|\exp\frac{\zeta(m,\mu,\sigma^2)}{2}\right| < \delta\epsilon/6A$, for all m, for all $\sigma^2 \epsilon [0,\delta]$ and for all μ such that $|\mu|>K$, for sufficiently large K.

Combining these results we have that $|f-fm|<\delta\epsilon/_3$, for all m, for all $\sigma^2\epsilon[0,\delta]$ and for all μ such that $|\mu|>K$, for sufficiently large K.

Now let us consider μ lying in any finite interval $(-K,K)$. From 4.5 and 4.6 we have that

$$|f-fm| \leq A \exp\frac{\mu_0{}^2\Sigma^{33}}{2}\left| 1-\left\{1+\frac{\sigma^2W(\mu)}{mnX(\mu)}+\frac{\sigma^4Z}{m^2nX(\mu)}\right\}^{-\frac{1}{2}}\exp\frac{\mu}{m}(\alpha_0\Sigma^{13}+\beta_0\Sigma^{23}) \right.$$

$$\left. \times \exp\frac{\xi(m,\mu,\sigma^2)}{m}\right| \qquad (4.7)$$

where $\xi(m,\mu,\sigma^2) = \frac{m}{2}\left\{\begin{bmatrix}a_m\\ b_m\end{bmatrix}^T \underset{\sim}{V}_m\begin{bmatrix}a_m\\ b_m\end{bmatrix} -\frac{n\bar{y}}{\sigma^2} - \frac{Y^2(\mu)}{\sigma^2X(\mu)}\right\}$.

It can easily be shown that

$$\xi(m,\mu,\sigma^2) = \frac{\frac{\sigma^4R(\mu)}{m^2}+\frac{\sigma^2S(\mu)}{m}+T(\mu)+Y^2(\mu)\left\{\frac{\sigma^2Z}{mX(\mu)}-\frac{W(\mu)}{X(\mu)}\right\}}{\frac{\sigma^4}{m^2}+\frac{\sigma^2}{m}W(\mu)+nX(\mu)} ,$$

where $R(\mu)$, $S(\mu)$ and $T(\mu)$ are polynomials in μ with coefficients independant of m and σ^2. If we consider $\mu \epsilon (-K,K)$, then $\xi(m,\mu,\sigma^2)$ will be bounded both above and below for all m and all $\sigma^2 \epsilon [0,\delta]$. Hence for sufficiently large m

$$\exp \frac{1}{m}\{\mu(\alpha_o \Sigma^{13} + \beta_o \Sigma^{23}) + \xi(m,\mu,\sigma^2)\}$$

will be arbitrarily close to 1 for all $\mu \epsilon (-K,K)$. The same applies to

$$\left\{1 + \frac{\sigma^2 W(\mu)}{mnX(\mu)} + \frac{\sigma^4 Z}{m^2 nX(\mu)}\right\}^{-\frac{1}{2}}$$

Consequently, by examining 4.7 we can see that for sufficiently large M $|f - f_m| < \frac{\delta\varepsilon}{3}$ for all $m > M$, all $\sigma^2 \epsilon [0,\delta]$ and all μ in any finite interval $(-K,K)$.

4.5 An Example: Tobramycin Data

We shall now try out our ideas on some genuine data. Table 4.1 contains data from four replicate assays of the antibiotic tobramycin. The assays are carried out in petrie dishes in which there is a layer of agar gel containing organisms. Wells are cut in the agar gel and filled with a dose of the preparation of antibiotic. The antibiotic will then diffuse into the gel in a zone around the well and the organisms will be inhibited from growing in this zone. The size of the inhibition zone will depend on the amount of antibiotic in the well and the response variable measured is the area of the inhibition zone. In this section we shall consider the data from the first assay in isolation. The first task is to decide on values for the parameters of the prior distributions. We have used the following values for the second stage parameters:

$$\begin{pmatrix} \alpha_0 \\ \beta_0 \\ \mu_0 \end{pmatrix} = \begin{pmatrix} .29 \times 10^5 \\ .84 \times 10^4 \\ .00 \end{pmatrix}, \quad \underset{\sim}{\Sigma} = \begin{bmatrix} .600\times 10^5 & .000 & .000 \\ .000 & .200 \times 10^5 & .000 \\ .000 & .000 & .4000 \times 10^{-3} \end{bmatrix}$$

The values of α_0, β_0 and μ_0 were obtained from the data for the remaining three tobramycin assays, and $\underset{\sim}{\Sigma}$ was chosen so that we would expect the prior information to carry about half as much weight as the data in the analysis. We have set $\nu=\lambda=0$ in the prior density of σ^2 as the data should contain a substantial amount of information about σ^2.

We have followed several of the suggestions made in section 4.3 for the estimation of log potency ratio and our results are summarized in Table 4.2 and Figures 4.1 - 3. The different estimates of μ are all very similar. The mean and mode of the marginal distribution of μ are a little higher than the other estimates and the large sample mean is somewhat lower. The marginal density of μ and the two approximate marginal densities obtained in the first case by ignoring the prior information about α and β, and in the second case by assuming σ^2 is known and equal to its value at the mode of the joint distribution of μ and σ^2, are illustrated in Figures 4.1 - 4.3. The three densities can be compared using Transparencies 5 and 6. In this case either of the approximations seems quite satisfactory.

The calculations involved in obtaining these results were quite simple, using only a small amount of programming and standard computer routines in all but one case. This was in the calculation of the marginal posterior density of μ. Analytically we can only find the joint density of μ and σ^2 up to a multiplicative constant. Let this be $f(\mu,\sigma^2)$. The constant must be calculated numerically and this requires a two-dimensional numerical integration. We performed the calculation by using two one-dimensional numerical integrations hierarchically. We wished to estimate

$$I = \int_{\mu=-\infty}^{\infty} \int_{\sigma^2=0}^{\infty} f(\mu,\sigma^2)\,d\sigma^2 d\mu .$$

If we let $J(\mu) = \int_{\sigma^2=0}^{\infty} f(\mu,\sigma^2)\,d\sigma^2$, then $I = \int_{\mu=-\infty}^{\infty} J(\mu)\,d\mu$.

We carried out a series of one-dimensional integrations to evaluate $J(\mu)$ at those values of μ required to estimate the one dimensional integral

$$I = \int_{\mu=-\infty}^{\infty} J(\mu)\,d\mu .$$

The marginal posterior density of μ is then $J(\mu)/_I$, and we can use those values of $J(\mu)$ which we have already calculated to plot the density and also in finding the marginal posterior mean of μ. This method proved straight forward to program and gave answers of the required accuracy quite quickly.

	Standard Preparation			Test Preparation		
Dose	.054	.090	.15	.054	.090	.15
Assay 1	10072.	13668.	16681.	10113.	13564.	16463.
	10068.	13712.	16426.	10004.	13395.	16570.
	10041.	13814.	16848.	10198.	13674.	16757.
	9956.	13712.	16444.	10053.	13340.	16427.
	10104.	13938.	17012.	10305.	13654.	16308.
	10082.	14051.	16762.	10434.	13459.	16812.
Assay 2	10053.	13833.	16619.	10161.	13592.	16704.
	10074.	13377.	16520.	9933.	13580.	16370.
	9997.	13757.	16640.	10228.	13457.	16681.
	10151.	13730.	16482.	10112.	13536.	16640.
	10052.	13612.	16549.	10140.	13436.	16532.
	10049.	13829.	16690.	10165.	13423.	16666.
Assay 3	10079.	13545.	16566.	10245.	13949.	16937.
	10213.	13610.	16917.	10515.	14340.	17080.
	10097.	13319.	16603.	10239.	13824.	16905.
	10102.	13517.	17012.	10528.	14136.	16843.
	10030.	13369.	16708.	10259.	14079.	16833.
	10089.	13115.	16633.	10179.	13968.	16478.
Assay 4.	9954.	13346.	16750.	10383.	13869.	16745.
	9965.	13446.	16582.	10208.	13915.	16856.
	10102.	13102.	16720.	10183.	14140.	16487.
	9905.	13370.	16834.	10420.	13966.	16891.
	9987.	13661.	17099.	10664.	13931.	16831.
	10110.	13196.	16524.	10229.	13858.	16610.

Table 4.1 Data from four replicate assays of the antibiotic tobramycin.

	μ	α	β	σ^2
Mean of $\pi(\mu\|y)$	-.00978			
Mode of $\pi(\mu\|y)$	-.00941			
Mode of $\pi(\alpha,\beta,\mu\sigma^2\|y)$	-.0128	28900.	6370.	49000.
Mode of $\pi(\alpha,\beta,\mu\|y)$	-.0128	28900.	6370.	
Mode of $\pi(\mu,\sigma^2\|y)$	-.0127			[illegible].
Mean of $\pi(\mu\|y)$ assuming $\Sigma_{11},\Sigma_{22}\to\infty$	-.0123			
Mode of $\pi(\mu\|y)$ assuming $\Sigma_{11},\Sigma_{22}\to\infty$	-.0128			
Mean of $\pi(\mu\|y,\hat{\sigma}^2)$ ($\hat{\sigma}^2$ is value of σ^2 at mode of $\pi(\mu,\sigma^2\|y)$)	-.0127			
Mean of Approximate Large Sample Distribution.	-.0173	28800.	6370.	52100.

Table 4.[illegible] <u>Results of analysis of first tobramycin assay with prior parameters</u>

$$\begin{pmatrix} \alpha_p \\ \beta_p \\ \mu_p \end{pmatrix} = \begin{bmatrix} .29 \times 10^5 \\ .84 \times 10^4 \\ .00 \end{bmatrix}, \underset{\sim}{\Sigma} = \begin{bmatrix} .6\times10^5 & 0 & 0 \\ 0 & .2\times10^5 & 0 \\ 0 & 0 & .4\times10^{-[illegible]} \end{bmatrix},$$

<u>$\nu = 0, \lambda = 0.$</u>

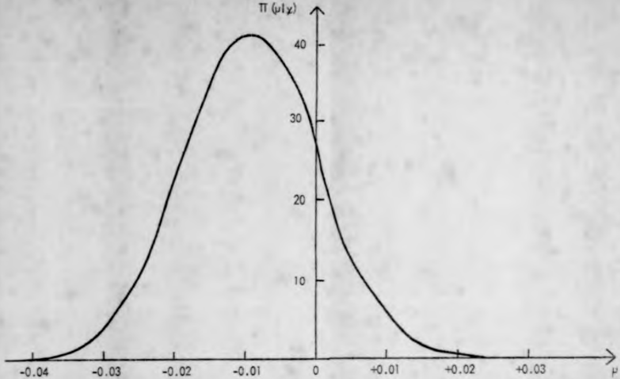


Figure 4.1 Marginal posterior density of μ for data from first tobramycin assay.

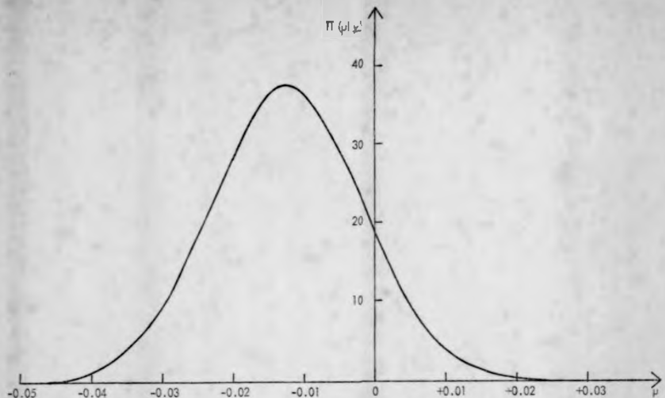


Figure 4.2 Approximate marginal posterior density of μ, neglecting prior information about α and β, for data from the first [illegible] assay.

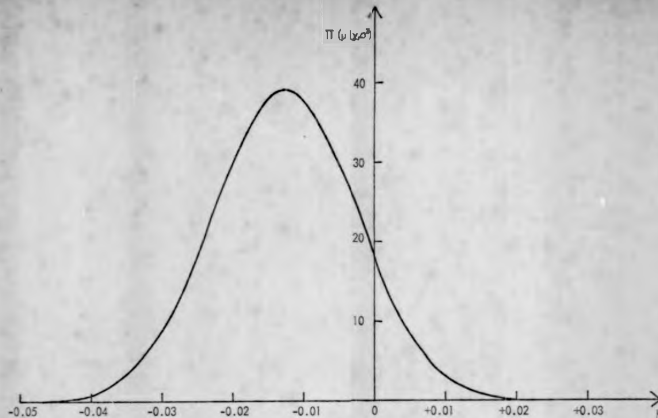


Figure 4.3 Approximate marginal posterior density of μ, assuming σ^2 to be known and equal to its value at the mode of the joint density of μ and σ^2, for data from the first tobramycin assay.

Chapter 5. Extension of the Model to Account for a More Complex Design Structure.

5.1 Introduction

Very commonly, experimental design features are incorporated into the design of assays. For example, with assays using live creatures such as rats a complete assay might consist of several identical assays each carried out on a set of litter mates. This type of design is a randomized block design.

In other types of assays such as free fat cell assays the experimental units may at some point be placed in a square configuration while undergoing some form of treatment. It may be thought likely that there are two sources of variation corresponding to the vertical and horizontal position of an experimental unit in the square. If this is the case then it may be possible to arrange the experimental units in a Latin square design. Suppose there are p^2 experimental units arranged in a pxp square, then there would be p dosage levels in the assay, each occuring once in each row in the square, once in each column in the square and p times altogether in the assay.

We have tried to extend our basic model, as described in chapters 2 and 4, in two separate ways to cover the two types of design described above.

For the randomized block design, assuming q blocks with m experimental units in each block we have used the following model for an observation in the k^{th} block:

1st stage: $y_{ik} \sim N\{(\alpha+\epsilon_k+\beta\mu z_i+\delta x_i),\sigma^2\}$,

independently for $i=1,\dots,m$, $k=1,\dots,q$,

2nd stage: $$\begin{pmatrix}\alpha\\ \beta\\ \mu\end{pmatrix} \sim N\left\{\begin{pmatrix}\alpha_0\\ \beta_0\\ \mu_0\end{pmatrix}, \underset{\sim}{\Sigma}\right\} \qquad (5.1)$$

$\epsilon_k \sim N(0,\sigma^2_\epsilon)$; independently for $k=1,\dots,q$.

The prior distribution for each ϵ is assumed independent of that for every other ϵ and also of the prior distributions for α, β and μ .

For the pxp Latin square design we have assumed the following model for an observation in the k^{th} vertical and the l^{th} horizontal position:

1st stage: $y_{kl(i)} \sim N((\alpha+\gamma_k+\delta_l+\beta x_{(i)}+\beta\epsilon_{(i)}), \sigma^2)$,

independently for $k=1,\ldots,p$, $l=1,\ldots,p$, $i=1,\ldots,p$.

$$\text{2nd stage:} \quad \begin{pmatrix}\alpha\\ \beta\\ \mu\end{pmatrix} \sim N\left\{\begin{pmatrix}\alpha_o\\ \beta_o\\ \mu_o\end{pmatrix}, \underset{\sim}{\Sigma}\right\} , \qquad (5.2)$$

$\gamma_k \sim N(0,\sigma^2_\gamma)$; independently for $k=1,\ldots,p$,

$\delta_l \sim N(0,\sigma^2_\delta)$; independently for $l=1,\ldots,p$,

where again independence of the prior distribution for each γ and δ from all other prior distributions is assumed.

Before proceding with calculating any posterior distributions one or two remarks seem appropriate.

Firstly, these two models are more complicated than our basic model in that more parameters are involved. Consequently we expect those posterior distributions which are obtainable analytically to be more complicated and in general to involve more parameters than in the previous case. In order to make inferences about the log-potency ratio we should therefore expect to have to rely more heavily than before on approximations and numerical techniques.

Secondly, we have assumed exchangeability between the individual ϵs, γs and δs respectively. We should like to stress that this assumption may not always be appropriate, especially in the case of the Latin square design where in many cases prior considerations would indicate $\gamma_1 < \gamma_2 < \ldots < \gamma_p$.

Lastly, if we had posed uniform prior distributions for means of the ϵs, γs and δs instead of fixing them at the particular value of zero, then we should have had to introduce constraints into the model of the type discussed by Smith (1973). This would have made the model conceptually more complicated. Given the exchangeability assumption, any prior information about the means of the ϵs, γs, and δs can be fully incorporated into the prior distribution of α . Hence there is no loss of generality in fixing the means.

5.2 Randomized Block Design With Known Variances

We shall first consider the randomized block design and in this section we shall assume that both the residual variance σ^2 and the between blocks variance σ^2_{ε} are known.

We can multiply together the likelihood and the prior densities as given by 5.1 to obtain, up to a multiplicative constant, the joint posterior density of all quantities involved:

$$\pi(\alpha,\beta,\mu,\varepsilon_1,\ldots,\varepsilon_q|y)\propto\exp-\tfrac{1}{2}\Bigg[\alpha^2\left(\Sigma^{11}+\frac{mq}{\sigma^2}\right)+\beta^2\left\{\Sigma^{22}+\frac{q}{\sigma^2}\sum_{i=1}^{m}(x_i^2+2\mu x_i z_i+\mu^2 z_i^2)\right\}$$

$$+\sum_{k=1}^{q}\varepsilon_k^2\left(\frac{1}{\sigma_\varepsilon^2}+\frac{m}{\sigma^2}\right)+2\alpha\beta\left\{\Sigma^{12}+\frac{q}{\sigma^2}\sum_{i=1}^{m}(x_i+\mu z_i)\right\}+\frac{2m}{\sigma^2}\alpha\sum_{k=1}^{q}\varepsilon_k+\frac{2\beta}{\sigma^2}\sum_{k=1}^{q}\varepsilon_k\sum_{i=1}^{m}(x_i+\mu z_i)$$

$$-2\alpha\left(\alpha_0\Sigma^{11}+\beta_0\Sigma^{12}-(\mu-\mu_0)\Sigma^{13}+\frac{mq\bar{y}_{..}}{\sigma^2}\right)$$

$$-2\beta\left\{\beta_0\Sigma^{22}+\alpha_0\Sigma^{12}-(\mu-\mu_0)\Sigma^{23}+\frac{q}{\sigma^2}\sum_{i=1}^{m}x_i\bar{y}_{i.}+\frac{q\mu}{\sigma^2}\sum_{i=1}^{m}(\bar{y}_{i}z_i)\right\}$$

$$-\frac{2m}{\sigma^2}\sum_{k=1}^{q}\bar{y}_{.k}\varepsilon_k+\mu^2\Sigma^{33}-2\mu(\alpha_0\Sigma^{13}+\beta_0\Sigma^{23}+\mu_0\Sigma^{33})\Bigg], \qquad (5.3)$$

where $\bar{y}_{..}=\frac{1}{mq}\sum_{k=1}^{q}\sum_{i=1}^{m}y_{ik}$, $\bar{y}_{i.}=\frac{1}{q}\sum_{k=1}^{q}y_{ik}$ and $\bar{y}_{.k}=\frac{1}{m}\sum_{i=1}^{m}y_{ik}$.

The mode of this density occurs at

$$\hat{\alpha}=\frac{\frac{mq}{\sigma^2}\bar{y}_{..}-\frac{mq}{\sigma^2}\bar{\varepsilon}_{.}-\frac{mq}{\sigma^2}\beta(\bar{x}+\mu\bar{z})+\alpha_0\Sigma^{11}-(\beta-\beta_0)\Sigma^{12}-(\mu-\mu_0)\Sigma^{13}}{\frac{mq}{\sigma^2}+\Sigma^{11}}.$$

$$\hat{\beta}=\frac{\frac{q}{\sigma^2}\sum_{i=1}^{m}(x_i\bar{y}_{i.}+\mu\bar{y}_{i.}z_i)-\frac{mq}{\sigma^2}(\alpha+\bar{\varepsilon}_{.})(\bar{x}+\mu\bar{z})+\beta_0\Sigma^{22}-(\alpha-\alpha_0)\Sigma^{12}-(\mu-\mu_0)\Sigma^{23}}{\frac{q}{\sigma^2}\sum_{i=1}^{m}(x_i^2+2\mu x_i z_i+\mu^2 z_i^2)+\Sigma^{22}}.$$

$$\frac{\frac{m\bar{y}_{\cdot k}}{\sigma^2} - \frac{m\alpha}{\sigma^2} - \frac{m\beta(\mu + \bar{z})}{\sigma^2} \; ; \; k=1,\ldots,q}{\frac{m}{\sigma^2} + \frac{1}{\sigma^2_{\epsilon}}} \qquad (5.4)$$

$$\mu = \frac{\frac{mq}{\sigma^2}\Sigma\bar{y}_{i\cdot} z_i - \frac{\beta^2 q}{\sigma^2}\Sigma x_i z_i - \frac{mq\beta}{\sigma^2}(\alpha+\bar{\epsilon}_{\cdot})\bar{z} + \mu_0\Sigma^{33} - (\alpha-\alpha_0)\Sigma^{13} - (\beta-\beta_0)\Sigma^{23}}{\frac{q\beta^2}{\sigma^2}\Sigma z_i^2 + \Sigma^{33}}$$

where $\bar{\epsilon}_{\cdot} = \frac{1}{q}\sum_{k=1}^{q} \epsilon_k$, $\bar{x} = \frac{1}{m}\sum_{i=1}^{m} x_i$ and $\bar{z} = \frac{1}{m}\sum_{i=1}^{m} z_i$

As in the case of the simple model, for given μ, the other parameters are jointly normally distributed and so we can obtain the marginal posterior density of μ up to a multiplicative constant:

$$\pi(\mu|\underline{y}) \propto |\underline{W}|^{\frac{1}{2}} \exp -\frac{1}{2}\left\{ \mu^2\Sigma^{33} - 2\mu(\alpha_0\Sigma^{13} + \beta_0\Sigma^{23} + \mu_0\Sigma^{33}) - \begin{bmatrix} c \\ d \\ e_1 \\ \vdots \\ e_q \end{bmatrix}^T \underline{W} \begin{bmatrix} c \\ d \\ e_1 \\ \vdots \\ e_q \end{bmatrix} \right\} \qquad (5.5)$$

where $c = \frac{mq\bar{y}_{\cdot\cdot}}{\sigma^2} + \alpha_0\Sigma^{11} + \beta_0\Sigma^{12} - (\mu-\mu_0)\Sigma^{13}$,

$$d = \frac{q}{\sigma^2}\sum_{i=1}^{m}(x_i\bar{y}_{i\cdot} + \mu\bar{y}_{i\cdot} z_i) + \beta_0\Sigma^{22} + \alpha_0\Sigma^{12} - (\mu-\mu_0)\Sigma^{23} \; ,$$

$$e_k = \frac{m\bar{y}_{\cdot k}}{\sigma^2} \; ; \quad k=1,\ldots,q$$

and $\underline{W}$ is the matrix whose inverse is

$$\underset{\sim}{W}^{-1}=\left[\begin{array}{cc|c} \Sigma^{11}+\dfrac{mq}{\sigma^2} & \Sigma^{12}+\dfrac{mq}{\sigma^2}(\bar{x}+\mu\bar{z}) & \dfrac{m}{\sigma^2}\underset{\sim}{1}_q^T \\ \Sigma^{12}+\dfrac{mq}{\sigma^2}(\bar{x}+\mu\bar{z}) & \left\{\Sigma^{22}+\dfrac{q}{\sigma^2}\sum_{i=1}^{m}(x_i^2+2\mu x_i z_i+\mu^2 z_i^2)\right\} & \dfrac{m}{\sigma^2}(\bar{x}+\mu\bar{z})\underset{\sim}{1}_q^T \\ \hline \dfrac{m}{\sigma^2}\underset{\sim}{1}_q & \dfrac{m}{\sigma^2}(\bar{x}+\mu\bar{z})\underset{\sim}{1}_q & \left(\dfrac{1}{\sigma^2_\varepsilon}+\dfrac{m}{\sigma^2}\right)\underset{\sim}{I}_q \end{array}\right].$$

where $\underset{\sim}{1}_q$ is the q x 1 matrix whose elements are all 1.

Since the calculation of $\underset{\sim}{W}$ and $|\underset{\sim}{W}|$ is a somewhat lengthy operation we give final forms here.

$$\underset{\sim}{W}=\frac{1}{\Delta}\left[\begin{array}{cc|c} W_{11} & W_{12} & W_{13}\underset{\sim}{1}_q^T \\ W_{12} & W_{22} & W_{23}\underset{\sim}{1}_q^T \\ \hline W_{13}\underset{\sim}{1}_q & W_{23}\underset{\sim}{1}_q & W_{33}^{(1)}\underset{\sim}{I}_q+W_{33}^{(2)}\underset{\sim}{J}_q \end{array}\right]$$

and $|\underset{\sim}{W}| = \Delta^{-1}(\sigma^2/\sigma^2_\varepsilon+m)^{-(q-1)}(\sigma^2)^q$,

$$\text{where } \Delta=\left(\frac{\sigma^2}{\sigma^2_\varepsilon}+m\right)\left[\left\{\Sigma^{11}+\frac{mq}{\sigma^2}\right\}\left\{\Sigma^{22}+\frac{q}{\sigma^2}\sum_{i=1}^{m}(x_i^2+2\mu x_i z_i+\mu^2 z_i^2)\right\}-\left\{\Sigma^{12}+\frac{mq}{\sigma^2}(\bar{x}+\mu\bar{z})\right\}^2\right]$$

$$-\frac{qm^2}{\sigma^2}\left[\Sigma^{11}(\bar{x}+\mu\bar{z})^2-2\Sigma^{12}(\bar{x}+\mu\bar{z})+\Sigma^{22}+\frac{q}{\sigma^2}(Sxx+2\mu Sxz+\mu^2 Szz)\right] .$$

$$W_{11}=\left\{\Sigma^{22}+\frac{q}{\sigma^2}\sum_{i=1}^{m}(x_i^2+2\mu x_i z_i+\mu^2 z_i^2)\right\}(m+\sigma^2/\sigma^2_\varepsilon)-\frac{m^2q}{\sigma^2}(\bar{x}+\mu\bar{z})^2,$$

$$W_{12}=-\left[\Sigma^{12}(\sigma^2/\sigma^2_\varepsilon+m)+\frac{mq}{\sigma^2_\varepsilon}(\bar{x}+\mu\bar{z})\right] .$$

$$W_{13}=-m\left[\Sigma^{22}-\Sigma^{12}(\bar{x}+\mu\bar{z})+\frac{q}{\sigma^2}(Szz+2\mu Sxz+\mu^2 Szz)\right] ,$$

$$W_{22}=\Sigma^{11}\left(\frac{\sigma^2}{\sigma^2_\varepsilon}+m\right)+\frac{mq}{\sigma^2_\varepsilon}\ ,$$

$$W_{23}=-m(\Sigma^{11}(\bar{x}+\mu\bar{z})-\Sigma^{12})\ ,$$

$$W^{(1)}_{33}=\sigma^2\Delta\ /(\sigma^2/_{\sigma^2_\varepsilon}+m)\ ,$$

$$W^{(2)}_{33}=\frac{m^2\left[\Sigma^{11}(\bar{x}+\mu\bar{z})^2-2\Sigma^{12}(\bar{x}+\mu\bar{z})+\Sigma^{22}+\frac{q}{\sigma^2}(Sxx+2\mu Sxz+\mu^2 Szz)\right]}{(\sigma^2/_{\sigma^2_\varepsilon}+m)}\ .$$

This density, although algebraically somewhat more complicated, is very similar in form to the corresponding density for the simple model given by 2.6.

In the case where we have uniform prior distributions for α and β a slight simplification occurs which will be useful in the next section. The uniform prior distributions imply that all the elements in $\underset{\sim}{\Sigma}^{-1}$ apart from Σ^{33} are zero, and consequently we can write $\underset{\sim}{W}$ in the form $\underset{\sim}{W}=\sigma^2\underset{\sim}{W}_o$, where σ^2 and σ^2_ε only occur in $\underset{\sim}{W}_o$ in the ratio $\sigma^2/_{\sigma^2_\varepsilon}$. Similarly we can write $c=c_o/\sigma^2$, $d=d_o/\sigma^2$ and $e_k=e_{k_o}/\sigma^2$, where c_o, d_o and e_{k_o} are independent of σ^2 and σ^2_ε. Hence we can write

$$\pi(\mu|\underset{\sim}{y};\ \Sigma_{11},\ \Sigma_{22}\to\infty)\propto|\sigma^2\underset{\sim}{W}_o|^{\frac{1}{2}}\exp-\tfrac{1}{2}\left\{(\mu^2-2\mu\mu_o)\Sigma^{33}-\frac{1}{\sigma^2}\begin{bmatrix}c_o\\ d_o\\ e_{1_o}\\ \vdots\\ e_{q_o}\end{bmatrix}^T\underset{\sim}{W}_o\begin{bmatrix}c_o\\ d_o\\ e_{1_o}\\ \vdots\\ e_{q_o}\end{bmatrix}\right\} \tag{5.6}$$

where c_o, d_o, $e_{1_o}\ldots e_{q_o}$, and $\underset{\sim}{W}_o$ only involve σ^2 and σ^2_ε in the ratio $\sigma^2/_{\sigma^2_\varepsilon}$.

5.3 Randomized Block Design With Unknown Variances

In practice the residual variance, σ^2 , and the between block variance, σ^2_ε, will be unknown and should be regarded as parameters in the model. Following section 4.1 we shall use the relevant conjugate prior distributions which are that $\frac{\nu\lambda}{\sigma^2}$ has a χ^2-distribution on ν degrees of freedom where ν, λ, ν_ε, λ_ε are known constants. If we call the expression on the right hand side of the $\propto$ sign in 5.3 $f_1(\alpha,\beta,\mu,\varepsilon_1,\ldots\varepsilon_q)$, then the joint posterior density of $\alpha,\beta,\mu,\varepsilon_1,\ldots\varepsilon_q,\sigma^2$ and σ^2_ε is

$$\pi(\alpha,\beta,\mu,\varepsilon_1,\ldots,\varepsilon_q,\sigma^2,\sigma^2_\varepsilon|y)\propto(\sigma^2)^{-\frac{(mq+\nu+2)}{2}}(\sigma^2_\varepsilon)^{-\frac{(q+\nu_\varepsilon+2)}{2}}$$

$$\times\ \exp-\tfrac{1}{2}\left\{\frac{\nu\lambda}{\sigma^2}+\frac{\nu_\varepsilon\lambda_\varepsilon}{\sigma^2_\varepsilon}+\sum_{i,k}\frac{y_{ik}^2}{\sigma^2}\right\}f_1(\alpha,\beta,\mu,\varepsilon_1,\ldots,\varepsilon_q) \qquad (5.7)$$

The mode of this density occurs at the point given by 5.4 with

$$\sigma^2=\frac{\nu\lambda+\sum_{k=1}^{q}\sum_{i=1}^{m}(y_{ik}-\alpha-\varepsilon_k-\beta\mu z_i-\beta x_i)^2}{mq+\nu+2}$$

$$\sigma^2_\varepsilon=\frac{\nu_\varepsilon\lambda_\varepsilon+\sum_{k=1}^{q}\varepsilon_k^2}{q+\nu_\varepsilon+2} \qquad (5.8)$$

We can integrate out from this density either $(\alpha,\beta,\varepsilon_1,\ldots\varepsilon_q)$, or σ^2 and σ^2_ε. Since the former possibility leaves a distribution of 3 parameters while the latter leaves a distribution of $(3+q)$ parameters we consider here only the former possibility.

Carrying out the integration we get

$$\pi(\mu,\sigma^2,\sigma^2_\varepsilon|\underline{y})\propto(\sigma^2)^{-\frac{(mq+\nu+2)}{2}}(\sigma^2_\varepsilon)^{-\frac{(q+\nu_\varepsilon+2)}{2}}\exp-\tfrac{1}{2}\left\{\frac{\nu\lambda}{\sigma^2}+\frac{\nu_\varepsilon\lambda_\varepsilon}{\sigma^2_\varepsilon}+\sum_{k=1}^{q}\sum_{i=1}^{m}\frac{y_{ik}^2}{\sigma^2}\right\}$$

$$\times\ f_2(\mu,\sigma^2,\sigma^2_\varepsilon) \qquad (5.9)$$

where $f_2(\mu,\sigma^2,\sigma^2_\varepsilon)$ is the expression on the right hand side of the = sign in 5.5.

Unfortunately, in the general case, we can proceed no further. The exact numerical treatment would require a three-dimensional numerical integration. Such an integration should be quite possible but we have not at present attempted it. It would probably be prohibitively expensive for routine analysis of data. Consequently we must resort to some approximations. Taking the approach suggested in the last paragraph of section 4.3 we could assign the values of σ^2 and σ^2_ε at the mode of $\pi(\mu,\sigma^2,\sigma^2_\varepsilon|y)$ to the marginal distribution of μ for known variances. This approximation should be quite good as regards σ^2 since the data should contain a substantial amount of information about the residual variance. Unfortunately the same cannot be said for σ^2_ε. This problem could be surmounted in part by assigning the value of σ^2 at the mode of $\pi(\mu,\sigma^2,\sigma^2_\varepsilon|y)$ to the joint distribution of μ and σ^2_ε for known σ^2, $\pi(\mu,\sigma^2_\varepsilon|\sigma^2,y)$ and then finding the mode of the marginal distribution of σ^2_ε, given the assigned value of σ^2, by a series of one-dimensional numerical integrations over μ.

In the case where we have uniform prior distributions for α and β we can proceed slightly further. From 5.8

$$\pi(\mu,\sigma^2,\sigma^2_\varepsilon|y;\Sigma_{11},\Sigma_{22}\rightarrow\infty)\propto(\sigma^2)^{-\frac{(m-pq+\nu)}{2}}(\sigma^2_\varepsilon)^{-\frac{(q+\nu_\varepsilon+2)}{2}}|W_o|^{\frac{1}{2}}$$

$$\times\ \exp-\frac{1}{2\sigma^2}\left\{\nu\lambda+\sum_{k=1}^{q}\sum_{i=1}^{m}y^2_{ik}-\begin{bmatrix}c_o\\d_o\\e_{1o}\\\vdots\\e_{qo}\end{bmatrix}^T\underset{\sim}{W}_o\begin{bmatrix}c_o\\d_o\\e_{1o}\\\vdots\\e_{qo}\end{bmatrix}\right. \tag{5.10}$$

$$\times\ \exp-\tfrac{1}{2}\left\{\frac{\nu_\varepsilon\lambda_\varepsilon}{\sigma^2_\varepsilon}+(\mu^2-2\mu\mu_o)\Sigma^{33}\right\}.$$

If we now make a transformation of variables from μ, σ^2 and σ^2_ε to μ,σ^2 and $S^2_\cdot$ where $S^2=\sigma^2/\sigma^2_\varepsilon$, we have

$$\pi(\mu,\sigma^2,S^2|\underset{\sim}{y};\Sigma_{11},\Sigma_{22}\to\infty)=(\sigma^2)^{-\frac{(mq+\nu_\varepsilon+\nu)}{2}}$$

$$\times\ \exp-\frac{1}{2\sigma^2}\left[\nu\lambda+\sum_{k=1}^{q}\sum_{i=1}^{m}y_{ik}^2-\begin{bmatrix}c_o\\d_o\\e_{1o}\\\vdots\\e_{qo}\end{bmatrix}^T\underset{\sim}{W}_o\begin{bmatrix}c_o\\d_o\\e_{1o}\\\vdots\\e_{qo}\end{bmatrix}+\nu_\varepsilon\lambda_\varepsilon S^2\right]$$

$$\times\ |\underset{\sim}{W}_o|^{\frac{1}{2}}\,(S^2)^{\frac{(q+\nu_\varepsilon)}{2}}\ \exp-\tfrac{1}{2}(\mu^2-2\mu\mu_o)\Sigma^{33}\ .$$

We can now integrate over σ^2 to obtain the bivariate density:

$$\pi(\mu,S^2|\underset{\sim}{y};\Sigma_{11},\Sigma_{22}\to\infty)\propto\left\{\nu\lambda+\sum_{k=1}^{q}\sum_{i=1}^{m}y_{ik}^2-\begin{bmatrix}c_o\\d_o\\e_{1o}\\\vdots\\e_{qo}\end{bmatrix}^T\underset{\sim}{W}_o\begin{bmatrix}c_o\\d_o\\e_{1o}\\\vdots\\e_{qo}\end{bmatrix}+\nu_\varepsilon\lambda_\varepsilon S^2\right\}^{-\frac{(mq+\nu_\varepsilon+\nu-2)}{2}}$$

$$\times\ |\underset{\sim}{W}_o|^{\frac{1}{2}}(S^2)^{\frac{q+\nu_\varepsilon}{2}}\ \exp-\tfrac{1}{2}(\mu^2-2\mu\mu_o)\Sigma^{33}, \qquad (5.11)$$

The same remarks can be made concerning the estimation of log potency ratio from this distribution as were made in section 4.3 concerning the joint distribution of μ and σ^2 in the basic model.

5.4 Latin Square Design.

To avoid repetition we shall consider the Latin square design with unknown residual, between row and between column variances straight away. We shall assume the relevant conjugate prior distributions and use a notation similar to that n section 5.3 .

Taking the model as stated in 5.2 the joint posterior density of all quantities is

$$\pi(\alpha,\beta,\mu,\gamma_1,\ldots,\gamma_p,\delta_1,\ldots,\delta_p,\sigma^2,\sigma^2_\gamma,\sigma^2_\delta|\underline{y})=(\sigma^2)^{-\frac{[\text{illegible}]}{2}}$$

$$\times\ (\sigma^2_\gamma)^{-\frac{(\nu_\gamma+p+2)}{2}}(\sigma^2_\delta)^{-\frac{(\nu_\delta+p+2)}{2}}$$

$$\times\ \exp-\tfrac{1}{2}\left\{\sum_{k=1}^{p}\sum_{l=1}^{p}\frac{y^2_{kl(i)}}{\sigma^2}+\frac{\nu\lambda}{\sigma^2}+\frac{\nu_\gamma\lambda_\gamma}{\sigma^2_\gamma}+\frac{\nu_\delta\lambda_\delta}{\sigma^2_\delta}\right\}$$

$$\times\ \exp-\tfrac{1}{2}\left[\alpha^2\left(\Sigma^{11}+\frac{p^2}{\sigma^2}\right)+\beta^2\left\{\Sigma^{22}+\frac{p}{\sigma^2}\sum_{i=1}^{p}(x_i^2+2\mu x_i z_i+\mu^2 z_i^2)\right\}+\sum_{k=1}^{p}\gamma_k^2\left(\frac{1}{\sigma^2_\gamma}+\frac{p}{\sigma^2}\right)\right.$$

$$+\sum_{l=1}^{p}\delta_l^2\left(\frac{1}{\sigma^2_l}+\frac{p}{\sigma^2}\right)+2\alpha\beta\left(\Sigma^{12}+\frac{p^2}{\sigma^2}(\bar{x}+\mu\bar{z})\right)+\frac{2p\alpha}{\sigma^2}\sum_{k=1}^{p}\gamma_k+\frac{2p\alpha}{\sigma^2}\sum_{l=1}^{p}\delta_l+\frac{2p\beta}{\sigma^2}\left(\sum_{l=1}^{p}\delta_l+\sum_{k=1}^{p}\gamma_k\right)(\bar{x}+\mu\bar{z})$$

$$+\frac{2}{\sigma^2}\sum_{k=1}^{p}\gamma_k\sum_{l=1}^{p}\delta_l-2\alpha\left\{\frac{p^2}{\sigma^2}\bar{y}..(.)+\alpha_o\Sigma^{11}+\beta_o\Sigma^{12}-(\mu-\mu_o)\Sigma^{13}\right\}$$

$$-2\beta\left\{\frac{p}{\sigma^2}\sum_{i=1}^{p}\bar{y}..(i)(x_i+\mu z_i)+\beta_o\Sigma^{22}+\alpha_o\Sigma^{12}-(\mu-\mu_o)\Sigma^{13}\right\}$$

$$-\frac{2p}{\sigma^2}\sum_{k=1}^{p}\bar{y}_{k}.(.)\gamma_k-\frac{2p}{\sigma^2}\sum_{l=1}^{p}y._l(.)\delta_l$$

$$+\mu^2\Sigma^{33}-2\mu(\alpha_o\Sigma^{13}+\beta_o\Sigma^{23}+\mu_o\Sigma^{33})\Bigg] , \qquad (5.12)$$

where $\bar{y}_{..}(.)=\frac{1}{p^2}\sum_{k=1}^{p}\sum_{l=1}^{p} y_{kl(i)}$, $\bar{y}_{k}.(.)=\frac{1}{p}\sum_{l=1}^{p} y_{kl(i)}$, $\bar{y}_{\cdot l}(.)=\frac{1}{p}\sum_{k=1}^{p} y_{kl(i)}$,

$$\bar{y}_{..(i)}=\frac{1}{p}\sum_{k=1}^{p}\sum_{l=1}^{p} y_{kl(i)} ,$$
(i fixed)

$\bar{x}=\frac{1}{p}\sum_{i=1}^{p} x_i$, $\bar{z}=\frac{1}{p}\sum_{i=1}^{p} z_i$.

The mode of this density occurs at

$$\alpha=\frac{\frac{p^2}{\sigma^2}\bar{y}_{..}(.)-\frac{p^2}{\sigma^2}\beta(\bar{x}+\mu\bar{z})-\frac{p^2}{\sigma^2}\bar{\delta}.-\frac{p^2}{\sigma^2}\bar{\gamma}.+\alpha_o\Sigma^{11}-(\beta-\beta_o)\Sigma^{12}-(\mu-\mu_o)\Sigma^{13}}{\frac{p^2}{\sigma^2}+\Sigma^{11}} ,$$

$$\beta=\frac{\frac{p}{\sigma^2}\sum_{i=1}^{p}\bar{y}_{..}(i)(x_i+\mu z_i)-\frac{p^2}{\sigma^2}(\alpha+\bar{\delta}.+\bar{\gamma}.)(\bar{x}+\mu\bar{z})+\beta_o\Sigma^{22}-(\alpha-\alpha_o)\Sigma^{12}-(\mu-\mu_o)\Sigma^{23}}{\frac{p}{\sigma^2}\sum_{i=1}^{p}(x_i^2+2\mu x_i z_i+\mu^2 z_i^2)+\Sigma^{22}} ,$$

$$\gamma_k=\frac{\frac{p}{\sigma^2}\bar{y}_{k}.(.)-\frac{p}{\sigma^2}(\alpha+\bar{\delta}.)-\frac{p}{\sigma^2}\beta(\bar{x}+\mu\bar{z})}{\frac{p}{\sigma^2}+\frac{1}{\sigma^2_\gamma}} , \quad k=1,\dots p , \qquad (5.13)$$

$$\mu=\frac{\frac{p}{\sigma^2}\beta\sum_{i=1}^{p} y_{\cdot\cdot(1)}z_i-\frac{p}{\sigma^2}\beta^2\sum z_i\,[\text{illegible}]+\mu_0\Sigma^{33}-(\beta-\beta_0)\Sigma^{23}-(\alpha-\alpha_0)\Sigma^{13}}{\frac{p}{\sigma^2}\beta\sum z_i^2+\Sigma^{33}} \; ,$$

$$\sigma^2=\frac{\nu\lambda+\sum_{k=1}^{p}\sum_{l=1}^{p}\left(y_{kl(1)}-[\text{illegible}]\right)^2}{p^2+\nu+2} \; ,$$

$$\sigma^2_{\gamma}=\frac{\nu_\gamma\lambda_\gamma+\sum_{k=1}^{p}\gamma_k^2}{\nu_\gamma+p+2} \; ;$$

$$\sigma^2_{\delta}=\frac{\nu_\delta\lambda_\delta+\sum_{l=1}^{p}\delta_l^2}{\nu_\delta+p+2} \; ;$$

where $\bar{\gamma}_{\cdot}=\frac{1}{p}\sum_{k=1}^{p}\gamma_k$ and $\bar{\delta}_{\cdot}=\frac{1}{p}\sum_{l=1}^{p}\delta_l$.

We can integrate over $\alpha,\beta,\gamma_1,\ldots,\gamma_p,\delta_1,\ldots,\delta_p$ giving the joint posterior density of $\mu,\sigma^2,\sigma^2_{\gamma}$, and σ^2_{δ} :

$$\pi(\mu,\sigma^2,\sigma^2_{\gamma},\sigma^2_{\delta}|y)=(\sigma^2)^{-\frac{(p^2+\nu+2)}{2}}(\sigma^2_{\gamma})^{-\frac{(p+\nu_\gamma+2)}{2}}(\sigma^2_{\delta})^{-\frac{(p+\nu_\delta+2)}{2}}|U|^{\frac{1}{2}}$$

$$\times\ \exp-\tfrac{1}{2}\left[\frac{\nu\lambda}{\sigma^2}+\frac{\nu_\gamma\lambda_\gamma}{\sigma^2_{\gamma}}+\frac{\nu_\delta\lambda_\delta}{\sigma^2_{\delta}}+\sum_{k=1}^{p}\sum_{l=1}^{p}\frac{y^2_{kl(1)}}{\sigma^2}\right]$$

$$\propto \exp -\tfrac{1}{2}\left[\mu^2\Sigma^{33} - 2\mu(\alpha_o\Sigma^{13} + \beta_o\Sigma^{23} + \mu_o\Sigma^{33}) - \begin{bmatrix} f \\ g \\ h_1 \\ \vdots \\ h_p \\ j_1 \\ \vdots \\ j_p \end{bmatrix}^T \underset{\sim}{U} \begin{bmatrix} f \\ g \\ h_1 \\ \vdots \\ h_p \\ j_1 \\ \vdots \\ j_p \end{bmatrix} \right] \qquad [5.44]$$

where $f = \frac{p^2}{\sigma^2}\bar{y}_{..}(.) + \alpha_o\Sigma^{11} + \beta_o\Sigma^{12} - (\mu-\mu_o)\Sigma^{13}$,

$$g = \frac{p}{\sigma^2}\sum_{i=1}^{p} \bar{y}_{..(i)}(x_i + \mu z_i) + \beta_o\Sigma^{22} + \alpha_o\Sigma^{12} - (\mu-\mu_o)\Sigma^{13} \ ,$$

$$h_k = \frac{p}{\sigma^2}\,\bar{y}_{k\cdot}(.) \ ; \quad k=1,\ldots p,$$

$$j_l = \frac{p}{\sigma^2}\,\bar{y}_{\cdot l}(.) \ ; \quad l=1,\ldots p,$$

and $\underset{\sim}{U}$ is the matrix whose inverse is

$$\underset{\sim}{U}^{-1} = \left[\begin{array}{cc:cc} \Sigma^{11} + \frac{p^2}{\sigma^2} & \Sigma^{12} + \frac{p^2}{\sigma^2}(\bar{x}+\mu\bar{z}) & \multicolumn{2}{c}{\frac{p}{\sigma^2}\underset{\sim}{1}_{2p}^T} \\ \Sigma^{12} + \frac{p^2}{\sigma^2}(\bar{x}+\mu\bar{z}) & \Sigma^{22} + \frac{p}{\sigma^2}\sum_{i=1}^{p}(x_i^2 + 2\mu x_i z_i + \mu^2 z_i^2) & \multicolumn{2}{c}{\frac{p}{\sigma^2}(\bar{x}+\mu\bar{z})\underset{\sim}{1}_{2p}^T} \\ \hdashline \frac{p}{\sigma^2}\underset{\sim}{1}_{2p} & \frac{p}{\sigma^2}(\bar{x}+\mu\bar{z})\underset{\sim}{1}_{2p} & \left(\frac{1}{\sigma_Y^2} + \frac{p}{\sigma^4}\right)\underset{\sim}{I}_p + \frac{1}{\sigma^2}\underset{\sim}{J}_p & \\ & & \frac{1}{\sigma^2}\underset{\sim}{J}_p & \left(\frac{1}{\sigma^2} + \frac{p}{\sigma^2}\right)\underset{\sim}{I}_p \end{array}\right]$$

Hence $\underset{\sim}{U}^{-1} = \dfrac{1}{\Gamma}$

$$\left[\begin{array}{cc:cc} U_{11} & U_{12} & U_{13}\underset{\sim}{1}_p^T & U_{14}\underset{\sim}{1}_p^T \\ U_{12} & U_{22} & U_{23}\underset{\sim}{1}_p^T & U_{24}\underset{\sim}{1}_p^T \\ \hdashline U_{13}\underset{\sim}{1}_p & U_{23}\underset{\sim}{1}_p & (U_{33}^{(1)}\underset{\sim}{I}_p + U_{33}^{(2)}\underset{\sim}{J}_p) & U_{34}\underset{\sim}{J}_p \\ U_{14}\underset{\sim}{1}_p & U_{24}\underset{\sim}{1}_p & U_{34}\underset{\sim}{J}_p & (U_{44}^{(1)}\underset{\sim}{I}_p + U_{44}^{(2)}\underset{\sim}{J}_p) \end{array}\right]$$

and $|\underset{\sim}{U}| = \left(\dfrac{\sigma^2}{\sigma^2_\gamma} + p\right)^{p-1} \left(\dfrac{\sigma^2}{\sigma^2_\delta} + p\right)^{p-1} (\sigma^2)^{-2p}\ \Gamma$,

where $\Gamma = \left[\left(\Sigma^{11} + \dfrac{p^2}{\sigma^2}\right)\left(\Sigma^{22} + \dfrac{p}{\sigma^2}\sum_{i=1}^{m} (x_i^2 + 2\mu x_i z_i + \mu^2 z_i^2)\right) - \left(\Sigma^{12} + \dfrac{p^2}{\sigma^2}(\bar{x} + \mu\bar{z})\right)^2\right]$

$$\times \left(\frac{\sigma^4}{\sigma^2_\gamma\sigma^2_\delta} + \frac{p\sigma^2}{\sigma^2_\gamma} + \frac{p\sigma^2}{\sigma^2_\delta}\right)$$

$= \left[\dfrac{p^2}{\sigma^2}\left(\Sigma^{22} + \dfrac{p}{\sigma^2}\sum_{i=1}^{m} (x_i^2 + 2\mu x_i z_i + \mu^2 z_i^2)\right) - 2\dfrac{p^2}{\sigma^2}\left(\Sigma^{12} + \dfrac{p^2}{\sigma^2}(\bar{x} + \mu\bar{z})\right)(\bar{x} + \mu\bar{z}) + \dfrac{p^2}{\sigma^2}\left(\Sigma^{11} + \dfrac{p^2}{\sigma^2}\right)\right.$

$$\left. \times\ (\bar{x} + \mu\bar{z})^2 \right] p\left(\frac{\sigma^2}{\sigma^2_\gamma} + \frac{\sigma^2}{\sigma^2_\delta}\right) ,$$

$U_{11} = \left(\dfrac{p\sigma^2}{\sigma^2_\gamma} + \dfrac{p\sigma^2}{\sigma^2_\delta} + \dfrac{\sigma^4}{\sigma^2_\gamma\sigma^2_\delta}\right)\left(\Sigma^{22} + \dfrac{p}{\sigma^2}\sum_{i=1}^{p} (x_i^2 + 2\mu x_i z_i + \mu^2 z_i^2)\right) - \dfrac{p^3}{\sigma^2}\left(\dfrac{\sigma^2}{\sigma^2_\gamma} + \dfrac{\sigma^2}{\sigma^2_\delta}\right)(\bar{x} + \mu\bar{z})^2$,

$U_{12} = -\left[\left(\dfrac{p\sigma^2}{\sigma^2_\gamma} + \dfrac{p\sigma^2}{\sigma^2_\delta} + \dfrac{\sigma^4}{\sigma^2_\gamma\sigma^2_\delta}\right)\Sigma^{12} + \dfrac{p^2}{\sigma^2} \cdot \dfrac{\sigma^4}{\sigma^2_\gamma\sigma^2_\delta}(\bar{x} + \mu\bar{z})\right]$,

$$U_{13} = -\frac{\rho\sigma^2}{\sigma^2_\delta}\left(\Sigma^{22} - \Sigma^{12}(\bar{x}+\mu\bar{z}) + \frac{\rho}{\sigma^2}\,(Sxx + 2\mu Sxz + \mu^2 Szz)\right),$$

$$U_{14} = -\frac{\rho\sigma^2}{\sigma^2_\gamma}\left(\Sigma^{22} - \Sigma^{12}(\bar{x}+\mu\bar{z}) + \frac{\rho}{\sigma^2}\,(Sxx + 2\mu Sxz + \mu^2 Szz)\right),$$

$$U_{22} = \left(\frac{\rho\sigma^2}{\sigma^2_\gamma} + \frac{\rho\sigma^2}{\sigma^2_\delta} + \frac{\sigma^4}{\sigma^2_\gamma\sigma^2_\delta}\right)\Sigma^{11} + \frac{\rho^2}{\sigma^2}\,\frac{\sigma^4}{\sigma^2_\gamma\sigma^2_\delta}$$

$$U_{23} = -\frac{\rho\sigma^2}{\sigma^2_\delta}(\Sigma^{11}(\bar{x}+\mu\bar{z}) - \Sigma^{12}),$$

$$U_{24} = -\frac{\rho\sigma^2}{\sigma^2_\gamma}(\Sigma^{11}(\bar{x}+\mu\bar{z}) - \Sigma^{12}),$$

$$U^{(1)}_{33} = \frac{\sigma^2\rho}{(\sigma^2/\sigma^2_\gamma + \rho)},$$

$$U^{(2)}_{33} = \left[\frac{\rho^2\sigma^2}{\sigma^2_\delta}\left(\Sigma^{11}(\bar{x}+\mu\bar{z})^2 - 2\Sigma^{12}(\bar{x}+\mu\bar{z}) + \Sigma^{22} + \frac{\rho}{\sigma^2}\,(Sxx + 2\mu Sxz + \mu^2 Szz)\right)\right.$$

$$\left. + \rho\sigma^2\left(\Sigma^{11}\Sigma^{22} - (\Sigma^{12})^2 + \Sigma^{11}\frac{\rho}{\sigma^2}\,(Sxx + 2\mu Sxz + \mu^2 Szz)\right)\right]\left(\frac{\sigma^2}{\sigma^2_\gamma} + \rho\right)^{-1},$$

$$U_{34} = -\left[\sigma^2\left(\Sigma^{11}\Sigma^{22} - (\Sigma^{12})^2\right) + \rho\Sigma^{11}(Sxx + 2\mu Sxz + \mu^2 Szz)\right],$$

(f)

$$U_{44} = \frac{\sigma^2 \Gamma}{(\sigma^2/_{\sigma^2_\delta} + p)}$$

$$U_{44}^{(k)} = \left[\frac{p^2\sigma^2}{\sigma^2_\gamma} \left(\Sigma^{11}(\bar{x}+\mu\bar{z})^2 - 2\Sigma^{12}(\bar{x}+\mu\bar{z}) + \Sigma^{22} + \frac{p}{\sigma^2}(Sxx + 2\mu Sxz + \mu^2 Szz) \right) \right.$$

$$\left. + p\sigma^2 \left(\Sigma^{11}\Sigma^{22} - (\Sigma^{12})^2 + \Sigma^{11} \frac{p}{\sigma^2}(Sxx + 2\mu Sxz + \mu^2 Szz) \right) \right] \left(\sigma^2/_{\sigma^2_\delta} + p \right)^{-1}$$

As in the case of the randomized block design we can proceed no further analytically. An approximate posterior density for μ can be obtained by assuming $\sigma^2, \sigma^2_\gamma$ and σ^2_δ are known and that they take the values at the mode of $\pi(\mu, \sigma^2, \sigma^2_\gamma, \sigma^2_\delta | y)$.

The case of uniform prior distributions for α and β is again similar to that of the randomized block design. We can write $U = \sigma^2 U_0$, $f = f_0/\sigma^2$, $g = g_0/\sigma^2$, h_{k0}/σ^2; $k = 1, \ldots, p$, $j_l = j_{l0}/\sigma^2$; $l = 1, \ldots, p$, where U_0, f_0, g_0, h_{k0}; $k = 1, \ldots, p$, j_{l0}; $l = 1, \ldots, p$, only involve $\sigma^2, \sigma^2_\gamma$ and σ^2_δ in the ratios $\sigma^2/_{\sigma^2_\gamma}$ and σ^2/σ^2_δ. We can transform to the variables $\mu, \sigma^2, S^2_\gamma = \sigma^2/\sigma^2_\gamma$ and $S^2_\delta = \sigma^2/\sigma^2_\delta$, and then integrate over σ^2 giving the posterior density of $\mu, S^2_\gamma, S^2_\delta$;

$$\pi(\mu, S^2_{\gamma}, S^2_{\delta} | \underset{\sim}{y}; \Sigma_{11}, \Sigma_{22} \to \infty) \propto$$

$$\left\{ \nu\lambda + \nu_{\gamma}\lambda_{\gamma}S^2_{\gamma} + \nu_{\delta}\lambda_{\delta}S^2_{\delta} + \sum_{l=1}^{p}\sum_{k=1}^{p} y^2_{kl(1)} - \begin{bmatrix} f_o \\ g_o \\ h1_o \\ \vdots \\ h_{po} \\ j1_o \\ \vdots \\ j_{po} \end{bmatrix}^T \underset{-}{U} \begin{bmatrix} f_o \\ g_o \\ h1_o \\ \vdots \\ h_{po} \\ j1_o \\ \vdots \\ j_{po} \end{bmatrix} \right\}^{-\frac{(p^2+p+\nu+\nu_{\gamma}+\nu_{\delta}-2)}{2}}$$

$$\times \ |\underset{\sim}{U}_o|^{\frac{1}{2}} (S^2_{\gamma})^{\frac{p+\nu_{\gamma}}{2}} (S^2_{\delta})^{\frac{p+\nu_{\delta}}{2}} \exp -\tfrac{1}{2}(\mu^2 - 2\mu\mu_o)\Sigma^{33} \ . \qquad (5.15)$$

Estimation of μ can be made after finding an approximate marginal posterior distribution of μ as suggested in the previous paragraph.

5.5 An Example: Factor VIII Data

In this section we analyse data from an assay of factor VIII. Factor VIII is one of the chain of enzymes responsible for blood clotting in man and deficiency of factor VIII leads to haemophilia. The response variable measured in the assay is the time taken for a clot to form after a dose of factor VIII is added to a set of reagents. The larger the dose the more quickly a clot is formed so the slope of the fitted regression lines will be negative. The data are given in Table 5.1. The assay was repeated on five consecutive days and so our theory for randomized block designs is appropriate.

Before analyzing the data we had very little idea of the likely results and so we have used uniform prior densities for α and β and let $\nu=0$ in our prior distribution for σ^2. We cannot put $\nu_\epsilon=0$ in the prior distribution for σ^2_ϵ since this implies that the block effects are all zero, a point which has been discussed by Lindley (1971 b), and so we have put $\nu_\epsilon=\lambda_\epsilon=1$. A uniform prior distribution for μ is not permissible for the reasons discussed in chapter 2 and so we have taken the prior distribution for μ to be $N(0.0,\ 1.5)$. This prior distribution and the one for σ^2_ϵ are based on introspection and rather arbitrary. It is clear from the posterior distributions that the prior distribution for μ carries very little information compared with the data, while the prior distribution for σ^2_ϵ carries as little information as possible and is not contradicted by the data.

The results of our analysis are summarized in Table 5.2 and Figure 5.1. The various estimates of log potency ratio are very similar indeed; however, there is a discrepancy between the modal estimates of σ^2 and σ^2_ϵ from $\pi(\alpha,\beta,\mu,c_1,\ldots,c_5,\sigma^2,\sigma^2_\epsilon|\underline{y})$ and the modal estimate of $S^2=\sigma^2/\sigma^2_\epsilon$ from $\pi(\mu,S^2|\underline{y})$.

	Standard Preparation			Test Preparation		
Dose	1/200	1/400	1/800	1/200	1/400	1/800
Day						
1	15.0	22.5	27.0	21.0	25.0	30.0
2	15.0	18.5	19.5	17.25	21.25	25.0
3	18.0	24.25	30.5	20.5	28.5	36.0
4	15.5	18.75	22.25	18.5	21.75	27.5
5	18.0	22.0	27.0	22.0	26.0	31.5

Table 5.1 Data from factor VIII [illegible]

<u>Mode of $\pi(\alpha,\beta,\mu,\epsilon_1,\ldots,\epsilon_5,\sigma^2,\sigma^2_\epsilon|y)$</u>

$\alpha = -20.5$

$\beta = -15.8$

$\mu = -.257$

$\epsilon_1 = .594$

$\epsilon_2 = -3.47$

$\epsilon_3 = 3.78$

$\epsilon_4 = -1.85$

$\epsilon_5 =$ [illegible]

$\sigma^2 = 1.52$

$\sigma^2_\epsilon = 3.78$

<u>Mode of $\pi(\mu,S^2|y)$</u>

$\mu = -.251$

$S^2 = .261$

<u>Mean of $\pi(\mu|y,\hat{S}^2)$</u>

$-.276$

($\hat{S}^2$ is the value of S^2 at the mode of $\pi(\mu,S^2|y)$)

Table 5.2 <u>Results of analysis of [illegible] VIII data with prior parameters $\mu_0=0.0$, $\Sigma_{33}=1.5$, $\nu=\lambda=0$, $\nu_\alpha=\nu_\beta=1$, $\Sigma_{11}\to\infty$, $\Sigma_{22}\to\infty$.</u>

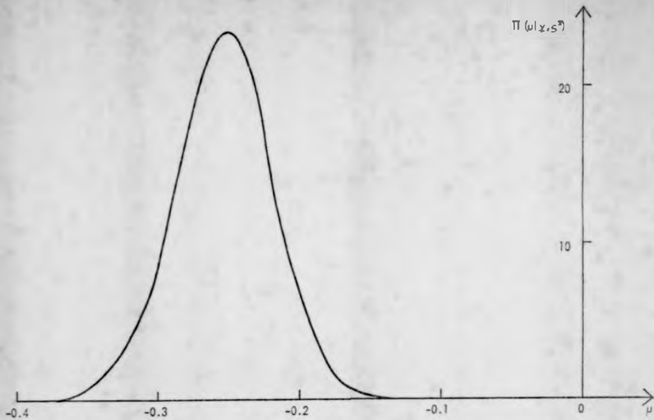


Figure 5.1 Approximate marginal posterior density of μ, assuming σ^2 to be known and equal to its value at the mode of the joint density of μ and σ^2, for the factor VIII data.

Chapter 8. A Model Combining Information From Several Assays.

8.1 Introduction

In many cases the need arises to combine information from several different assays, and we shall devote the next few chapters to considering this problem. The model that we shall consider first is a model combining information from several assays and we shall assume our prior knowledge of the parameters of every assay to be exchangeable. This model is a straight forward extension of the two stage model for the analysis of a single assay that was discussed in chapter 2 to a three stage model. The extra stage is necessary since the data will now contain some information about the parameters in the second stage of the model. Suppose we wish to combine information from m assays, then the model is as follows:

1st stage: $\underset{\sim}{y}_j \sim N\left[\underset{\sim}{X}_j \begin{bmatrix} \alpha_j \\ \beta_j\mu_j \\ \beta_j \end{bmatrix}, \sigma_j^2 \underset{\sim}{I}_{n_j}\right]$ independently for $j=1,\ldots,m$,

2nd stage: $\begin{pmatrix} \alpha_j \\ \beta_j \\ \mu_j \end{pmatrix} \sim N\left[\begin{pmatrix} \alpha_0 \\ \beta_0 \\ \mu_0 \end{pmatrix}, \underset{\sim}{\Sigma}\right]$ independently for $j=1,\ldots,m$,

3rd stage: $\begin{pmatrix} \alpha_0 \\ \beta_0 \\ \mu_0 \end{pmatrix} \sim N\left[\begin{pmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \end{pmatrix}, \underset{\sim}{\Phi}\right]$.

Where the suffix j refers to the j^{th} assay in the series. $\underset{\sim}{X}_j$ is a matrix of the form

$$\underset{\sim}{X}_j = \begin{bmatrix} 1 & z_{1j} & x_{1j} \\ 1 & z_{2j} & x_{2j} \\ \vdots & \vdots & \vdots \\ 1 & z_{n_j j} & x_{n_j j} \end{bmatrix}$$

and nj is the number of responses in the j^{th} assay. For the moment we assume all variances and covariances to be known.

There are two main situations where this model may be appropriate. The first is where a manufacturer has made several batches of a preparation, has calibrated them all against the same standard using the same assay medium, and wishes to make inferences about the manufacturing process in general. The second is in collaborative assays where several laboratories carry out assays using the same pair of substances and wish to combine their results. In this latter case one could argue that the true potency ratio will be the same in each assay and therefore the model should stipulate that the μ_j are all identical. However, each assay will be carried out by a different person in a different laboratory, and it may be that for some types of assay the effect of variation in personal technique is great enough to make such an assumption unreasonable. A model which does stipulate that the μ_j are all identical is discussed in chapter 7.

For both the cases described above the model is rather crude; in the first case we have not allowed for any trends in the parameters, and in the second case we have assumed that all the assays are carried out on the same medium. The model could be extended to cover either of these refinements.

In both the cases described above interest will centre on the second stage parameters. In the case of the manufacturer carrying out assays on different batches of a preparation, estimates of the second stage parameters could be used in estimating the parameters of a prior distribution for the analysis of an assay on a further batch of preparation. In the collaborative assay, inferences about the log potency ratio of the two substances under investigation would ideally be based on the marginal posterior distribution $\mu_*|y$.

6.2 Posterior Distributions for Known Covariance S[illegible]

Before combining the information from the data with our prior information, we need to combine the information in the second and third stages of the model. We get the following prior density for the first and second stage parameters:

$$\pi(\alpha_o,\beta_o,\mu_o,\alpha_1,\beta_1,\mu_1,\ldots\alpha_m,\beta_m,\mu_m|\eta_1,\eta_2,\eta_3) =$$

$$\exp-\tfrac{1}{2}\left[\sum_{j=1}^{m}\begin{pmatrix}\alpha_j\\ \beta_j\\ \mu_j\end{pmatrix}^T \underset{\sim}{\Sigma}^{-1}\begin{pmatrix}\alpha_j\\ \beta_j\\ \mu_j\end{pmatrix} + \begin{pmatrix}\alpha_o\\ \beta_o\\ \mu_o\end{pmatrix}^T (m\underset{\sim}{\Sigma}^{-1}+\underset{\sim}{\phi}^{-1})\begin{pmatrix}\alpha_o\\ \beta_o\\ \mu_o\end{pmatrix}\right.$$

$$\left. -2\left\{\begin{pmatrix}\bar{\alpha}\\ \bar{\beta}\\ \bar{\mu}\end{pmatrix}^T m\underset{\sim}{\Sigma}^{-1} + \begin{pmatrix}\eta_1\\ \eta_2\\ \eta_3\end{pmatrix}^T\phi^{-1}\right\}\right] , \qquad (6.2)$$

where $\bar{\alpha} = \frac{1}{m}\sum_{j=1}^{m}\alpha_j$ and similarly for $\bar{\beta}$ and $\bar{\mu}$.

Combining the above density with the likelihood, the joint posterior density of the first and second stage parameters is:

$$\pi(\alpha_o,\beta_o,\mu_o,\alpha_1,\beta_1,\mu_1,\ldots\alpha_m,\beta_m,\mu_m|\underset{\sim}{y},\eta_1,\eta_2,\eta_3) =$$

$$\exp-\tfrac{1}{2}\left[\sum_{j=1}^{m}\left\{\frac{1}{\sigma_j^2}\begin{pmatrix}\alpha_j\\ \beta_j\mu_j\\ \beta_j\end{pmatrix}^T x_j^T x_j\begin{pmatrix}\alpha_j\\ \beta_j\mu_j\\ \beta_j\end{pmatrix} + \begin{pmatrix}\alpha_j\\ \beta_j\\ \mu_j\end{pmatrix}^T \underset{\sim}{\Sigma}^{-1}\begin{pmatrix}\alpha_j\\ \beta_j\\ \mu_j\end{pmatrix}\right\}\right.$$

$$
-2\sum_{j=1}^{m}\left\{\frac{1}{\sigma_j^2}\,\underset{\sim}{y}_j^{T}\,\underset{\sim}{X}_j\begin{pmatrix}\alpha_j\\ \beta_j\mu_j\\ \beta_j\end{pmatrix}+\begin{pmatrix}\alpha_j\\ \beta_j\\ \mu_j\end{pmatrix}^{T}\underset{\sim}{\Sigma}^{-1}\begin{pmatrix}\alpha_o\\ \beta_o\\ \mu_o\end{pmatrix}\right\}+\begin{pmatrix}\alpha_o\\ \beta_o\\ \mu_o\end{pmatrix}^{T}(m\underset{\sim}{\Sigma}^{-1}+\Phi^{-1})\begin{pmatrix}\alpha_o\\ \beta_o\\ \mu_o\end{pmatrix}-2\begin{pmatrix}\eta_1\\ \eta_2\\ \eta_3\end{pmatrix}^{T}\Phi^{-1}\begin{pmatrix}\alpha_o\\ \beta_o\\ \mu_o\end{pmatrix}\Bigg]
$$

(8.3)

If our prior knowledge at the third stage of the model is extremely weak, the elements of $\underset{\sim}{\Phi}^{-1}$ will be zero, and those terms involving $\underset{\sim}{\Phi}^{-1}$ in the exponent of (8.3) will disappear. The conditions under which (8.3) is a normed density when $\underset{\sim}{\Phi}^{-1}=\underset{\sim}{0}$ will be investigated at the end of this section.

The mode of (8.3) occurs at the point:

$$
\begin{pmatrix}\alpha_o\\ \beta_o\\ \mu_o\end{pmatrix}=\left(m\underset{\sim}{\Sigma}^{-1}+\underset{\sim}{\Phi}^{-1}\right)^{-1}\left[m\underset{\sim}{\Sigma}^{-1}\begin{pmatrix}\text{[illegible]}\end{pmatrix}+\underset{\sim}{\Phi}^{-1}\begin{pmatrix}\eta_1\\ \eta_2\\ \eta_3\end{pmatrix}\right],
$$

$$
\alpha_j=\frac{\dfrac{1}{\sigma_j^2}\displaystyle\sum_{k=1}^{n_j}(y_{kj}-\beta_j\mu_j z_{kj}-\beta_j x_{kj})+\alpha_o\Sigma^{11}-(\beta_j-\beta_o)\Sigma^{12}-(\mu_j-\mu_o)\Sigma^{13}}{\dfrac{n_j}{\sigma_j^2}+\Sigma^{11}},\quad j=1,\ldots,m,
$$

$$
\beta_j=\frac{\dfrac{1}{\sigma_j^2}\displaystyle\sum_{k=1}^{n_j}(y_{kj}-\alpha_j)(x_{kj}+\mu_j z_{kj})+\beta_o\Sigma^{22}-(\alpha_j-\alpha_o)\Sigma^{12}-(\mu_j-\mu_o)\Sigma^{23}}{\dfrac{1}{\sigma_j^2}\displaystyle\sum_{k=1}^{n_j}(x_{kj}+\mu_j z_{kj})^2+\Sigma^{22}},\quad j=1,\ldots,m,
$$

$$\mu_j = \frac{\dfrac{\beta_j}{\sigma_j^2}\displaystyle\sum_{k=1}^{nj} z_{kj}(y_{kj}-\alpha_j-\beta_j x_{kj}) + \mu_o\Sigma^{33} - (\alpha_j-\alpha_o)\Sigma^{13} - (\beta_j-\beta_o)\Sigma^{33}}{\dfrac{\beta_j^{\,2}}{\sigma^2}\displaystyle\sum_{k=1}^{nj} z_{kj}^2 + \Sigma^{33}}, \quad j=1\ldots m, \tag{6.4}$$

The modal values for the first stage parameters of an individual assay are very similar to the mode of the joint posterior density of the first stage parameters in the analysis of a single assay as given by **2.3**. There are two slight differences. Firstly, in this case, the second stage parameters α_o, β_o, and μ_o are themselves modal values whereas in the single assay case they were known, and secondly, the second stage variance Σ has a slightly different status in the two cases. In the multiple assay case Σ expresses our opinion about the similarity of the parameters of the different assays, while the strength of our opinion about the likely location of the parameters is expressed in the third stage variance $\underset{\sim}{C}$. By integrating over the second stage parameters α_o, β_o and μ_o in (6.2) we have that the prior density for $\alpha_1, \beta_1, \mu_1, \ldots, \alpha_m, \beta_m, \mu_m$ is

$$\pi(\alpha_1,\beta_1,\mu_1,\ldots,\alpha_m,\beta_m,\mu_m \mid \text{[illegible]}) \ldots \begin{bmatrix} \begin{pmatrix}\alpha_1\\ \beta_1\\ \mu_1\end{pmatrix} - \begin{pmatrix}\alpha^*\\ \beta^*\\ \mu^*\end{pmatrix} \\ \vdots \\ \begin{pmatrix}\alpha_m\\ \beta_m\\ \mu_m\end{pmatrix} - \begin{pmatrix}\alpha^*\\ \beta^*\\ \mu^*\end{pmatrix} \end{bmatrix}^T (\underset{\sim}{V}^*)^{-1} \begin{bmatrix} \begin{pmatrix}\alpha_1\\ \beta_1\\ \mu_1\end{pmatrix} - \begin{pmatrix}\alpha^*\\ \beta^*\\ \mu^*\end{pmatrix} \\ \vdots \\ \begin{pmatrix}\alpha_m\\ \beta_m\\ \mu_m\end{pmatrix} - \begin{pmatrix}\alpha^*\\ \beta^*\\ \mu^*\end{pmatrix} \end{bmatrix} \tag{6.5}$$

where $\underset{\sim}{V}^*$ is the $3m \times 3m$ matrix $\begin{bmatrix} \Sigma & 0 & \cdots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \Sigma \end{bmatrix} + \begin{bmatrix} I_2 \\ \vdots \\ I_2 \end{bmatrix} \Phi \begin{bmatrix} I_2 \\ \vdots \\ I_2 \end{bmatrix}^T$

and $$\begin{pmatrix} \alpha^* \\ \beta^* \\ \mu^* \end{pmatrix} = (\Sigma + m\Phi)\Sigma^{-1}(m\Sigma^{-1} + \Phi^{-1})^{-1} \Phi^{-1} \begin{pmatrix} n_1 \\ n_2 \\ n_3 \end{pmatrix}$$

Hence the prior distribution for the first stage parameters of an individual assay, say the j^{th} is

$$\begin{pmatrix} \alpha_j \\ \beta_j \\ \mu_j \end{pmatrix} \sim N\left\{ \begin{pmatrix} \alpha^* \\ \beta^* \\ \mu^* \end{pmatrix}, (\Sigma + \Phi) \right\} .$$

In the single assay case Σ expresses the strength of our opinion on the two sources of variation and will be comparable with $(\Sigma + \Phi)$ in the multiple assay case.

By integrating over $\alpha_1 \beta_1, \ldots, \alpha_m, \beta_m$ in 8.3 we can find the joint posterior density of $\alpha_o, \beta_o, \mu_o, \mu_1, \ldots, \mu_m$:

$$\pi(\alpha_o, \beta_o, \mu_o, \mu_1 \ldots \mu_m | \underset{\sim}{y}, n_1, n_2, n_3) = \prod_{j=1}^{m} |A_j|^{\frac{1}{2}} \exp -\tfrac{1}{2} \left[\begin{pmatrix} \alpha_o \\ \beta_o \\ \mu_o \end{pmatrix}^T (m\Sigma^{-1} + \Phi^{-1}) \begin{pmatrix} \alpha_m \\ \beta_m \\ \mu_m \end{pmatrix} \right.$$

$$\left. -2 \begin{pmatrix} n_1 \\ n_2 \\ n_3 \end{pmatrix}^T \Phi^{-1} \begin{pmatrix} \alpha_o \\ \beta_o \\ \mu_o \end{pmatrix} + \sum_{j=1}^{m} \{ \mu_j^2 \Sigma^{33} - 2\mu_j (\alpha_o \Sigma^{13} + \beta_o \Sigma^{23} + \mu_o \Sigma^{33}) \} \right.$$

$$-\sum_{j=1}^{m}\left[\begin{pmatrix}a_j\\ b_j\end{pmatrix}+\begin{bmatrix}\Sigma^{11}\Sigma^{12}\\ \Sigma^{12}\Sigma^{22}\\ \Sigma^{13}\Sigma^{23}\end{bmatrix}^T\begin{pmatrix}\alpha_o\\ \beta_o\\ \mu_o-\mu_j\end{pmatrix}\right]^T \underset{\sim}{V}_j\left[\begin{pmatrix}a_j\\ b_j\end{pmatrix}+\begin{bmatrix}\Sigma^{11}\Sigma^{12}\\ \Sigma^{12}\Sigma^{22}\\ \Sigma^{13}\Sigma^{23}\end{bmatrix}^T\begin{pmatrix}\alpha_o\\ \beta_o\\ \mu_o-\mu_j\end{pmatrix}\right]\right] \quad (6.6)$$

where
$$\underset{\sim}{V}_j=\begin{bmatrix}\dfrac{n_j}{\sigma_j^2}+\Sigma^{11} & \dfrac{1}{\sigma_j^2}\displaystyle\sum_{k=1}^{n_j}(x_{kj}+\mu_j z_{kj})+\Sigma^{12}\\ \dfrac{1}{\sigma_j^2}\displaystyle\sum_{k=1}^{n_j}(x_{kj}+\mu_j z_{kj})+\Sigma^{12} & \dfrac{1}{\sigma_j^2}\displaystyle\sum_{k=1}^{n_j}(\mu_j^2 z_{kj}^2+2\mu_j x_{kj}z_{kj}+x_{kj}^2)+\Sigma^{22}\end{bmatrix}^{-1},$$

$$a_j=\frac{1}{\sigma_j^2}\sum_{k=1}^{n_j}y_{kj}\quad,$$

and
$$b_j=\frac{1}{\sigma_j^2}\sum_{k=1}^{n_j}(x_{kj}y_{kj}+\mu_j z_{kj}y_{kj})\quad.$$

Unfortunately we cannot integrate over $\mu_1,\dots,\mu_m$ in (6.6). This means we cannot obtain the marginal distributions of the parameters we are interested in analytically. In general we will not be able to find these distributions numerically either, since to do so would involve carrying out numerical integrations in m dimensions. If we are interested in the three second stage parameters α_o,β_o,μ_o, we could estimate them by the mode of 6.6. Even this mode cannot be found analytically but must be obtained numerically. If, as in collaborative assays, we are interested in μ_o but not in α_o or β_o, we can integrate over α_o and β_o in 6.6 to obtain the posterior distribution of $\mu_o,\mu_1,\dots,\mu_m$:

$$\pi(\mu_o,\mu_1\dots\mu_m|n_1,n_2,n_3)=\left|\underset{\sim}{V}^{-1}+\underset{\sim}{S}^{-1}\sum_{j=1}^{m}\underset{\sim}{V}_j\underset{\sim}{D}_j\right|^{-\frac{1}{2}}\left(\prod_{j=1}^{m}|\underset{\sim}{V}_j|^{\frac{1}{2}}\right)\exp-\tfrac{1}{2}\sum_{j=1}^{m}\Big[(\mu_j-\mu_o)^2 \text{[illegible]}$$

$$+\mu_o^2\phi^{33}-2\mu_o\begin{pmatrix}\phi^{13}\\ \phi^{23}\\ \phi^{33}\end{pmatrix}^T\begin{pmatrix}n_1\\ n_2\\ n_3\end{pmatrix}-\sum_{j=1}^{m}\left\{\begin{pmatrix}a_j\\ b_j\end{pmatrix}+(\mu_o-\mu_j)\begin{pmatrix}\Sigma^{13}\\ \Sigma^{23}\end{pmatrix}\right\}^T \underset{\sim}{V}_j\left\{\begin{pmatrix}a_j\\ b_j\end{pmatrix}+(\mu_o-\mu_j)\begin{pmatrix}\Sigma^{13}\\ \Sigma^{23}\end{pmatrix}\right\}$$

$$\left.-\underset{\sim}{X}^T\left(\underset{\sim}{\Psi}^{-1}+\underset{\sim}{S}^{-1}\sum_{j=1}^{m}\underset{\sim}{V}_j\underset{\sim}{D}_j\right)^{-1}\underset{\sim}{X}\right] \qquad (6.7)$$

where $\underset{\sim}{\Psi}^{-1}=\begin{bmatrix}\phi^{11} & \phi^{12}\\ \phi^{12} & \phi^{22}\end{bmatrix}$, ϕ^{ij} being the $(ij)^{th}$ element of $\underset{\sim}{\Phi}^{-1}$,

$$\underset{\sim}{S}^{-1}=\begin{bmatrix}\Sigma^{11} & \Sigma^{12}\\ \Sigma^{12} & \Sigma^{22}\end{bmatrix},$$

$$\underset{\sim}{D}_j=\frac{1}{\sigma_j^2}\begin{bmatrix}n_j & \sum_{k=1}^{n_j}(x_{kj}+\mu_j z_{kj})\\ \sum_{k=1}^{n_j}(x_{kj}+\mu_j z_{kj}) & \sum_{k=1}^{n_j}(\mu_j^2 z_{kj}^2+2\mu_j x_{kj}z_{kj}+x_{kj}^2)\end{bmatrix},$$

and $$\underset{\sim}{X}=\underset{\sim}{S}^{-1}\sum_{j=1}^{m}\underset{\sim}{V}_j\begin{pmatrix}a_j\\ b_j\end{pmatrix}+\sum_{j=1}^{m}(\mu_j-\mu_o)\underset{\sim}{D}_j\underset{\sim}{V}_j\begin{pmatrix}\Sigma^{13}\\ \Sigma^{23}\end{pmatrix}+\begin{bmatrix}\phi^{11} & \phi^{12}\\ \phi^{12} & \phi^{22}\\ \phi^{13} & \phi^{23}\end{bmatrix}^T\begin{pmatrix}n_1\\ n_2\\ n_3\end{pmatrix}-\begin{pmatrix}\phi^{13}\\ \phi^{23}\end{pmatrix}\mu_o.$$

One could estimate μ_o by the mode of 6.7. This can be found numerically.

We can proceed one step further and find the joint density of $\mu_1,\ldots,\mu_m$ by integrating over μ_o in 6.7. This density will rarely be of any practical interest but it is useful in investigating the conditions under which it is permissible to consider uniform priors for all three third stage parameters. If we set $\underset{\sim}{\Phi}^{-1}=0$ in $\pi(\mu_1,\ldots,\mu_m|\underset{\sim}{y})$, then

$$\pi(\mu_1,\ldots,\mu_m|\underset{\sim}{y}) \propto \left|\underset{\sim}{S}^{-1}\sum_{j=1}^{m}\underset{\sim}{V}_j\underset{\sim}{D}_j\right|^{-\frac{1}{2}}\left\{\prod_{j=1}^{m}|\underset{\sim}{V}_j|^{\frac{1}{2}}\right\}\exp-\tfrac{1}{2}\left[\Sigma^{33}\sum_{j=1}^{m}\mu_j^2-\frac{|\Sigma^{-1}|}{m|\underset{\sim}{S}^{-1}|}\left(\sum_{j=1}^{m}\mu_j\right)^2\right.$$

$$-\left\{\underset{\sim}{S}^{-1}\sum_{j=1}^{m}\underset{\sim}{V}_j\begin{pmatrix}a_j\\b_j\end{pmatrix}+\sum_{j=1}^{m}\mu_j\underset{\sim}{D}_j\underset{\sim}{V}_j\begin{pmatrix}\Sigma^{13}\\\Sigma^{23}\end{pmatrix}\right\}^T\left[\underset{\sim}{S}^{-1}\sum_{j=1}^{m}\underset{\sim}{V}_j\underset{\sim}{D}_j\right]^{-1}\left\{\underset{\sim}{S}^{-1}\sum_{j=1}^{m}\underset{\sim}{V}_j\begin{pmatrix}a_j\\b_j\end{pmatrix}+\sum_{j=1}^{m}\mu_j\underset{\sim}{D}_j\underset{\sim}{V}_j\begin{pmatrix}\Sigma^{13}\\\Sigma^{23}\end{pmatrix}\right\}$$

$$\left.-\sum_{j=1}^{m}\left\{\begin{pmatrix}a_j\\b_j\end{pmatrix}-\mu_j\begin{pmatrix}\Sigma^{13}\\\Sigma^{23}\end{pmatrix}\right\}^T\underset{\sim}{V}_j\left\{\begin{pmatrix}a_j\\b_j\end{pmatrix}-\mu_j\begin{pmatrix}\Sigma^{13}\\\Sigma^{23}\end{pmatrix}\right\}\right] . \qquad (6.8)$$

If we call the expression on the right hand side of the $\propto$ sign $g(\mu_1,\ldots\mu_m)$ then the posterior density of $\mu_1,\ldots,\mu_m$, and consequently all the other posterior densities given in this section, will be normed when $\underset{\sim}{\Phi}^{-1}=\underset{\sim}{0}$ only if the m-dimensional integral $\int\ldots\int g(\mu_1,\ldots,\mu_m)d\mu_1\ldots d\mu_m$ is finite. In the following paragraphs we give a loose argument indicating when this integral will be finite. We have not given a rigorous proof since such a proof, although straightforward, would be very lengthy.

We assume that there are at least two assays under consideration, and that for each of them at least two different doses have been administered for at least one preparation, and at least one dose of each preparation has been administered. We also assume that $\underset{\sim}{\Sigma}$ is a positive definite symmetric matrix. Examination of the expressions

$$\sum_{j=1}^{m}\begin{pmatrix}a_j\\b_j\end{pmatrix}\underset{\sim}{V}_j\begin{pmatrix}a_j\\b_j\end{pmatrix}\cdot\left[\underset{\sim}{S}^{-1}\sum_{j=1}^{m}\underset{\sim}{V}_j\begin{pmatrix}a_j\\b_j\end{pmatrix}\right]^T\left[\underset{\sim}{S}^{-1}\sum_{j=1}^{m}\underset{\sim}{V}_j\underset{\sim}{D}_j\right]^{-1}\left[\underset{\sim}{S}^{-1}\sum_{j=1}^{m}\underset{\sim}{V}_j\begin{pmatrix}a_j\\b_j\end{pmatrix}\right] ,$$

and $\left|\underset{\sim}{S}^{-1}\sum_{j=1}^{m}\underset{\sim}{V}_j\underset{\sim}{D}_j\right|^{-\frac{1}{2}}$ shows them to be bounded above and below for all μ_j, $j=1,\ldots m$, and to tend to finite limits as all the μ_j become simultaneously large in absolute value. Also $\left|\underset{\sim}{S}^{-1}\sum_{j=1}^{m}\underset{\sim}{V}_j\underset{\sim}{D}_j\right|^{-\frac{1}{2}}$ is always strictly greater than zero since

$\left[\underset{\sim}{S}^{-1}\sum_{j=1}^{m}\underset{\sim}{V}_j\underset{\sim}{D}_j\right]$ is a symmetric positive definite matrix. Hence if all

the μ_j are large in absolute value

$$g(\mu_1\ldots\mu_m) \simeq k\prod_{j=1}^{m}|\underset{\sim}{V}_j|^{\frac{1}{2}}\exp-\tfrac{1}{2}\left[\sum_{j=1}^{m}\mu_j^2\left\{\Sigma^{33}-\begin{pmatrix}\Sigma^{13}\\ \Sigma^{23}\end{pmatrix}^T\underset{\sim}{V}_j\begin{pmatrix}\Sigma^{13}\\ \Sigma^{23}\end{pmatrix}\right\}-\left(\sum_{j=1}^{m}\mu_j\right)^2\frac{|\underset{\sim}{\Sigma}^{1}|}{m|\underset{\sim}{S}^{-1}|}\right.$$

$$-\begin{pmatrix}\Sigma^{13}\\ \Sigma^{23}\end{pmatrix}^T\left[\sum_{j=1}^{m}\underset{\sim}{V}_j\underset{\sim}{D}_j\mu_j\right]\left[\underset{\sim}{S}^{-1}\sum_{j=1}^{m}\underset{\sim}{V}_j\underset{\sim}{D}_j\right]^{-1}\left[\sum_{j=1}^{m}\mu_j\underset{\sim}{D}_j\underset{\sim}{V}_j\right]\begin{pmatrix}\Sigma^{13}\\ \Sigma^{23}\end{pmatrix}$$

$$\left.+2\sum_{j=1}^{m}\mu_j\left\{\begin{pmatrix}\Sigma^{13}\\ \Sigma^{23}\end{pmatrix}^T\underset{\sim}{V}_j\begin{pmatrix}a_j\\ b_j\end{pmatrix}+\begin{pmatrix}\Sigma^{13}\\ \Sigma^{23}\end{pmatrix}^T\underset{\sim}{V}_j\underset{\sim}{D}_j\left[\underset{\sim}{S}^{-1}\sum_{j=1}^{m}\underset{\sim}{V}_j\underset{\sim}{D}_j\right]^{-1}\underset{\sim}{S}^{-1}\sum_{j=1}^{m}\underset{\sim}{V}_j\begin{pmatrix}a_j\\ b_j\end{pmatrix}\right\}\right] .$$

for some positive constant k.

Also, we have the following limiting results:

$$\lim_{\mu_j\to\pm\infty}\underset{\sim}{V}_j=\begin{bmatrix}\left(\dfrac{\sum_{k=1}^{n_j} z_{kj}^2}{\frac{n_j}{\sigma_j^2}Szz^j+\Sigma^{11}\sum_{k=1}^{n_j} z_{kj}^2}\right) & 0\\ 0 & 0\end{bmatrix},$$

$$\text{and }\lim_{\mu_j\to\pm\infty}\underset{\sim}{V}_j\underset{\sim}{D}_j=\begin{pmatrix}\dfrac{\frac{n_j}{\sigma_j^2}Szy^j}{\frac{n_j}{\sigma_j^2}Szz^j+\Sigma^{11}\sum_{k=1}^{n_j} z_{kj}^2}\\ 0\end{pmatrix}.$$

So if all the μ_j are large in absolute value,

$$g(\mu_1\ldots\mu_m) \simeq k \prod_{j=1}^{m} |\underset{\sim}{V}_j|^{\frac{1}{2}} \exp -\tfrac{1}{2}\left\{\begin{pmatrix}\mu_1\\ \vdots \\ \mu_m\end{pmatrix}^T \underset{\sim}{W} \begin{pmatrix}\mu_1\\ \vdots \\ \mu_m\end{pmatrix} + 2\sum_{j=1}^{m} c_j \mu_j\right\} .$$

where $c_1 \ldots c_m$ are constants independent of $\mu_1 \ldots \mu_m$,

$$\text{and } \underset{\sim}{W} = \begin{bmatrix} \Sigma^{33}-(\Sigma^{13})^2 d_1 & 0 & \cdots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \Sigma^{33}-(\Sigma^{13})^2 d_m \end{bmatrix} - \frac{1}{m}\frac{|\underset{\sim}{\Sigma}^{-1}|}{|\underset{\sim}{S}^{-1}|}\underset{\sim}{J}_m - \begin{pmatrix}\Sigma^{13}\\ \Sigma^{23}\end{pmatrix}^T \underset{\sim}{S} \begin{pmatrix}\Sigma^{13}\\ \Sigma^{23}\end{pmatrix} \frac{1}{\sum_{j=1}^{m} e_j} \cdot \begin{bmatrix} e_1 \\ \vdots \\ e_m \end{bmatrix} \begin{bmatrix} e_1 \\ \vdots \\ e_m \end{bmatrix}^T$$

$$\text{where } d_j = \frac{\sum_{k=1}^{n_j} z_{kj}^2}{\frac{n_j}{\sigma_j^2} Szz^j + \Sigma^{11} \sum_{k=1}^{n_j} z_{kj}^2} \qquad \text{and } e_j = \frac{\frac{n_j}{\sigma_j^2} Szz^j}{\frac{n_j}{\sigma_j^2} Szz^j + \Sigma^{11} \sum_{k=1}^{n_j} z_{kj}^2} .$$

If $\underset{\sim}{W}$ is positive definite the integral $\int \ldots \int g(\mu_1 \ldots \mu_m) d\mu_1 \ldots d\mu_m$ will be finite, otherwise it will not, the term $\prod_{j=1}^{m} |\underset{\sim}{V}_j|^{\frac{1}{2}}$ playing a similar role to $\{A(\mu)\}^{-\frac{1}{2}}$ in the discussion surrounding 2.7. In order for $\underset{\sim}{W}$ to be positive definite, all its principal minors must be positive. For this we need

$$\prod_{i=1}^{p} f_i \left[\left\{\left(\frac{1}{p}-\frac{1}{m}\right)\left(\Sigma^{33}-\frac{(\Sigma^{13})^2}{\Sigma^{11}}\right)+\frac{1}{m}\left[\begin{pmatrix}\Sigma^{13}\\ \Sigma^{23}\end{pmatrix}^T \underset{\sim}{S} \begin{pmatrix}\Sigma^{13}\\ \Sigma^{23}\end{pmatrix} - \frac{(\Sigma^{13})^2}{\Sigma^{11}}\right]\right\} \frac{1}{\sum_{j=1}^{m} e_j}\right.$$

$$\left. \times \left\{\Sigma^{11} \sum_{j=1}^{m} e_j \sum_{i=1}^{p} \frac{d_i}{f_i} - \begin{pmatrix}\Sigma^{13}\\ \Sigma^{23}\end{pmatrix}^T \underset{\sim}{S} \begin{pmatrix}\Sigma^{13}\\ \Sigma^{23}\end{pmatrix} \left[\sum_{i=1}^{p} \frac{e_i^2}{f_i} \sum_{i=1}^{p} \frac{1}{f_i} - \left(\sum_{i=1}^{p} \frac{e_i}{f_i}\right)^2\right]\right\}\right]$$

$$+\sum_{i=1}^{p}\frac{a_i}{f_i}\left(\frac{1}{p}-\frac{1}{m}\right)\Sigma^{33}-\begin{pmatrix}\Sigma^{13}\\ \Sigma^{23}\end{pmatrix}^{T}\underset{\sim}{S}\begin{pmatrix}\Sigma^{13}\\ \Sigma^{23}\end{pmatrix}\left\{\frac{\sum_{i=1}^{p}a_i}{\sum_{j=1}^{m}a_j}\cdot\frac{1}{p}-\frac{1}{m}\right\}\Bigg]>0,\quad p=1,\ldots,m, \tag{6.9}$$

where $f_i=\Sigma^{33}-(\Sigma^{13})^2 d_i$.

After some algebra we can show that 6.9.holds precisely when

$\begin{pmatrix}\Sigma^{13}\\ \Sigma^{23}\end{pmatrix}^{T}\underset{\sim}{S}\begin{pmatrix}\Sigma^{13}\\ \Sigma^{23}\end{pmatrix}-\dfrac{(\Sigma^{13})^2}{\Sigma^{11}}>0.$ We can also show that

$$\begin{pmatrix}\Sigma^{13}\\ \Sigma^{23}\end{pmatrix}^{T}\underset{\sim}{S}\begin{pmatrix}\Sigma^{13}\\ \Sigma^{23}\end{pmatrix}-\frac{(\Sigma^{13})^2}{\Sigma^{11}}=\frac{\Sigma_{23}^{2}\,|\underset{\sim}{\Sigma}|^2}{\Sigma_{33}\left(\Sigma_{22}\Sigma_{33}-\Sigma_{23}^{2}\right)}.$$

Consequently we can set $\underset{\sim}{\phi}^{-1}=\underset{\sim}{0}$ provided Σ_{23} is not equal to zero Unfortunately we have not been very successful in our attempts to interpret this condition. Suppose in (6.1) that $\underset{\sim}{\phi}^{-1}=\underset{\sim}{0}$ and $\Sigma_{13}=\Sigma_{23}=0$, then we have effectively a uniform prior distribution for each μ_j at the second stage, independently of the prior distributions for any α_j or β_j. This situation is very similar to having a uniform prior distribution for μ in the single assay case and so it seems quite reasonable that the posterior distributions are unnormed. It now remains to explain why a non-zero Σ_{13} does not affect the above situation, while a non-zero Σ_{23} does. We feel that this must be due to the asymmetry in the first stage of the model but we have been unable to make any precise statements about it.

6.3 Unknown Variances and Large Sample Theory

We now remove the assumptions, made in the last section, that the first stage residual variances σ^2_j, $j=1,\ldots,m$ and the second stage covariance matrix $\underset{\sim}{\Sigma}$ are all known. We shall use the relevant conjugate prior distributions for each of these parameters: that is the inverse χ^2-distribution for the residual variances and the Wishart distribution for $\underset{\sim}{\Sigma}^{-1}$. In the line with our assumption of exchangeable prior knowledge about the other parameters it would be most reasonable to assume exchangeable prior knowledge about the residual variances of the assays, however for simplicity we have taken identical independent prior distributions for these. Our prior densities will be:

$$\pi(\sigma^2_j|\nu,\lambda) \propto (\sigma^2_j)^{-\frac{(\nu+2)}{2}} \exp\left\{-\frac{\nu\lambda}{2\sigma^2_j}\right\}, \quad (\sigma^2_j > 0); \text{ independently for}$$

$j=1,\ldots m$, and independent of the above densities,

$$\pi(\underset{\sim}{\Sigma}^{-1}|\underset{\sim}{R},\rho) \propto |\underset{\sim}{\Sigma}|^{-\frac{1}{2}(\rho-4)} \exp -\tfrac{1}{2}\mathrm{tr}(\underset{\sim}{\Sigma}^{-1}\underset{\sim}{R}) \quad ; \quad \underset{\sim}{\Sigma} > \underset{\sim}{0} .$$

$\underset{\sim}{R}$ is a 3×3 matrix, ρ is an integer, and the values of these two together with the values of ν and λ depend on the nature and precision of our prior knowledge about the parameters concerned. We can now write down the joint posterior distribution of all the parameters in the model:

$$\pi(\alpha_o,\beta_o,\mu_o,\underset{\sim}{\Sigma}^{-1},\alpha_1,\beta_1,\mu_1,\sigma^2_1,\ldots,\alpha_m,\beta_m,\mu_m,\sigma^2_m|\underset{\sim}{y}_1,\ldots\underset{\sim}{y}_m,n_1,n_2,n_3,\underset{\sim}{\phi},\nu,\lambda,\underset{\sim}{R},\rho) \propto$$

$$(\sigma^2_1)^{-\frac{n_1}{2}} \ldots\ldots(\sigma^2_m)^{-\frac{n_m}{2}} \exp -\tfrac{1}{2}\sum_{j=1}^{m} \frac{1}{\sigma^2_j}\left\{\underset{\sim}{y}_j-\underset{\sim}{X}_j\begin{pmatrix}\alpha_j\\ \beta_j\mu_j\\ \beta_j\end{pmatrix}\right\}^T\left\{\underset{\sim}{y}_j-\underset{\sim}{X}_j\begin{pmatrix}\alpha_j\\ \beta_j\mu_j\\ \beta_j\end{pmatrix}\right\}$$

$$\times \quad |\Sigma|^{-\frac{m}{2}} \exp -\tfrac{1}{2} \sum_{j=1}^{m} \begin{bmatrix} \alpha_j-\alpha_0 \\ \beta_j-\beta_0 \\ \mu_j-\mu_0 \end{bmatrix}^T \Sigma^{-1} \begin{bmatrix} \alpha_j-\alpha_0 \\ \beta_j-\beta_0 \\ \mu_j-\mu_0 \end{bmatrix}$$

$$\times \quad \exp -\tfrac{1}{2} \begin{bmatrix} \alpha_0-\eta_1 \\ \beta_0-\eta_2 \\ \mu_0-\eta_3 \end{bmatrix}^T \Phi^{-1} \begin{bmatrix} \alpha_0-\eta_1 \\ \beta_0-\eta_2 \\ \mu_0-\eta_3 \end{bmatrix} \tag{6.11}$$

$$\times \quad (\sigma^2{}_1 \ldots \sigma^2{}_m)^{-\frac{(\nu+2)}{2}} \exp -\frac{\nu\lambda}{2}\left(\frac{1}{\sigma^2{}_1} + \ldots + \frac{1}{\sigma^2{}_m}\right)$$

$$\times \quad |\Sigma|^{-\frac{(\rho-4)}{2}} \exp\{-\tfrac{1}{2}\mathrm{tr}(\Sigma^{-1}R)\} \quad .$$

The mode of this distribution occurs at the point given by 6.4 except that $\sigma^2{}_j$, $j=1,\ldots m$ and the elements of Σ^{-1}, instead of being constants are now given by

$$\sigma^2{}_j = (\nu+n_j+2)^{-1}\left\{\left[y_j - X_j\begin{pmatrix}\alpha_j \\ \beta_j\mu_j \\ \beta_j\end{pmatrix}\right]^T\left[y_j - X_j\begin{pmatrix}\alpha_j \\ \beta_j\mu_j \\ \beta_j\end{pmatrix}\right] + \nu\lambda\right\}; \; j=1,\ldots m$$

and

$$\Sigma = (m+\rho-4)^{-1}\left[R + \sum_{j=1}^{m}\begin{bmatrix}\alpha_j-\alpha_0 \\ \beta_j-\beta_0 \\ \mu_j-\mu_0\end{bmatrix}\begin{bmatrix}\alpha_j-\alpha_0 \\ \beta_j-\beta_0 \\ \mu_j-\mu_0\end{bmatrix}^T\right] \; .$$

Integrating over $\alpha_1 \ldots \alpha_m$, $\beta_1 \ldots \beta_m$ in 6.11 we obtain

$$p(\alpha_0, \beta_0, \mu_0, \Sigma^{-1}, \mu_1, \sigma^2{}_1, \ldots, \mu_m, \sigma^2{}_m \mid y_1, \ldots y_m, \eta_1, \eta_2, \eta_3, \Phi, \nu, \lambda, R, \rho) =$$

$$(\sigma^2_1)^{-\frac{(n_1+\nu+2)}{2}} \ldots (\sigma^2_m)^{-\frac{(n_m+\nu+2)}{2}} |\underset{\sim}{\Sigma}|^{-\frac{(m+\rho-4)}{2}} \left(\prod_{j=1}^{m} |\underset{\sim}{V}_j|^{\frac{1}{2}}\right)$$

$$\times \exp -\tfrac{1}{2}\left[tr(\underset{\sim}{\Sigma}^{-1}\underset{\sim}{R}) + \sum_{j=1}^{m}\left(\frac{\underset{\sim}{y}_j^T\underset{\sim}{y}_j+\nu\lambda}{\sigma^2_j}\right) + \begin{pmatrix}\alpha_o\\ \beta_o\\ \mu_o\end{pmatrix}^T (m\underset{\sim}{\Sigma}^{-1}+\underset{\sim}{\phi}^{-1})\begin{pmatrix}\alpha_o\\ \beta_o\\ \mu_o\end{pmatrix} - 2\begin{pmatrix}\eta_1\\ \eta_2\\ \eta_3\end{pmatrix}^T \underset{\sim}{\phi}^{-1}\begin{pmatrix}\alpha_o\\ \beta_o\\ \mu_o\end{pmatrix}\right.$$

$$+\sum_{j=1}^{m}\{\mu_j^2\Sigma^{33}-2\mu_j(\alpha_o\Sigma^{13}+\beta_o\Sigma^{23}+\mu_o\Sigma^{33})\}$$

$$\left. -\sum_{j=1}^{m}\left\{\begin{pmatrix}a_j\\ b_j\end{pmatrix}+\begin{bmatrix}\Sigma^{11}\Sigma^{12}\\ \Sigma^{12}\Sigma^{22}\\ \Sigma^{13}\Sigma^{23}\end{bmatrix}^T\begin{pmatrix}\alpha_o\\ \beta_o\\ \mu_o-\mu_j\end{pmatrix}\right\}^T \underset{\sim}{V}_j \left\{\begin{pmatrix}a_j\\ b_j\end{pmatrix}+\begin{bmatrix}\Sigma^{11}\Sigma^{12}\\ \Sigma^{12}\Sigma^{22}\\ \Sigma^{13}\Sigma^{23}\end{bmatrix}^T\begin{pmatrix}\alpha_o\\ \beta_o\\ \mu_o-\mu_j\end{pmatrix}\right\}\right] \quad (6.13)$$

and integrating over α_o and β_o in 6.13 we obtain

$$\pi(\mu_o,\underset{\sim}{\Sigma}^{-1},\mu_1,\sigma^2_1,\ldots,\mu_m,\sigma^2_m|\underset{\sim}{y}_1\ldots\underset{\sim}{y}_m,\eta_1,\eta_2,\eta_3,\underset{\sim}{\phi},\nu,\lambda,\underset{\sim}{R},\rho) \propto$$

$$(\sigma^2_1)^{-\frac{(n_1+\nu+2)}{2}} \ldots (\sigma^2_m)^{-\frac{(n_m+\nu+2)}{2}} |\underset{\sim}{\Sigma}|^{-\frac{(m+\rho-4)}{2}} |\underset{\sim}{\Psi}^{-1}+\underset{\sim}{S}^{-1}\sum_{j=1}^{m}\underset{\sim}{V}_j\underset{\sim}{Q}_j|^{-\frac{1}{2}}\left(\prod_{j=1}^{m}|\underset{\sim}{V}_j|^{\frac{1}{2}}\right)$$

$$\times \exp -\tfrac{1}{2}\left[tr(\underset{\sim}{\Sigma}^{-1}\underset{\sim}{R})+\nu\lambda\left(\frac{1}{\sigma^2_1}+\ldots+\frac{1}{\sigma^2_m}\right)+\sum_{j=1}^{m}\text{[illegible]}\begin{pmatrix}\phi^{13}\\ \phi^{23}\\ \phi^{33}\end{pmatrix}^T\begin{pmatrix}\eta_1\\ \eta_2\\ \eta_3\end{pmatrix}\right.$$

$$\left.+\text{[illegible]}\sum_{j=1}^{m}\text{[illegible]}\sum_{j=1}^{m}\left\{\begin{pmatrix}a_j\\ b_j\end{pmatrix}+\text{[illegible]}\right\}^T\underset{\sim}{V}_j\left\{\begin{pmatrix}a_j\\ b_j\end{pmatrix}+\text{[illegible]}\right\}\right] \quad (6.14)$$

where $\underset{\sim}{\Psi}$, $\underset{\sim}{S}$ and $\underset{\sim}{X}$ are as defined in section 6.2 .

If estimates of all the second stage parameters are required we suggest using the mode of 6.13. Alternatively if only μ_* is of interest we suggest using the mode of 6.14 . We do not feel altogether happy about these suggestions since there are so many nuisance parameters in both 6.13 and 6.14 In the types of situation where the present model is appropriate there may well be fairly large amounts of data available. In spite of this, unless an enormous number of assays are involved, the amount of information about the second stage parameters may not be very great; not enough to assume that either 6.13 or 6.14 approximates to a multivariate normal density. An improvement on the modal estimates would be to find an approximation to the marginal distribution of the parameters of interest. An attempt to do this might be made along the lines suggested in the last paragraph of section 4.3.

Suppose we have data from m similar assays, and suppose we have, by whatever method, obtained estimates of α_*, β_*, μ_* and Σ. We now wish to use these estimates in deciding on the parameters of a prior distribution for the analysis, using the model of chapter 4, of one further assay which we expect to be similar to our previous assays. We can use our estimates of α_o, β_o and μ_o directly as the second stage means, but we should not use Σ directly as the second stage variance. There are two reasons for this. Firstly we must remember that Σ plays a different role in the two models, and the appropriate prior variance of $\begin{pmatrix}\alpha\\ \beta\\ \mu\end{pmatrix}$ will be Σ (in the second model) plus the posterior variance of $\begin{pmatrix}\alpha_o\\ \beta_o\\ \mu_o\end{pmatrix}$; secondly we cannot be absolutely certain that the assay we are about to analyse is comparable with our previous assays. Experimental conditions may have changed in some way without our knowledge. In principle, one could cope with the first of these points theoretically by finding the approximate variance of the estimates of the means. However the distributions involved are very complicated and we suggest that the experimenter take the pragmatic approach of adding on to Σ a matrix, possibly a diagonal one, that represents a subjective view of the uncertainty from these two sources.

Finally, a word about the large sample theory for this model. Let m be a fixed integer greater than one, and suppose the number of responses available for each of m similar assays tends to infinity. The posterior distributions of the assay parameters will now depend entirely on the likelihood, given by the first stage of the model, and the form of the prior densities, given by the second and third stages of the model, are irrelevant. The likelihood of the m assays combined is the product of the likelihoods of the m individual assays. Consequently we cease to regard the assays as similar or dependent in any way and we treat them as m independent single assays. The large sample theory for single assays has already been given in sections 2.3 and 4.1.

6.4 An Example: Insulin Data

In Tables 6.1, 6.2, 6.3 we have data for 11 assays of A_1-B_{29} diacetyl insulin against standard insulin. The 11 test preparations of A_1-B_{29} diacetyl insulin are repeated dilutions of the same stock solution. It is unlikely that the stock solution changed appreciably during the period in which the dilutions were made, however, we expect there to be some variation in the strength of the test preparations due to inaccuracies in the dilution process.

Before analysing the data we have to choose values for the parameters of our prior distributions. We have put $v=0$ and $\phi^{-1}=0$. This should not cause any difficulties provided we allow Σ_{23} to be non-zero. It remains to choose values for ρ and R. Letting $R=0$ and $\rho=0$ would give the Jeffreys' ignorance prior distribution, but use of such a prior distribution causes the joint posterior density of all the parameters (6.11) to be infinite when $\alpha_j=\alpha_o$, $\beta_j=\beta_o$, $\mu_j=\mu_o$; $j=1,\ldots,m$, and [illegible]. To avoid this we have set $\rho=3$, the smallest value consistent with the convergence of the prior distribution of Σ^{-1}. In the prior distribution of Σ^{-1}, $E(\Sigma^{-1})=\rho R^{-1}$, so we can choose a value for R by making a guess at Σ and multiplying it by 3. Since we have very little idea of what Σ may be, we have taken as our guess its unbiased estimate obtained from the maximum likelihood estimates of the parameters. The maximum likelihood estimates have the same values as the large sample means and are given in Table 6.4. Using the resulting value of R we have calculated the mode of the joint posterior density of all the parameters, given by (6.11), and the mode of the posterior density of $(\alpha_o,\beta_o,\mu_o,\Sigma^{-1},\mu_1,\sigma^2_1,\ldots,\mu_m,\sigma^2_m)$ given by (6.13). In order to check the sensitivity of the procedure to our guess at Σ we have repeated the procedure with a guess ten times and one tenth our original one. The results are shown in Tables 6.5 - 6.7.

If we compare the two modes in Table 6.5 with the large sample means in Table 6.4, the α_i's, β_i's and the two sets of μ_i's and σ_i^2's in Table 6.5 are all pulled together compared with their large sample counterparts as one might expect. Comparing the two modes in Table 6.5 with each other, the estimates

	Dose (pmol l^{-1})		Responses		
	Insulin	A_1-B_{29} Diacetyl Insulin			
Assay 1	14.54		[illegible]	25.09	32.74
	21.80		[illegible]	33.73	46.42
	34.88		[illegible]	47.54	50.15
	58.14		[illegible]	57.38	59.32
		48.45	[illegible]	35.22	33.78
		78.68	[illegible]	39.27	37.11
		118.29	[illegible]	44.78	51.72
		[illegible]	[illegible]	55.27	57.18
2	14.54		[illegible]	33.41	39.01
	21.80		[illegible]	46.74	48.25
	[illegible]		[illegible]	58.54	57.13
	[illegible]		[illegible]	87.95	61.17
		48.85	[illegible]	27.49	35.19
		72.68	[illegible]	39.06	33.38
		183.81	[illegible]	58.24	58.78
3	14.54		[illegible]	8.84	12.57
	21.8C		[illegible]	13.09	14.71
	43.61		[illegible]	24.41	22.57
	87.21		[illegible]	26.52	32.39
	48.45	48.45	[illegible]	10.02	10.24
		72.68	[illegible]	15.46	11.38
		145.36	[illegible]	21.89	20.65
		290.72	[illegible]	28.85	28.85
4	17.44		[illegible]	17.73	12.78
	21.80		[illegible]	19.18	19.28
	29.07		[illegible]	38.36	34.54
		[illegible]	[illegible]	13.48	23.91
		[illegible]	[illegible]	37.42	33.92
5	29.07		[illegible]	11.40	9.69
	43.61		[illegible]	22.12	23.98
		[illegible]	[illegible]	12.95	12.75
		[illegible]	[illegible]	20.25	22.33

Table 8.1 Data from several assays of A_1-B_{29} diacetyl insulin against insulin.

Assay	Dose (pmol l^{-1}): Insulin	Dose (pmol l^{-1}): A1-B29 Diacetyl Insulin	Response		
6	21.80		24.26	25.82	24.25
	29.07		21.89	29.93	26.81
	43.61		32.13	34.67	35.43
		72.68	19.66	18.06	15.65
		96.90	24.02	22.14	21.04
		145.35	30.81	31.08	30.95
7	14.54		11.34	11.06	14.17
	43.61		19.11	21.48	21.14
	87.21		25.48	25.66	23.22
		48.45	9.25	12.31	10.56
		145.36	18.94	19.22	17.30
8	21.80		15.74	15.08	15.58
	29.07		23.07	27.13	28.63
	43.61		41.13	45.47	40.84
		96.90	24.49	24.26	29.38
		145.35	41.31	40.86	41.40
9	17.44		4.48	5.88	7.76
	34.89		18.90	17.23	19.57
		58.14	4.86	4.98	6.99
		87.21	7.95	13.46	13.04
		116.29	17.37	13.87	14.03
10	13.08		7.04	7.04	8.99
	21.80		13.02	16.69	14.90
	34.89		20.78	23.18	20.88
	58.14		29.18	30.23	28.55
		34.38	9.12	9.22	10.17
		57.30	15.13	11.63	12.81
		85.96	19.29	22.86	24.77
		128.93	24.81	29.08	23.74

Table 6.2 Data from several assays of A_1-B_{29} diacetyl insulin against insulin (continued)

	Dose (pmol l^{-1})		Response		
	Insulin	A_1-B_{29} Diacetyl Insulin			
Assay 6	21.80		24.26	25.82	24.25
	29.07		21.95	29.93	26.81
	43.61		32.13	34.67	35.43
		72.68	19.66	18.08	15.85
		96.90	24.02	22.14	21.04
		145.35	30.81	31.08	30.95
7	14.54		11.34	11.08	14.17
	43.61		19.11	21.48	21.14
	87.21		25.48	25.66	23.22
		48.45	9.25	12.31	10.56
		145.36	18.94	19.22	17.30
8	21.80		15.74	15.08	15.58
	29.07		23.07	27.13	26.63
	43.61		41.13	45.47	40.84
		96.90	24.49	24.26	29.38
		145.35	41.31	40.86	41.40
9	17.44		4.48	5.68	7.76
	34.89		18.90	17.23	19.57
		58.14	4.86	4.96	6.99
		87.21	7.95	13.46	13.04
		116.29	17.37	13.87	14.03
10	13.08		7.04	7.04	8.99
	21.80		13.02	16.66	14.90
	34.89		20.78	23.18	20.88
	58.14		29.18	30.23	26.55
		34.38	8.12	9.22	10.17
		57.30	15.13	11.63	12.81
		85.96	19.29	22.86	24.77
		128.93	24.81	29.08	23.74

Table 6.2 Data from several assays of A_1-B_{29} diacetyl insulin against insulin (continued)

	Dose (pmol l^{-1})		Response	
	Insulin	A_1-B_{29} Diacetyl Insulin		
Assay 11	13.08		19.84	14.10
	34.89		52.57	60.62
	58.14		73.15	72.22
		34.38	13.37	19.15
		57.30	29.17	37.87
		85.95	45.91	50.17
		128.83	64.35	59.59

Table 8.3 Data from several assays of A_1-B_{29} diacetyl insulin against insulin (continued).

	α	β	μ	σ^2
Assay 1	37.9	.200	-74.7	42.8
2	42.3	.219	-101.	45.4
3	15.5	.0951	-109.	16.0
4	5.69	.706	-54.9	38.0
5	8.42	.234	-83.5	9.50
6	22.0	.201	-96.5	5.76
7	13.8	.112	-89.5	5.37
8	13.0	.471	-77.4	37.0
9	6.84	.209	-68.6	11.3
10	10.9	.234	-47.4	14.1
11	27.5	.600	-55.9	69.3

Table 6.4 Mean of approximate large sample distribution using insulin assay data.

Mode of $\pi(\alpha_o,\beta_o,\mu_o,\underset{\sim}{\Sigma}^{-1},\alpha_1,\beta_1,\mu_1,\sigma^2{}_1\ldots\alpha_m,\beta_m,\mu_m\ \sigma_m^2|\underset{\sim}{y}_1\ldots\underset{\sim}{y}_m,\underset{\sim}{\Phi},\nu,\underset{\sim}{R},\rho)$

	α	β	μ	σ^2
Assay 1	37.9	.205	-77.3	35.8
2	41.0	.227	-92.4	57.6
3	14.8	.0947	-92.1	21.7
4	8.98	.572	-54.0	28.6
5	6.55	.226	-81.7	11.3
6	21.6	.206	-93.4	9.89
7	13.5	.114	-85.7	6.17
8	14.4	.423	-74.9	39.4
9	7.00	.209	-69.7	8.59
10	11.2	.235	-50.0	9.85
11	28.3	.578	-57.6	78.0

$$\begin{pmatrix}\alpha_o\\ \beta_o\\ \mu_o\end{pmatrix} = \begin{pmatrix}18.8\\ .281\\ -75.3\end{pmatrix} \qquad \underset{\sim}{\Sigma} = \begin{bmatrix}188 & -.194 & -81.5\\ -.194 & .0490 & 1.73\\ -81.5 & 1.73 & 369.\end{bmatrix}$$

Mode of $\pi(\alpha_o,\beta_o,\mu_o,\underset{\sim}{\Sigma}^{-1},\mu_1,\sigma^2{}_1\ldots\mu_m,\sigma^2{}_m|\underset{\sim}{y}_1\ldots\underset{\sim}{y}_m,\Phi,\nu,\underset{\sim}{R},\rho)$

	μ	σ^2
Assay 1	-77.5	42.9
2	-93.2	46.0
3	-94.6	16.4
4	-54.7	38.7
5	-82.3	9.47
6	-94.3	5.80
7	-85.7	5.40
8	-75.6	37.0
9	-69.9	11.3
10	-50.9	14.3
11	-57.9	98.9

$$\begin{pmatrix}\alpha_o\\ \beta_o\\ \mu_o\end{pmatrix} = \begin{pmatrix}18.9\\ .280\\ -76.1\end{pmatrix}$$

$$\underset{\sim}{\Sigma} = \begin{bmatrix}192. & -.278 & -79.5\\ -.278 & .0517 & 1.72\\ -79.5 & 1.72 & 376.\end{bmatrix}$$

Table 6.5 Modes of joint posterior densities using assay data with prior parameters $\nu=0,\ \underset{\sim}{\Phi}^{-1}=\underset{\sim}{0},\ \rho=3,\ \underset{\sim}{R}=\begin{bmatrix}460 & -1.8 & -230.\\ -1.8 & .21 & -.23\\ -230. & -.23 & 1200.\end{bmatrix}$

of α_o, β_o, μ_o and the μ_i's are almost the same although the σ_i^2's are less dispersed in the first mode compared with the second mode. The estimates of Σ are very similar with the exception of Σ_{12} where there is a considerable difference. Our initial guess at Σ was

$$\Sigma = \begin{bmatrix} 150. & -.61 & -78. \\ -.61 & [illegible] & -.078 \\ -78. & -.078 & 330. \end{bmatrix}$$

The diagonal elements and Σ_{13} are similar to our estimates but Σ_{12} differs from either of the estimates. Σ_{23} also differs from our estimates although the two estimates are very similar in this case.

Comparing the two modes in Table 6.6 with their counterpart in Table 6.5, the estimates of α_o, β_o, μ_o, the α_i's, the β_i's and the μ_i's have scarcely changed. There have been small changes in the estimates of the σ_i^2's and substantial ones in the estimate of Σ. The discrepancy seems to be greater for the larger R than the smaller R. Very similar remarks apply when comparing Table 6.7 with its counterpart in Table 6.5 except here the σ_i^2's do not seem to be so sensitive to changes in R.

a) $R = \begin{bmatrix} 4600. & [illegible] & -2300. \\ -18. & [illegible] & -2.3 \\ -2300. & -2.3 & 12000. \end{bmatrix}$

	α	β	μ	σ^2
Assay 1	37.9	.201	-75.5	38.5
2	41.8	.222	-88.4	45.5
3	15.2	.0952	-103.	17.3
4	7.11	.651	-54.7	28.8
5	8.47	.232	-83.2	8.60
6	22.0	.202	[illegible]	5.89
7	13.7	.113	-88.2	5.20
8	13.3	.482	-77.2	32.4
9	8.89	.209	-88.9	8.60
10	11.0	.234	-48.1	12.2
11	27.7	.585	[illegible]	82.4

$$\begin{pmatrix} \alpha_0 \\ \beta_0 \\ \mu_0 \end{pmatrix} = \begin{pmatrix} 18.7 \\ .292 \\ -77.[illegible] \end{pmatrix} \qquad \Sigma = \begin{bmatrix} 611. & -2.04 & -301. \\ -2.04 & .246 & 2.01 \\ -301. & 2.01 & 1547. \end{bmatrix}$$

b) $R = \begin{bmatrix} 48. & -.18 & -23. \\ -.18 & .021 & -.023 \\ -23. & -.023 & 120. \end{bmatrix}$

	α	β	μ	σ^2
Assay 1	37.9	[illegible]	-80.1	31.4
2	39.9	.233	-84.3	74.3
3	14.3	.0927	-79.0	26.0
4	12.4	.432	52.3	30.3
5	8.54	.178	-70.8	21.8
6	20.5	.193	-77.9	33.9
7	13.0	.118	-76.4	10.1
8	17.3	.317	-65.8	55.2
9	8.87	.201	-88.7	12.5
10	11.3	.238	-52.8	8.23
11	30.2	.513	-58.4	92.8

Table 6.8 Mode of $\pi(\alpha_0, \beta_0, \mu_0, \Sigma^{-1}, \alpha_1, \beta_1, \mu_1, \sigma_1^2, \ldots, \alpha_m, \beta_m, \mu_m, \sigma_m^2 \mid y_1, \ldots, y_m, \delta, \nu, R, \rho)$ for insulin assay data with prior parameters $\delta^{-1} = 0$, $\rho = 3$ and R as indicated.

a) $\underset{\sim}{R} = \begin{bmatrix} 4500. & -18. & -2300. \\ -18. & 2.1 & -2.3 \\ -2300. & -2.3 & 12000. \end{bmatrix}$

	μ	σ^2
Assay 1	-75.5	42.8
2	-97.9	45.5
3	-103.	16.1
4	-55.2	37.9
5	-83.2	9.49
6	-95.7	5.77
7	-87.5	5.38
8	-77.6	36.9
9	-88.6	11.3
10	-48.2	14.1
11	-56.4	99.1

$\begin{pmatrix} \alpha_o \\ \beta_o \\ \mu_o \end{pmatrix} = \begin{pmatrix} 18.6 \\ .294 \\ -77.2 \end{pmatrix}$ $\quad \underset{\sim}{\Sigma} = \begin{bmatrix} 617. & -2.20 & -300. \\ -2.20 & .251 & 2.00 \\ -300. & 2.00 & 1540. \end{bmatrix}$

b) $\underset{\sim}{R} = \begin{bmatrix} 48. & -.18 & -23. \\ -.18 & .021 & -.023 \\ -23. & -.023 & 120. \end{bmatrix}$

	μ	σ^2
Assay 1	-80.9	43.3
2	-90.4	46.4
3	-90.7	16.7
4	-54.4	40.3
5	-80.3	9.53
6	-92.3	5.91
7	-85.2	5.41
8	-71.6	36.7
9	-70.9	11.2
10	-55.2	15.0
11	-56.6	99.5

$\begin{pmatrix} \alpha_o \\ \beta_o \\ \mu_o \end{pmatrix} = \begin{pmatrix} 19.4 \\ .264 \\ -75.5 \end{pmatrix}$

$\underset{\sim}{\Sigma} = \begin{bmatrix} 144. & .111 & -51.8 \\ .111 & .0238 & 1.47 \\ -51.8 & 1.47 & 218. \end{bmatrix}$

Table 6.7 Mode of $\pi(\alpha_o, \beta_o, \mu_o, \Sigma^{-1}, \mu_1, \sigma^2_1, \ldots, \mu_m, \sigma^2_m | \underset{\sim}{y}_1, \ldots, \underset{\sim}{y}_m, \phi, \nu, \underset{\sim}{R}, \rho)$ for insulin assay data with prior parameters $\nu = 0, \underset{\sim}{\phi}^{-1} = \underset{\sim}{0}, \rho = 3$ and $\underset{\sim}{R}$ as indicated.

Chapter 7 A More Specialized Model Combining Information from Several Very Similar Assays.

7.1 Introduction

Suppose that one wishes to assay a particular preparation, and that using the relevant assay method and apparatus one is limited to a certain size of assay. If the amount of information that can be gained from one such assay is not sufficient, then several assays will be carried out and the information from them all will need to be combined. Replicate assays of this type will be very similar to one another in several respects. Firstly the true potency ratio will be the same throughout, although biological variation will cause the pairs of log dose-response curves to vary in other respects. Secondly the assays will be carried out in the same laboratory and probably also by the same person using the same apparatus. As a result of this we conjecture that a suitable model for the analysis of such replicate assays stipulates that the log potency ratio remains unchanged throughout. Another, more minor, stipulation is that the residual variance for all the assays is the same. These two assumptions give the following model:

1st stage: $\underset{\sim}{y}_j \sim N\left[\underset{\sim}{X}_j\begin{pmatrix}\alpha_j\\ \beta_j\mu\\ \beta_j\end{pmatrix}, \sigma^2\underset{\sim}{I}_{nj}\right]$; independently for $j=1\ldots m$,

2nd stage: $\begin{pmatrix}\alpha_j\\ \beta_j\end{pmatrix} \sim N\left[\begin{pmatrix}\alpha_o\\ \beta_o\end{pmatrix}, \begin{pmatrix}\Sigma_{11} & \Sigma_{12}\\ \Sigma_{12} & \Sigma_{22}\end{pmatrix}\right]$; independently for $j=1\ldots m$,

(7.1)

$\mu \sim N(\mu_o, \Sigma_{33})$; independent of the distributions of $\begin{pmatrix}\alpha_j\\ \beta_j\end{pmatrix}$, $j=1\ldots m$.

3rd stage: $\begin{pmatrix} \alpha_o \\ \beta_o \end{pmatrix} \sim N\left[\begin{pmatrix} \eta_1 \\ \eta_2 \end{pmatrix}, \underset{\sim}{\Phi}\right].$

Again, we assume for the moment that all variances and covariances are known. This model can in a sense be derived from the model described in section 6.1 by setting $\sigma^2_j = \sigma^2$; $j=1...m$, and by setting the (3,3) element of the covariance matrix in the second stage of equation 6.1 to zero. The prior information about μ in the second stage of equation 7.1 is comparable with the prior information about μ_o in the third stage of equation 6.1.

In addition to the analysis of replicate assays this model may be the correct one for certain collaborative assay where variation in personal assay technique is thought to be unimportant. Also, as will be apparent in the following sections, this model is considerably more tractable than the model described in chapter 6, and so it may be a useful approximate model even in cases where the assumptions do not hold precisely.

7.2 Posterior Distributions for Known Covariance Structure

Combining the likelihood with the second and third stage prior densities, the joint posterior density of the first and second stage parameters is

$$\pi(\alpha_o,\beta_o,\mu,\alpha_1,\beta_1,\ldots,\alpha_m,\beta_m|\underset{\sim}{y}_1,\ldots y_m,\mu_o,\eta_1,\eta_2)$$

$$\exp-\tfrac{1}{2}\left[\sum_{j=1}^{m}\left[\frac{1}{\sigma^2}\begin{pmatrix}\alpha_j\\ \beta_j\mu\\ \beta_j\end{pmatrix}^T \underset{\sim}{X}_j^T \underset{\sim}{X}_j\begin{pmatrix}\alpha_j\\ \beta_j\mu\\ \beta_j\end{pmatrix}+\begin{pmatrix}\alpha_j\\ \beta_j\end{pmatrix}^T\underset{\sim}{\Sigma}^{-1}\begin{pmatrix}\alpha_j\\ \beta_j\end{pmatrix}-\frac{2}{\sigma^2}\underset{\sim}{y}_j^T\underset{\sim}{X}_j\begin{pmatrix}\alpha_j\\ \beta_j\mu\\ \mu_j\end{pmatrix}-2\begin{pmatrix}\alpha_o\\ \beta_o\end{pmatrix}^T\underset{\sim}{\Sigma}^{-1}\begin{pmatrix}\alpha_j\\ \beta_j\end{pmatrix}\right\}\right.$$

$$\left.+\frac{(\mu-\mu_o)^2}{\Sigma_{33}}+\begin{pmatrix}\alpha_o\\ \beta_o\end{pmatrix}^T(m\underset{\sim}{\Sigma}^{-1}+\underset{\sim}{\phi}^{-1})\begin{pmatrix}\alpha_o\\ \beta_o\end{pmatrix}-2\begin{pmatrix}\alpha_o\\ \beta_o\end{pmatrix}^T\underset{\sim}{\phi}^{-1}\begin{pmatrix}\eta_1\\ \eta_2\end{pmatrix}\right]. \qquad (7.2)$$

where we now let $\underset{\sim}{\Sigma}=\begin{pmatrix}\Sigma_{11}\Sigma_{12}\\ \Sigma_{12}\Sigma_{22}\end{pmatrix}$. This is a change in notation from the preceding chapters. Integrating over the second stage means α_o and β_o, the joint distribution of the remaining parameters is

α /

$$\pi(\mu,\alpha_1,\beta_1,\ldots\cancel{\mu}_m,\beta_m|\underset{\sim}{y}_1,\ldots\underset{\sim}{y}_m,\mu_o,\eta_1,\eta_2)\propto$$

$$\exp-\tfrac{1}{2}\left[\sum_{j=1}^{m}\left\{\frac{1}{\sigma^2}\begin{pmatrix}\alpha_j\\ \beta_j\mu\\ \beta_j\end{pmatrix}^T\underset{\sim}{X}_j^T\underset{\sim}{X}_j\begin{pmatrix}\alpha_j\\ \beta_j\mu\\ \beta_j\end{pmatrix}+\begin{pmatrix}\alpha_j\\ \beta_j\end{pmatrix}^T\underset{\sim}{\Sigma}^{-1}\begin{pmatrix}\alpha_j\\ \beta_j\end{pmatrix}-\frac{2}{\sigma^2}\underset{\sim}{y}_j^T\underset{\sim}{X}_j\begin{pmatrix}\alpha_j\\ \beta_j\mu\\ \beta_j\end{pmatrix}\right\}+\frac{(\mu-\mu_o)^2}{\Sigma_{33}}\right.$$

$$\left.-\left\{\sum_{j=1}^{m}\begin{pmatrix}\alpha_j\\ \beta_j\end{pmatrix}^T\underset{\sim}{\Sigma}^{-1}+\begin{pmatrix}\eta_1\\ \eta_2\end{pmatrix}^T\underset{\sim}{\phi}^{-1}\right\}(m\underset{\sim}{\Sigma}^{-1}+\phi^{-1})^{-1}\left\{\underset{\sim}{\Sigma}^{-1}\sum_{j=1}^{m}\begin{pmatrix}\alpha_j\\ \beta_j\end{pmatrix}+\underset{\sim}{\phi}^{-1}\begin{pmatrix}\eta_1\\ \eta_2\end{pmatrix}\right\}\right] \qquad (7.3)$$

The mode of this density occurs at the point

$$\alpha_j = \frac{\frac{1}{\sigma^2}\sum_{k=1}^{n_j}(y_{kj}-\beta_j x_{kj}-\mu z_{kj})-\Sigma^{12}\beta_j+\begin{pmatrix}1\\0\end{pmatrix}^T\Sigma^{-1}(m\Sigma^{-1}+\Phi^{-1})^{-1}\left\{\Sigma^{-1}\sum_{j=1}^{m}\begin{pmatrix}\alpha_j\\ \beta_j\end{pmatrix}+\Phi^{-1}\begin{pmatrix}m_1\\ m_2\end{pmatrix}\right\}}{\frac{n_j}{\sigma^2}+\Sigma^{11}}, \quad j=1,\ldots,m;$$

$$\beta_j = \frac{\frac{1}{\sigma^2}\sum_{k=1}^{n_j}(x_{kj}+\mu z_{kj})(y_{kj}-\alpha_j)-\Sigma^{12}\alpha_j+\begin{pmatrix}0\\1\end{pmatrix}^T\Sigma^{-1}(m\Sigma^{-1}+\Phi^{-1})^{-1}\left\{\Sigma^{-1}\sum_{j=1}^{m}\begin{pmatrix}\alpha_j\\ \beta_j\end{pmatrix}+\Phi^{-1}\begin{pmatrix}m_1\\ m_2\end{pmatrix}\right\}}{\frac{1}{\sigma^2}\sum_{k=1}^{n_j}(x_{kj}+\mu z_{kj})^2+\Sigma^{22}}, \quad j=1,\ldots,m;$$

$$\mu = \frac{\frac{1}{\sigma^2}\left[\sum_{j=1}^{m}\beta_j\sum_{k=1}^{n_j}(y_{kj}-\alpha_j-\beta_j x_{kj})z_{kj}\right]+\frac{\mu_0}{\Sigma_{33}}}{\frac{1}{\sigma^2}\sum_{j=1}^{m}\beta_j\sum_{k=1}^{n_j}z_{kj}^2+\frac{1}{\Sigma_{33}}} \qquad (7.4)$$

Suppose we have very little prior knowledge of the location of either α_0, β_0 or μ. We will then have $\Phi^{-1}=0$ and $\frac{1}{\Sigma_{33}}=0$. It will be shown at the end of this section that such improper prior distributions do not cause the posterior distributions to be unnormed. In this case the mode of the joint posterior distribution of $\mu, \alpha_1, \beta_1, \ldots \alpha_m, \beta_m$ occurs at the point

$$\alpha_j = \frac{\frac{1}{\sigma^2}\sum_{k=1}^{n_j}(y_{kj}-\beta_j x_{kj}-\mu z_{kj})+\Sigma^{11}\left(\bar{\alpha}-\frac{\Sigma^{12}}{\Sigma^{11}}(\beta_j-\bar{\beta})\right)}{\frac{n_j}{\sigma^2}+\Sigma^{11}}, \quad j=1,\ldots,m;$$

$$\beta_j = \frac{\frac{1}{\sigma^2}\sum_{k=1}^{n_j}(x_{kj}+\mu z_{kj})(y_{kj}-\alpha_j)+\Sigma^{22}\left(\bar{\beta}-\frac{\Sigma^{12}}{\Sigma^{22}}(\alpha_j-\bar{\alpha})\right)}{\frac{1}{\sigma^2}\sum_{k=1}^{n_j}(x_{kj}+\mu z_{kj})^2+\Sigma^{22}}; \quad j=1,\ldots,m, \qquad (7.5)$$

$$\frac{\mu - \sum_{j=1}^{m} \beta_j \sum_{k=1}^{n_j} (y_{kj} - \alpha_j - \beta_j x_{kj}) z_{kj}}{\sum_{j=1}^{m} \beta_j^2 \sum_{k=1}^{n_j} z_{kj}^2}$$

This model is rather more tractable than the model described in the previous chapter in that we can now integrate over $\alpha_1, \beta_1, \ldots \alpha_m, \beta_m$ in 7.3 and obtain the marginal posterior distribution of μ:

$$\pi(\mu | \underset{\sim}{y}_1, \ldots \underset{\sim}{y}_m, \mu_0, n_1, n_2) = \left\{ \prod_{j=1}^{m} |\underset{\sim}{D}_j + \underset{\sim}{\Sigma}^{-1}|^{-\frac{1}{2}} \right\} \left| m\underset{\sim}{\Sigma}^{-1} + \underset{\sim}{\phi}^{-1} - \sum_{j=1}^{m} \underset{\sim}{\Sigma}^{-1} (\underset{\sim}{D}_j + \underset{\sim}{\Sigma}^{-1})^{-1} \underset{\sim}{\Sigma}^{-1} \right|^{-\frac{1}{2}}$$

$$\times \exp -\tfrac{1}{2} \left[\frac{(\mu - \mu_0)^2}{\Sigma_{33}} - \sum_{j=1}^{m} \begin{pmatrix} a_j \\ b_j \end{pmatrix}^T (\underset{\sim}{D}_j + \underset{\sim}{\Sigma}^{-1})^{-1} \begin{pmatrix} a_j \\ b_j \end{pmatrix} \right.$$

$$- \left\{ \sum_{j=1}^{m} \underset{\sim}{\Sigma}^{-1} (\underset{\sim}{D}_j + \underset{\sim}{\Sigma}^{-1})^{-1} \begin{pmatrix} a_j \\ b_j \end{pmatrix} + \underset{\sim}{\phi}^{-1} \begin{pmatrix} n_1 \\ n_2 \end{pmatrix} \right\}^T \left[m\underset{\sim}{\Sigma}^{-1} + \underset{\sim}{\phi}^{-1} - \sum_{j=1}^{m} \underset{\sim}{\Sigma}^{-1} (\underset{\sim}{D}_j + \underset{\sim}{\Sigma}^{-1})^{-1} \underset{\sim}{\Sigma}^{-1} \right]^{-1}$$

$$\left. \times \left\{ \sum_{j=1}^{m} \underset{\sim}{\Sigma}^{-1} (\underset{\sim}{D}_j + \underset{\sim}{\Sigma}^{-1})^{-1} \begin{pmatrix} a_j \\ b_j \end{pmatrix} + \underset{\sim}{\phi}^{-1} \begin{pmatrix} n_1 \\ n_2 \end{pmatrix} \right\} \right] \qquad (7.6)$$

where $a_j = \dfrac{n_j \bar{x}_j}{\sigma^2}$,

$$b_j = \frac{1}{\sigma^2} \sum_{k=1}^{n_j} (x_{kj} + z_{kj}) y_{kj} \quad .$$

and $\underline{D}_j = \frac{1}{\sigma^2}\begin{bmatrix} n_j & n_j(\bar{x}_{\cdot j}+\mu\bar{z}_{\cdot j}) \\ n_j(\bar{x}_{\cdot j}+\mu\bar{z}_{\cdot j}) & \sum_{k=1}^{n_j}(x_{kj}+z_{kj})^2 \end{bmatrix}$. This notation is

slightly different from that of chapter 8. In the case $\underline{\Phi}^{-1}=\underline{0}$, $\frac{1}{\Sigma_{33}}=0$ the marginal distribuiton of μ simplifies to

$$\pi(\mu|\underline{y}_1,\dots,\underline{y}_m) \propto \left|\sum_{j=1}^{m}(\underline{D}_j+\underline{\Sigma}^{-1})^{-1}\right|^{-\frac{1}{2}}\left|\sum_{j=1}^{m}\underline{\Sigma}^{-1}(\underline{D}_j+\underline{\Sigma}^{-1})^{-1}\underline{D}_j\right|^{-\frac{1}{2}}$$

$$\times \exp\frac{1}{2}\left[\sum_{j=1}^{m}\begin{pmatrix}a_j\\b_j\end{pmatrix}^T(\underline{D}_j+\underline{\Sigma}^{-1})^{-1}\begin{pmatrix}a_j\\b_j\end{pmatrix}+\left\{\sum_{j=1}^{m}\underline{\Sigma}^{-1}(\underline{D}_j+\underline{\Sigma}^{-1})^{-1}\begin{pmatrix}a_j\\b_j\end{pmatrix}\right\}^T\right.$$

$$\left.\times\left\{\sum_{j=1}^{m}\underline{\Sigma}^{-1}(\underline{D}_j+\underline{\Sigma}^{-1})^{-1}\underline{D}_j\right\}^{-1}\left\{\sum_{j=1}^{m}\underline{\Sigma}^{-1}(\underline{D}_j+\underline{\Sigma}^{-1})^{-1}\begin{pmatrix}a_j\\b_j\end{pmatrix}\right\}\right]. \qquad (7.7)$$

In order to see if this density is normed we need to examine the expression on the right hand side of the = sign in 7.7. If the integral of this expression with respect to μ is finite then we can safely put $\underline{\Phi}^{-1}=\underline{0}$ and $\frac{1}{\Sigma_{33}}=0$, and we can easily show that this is so. If we make the same assumptions about the assays as in section 8.2, examination of the terms inside the exponent shows them both to be bounded above and below for all μ and to tend to finite limits as μ becomes large in absolute value. The same applies to the term

$\left|\sum_{j=1}^{m}\underline{\Sigma}^{-1}(\underline{D}_j+\underline{\Sigma}^{-1})^{-1}\underline{D}_j\right|^{-\frac{1}{2}}$. That the integral is finite now follows

from the fact that, providud m is at least 2,

$\int_{-\infty}^{\infty} \prod_{j=1}^{m} \left| \underset{\sim}{D}_j + \underset{\sim}{\Sigma}^{-1} \right|^{-\frac{1}{2}} d\mu$ is finite.

7.3 Large Sample Distributions

Using the theory described in section 2.3 we can show that the distribution of $\mu, \alpha_1, \beta_1, \ldots \alpha_m, \beta_m$ as the number of responses in each assay becomes very large is asymptotically

$$f(\mu, \alpha_1, \beta_1, \ldots \alpha_m, \beta_m \mid y \ldots y_m) \sim N\left[\begin{pmatrix} \hat{\mu} \\ \hat{\alpha}_1 \\ \hat{\beta}_1 \\ \vdots \\ \hat{\alpha}_m \\ \hat{\beta}_m \end{pmatrix}, \tilde{V}\right]$$

$$\text{where } \hat{\mu} = \frac{\sum_{j=1}^{m} \hat{\beta}_j \sum_{k=1}^{n_j} (y_{kj} - \hat{\alpha}_j - \hat{\beta}_j x_{kj}) z_{kj}}{\sum_{j=1}^{m} \hat{\beta}_j^2 \sum_{k=1}^{n_j} z_{kj}^2},$$

$$\hat{\alpha}_j = \bar{y}_{.j} - \hat{\beta}_j \hat{\mu} \bar{z}_{.j} - \hat{\beta}_j \bar{x}_{.j}, \quad j = 1, \ldots, m, \tag{7.8}$$

$$\hat{\beta}_j = \frac{\sum_{k=1}^{n_j} (x_{kj} + \hat{\mu} z_{kj})(y_{kj} - \hat{\alpha}_j)}{\sum_{k=1}^{n_j} (x_{kj} + \hat{\mu} z_{kj})^2}, \quad j = 1, \ldots, m,$$

and

$$\tilde{V} = \sigma^2 \left[\begin{array}{c|cccc} A & -A\tilde{C}_1^T & \cdots & \cdots & -A\tilde{C}_m^T \\ \hline -A\tilde{C}_1 & (\tilde{B}^{-1} + A\tilde{C}_1\tilde{C}_1^T) & A\tilde{C}_1\tilde{C}_2^T & \cdots & -A\tilde{C}_1\tilde{C}_m^T \\ \vdots & A\tilde{C}_2\tilde{C}_1^T & \ddots & & \vdots \\ \vdots & \vdots & & \ddots & A\tilde{C}_{m-1}\tilde{C}_m^T \\ -A\tilde{C}_m & A\tilde{C}_m\tilde{C}_1^T & \cdots & A\tilde{C}_m\tilde{C}_{m-1}^T & (\tilde{B}_m^{-1} + A\tilde{C}_m\tilde{C}_m^T) \end{array}\right]$$

where $\hat{\mu} = \left\{ \sum_{j=1}^{m} \left[\hat{\beta}_j^2 Szz^j - \left(\frac{(Syz^j - 2\hat{\beta}_j Sxz^j - 2\hat{\beta}_j \hat{\mu} Szz^j)^2}{Sxx^j + 2\hat{\mu} Sxz + \hat{\mu}^2 Szz^j} \right) \right] \right\}^{-1}$,

$$\hat{\alpha}_j = \begin{bmatrix} n_j & n_j(\bar{x}_{.j} + \hat{\mu}\bar{z}_{.j}) \\ n_j(\bar{x}_{.j} + \hat{\mu}\bar{z}_{.j}) & \sum_{i=1}^{n_j} (x_{ij} + \hat{\mu} z_{ij})^2 \end{bmatrix} .$$

and

$$\hat{\beta}_j = (Sxx^j + 2\hat{\mu} Sxz^j + \hat{\mu}^2 Szz^j)^{-1} \begin{bmatrix} (\bar{x}_{.j} + \hat{\mu}\bar{z}_{.j})(Syz^j - \beta_j \hat{\mu} Szz^j) \\ \qquad + \hat{\beta}_j(\bar{z}_{.j} Sxx^j - 2\bar{x}_{.j} Sxz^j - \bar{z}_{.j}\hat{\mu} Szz^j) \\ 2\hat{\beta}_j \hat{\mu} Szz^j + 2\hat{\beta}_j Sxz^j - Syz^j \end{bmatrix} .$$

If we now turn back to the mode of the joint posterior distribution of $\mu, \alpha_1, \beta_1, \ldots \alpha_m, \beta_m$ we can see that in the case where $\Phi^{-1} = 0$ and $\frac{1}{\Sigma_{33}} = 0$, given by 7.5, the mode occurs at a point where the $\hat{\alpha}_j$ are weighted averages of the large sample means and the overall average of the α_j, adjusted for dependence on the β_j. The weights depend on the size of the assays, the residual variance and the second stage covariance matrix Σ. Weighted averages of this type occur frequently in expressions for posterior means using linear models, see for example Lindley (1971 b). Parallel remarks apply to the value of the β_j at this mode. The expression for μ at the mode has a similar form to the large sample mean, however after substitution for α_j, β_j in the one case and $\hat{\alpha}_j, \hat{\beta}_j$ in the other, the two values will not be identical.

The equations for the mode of the joint density of $\mu, \alpha_1, \beta_1, \ldots \alpha_m, \beta_m$ when the elements of Φ^{-1} and $\frac{1}{\Sigma_{33}}$ are not all zero, given by 7.4, are more complicated weighted averages involving the prior knowledge about the location of the parameters.

We can eliminate $\hat{\alpha}_j$, $j=1,\dots m$ from the expressions for $\hat{\mu}$ and $\hat{\beta}_j$, $j=1,\dots m$ in 7.8. This gives the following expressions for the large sample posterior means for $\hat{\mu}$ and $\hat{\beta}_j$, $j=1,\dots m$:

$$\hat{\mu}=\frac{\sum_{j=1}^{m} \hat{\beta}_j (S^j_{yz}-\hat{\beta}_j S^j_{xz})}{\sum_{j=1}^{m} \hat{\beta}_j^2 S^j_{zz}} \quad , \qquad (7.9)$$

$$\hat{\beta}_j=\frac{\hat{\mu} S^j_{yz}+S^j_{xy}}{\hat{\mu}^2 S^j_{zz}+2\hat{\mu} S^j_{xz}+S^j_{xx}} \; ; \; j=1,\dots m .$$

A sampling theory approach to the situation under consideration has been investigated by Armitage et al (1976). It is interesting to note that although their model has been set up very differently from ours, they obtain maximum likelihood estimates of the log potency ratio and the slopes of the individual assay identical to those in 7.9. The asymptotic sampling variance of their maximum likelihood estimate of log potency ratio is

$$\sigma^2\left\{\sum_{j=1}^{m}\left[\beta_j^2 S^j_{zz} - \frac{\beta_j^2 (S^j_{xz}+\mu S^j_{zz})^2}{S^j_{xx}+2\mu S^j_{xz}+\mu^2 S^j_{zz}}\right]\right\}^{-1} .$$

7.4 A Pathological Example

We have had very little success in trying to examine the form of the posterior distribution of μ analytically; the algebra is too complicated. We have concentrated instead on two special cases; in section 7.6 we attempt to combine genuine data from several assays which are in good agreement with one another, and in this section we examine highly artificial data from two assays which disagree violently with one another.

Suppose we carry out two four-point assays, in both of which log-doses of $+\frac{1}{2}$ and $-\frac{1}{2}$ are administered for both test and standard preparations. Suppose that in the first assay each point is replicated just once, and in the second assay each point is replicated a times, the same response occuring for each dose throughout the replications. The responses are as given in Table 7.1. We assume d to be non-negative, ε to be small, and the residual variance to be the same for both assays and equal to σ^2. The sufficient statistics from these two assays are:

$\bar{x}_{\cdot 1}=0$,	$\bar{x}_{\cdot 2}=0$,
$\bar{y}_{\cdot 1}=0$,	$\bar{y}_{\cdot 2}=0$,
$\bar{z}_{\cdot 1}=\frac{1}{2}$,	$\bar{z}_{\cdot 2}=\frac{1}{2}$,
$S^1_{xx}=1$,	$S^2_{xx}=a$,
$S^1_{xy}=1$,	$S^2_{xy}=a$,
$S^1_{xz}=0$,	$S^2_{xz}=0$,
$S^1_{yz}=-d$,	$S^2_{yz}=ad$,
$S^1_{zz}=1$,	$S^2_{zz}=a$.

These assays are intended to provide completely contradictory information about μ, with the second assay containing a times as much information as the first. In addition to values of a greater than 1 we shall also consider values of a lying between 0 & 1. This corresponds to the first assay being replicated and not the second.

Looking at the first assay by itself we have the following large sample results:

	y	x	z
Assay 1	$\frac{d}{2} - \frac{1}{2} + \varepsilon$	$\frac{-1}{2}$	0
	$\frac{d}{2} + \frac{1}{2} - \varepsilon$	$\frac{+1}{2}$	0
	$\frac{-d}{2} - \frac{1}{2} - \varepsilon$	$\frac{-1}{2}$	1
	$\frac{-d}{2} + \frac{1}{2} + \varepsilon$	$\frac{+1}{2}$	1
Assay 2	$\frac{-d}{2} - \frac{1}{2} + \varepsilon$	$\frac{-1}{2}$	0
	$\frac{-d}{2} + \frac{1}{2} - \varepsilon$	$\frac{+1}{2}$	0
	$\frac{d}{2} - \frac{1}{2} - \varepsilon$	$\frac{-1}{2}$	1
	$\frac{d}{2} + \frac{1}{2} + \varepsilon$	$\frac{+1}{2}$	1

(Each dose and response in assay 2 is replicated a times)

Table 7.1 Results of two hypothetical assays.

$$\begin{bmatrix}\hat{\alpha}\\ \hat{\beta}\\ \hat{\gamma}\end{bmatrix} \sim N\left(\begin{bmatrix}\frac{a}{2}\\ 1\\ -d\end{bmatrix}, \sigma^2\begin{bmatrix}\frac{1}{4} & 0 & \frac{1}{2}\\ 0 & 1 & d\\ \frac{1}{2} & d & (1+d^2)\end{bmatrix}\right) ,$$

and similarly, looking at the second assay by itself:

$$\begin{bmatrix}\hat{\alpha}\\ \hat{\beta}\\ \hat{\gamma}\end{bmatrix} \sim N\left(\begin{bmatrix}\frac{d}{2}\\ 1\\ d\end{bmatrix}, \frac{\sigma^2}{a}\begin{bmatrix}\frac{1}{4} & 0 & -\frac{1}{2}\\ 0 & 1 & -d\\ -\frac{1}{2} & -d & (1+d^2)\end{bmatrix}\right) .$$

If we combine the information from the two assays we have the following equations for the large sample means:

$$\begin{aligned}
\hat{\alpha}_1 &= -\tfrac{1}{2}\hat{\beta}_1\hat{\mu} ,\\
\hat{\beta}_1 &= \frac{-d\hat{\mu}+1}{\hat{\mu}^2+1} ,\\
\hat{\alpha}_2 &= -\tfrac{1}{2}\hat{\beta}_2\hat{\mu} ,\\
\hat{\beta}_2 &= \frac{d\hat{\mu}+1}{\hat{\mu}^2+1} ,\\
\hat{\mu} &= \frac{d(-\hat{\beta}_1+a\hat{\beta}_2)}{\hat{\beta}_1^2+a\hat{\beta}_2^2} .
\end{aligned} \qquad (7.10)$$

Eliminating $\hat{\beta}_1$ and $\hat{\beta}_2$ from the expression for $\hat{\mu}$ we have the following quadratic for $\hat{\mu}$

$$d(a-1)\hat{\mu}^2 + (1-d^2)(1+a)\hat{\mu}-d(a-1) = 0 . \qquad (7.11)$$

If $a=1$ and $d\neq1$ then $\hat{\mu}=0$, and if $a=d=1$ then any value of $\hat{\mu}$ satisfies the equation. If $a\neq1$ then we have the following two solutions for $\hat{\mu}$:

$$\hat{\mu} = -b \pm \sqrt{1+b^2} \qquad (7.12)$$

where $b = \dfrac{(1-d^2)(1+a)}{2d(a-1)}$.

In order to see which of these solutions occurs at a maximum in the likelihood we need to examine the matrix of second derivatives of the log-likelihood. A solution to the equations 7.10 will be a maximum if the following matrix is positive definite:

$$\begin{bmatrix} 4 & 2\hat{\mu} & 0 & 0 & 2\hat{\beta}_1 \\ 2\hat{\mu} & 2\hat{\mu}^2+1 & 0 & 0 & 3\hat{\beta}_1\hat{\mu}+d \\ 0 & 0 & 4a & 2a\hat{\mu} & 2a\hat{\beta}_2 \\ 0 & 0 & 2a\hat{\mu} & a(2\hat{\mu}^2+1) & a(3\hat{\beta}_2\hat{\mu}-d) \\ 2\hat{\beta}_1 & (3\hat{\beta}_1\hat{\mu}+d) & 2a\hat{\beta}_2 & a(3\hat{\beta}_2\hat{\mu}-d) & 2(\hat{\beta}_1^2+a\hat{\beta}_2^2) \end{bmatrix}.$$

The matrix will be positive definite if all its principal minors are strictly positive. If a is strictly positive the first four principal minors are always strictly positive, and after a little algebra it can be shown that the fifth principle minor is strictly positive if

$$2d\hat{\mu}(a-1) + (1-d^2)(1+a) > 0 \qquad (7.13)$$

If $a=1$ then 7.13 is satisfied if $d<1$. If $a<1$ then 7.13 is satisfied if $\hat{\mu} = -b-\sqrt{b^2+1}$, and if $a>1$ we need $\hat{\mu} = -b+\sqrt{b^2+1}$. It can easily be shown that where there are two solutions to 7.11 the second solution is at a point which is neither a maximum nor a minimum in the likelihood. We can investigate the behaviour of the solutions to 7.11 for varying a and this is summarised schematically in Figure 7.[illegible] First let us consider the case $d<1$. This is intuitively a very pleasing result. The maximum likelihood value always falls in the range $[-d,+d]$ and it lies near $-d$ when the first assay contains much more information than the second, near $+d$ when the second assay contains much more

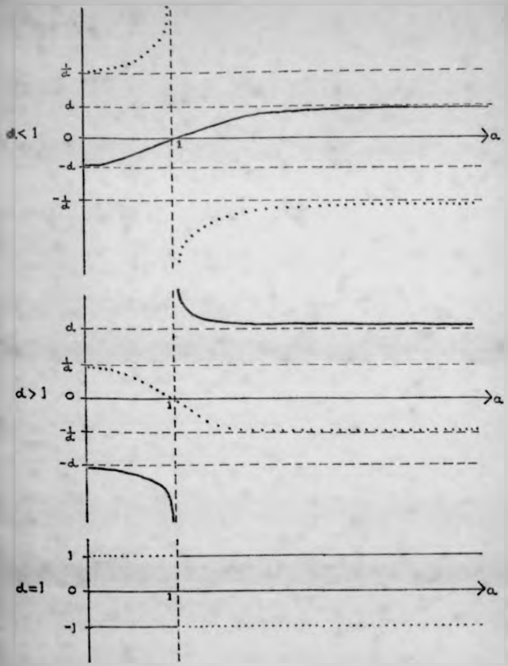


Figure 7.1 Schematic representation of the solutions to equation 7.11 for varying a. An unbroken line represents a maximum in the likelihood and a dotted line a second stationary point in the likelihood.

information than the first, and it equals 0 when the two assays contain equal amounts of information. In the case [illegible] > 1, the maximum likelihood value always lies outside the range $(-d, +d)$. This can be explained as follows. The data are now better explained if β_1 and β_2 lie near zero with opposite signs, than if μ lies near zero. If [illegible] and [illegible] lie near zero so will $\hat{\beta}_1$ and [illegible] and small values of $\hat{\beta}_1$ and $\hat{\beta}_2$ imply large values of the maximum likelihood value for μ. The case $d=1$ is the borderline between the two previous cases. The maximum likelihood value takes the value -1 when $a < 1$ and $+1$ when $a > 1$. When $a=1$ the likelihood has no maximum.

The asymptotic variance of μ is

$$\frac{\sigma^2(1+\hat{\mu})^2}{(1-d^2)(1+a)-2\hat{\mu}d(1-a)} \quad ,$$

where $\hat{\mu}$ is the relevant solution to 7.11.

We have examined the small sample case by plotting the posterior density of μ for various values of d and a. In each case we have let μ_*, the prior mean of μ, equal d, so that the second assay supports out prior beliefs while the first one contradicts them. We have let $\Phi^{-1}=0$ and changed the second stage variances according to our value of d so that the discrepancy between the assays when compared with the strength of the prior information remains roughly the same. For illustration we have taken $\sigma^2=\frac{1}{[illegible]}$ throughout. In our first example $d=\frac{1}{2}$ with second stage variances $\Sigma = \begin{bmatrix} \frac{1}{[illegible]} & 0 \\ 0 & \frac{1}{[illegible]} \end{bmatrix}$ and $\Sigma_{33}=\frac{1}{[illegible]}$. The resulting posterior density of μ when $a=\frac{1}{3}$ and $a=5$ are illustrated in Figure 7.2. As we might expect from the large sample results the density is unimodal, with mode lying near $-d$ when $a=\frac{2}{3}$ and near $+d$ when $a=5$. The densities are both slightly skewed to the right because we have taken $\mu_0=+1$. The case $d=1$,

$\underset{\sim}{\Sigma} = \begin{bmatrix} \frac{1}{4} & 0 \\ 0 & \frac{1}{3} \end{bmatrix}$, $\Sigma_{33} = \frac{2}{3}$ is illustrated in Figure 7.3 for $a = \frac{2}{3}$ and $a = 5$, and it is very similar to t e case $d = 1$. The posterior densities for these two values of d remain unimodal and of similar shapes even when the residual variance is very small; we have examined cases down to $\sigma^2 = 1/10{,}000$. Finally we have taken $d = 4$, $\underset{\sim}{\Sigma} = \begin{bmatrix} 4 & 0 \\ 0 & \frac{1}{3} \end{bmatrix}$, $\Sigma_{33} = 6$. This is illustrated in Figure 7.4 for $a = \frac{1}{3}$, 1 and 3. In the case $a = \frac{1}{3}$ the density is bimodal, the modes occuring at $\mu = -4.2$ and $\mu = 7.8$, while values of μ in the range (-2,4) are extremely improbable. When $a = 1$ the density is unimodal, with mode at $\mu = 7.4$, while negative values of μ are extremely improbable. The asymmetry in the situation is caused by the prior information. When $a = 3$ the density is again unimodal with mode at $\mu = 6.4$. Although this mode is at a value of substantially greater than 4, it is closer to 4 than in the case $a = 1$, thus following the behaviour of the large sample case.

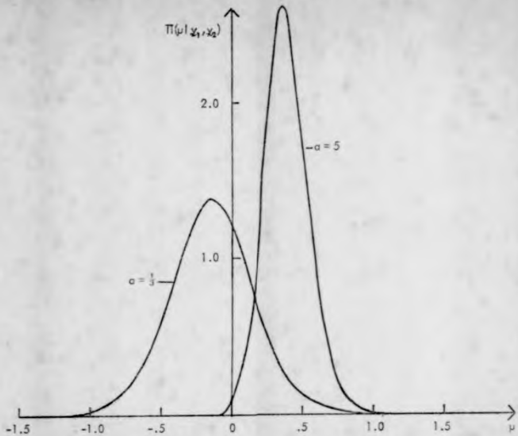


Figure 7.2 Posterior density of μ for the data given in Table 7.1 with parameters $d=\frac{1}{2}$ $\sigma^2=\frac{1}{9}$ $\underline{\Sigma}=\begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{3} \end{pmatrix}$, $\mu_0 = \frac{1}{2}$, $\Sigma_{33} = 1$, $\underset{\sim}{\hat{\phi}}^{-1} = \underset{\sim}{0}$.

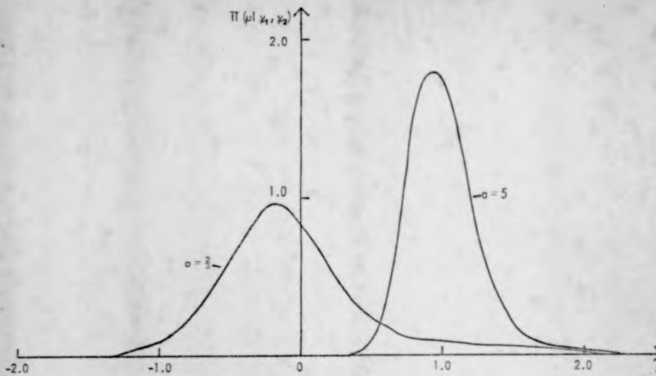


Figure 7.3 Posterior density of μ for the data given in Table 7.1 with parameters $d=1$, $\sigma^2=\frac{1}{9}$

$\Sigma = \begin{pmatrix} \frac{1}{4} & 0 \\ 0 & \frac{1}{3} \end{pmatrix}$, $\mu_0=1$ [illegible].

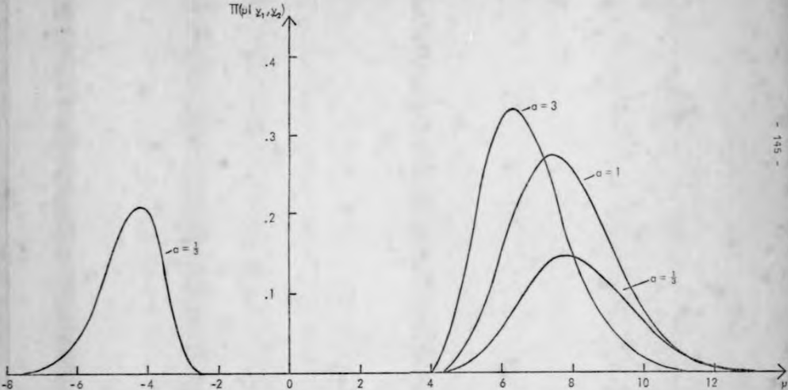


Figure 7.4 Posterior density of for the data given in Table 7.1 with parameters $d=4$, $\sigma^2=\frac{1}{9}$

[illegible] $^{-1} = \underset{\sim}{0}$.

7.5 Unknown Variances

We now consider the residual variance σ^2 and the second stage covariance matrix Σ as parameters in the model. We shall assume that our prior knowledge about each of them is independent and follows the relevent conjugate distribution, and so we have the following prior densities:

$$\pi(\sigma^2|\nu,\lambda) = (\sigma^2)^{-\frac{(\nu+2)}{2}} \exp - \frac{\nu\lambda}{2\sigma^2} \; ; \; \sigma^2 > 0 \; ,$$

$$\text{and } \pi(\underset{\sim}{\Sigma}^{-1}|\underset{\sim}{R},\rho) = |\underset{\sim}{\Sigma}|^{-\frac{(\rho-3)}{2}} \exp - \tfrac{1}{2}\, \mathrm{tr}\underset{\sim}{\Sigma}^{-1}\underset{\sim}{R} \; , \; \underset{\sim}{\Sigma} > 0 \; .$$

where $\underset{\sim}{R}$ is a 2 x 2 matrix, ρ is an integer and the values of $\underset{\sim}{R}$, ρ, ν and λ depend on the nature and precision of our prior knowledge.

In this section we shall assume that our prior knowledge of the location of α_o, β_o and μ is vague, and consequently $\underset{\sim}{\phi}^{-1} = \underset{\sim}{0}$ and $\frac{1}{\Sigma_{33}} = 0$. This may not be a valid assumption for any particular application, but our arguments can easily be adjusted if necessary.

The joint posterior density of all the parameters in the model is

$$\pi(\alpha_o,\beta_o,\mu,\alpha_1,\beta_1,\ldots,\alpha_m,\beta_m,\sigma^2,\underset{\sim}{\Sigma}^{-1}|\underset{\sim}{y}_1,\ldots,\underset{\sim}{y}_m,\nu,\lambda,\underset{\sim}{R},\rho) =$$

$$(\sigma^2)^{-\frac{\sum_{j=1}^{m} n_j}{2}} \exp - \frac{1}{2\sigma^2} \sum_{j=1}^{m} \left\{ \underset{\sim}{y}_j - \underset{\sim}{X}_j \begin{pmatrix} \alpha_j \\ \beta_j\mu \\ \beta_j \end{pmatrix} \right\}^T \left\{ \underset{\sim}{y}_j - \underset{\sim}{X}_j \begin{pmatrix} \alpha_j \\ \beta_j\mu \\ \beta_j \end{pmatrix} \right\}$$

$$\times \; |\underset{\sim}{\Sigma}|^{-\frac{m}{2}} \exp - \tfrac{1}{2} \sum_{j=1}^{m} \begin{pmatrix} \alpha_j - \alpha_o \\ \beta_j - \beta_o \end{pmatrix}^T \underset{\sim}{\Sigma}^{-1} \begin{pmatrix} \alpha_j \; \alpha_o \\ \beta_j \; \beta_o \end{pmatrix} \qquad ([illegible])$$

$$\times \quad (\sigma^2)^{-\frac{(\nu+2)}{2}} \exp - \frac{\nu\lambda}{2\sigma^2}$$

$$\times \quad |\underset{\sim}{\Sigma}|^{-\frac{(\rho-3)}{2}} \exp -\tfrac{1}{2}\mathrm{tr}\underset{\sim}{\Sigma}^{-1}\underset{\sim}{R}$$

Integrating over α_o and β_o in 7.14 the joint posterior density of the remaining parameters is

$$\pi\,(\mu,\alpha_1,\beta_1,\ldots\alpha_m,\beta_m,\sigma^2,\underset{\sim}{\Sigma}^{-1}\,|\,\underset{\sim}{y}_1,\ldots\underset{\sim}{y}_m,\nu,\lambda,\underset{\sim}{R},\rho) \propto$$

$$(\sigma^2)^{-\frac{\left(\sum\limits_{j=1}^{m} n_j+\nu+2\right)}{2}} \; |\underset{\sim}{\Sigma}|^{-\frac{(m+\rho-4)}{2}} \exp -\frac{1}{2\sigma^2}\left[\nu\lambda+\sum_{j=1}^{m}\left\{\underset{\sim}{y}_j-\underset{\sim}{X}_j\begin{pmatrix}\alpha_j\\ \beta_j\mu\\ \beta_j\end{pmatrix}\right\}^T\left\{\underset{\sim}{y}_j-\underset{\sim}{X}_j\begin{pmatrix}\alpha_j\\ \beta_j\mu\\ \beta_j\end{pmatrix}\right\}\right]$$

$$\times \; \exp -\tfrac{1}{2}\mathrm{tr}\underset{\sim}{\Sigma}^{-1}\left\{\underset{\sim}{R}+\sum_{j=1}^{m}\begin{pmatrix}\alpha_j-\bar{\alpha}\\ \beta_j-\bar{\beta}\end{pmatrix}\begin{pmatrix}\alpha_j-\bar{\alpha}\\ \beta_j-\bar{\beta}\end{pmatrix}^T\right\} . \qquad (7.15)$$

The mode of this density occurs at the point given by 7.5, but where σ^2 and $\underset{\sim}{\Sigma}$, instead of being known, are given by

$$\sigma^2 = \left(\sum_{j=1}^{m} n_j+\nu+2\right)^{-1}\left[\nu\lambda+\sum_{j=1}^{m}\left\{\underset{\sim}{y}_j-\underset{\sim}{X}_j\begin{pmatrix}\alpha_j\\ \beta_j\mu\\ \beta_j\end{pmatrix}\right\}^T\left\{\underset{\sim}{y}_j-\underset{\sim}{X}_j\begin{pmatrix}\alpha_j\\ \beta_j\mu\\ \beta_j\end{pmatrix}\right\}\right] , \qquad (7.16)$$

and

$$\underset{\sim}{\Sigma} = (m+\rho-4)^{-1}\left[\underset{\sim}{R}+\sum_{j=1}^{m}\begin{pmatrix}\alpha_j-\bar{\alpha}\\ \beta_j-\bar{\beta}\end{pmatrix}\begin{pmatrix}\alpha_j-\bar{\alpha}\\ \beta_j-\bar{\beta}\end{pmatrix}^T\right] .$$

Integrating over σ^2 and $\underset{\sim}{\Sigma}^{-1}$ in 7.15 the joint posterior density of $\mu, \alpha_1, \beta_1, \ldots, \alpha_m, \beta_m$ is

$$\pi(\mu, \alpha_1, \beta_1, \ldots, \alpha_m, \beta_m \mid y_1, \ldots, y_m, \nu, \lambda, \underset{\sim}{R}, \rho) \propto$$

$$\left[\nu\lambda + \sum_{j=1}^{m} \left\{ \underset{\sim}{y}_j - \underset{\sim}{X}_j \begin{pmatrix} \alpha_j \\ \beta_j \mu \\ \beta_j \end{pmatrix} \right\}^T \left\{ \underset{\sim}{y}_j - \underset{\sim}{X}_j \begin{pmatrix} \alpha_j \\ \beta_j \mu \\ \beta_j \end{pmatrix} \right\} \right]^{-\left(\frac{\Sigma n_j + \nu}{2}\right)}$$

$$\times \left| \underset{\sim}{R} + \sum_{j=1}^{m} \begin{pmatrix} \alpha_j - \bar{\alpha} \\ \beta_j - \bar{\beta} \end{pmatrix} \begin{pmatrix} \alpha_j - \bar{\alpha} \\ \beta_j - \bar{\beta} \end{pmatrix}^T \right|^{-\frac{(m+\rho-1)}{2}} . \qquad (7.17)$$

The mode of this density also occurs at the point 7.5, except σ^2 and $\underset{\sim}{\Sigma}$ are now estimated by

$$\sigma^2 = \left(\sum_{j=1}^{m} n_j + \nu\right)^{-1} \left[\nu\lambda + \sum_{j=1}^{m} \left\{ \underset{\sim}{y}_j - \underset{\sim}{X}_j \begin{pmatrix} \alpha_j \\ \beta_j \mu \\ \beta_j \end{pmatrix} \right\}^T \left\{ \underset{\sim}{y}_j - \underset{\sim}{X}_j \begin{pmatrix} \alpha_j \\ \beta_j \mu \\ \beta_j \end{pmatrix} \right\} \right] .$$

and

$$\underset{\sim}{\Sigma} = (m+\rho-1)^{-1} \left[\underset{\sim}{R} + \sum_{j=1}^{m} \begin{pmatrix} \alpha_j - \bar{\alpha} \\ \beta_j - \bar{\beta} \end{pmatrix} \begin{pmatrix} \alpha_j - \bar{\beta} \\ \beta_j - \bar{\beta} \end{pmatrix}^T \right] . \qquad (7.18)$$

Apart from the denominators these are the same equations as 7.16.

Returning to 7.15 and integrating over $\alpha_1, \beta_1, \ldots, \alpha_m, \beta_m$ the joint posterior density of μ, σ^2 and $\underset{\sim}{\Sigma}^{-1}$ is

$$\pi(\mu, \sigma^2, \underset{\sim}{\Sigma}^{-1} \mid \underset{\sim}{y}_1, \ldots, \underset{\sim}{y}_m, \nu, \lambda, \underset{\sim}{R}, \rho) \propto (\sigma^2)^{-\left(\frac{\sum_{j=1}^{m} n_j + \nu + 2}{2}\right)} |\underset{\sim}{\Sigma}|^{-\frac{(m+\rho-3)}{2}}$$

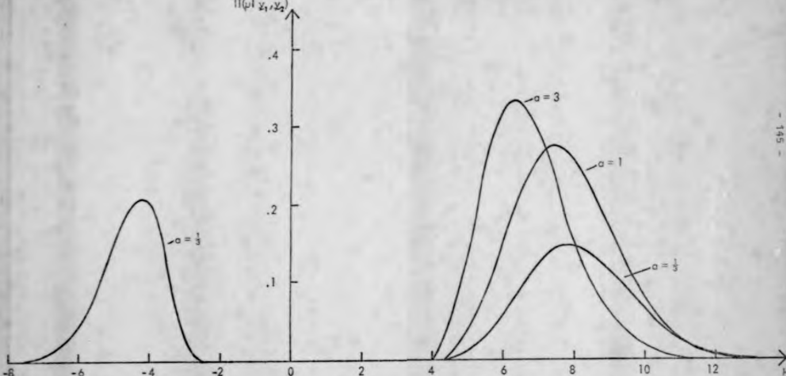


Figure 7.4 Posterior density of for the data given in Table 7.1 with parameters $d=4$, $\sigma^2 = \frac{1}{9}$

$\Sigma = \begin{pmatrix} 4 & 0 \\ 0 & \frac{1}{3} \end{pmatrix}$, $\nu_0 = 4$, [illegible]

7.5 Unknown Variances

We now consider the residual variance σ^2 and the second stage covariance matrix Σ as parameters in the model. We shall assume that our prior knowledge about each of them is independent and follows the relevent conjugate distribution, and so we have the following prior densities:

$$\pi(\sigma^2|\nu,\lambda) \propto (\sigma^2)^{-\frac{(\nu+2)}{2}} \exp_{-} \frac{\nu\lambda}{2\sigma^2} \quad ; \sigma^2 > 0 \; ,$$

$$\text{and } \pi(\underset{\sim}{\Sigma}^{-1}|\underset{\sim}{R},\rho) \propto |\underset{\sim}{\Sigma}|^{-\frac{(\rho-3)}{2}} \exp -\tfrac{1}{2} \, \mathrm{tr}\underset{\sim}{\Sigma}^{-1}\underset{\sim}{R} \; ; \; \underset{\sim}{\Sigma} > 0 \; .$$

where $\underset{\sim}{R}$ is a 2 x 2 matrix, ρ is an integer and the values of $\underset{\sim}{R}$, ρ, ν and λ depend on the nature and precision of our prior knowledge.

In this section we shall assume that our prior knowledge of the location of α_o, β_o and μ is vague, and consequently $\underset{\sim}{\phi}^{-1} = \underset{\sim}{0}$ and $\frac{1}{\Sigma_{33}} = 0$. This may not be a valid assumption for any particular application, but our arguments can easily be adjusted if necessary.

The joint posterior density of all the parameters in the model is

$$\pi(\alpha_o, \beta_o, \mu, \alpha_1, \beta_1, \ldots, \alpha_m, \beta_m, \sigma^2, \underset{\sim}{\Sigma}^{-1} | \underset{\sim}{y}_1, \ldots, \underset{\sim}{y}_m, \nu, \lambda, \underset{\sim}{R}, \rho) \propto$$

$$(\sigma^2)^{-\frac{\sum_{j=1}^{m} n_j}{2}} \exp -\frac{1}{2\sigma^2} \sum_{j=1}^{m} \left\{ \underset{\sim}{y}_j - \underset{\sim}{X}_j \begin{pmatrix} \alpha_j \\ \beta_j \mu \\ \beta_j \end{pmatrix} \right\}^T \left\{ \underset{\sim}{y}_j - \underset{\sim}{X}_j \begin{pmatrix} \alpha_j \\ \beta_j \mu \\ \beta_j \end{pmatrix} \right\}$$

$$\times \; |\underset{\sim}{\Sigma}|^{-\frac{m}{2}} \exp -\tfrac{1}{2} \sum_{j=1}^{m} \begin{pmatrix} \alpha_j - \alpha_o \\ \beta_j - \beta_o \end{pmatrix}^T \underset{\sim}{\Sigma}^{-1} \begin{pmatrix} \alpha_j & \alpha_o \\ \beta_j & \beta_o \end{pmatrix} \qquad (7.14)$$

$$\times \quad (\sigma^2)^{-\frac{(\nu+2)}{2}} \exp - \frac{\nu\lambda}{2\sigma^2}$$

$$\times \quad |\underset{\sim}{\Sigma}|^{-\frac{(\rho-3)}{2}} \exp - \tfrac{1}{2}\mathrm{tr}\underset{\sim}{\Sigma}^{-1}\underset{\sim}{R}$$

Integrating over α_o and β_o in 7.14 the joint posterior density of the remaining parameters is

$$\pi\,(\mu, \alpha_1, \beta_1, \ldots, \alpha_m, \beta_m, \sigma^2, \underset{\sim}{\Sigma}^{-1} \mid y_1, \ldots, y_m, \nu, \lambda, \underset{\sim}{R}, \rho) =$$

$$(\sigma^2)^{-\frac{\left(\sum_{j=1}^{m} n_j+\nu+2\right)}{2}} |\underset{\sim}{\Sigma}|^{-\frac{(m+\rho-4)}{2}} \exp - \frac{1}{2\sigma^2}\left[\nu\lambda + \sum_{j=1}^{m}\left\{\underset{\sim}{y}_j - \underset{\sim}{X}_j\begin{pmatrix}\alpha_j\\ \beta_j\mu\\ \beta_j\end{pmatrix}\right\}^T\left\{\underset{\sim}{y}_j - \underset{\sim}{X}_j\begin{pmatrix}\alpha_j\\ \beta_j\mu\\ \beta_j\end{pmatrix}\right\}\right]$$

$$\times \quad \exp - \tfrac{1}{2}\mathrm{tr}\underset{\sim}{\Sigma}^{-1}\left\{\underset{\sim}{R} + \sum_{j=1}^{m}\begin{pmatrix}\alpha_j-\bar{\alpha}\\ \beta_j-\bar{\beta}\end{pmatrix}\begin{pmatrix}\alpha_j-\bar{\alpha}\\ \beta_j-\bar{\beta}\end{pmatrix}^T\right\}. \qquad (7.15)$$

The mode of this density occurs at the point given by [illegible], but where σ^2 and $\underset{\sim}{\Sigma}$, [illegible], are given by

$$\sigma^2 = \left(\sum_{j=1}^{m} n_j+\nu+2\right)^{-1}\left[\nu\lambda + \sum_{j=1}^{m}\left\{\underset{\sim}{y}_j - \underset{\sim}{X}_j\begin{pmatrix}\alpha_j\\ \beta_j\mu\\ \beta_j\end{pmatrix}\right\}^T\left\{\underset{\sim}{y}_j - \underset{\sim}{X}_j\begin{pmatrix}\alpha_j\\ \beta_j\mu\\ \beta_j\end{pmatrix}\right\}\right], \qquad (7.16)$$

and

$$\underset{\sim}{\Sigma} = (m+\rho-4)^{-1}\left[\underset{\sim}{R} + \sum_{j=1}^{m}\begin{pmatrix}\alpha_j-\bar{\alpha}\\ \beta_j-\bar{\beta}\end{pmatrix}\begin{pmatrix}\alpha_j-\bar{\alpha}\\ \beta_j-\bar{\beta}\end{pmatrix}^T\right].$$

PAGE

MISSING

Integrating over σ^2 and $\underset{\sim}{\Sigma}^{-1}$ in 7.15 the joint posterior density of $\mu,\alpha_1,\beta_1,\ldots,\alpha_m,\beta_m$ is

$$\pi(\mu,\alpha_1,\beta_1,\ldots,\alpha_m,\beta_m|\underset{\sim}{y}_1\ldots\underset{\sim}{y}_m,\nu,\lambda,\underset{\sim}{R},\rho) \propto$$

$$\left[\nu\lambda+\sum_{j=1}^{m}\left\{\underset{\sim}{y}_j-\underset{\sim}{X}_j\begin{pmatrix}\alpha_j\\ \beta_j\mu\\ \beta_j\end{pmatrix}\right\}^T\left\{\underset{\sim}{y}_j-\underset{\sim}{X}_j\begin{pmatrix}\alpha_j\\ \beta_j\mu\\ \beta_j\end{pmatrix}\right\}\right]^{-\left(\frac{\Sigma n_j+\nu}{2}\right)}$$

$$\times\quad\left|\underset{\sim}{R}+\sum_{j=1}^{m}\begin{pmatrix}\alpha_j-\bar{\alpha}\\ \beta_j-\bar{\beta}\end{pmatrix}\begin{pmatrix}\alpha_j-\bar{\alpha}\\ \beta_j-\bar{\beta}\end{pmatrix}^T\right|^{-\frac{(m+\rho-1)}{2}} . \qquad (7.17)$$

The mode of this density also occurs at the point 7.5, except σ^2 and $\underset{\sim}{\Sigma}$ are now estimated by

$$\sigma^2=\left(\sum_{j=1}^{m}n_j+\nu\right)^{-1}\left[\nu\lambda+\sum_{j=1}^{m}\left\{\underset{\sim}{y}_j-\underset{\sim}{X}_j\begin{pmatrix}\alpha_j\\ \beta_j\mu\\ \beta_j\end{pmatrix}\right\}^T\left\{\underset{\sim}{y}_j-\underset{\sim}{X}_j\begin{pmatrix}\alpha_j\\ \beta_j\mu\\ \beta_j\end{pmatrix}\right\}\right] . \qquad (7.18)$$

and

$$\underset{\sim}{\Sigma}=(m+\rho-1)^{-1}\left[\underset{\sim}{R}+\sum_{j=1}^{m}\begin{pmatrix}\alpha_j-\bar{\alpha}\\ \beta_j-\bar{\beta}\end{pmatrix}\begin{pmatrix}\alpha_j-\bar{\beta}\\ \beta_j-\bar{\beta}\end{pmatrix}^T\right] .$$

Apart from the denominators these are the same equations as 7.16.

Returning to 7.15 and integrating over $\alpha_1,\beta_1,\ldots,\alpha_m,\beta_m$ the joint posterior density of μ,σ^2 and $\underset{\sim}{\Sigma}^{-1}$ is

$$\pi(\mu,\sigma^2,\underset{\sim}{\Sigma}^{-1}|\underset{\sim}{y}_1,\ldots,\underset{\sim}{y}_m,\nu,\lambda,\underset{\sim}{R},\rho)=(\sigma^2)^{-\left(\frac{\sum_{j=1}^{m}n_j+\nu+2}{2}\right)}|\underset{\sim}{\Sigma}|^{-\frac{(m+\rho-2)}{2}}$$

$$\times \left\{\prod_{j=1}^{m} |\underset{\sim}{D}_j+\underset{\sim}{\Sigma}^{-1}|^{-\frac{1}{2}}\right\} \left|\sum_{j=1}^{m} \underset{\sim}{\Sigma}^{-1}(\underset{\sim}{D}_j+\underset{\sim}{\Sigma}^{-1})^{-1}\underset{\sim}{D}_j\right|^{-\frac{1}{2}} \exp-\frac{1}{2}\left\{\frac{\nu\lambda+\mathrm{tr}\,\underset{\sim}{\Sigma}^{-1}\underset{\sim}{R}}{\sigma^2}\right\}$$

$$\exp-\frac{1}{2}\left[\frac{1}{\sigma^2}\sum_{j=1}^{m} \underset{\sim}{y}_j^T \underset{\sim}{y}_j - \sum_{j=1}^{m}\begin{pmatrix}a_j\\ b_j\end{pmatrix}^T (\underset{\sim}{D}_j+\underset{\sim}{\Sigma}^{-1})^{-1}\begin{pmatrix}a_j\\ b_j\end{pmatrix}\right.$$

$$\left.-\left\{\sum_{j=1}^{m} \underset{\sim}{\Sigma}^{-1}(\underset{\sim}{D}_j+\underset{\sim}{\Sigma}^{-1})^{-1}\begin{pmatrix}a_j\\ b_j\end{pmatrix}\right\}^T \left\{\sum_{j=1}^{m} \underset{\sim}{\Sigma}^{-1}(\underset{\sim}{D}_j+\underset{\sim}{\Sigma}^{-1})^{-1}\underset{\sim}{D}_j\right\}^{-1} \left\{\sum_{j=1}^{m} \underset{\sim}{\Sigma}^{-1}(\underset{\sim}{D}_j+\underset{\sim}{\Sigma}^{-1})^{-1}\begin{pmatrix}a_j\\ b_j\end{pmatrix}\right\}\right] \quad (7.19)$$

where a_j, b_j and $\underset{\sim}{D}_j$, $j=1,\ldots,m$ are as defined in section 7.2. The mode of this density cannot be found analytically.

In the case $\underset{\sim}{\Phi}^{-1}=0$, although not otherwise, we can proceed one step further by transforming from the variables μ,σ^2 and $\underset{\sim}{\Sigma}^{-1}$ to μ,σ^2 and $\underset{\sim}{S}^{-1}$ where $\underset{\sim}{S}^{-1}=\sigma^2\underset{\sim}{\Sigma}^{-1}$, and integrating over σ^2. This gives the posterior density of μ and $\underset{\sim}{S}^{-1}$

$$s(\mu,\underset{\sim}{S}^{-1}|\underset{\sim}{y}_1\ldots\underset{\sim}{y}_m,\nu,\lambda,\underset{\sim}{R},\rho) \propto |\underset{\sim}{S}|^{-\frac{(m-\rho-3)}{2}}\left\{\prod_{j=1}^{m}|\underset{\sim}{D}_j+\underset{\sim}{S}^{-1}|^{-\frac{1}{2}}\right\}$$

$$\times \left|\sum_{j=1}^{m} \underset{\sim}{S}^{-1}(\underset{\sim}{D}_{jo}+\underset{\sim}{S}^{-1})^{-1}\underset{\sim}{D}_{jo}\right|^{-\frac{1}{2}}$$

$$\times \left[\sum_{j=1}^{m} \underset{\sim}{y}_j^T \underset{\sim}{y}_j+\nu\lambda+\mathrm{tr}\,\underset{\sim}{S}^{-1}\underset{\sim}{R}-\sum_{j=1}^{m}\begin{pmatrix}a_{jo}\\ b_{jo}\end{pmatrix}^T (\underset{\sim}{D}_{jo}+\underset{\sim}{S}^{-1})^{-1}\begin{pmatrix}a_{jo}\\ b_{jo}\end{pmatrix}\right.$$

$$\left.-\left\{\sum_{j=1}^{m} \underset{\sim}{S}^{-1}(\underset{\sim}{D}_{jo}+\underset{\sim}{S}^{-1})^{-1}\begin{pmatrix}a_{jo}\\ b_{jo}\end{pmatrix}\right\}^T \left\{\sum_{j=1}^{m} \underset{\sim}{S}^{-1}(\underset{\sim}{D}_{jo}+\underset{\sim}{S}^{-1})^{-1}\underset{\sim}{D}_{jo}\right\}^{-1} \left\{\sum_{j=1}^{m} \underset{\sim}{S}^{-1}(\underset{\sim}{D}_{jo}+\underset{\sim}{S}^{-1})^{-1}\begin{pmatrix}a_{jo}\\ b_{jo}\end{pmatrix}\right\}\right]^{-\left(\frac{\sum n_j+2\rho+\nu-2}{2}\right)} \quad (7.20)$$

where $a_{jo}=\sigma^2 a_j$, $b_{jo}=\sigma^2 b_j$ and $\underset{\sim}{D}_{jo}=\sigma^2\underset{\sim}{D}_j$, $j=1,\ldots,m$. Again the mode of this density cannot be found analytically.

As in several of our previous models we cannot find the marginal posterior density of μ analytically. We could find an approximation to it by substituting an estimate of $\underset{\sim}{S}$ in 7.20. Alternatively, with only three nuisance parameters involved, calculation of the density numerically is not out of the question. However, in contrast to the previous cases, if we are combining a fairly large number of assays, we may have available a substantial amount of information about both μ and $\underset{\sim}{S}$. Consequently the joint posterior distribution of μ and $\underset{\sim}{S}$ may not be very different from a multivariate normal distribution. In this case the value of μ at the mode of the joint density would be approximately equal to the mean of its marginal posterior distribution, and an estimate of the precision of our information about μ could also be made by looking at the curvature of the joint density at its mode.

The theory described in chapter 5 to take account of experimental design features in a single assay extends straightforwardly both to the present model and to the model described in chapter 6. We have not repeated the theory for either of these two cases since the algebra is cumbersome and no new ideas are involved.

7.8 An Example: Tobramycin Data

We shall now analyse the data from four replicate assays of the antibiotic tobramycin given in Table 4.1. We have assumed that our prior knowledge of the likely values of the parameters is vague and so we have set $v=0$, $\frac{1}{\Sigma_{33}}=0$, and $\underline{\Phi}^{-1}=\underline{0}$ in our prior distributions. If we let $\underline{R}=0$ and $\rho=0$ the joint posterior density of all the parameters (7.14) is infinite when $\alpha_j=\alpha_0$, $\beta_j=\beta_0$, $j=1,\ldots,m$ and $\underline{\Sigma}=\underline{0}$, so, following section 6.4 we have put $\rho=2$ and chosen our value of $\underline{R}$ by estimating $\underline{\Sigma}$ from the large sample means, which are given in Table 7.2. The unbiased estimate of $\underline{\Sigma}$ in this case is not positive definite so we have taken as our estimate the sums of squares and cross-products of the large sample means divided by 3.

Using the above parameters in our prior distributions we have estimated μ in several different ways. We have then repeated the exercise with $\underline{R}$ ten times and one tenth our original value. The results are given in Table 7.3. The most obvious feature of these results is that all our estimates of μ are almost identical, whatever distribution they are based on, and regardless of $\underline{R}$. An approximate posterior density of μ is given in Figure 7.5. Transparencies 7 and 8 reveal this to be almost unchanged both for the smaller and for the larger $\underline{R}$.

As regards the other parameters, in the mode of the joint density of $\mu,\alpha_1,\beta_1,\ldots,\alpha_m,\beta_m,\sigma^2,\underline{\Sigma}^{-1}$, the α_i's and the β_i's are pulled together compared with the large sample means, but are largely independent of our choice of $\underline{R}$. The estimates of $\underline{\Sigma}$ depend quite heavily on our choice of $\underline{R}$. Our original guess at $\underline{\Sigma}$ was $\underline{\Sigma}=\begin{bmatrix} 28100. & 7230. \\ & 1860. \end{bmatrix}$ and this is consistent with our estimate of $\underline{\Sigma}$ based on the middle value of $\underline{R}$.

In the mode of the joint density of μ and $\underline{S}^{-1}$, the estimate of $\underline{S}$ again changes with our value of $\underline{R}$, and there are some inconsistencies between our estimates and our estimates of $\underline{\Sigma}$ and σ^2 in the previous case.

μ = .0186

α_1 = 28700.

α_2 = 28700.

α_3 = 29000.

α_4 = 28900.

β_1 = 6360.

β_2 = 6360.

β_3 = 6450.

β_4 = 6420.

σ^2 = 52500.

Table 7.2 Mean of approximate large sample distribution using data of four replicate tobramycin assays.

Mode of $\pi(\mu,\alpha_1,\beta_1,\ldots,\alpha_m,\beta_m,\sigma^2,\Sigma^{-1}|y_1,\ldots,y_m,v,\lambda,R,\rho)$

	a) $R=\begin{bmatrix}.4\times10^5 & .1\times10^5\\ .1\times10^5 & .4\times10^4\end{bmatrix}$	b) $R=\begin{bmatrix}[illegible] & .1\times10^4\\ [illegible] & .4\times10^3\end{bmatrix}$	c) $R=\begin{bmatrix}.4\times10^6 & .1\times10^6\\ .1\times10^5 & .4\times10^5\end{bmatrix}$
μ	.0185	.018[illegible]	.0186.
α_1	28800.	28800.	28600.
α_2	28700.	28800.	28700.
α_3	28000.	28900.	29000.
α_4	28900.	28900.	28900.
β_1	6370.	6380.	6370.
β_2	6380.	6390.	6370.
β_3	6410.	6400.	6430.
β_4	6400.	6400.	[illegible]
Σ_{11}	32800.	8430.	230000.
Σ_{12}	6230.	1140.	56700.
Σ_{22}	2280.	264.	21600.
σ^2	52000.	52200.	[illegible]

Mode of $\pi(\mu,\Sigma^{-1}|y_1,\ldots,y_m,v,\lambda,R,\rho)$

	a) $R=\begin{bmatrix}.4\times10^5 & .1\times10^5\\ .1\times10^5 & .4\times10^4\end{bmatrix}$	b) $R=\begin{bmatrix}.4\times10^4 & .1\times10^4\\ .1\times10^4 & .4\times10^3\end{bmatrix}$	c) $R=\begin{bmatrix}.4\times10^6 & .1\times10^6\\ .1\times10^5 & .4\times10^5\end{bmatrix}$
μ	.0185	.0185	.0185
S_{11}	1.61	1.14	5.51
S_{12}	.500	.362	1.54
S_{22}	[illegible]	.127	.601

$\pi(\mu|y_1,\ldots,y_m,v,\lambda,R,\rho,S)$

	a) $R=\begin{bmatrix}.4\times10^5 & .1\times10^5\\ .1\times10^5 & .4\times10^4\end{bmatrix}$	b) $R=\begin{bmatrix}.4\times10^4 & .1\times10^4\\ .1\times10^4 & .4\times10^3\end{bmatrix}$	c) $R=\begin{bmatrix}.4\times10^6 & .1\times10^6\\ .1\times10^6 & .4\times10^5\end{bmatrix}$
	$S=\begin{bmatrix}1.61 & .500\\ .500 & .186\end{bmatrix}$	$S=\begin{bmatrix}1.14 & .362\\ .362 & .127\end{bmatrix}$	$S=\begin{bmatrix}5.51 & 1.54\\ 1.54 & .601\end{bmatrix}$
Mean	.0185	.0185	.0188
Mode	.0185	.0185	.0188

Table 7.1 Results of analysis of [illegible] with prior parameters $v=0$, $\Phi^{-1}=0$, $\rho=2$ and R as indicated.

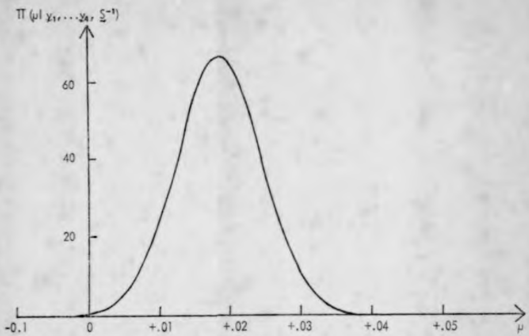


Figure 7.5 Approximate marginal posterior density of μ for tobramycin data assuming $\underset{\sim}{S}$ to be known and equal to its value at the mode of the joint density of μ and $\underset{\sim}{S}^{-1}$. Prior parameters are $v=0$, $\underset{\sim}{\phi}^{-1}=\underset{\sim}{0}$, $\rho=2$

$\underset{\sim}{R} = \begin{pmatrix} [illegible] & [illegible] \\ & [illegible] \end{pmatrix}$

Chapter 8. Conclusions.

8.1 General Remarks

We feel that we have been on the whole successful in our attempts to look at parallel line bioassay from a Bayesian point of view. We feel also that despite the algebraic complexities involved there are advantages to be gained from our nonlinear formulation of the problem, and we are satisfied that the major ideas behind the theory for the Bayesian linear model set out by Lindley & Smith (1972) carry over to this non-linear case.

A major advantage of our approach when compared with the standard sampling theory approach is that logically the way to proceed is very straightforward; the marginal posterior density of the log potency ratio should be calculated. This contrasts markedly with the theoretical complexities of combining information from several different assays using the sampling theory approach in its standard linear formulation.

A second advantage of our approach is that full use can be made of any available prior information. Biological assay is perhaps rather unusual in that fairly precise information about the potencies of both test and standard preparation is normally available before an assay is carried out. This is because the experimenter is restricted to estimating potency from doses which lie in the linear section of the log-dose response curve, and the range of doses for which this is so will depend critically on the potencies of the preparations concerned. In the absence of previous data a pilot study in the form of a small assay is often carried out before the main assay. Typically the results of this pilot study are used only to determine the doses for the main assay and are then ignored. In our present approach further use could be made of the results of such a pilot study in estimating the parameters of the prior distributions to be used for analyzing the results of the main assay.

A third advantage of our approach when considering several assays together is that we can make use of the fact that the results of the separate assays are likely to be similar to one another. This fact is ignored in all the sampling theory approaches to the problem that we have seen.

8.2 Possibilities for Further Work

We do not feel that this thesis is in any sense a complete treatment of the problem in hand. One particular point which deserves further theoretical investigation is the estimation of log potency ratio in cases where its marginal distribution is not obtainable analytically. Multidimensional numerical integrations provide a partial answer to the problem, and facilities for carrying these out are likely to be better in the future than they have been in the past. The ability to carry out such integrations in up to five dimensions would enable numerical estimation of the marginal density of μ in all the cases considered except that of chapter 6. In this case the dimension of the integration necessary to estimate the posterior mean of μ is $7+2m$ where m is the number of assays for which information is available.

There are two other points which we feel deserve a fuller treatment than we have given them. The first is the possibility of using a loss function other than a quadratic one in the point estimation of log potency ratio. For drugs such as antibiotics an overestimate of the potency is a more serious fault than an underestimate, and this indicates that an asymmetric loss function might be more appropriate than a symmetric one. We feel that this topic would be best approached by a detailed consideration of one or two particular drugs.

The other point which would be worth pursuing is a more sophisticated approach to the estimation of prior distributions from past assays. Trends in both the assay medium and the test preparation may occur and allowance should be made for this.

We feel that an approach very similar to our approach to parallel line assays could be made to slope-ratio assays. Slope ratio assays are similar to parallel line assays except that the response in the biological system is now linearly related to the dose of preparation administered rather than the log-dose. The residual variance is again assumed approximately normal. Suppose the slope of the linear section of the dose-response curve for the standard preparation is β, then the slope of the corresponding line for the test preparation is $\rho\beta$ where ρ is the potency ratio of the test preparation in terms of the standard. The first stage of a model for the analysis of a slope ratio assay would thus be

$$y \sim N((\alpha+\beta\rho xz+\beta x(1-z)),\ \sigma^2 \quad ,$$

where ..α,z and σ^2 have the same interpretation as in the parallel line case, and x is now the dose administered rather than the log-dose. Other aspects of the problem are identical with the parallel line case and much of our theory can easily be adapted by replacing x and z in the parallel-line case by xz and x(1-z) in the slope-ratio case.

8.3 A Note on Hypothesis Tests.

In this thesis we have made no mention of testing models to see if they are adequate descriptions of the data. There is definitely a need for a Bayesian equivalent to the sampling theory tests for linearity and parallelism in a single assay and also a test to detect outliers in a group of assays. The reason for this omission is that we have found there to be no general concensus of opinion on the subject of hypothesis testing in the literature, which in many cases is of a very abstract nature.

In the appendix we have included a short paper written in response to a request for a test for synergism between mixtures of drugs in parallel-line bioassays. The paper is written entirely from a sampling theory point of view since we were unsuccessful in producing a Bayesian test.

A Test for Synergism Between Two Drugs

S.C. Darby - Department of Medical Statistics and Epidemiology, London School of Hygiene and Tropical Medicine.

and

M.J. Ellis - Department of Medicine, St. Thomas's Hospital Medical School, London.

Summary

A likelihood ratio test is devised to detect the presence of synergism between two drugs which have similar actions. An example is given.

Keywords

BIOASSAY, INSULIN, LIKELIHOOD RATIO, MAXIMUM LIKELIHOOD, SYNERGISM.

1. Introduction

Suppose two drugs produce quantitative responses which are qualitatively similar. If mixtures of the drugs are applied, the question arises as to whether the drugs are additive or synergistic. By additive we mean that one drug can be replaced at a constant proportion by the other without affecting the response, and by synergistic we mean that the potency of a mixture of the drugs depends not only on the potency of the individual drugs but also on the proportions in which they are mixed. The type of joint action described by the additive model is often called simple similar action, see for example Finney (1971) and Ashford and Cobby (1974). We use the word synergism to denote any kind of deviation from additivity, including both potentiation and antagonism. The model that we use is a mathematical one. We have not attempted to represent the underlying mode of pharmacological or biological action of the drugs as Ashford and Cobby (1974) have done. Finney (1971) has considered the equivalent qualitative case. We devise a test to detect the presence of such synergism between the two drugs. The direction of the synergism can be determined graphically.

2. The Test

The two drugs, A and B, and all mixtures of them are assumed to have parallel log-dose response curves which are linear over the same range of responses. We assume that an assay has been carried out on q mixtures of the drugs, one mixture being pure A. We place no restriction on the number of doses of each mixture assayed, except that more than one dose must be used in at least one mixture. This is necessary in order to be able to estimate the slope of the linear part of the log dose response curve, and hence to obtain the residual sums of squares. We have also assumed that each point in the assay is replicated n times although very similar theory holds when different points are replicated differing numbers of times.

We test the null hypothesis, Ho, that the effect of the drugs is additive against the alternative, H_A, that the strength of any particular mixture is a property of that mixture alone. This general alternative will cover most types of synergism between the drugs.

Under the null hypothesis we assume that a dose of x units of A and z units of B is equivalent to a dose of x+μz units of A. Let the j^{th} dose of the i^{th} mixture be (x_{ij}, z_{ij}) and the k^{th} replicate response be y_{ijk}. The model is

$$E\ (y_{ijk}) = \alpha + \beta \log\ (x_{ij} + \mu\ z_{ij}).$$

Errors are assumed independently normally distributed. For any fixed μ the regression parameters can be estimated using maximum likelihood. This gives residual sum of squares:

$$\sum_{i,j,k} (y_{ijk} - \bar{y}_{\cdot\cdot\cdot})^2 - \frac{\left[\sum_{i,j} (\bar{y}_{ij\cdot} - \bar{y}_{\cdots}) \left(\log(x_{ij} + \mu z_{ij}) - \sum_{i,j} \frac{\log(x_{ij} + \mu z_{ij})}{m}\right)\right]^2}{\sum_{i,j} \left(\log(x_{ij} + \mu z_{ij}) - \sum_{i,j} \frac{\log(x_{ij} + \mu z_{ij})}{m}\right)^2},$$

where m is the total number of different doses in the assay, $\bar{y}_{\cdots}$ is the mean response for the entire assay, and $\bar{y}_{ij\cdot}$ is the mean response for the j^{th} dose of the i^{th} mixture. This residual sum of squares has mn-2 degrees of freedom. In order to find the maximum likelihood estimate of μ we minimize the above expression numerically with respect to μ. This minimum is the residual sum of squares under Ho, RSS_{Ho}, with mn-3 degrees of freedom.

Under the alternative hypothesis we assume that in the i^{th} mixture a dose of x units of A and z units of B are equivalent to a dose of $x + \mu_i z$ units of A. The model

$$E(y_{ijk}) = \alpha + \beta \log(x_{ij} + \mu_i z_{ij}).$$

Errors are again assumed independently normally distributed.

In the i^{th} mixture let $z_{ij} = p_i x_{ij}$, then the model becomes

$$E(y_{ijk}) = \gamma_i + \beta \log(x_{ij} + z_{ij}),$$

where $\gamma_i = \alpha + \beta \log\left\{(1 + \mu_i p_i) / (1 + p_i)\right\}$.

For that mixture which is pure A the corresponding μ_i is not defined since p_i is zero. From this formulation the alternative hypothesis can be seen to be symmetric with respect to the two drugs A and B. This model is linear, and so straightforward estimation of the parameters by maximum likelihood is possible. The residual sum of squares, RSS_{H_1}, is

$$\sum_{i,j,k}(y_{ijk} - \bar{y}_{...})^2 - \frac{\left[\sum_{i,j}(\bar{y}_{ij.} - \bar{y}_{i..})\left(\log(x_{ij} + z_{ij}) - \sum_j \frac{\log(x_{ij} + z_{ij})}{r_i}\right)\right]^2}{\sum_{i,j}\left(\log(x_{ij} + z_{ij}) - \sum_j \frac{\log(x_{ij} + z_{ij})}{r_i}\right)^2},$$

on mn-q-1 degrees of freedom, where r_i is the number of different doses of the i^{th} mixture that occur, and $\bar{y}_{i..}$ is the average response for the i^{th} mixture.

The test of Ho against H_A is made by considering the ratio

$$\frac{\dfrac{RSS_{Ho} - RSS_{H_A}}{q-2}}{\dfrac{RSS_{H_A}}{mn-q-1}}$$

and referring it to the F (q-2, mn-q-1) distribution. Asymptotic theory would suggest use of the likelihood ratio test statistic and the X^2 distribution here, however we conjecture that for finite samples, by analogy with the theory for linear models, use of the above test statistic and the F distribution will be a better approximation. The authors feel this point merits further investigation.

If there is evidence of synergism, a simple graphical method of determining its direction can be made by drawing an isobol or plot of the doses (x_{ij}, z_{ij}) which, under the alternative hypothesis, are estimated to produce the same response for each mixture assayed (Loewe, 1957). This can be done without calculating the estimated values for the μ_i. These values can of course be obtained if they are needed for further study.

The test described above may lack power due to considering arbitrary μ_i in the alternative hypothesis. Potentially more powerful tests for synergism might be developed for particular drugs by considering a more restricted class of alternatives. For example one could write the alternative model in the equivalent form

$$E(y_{ijk}) = \alpha + \beta \log [f(\pi_i) \times (x_{ij} + \mu z_{ij})],$$

with $\pi_i = x_{ij} / (x_{ij} + \mu z_{ij})$, and μ is the potency ratio of B in terms of A. In the above discussion $f(\pi_i)$ is completely general except that $f(0) = f(1) = 1$, but a parametric form could be posed for it. A point estimate of $f(\pi_i)$ for each of the various mixtures can be obtained from the isobol.

3. An Example

The topic under investigation is the interaction of insulin and a chemically modified insulin, A1-B29 suberoyl insulin, at the cellular level. The response measured is the conversion of (3-3H) glucose to toluene extractable lipids in isolated rat fat cells (Moody et al, 1974). The two drugs produce parallel log dose response curves which are linear over the range under consideration. The data are given in Table 1.

Table 1 here

The residual sums of squares for these data are $RSS_{H_0} = 260.1$ with 53 degrees of freedom, and $RSS_{H_A} = 194.4$ with 48 degrees of freedom. The test statistic is [illegible] with 5 and 48 degrees of freedom, and is significant at the [illegible] level. Hence this assay provides strong evidence that the effects of the two drugs are not additive.

Figure 1 here

An isobol (see Figure 1) indicates that greater amounts of the two substances are required when they are in combination than when applied independently, thus suggesting antagonism. The reproducibility of these results in further assays will be reported elsewhere.

Acknowledgements

The authors would like to thank Professor P. Armitage and Dr. P.H. Sönksen for their guidance and supervision during the research. S.C.D. is the recipient of a Medical Research Council studentship, and M.J.E. is a New Zealand recipient of a Commonwealth Postgraduate Scholarship in the United Kingdom.

References

ASHFORD, J.R., and COBBY, J.M. (1974). A system of models for the action of drugs applied singly or jointly to biological organisms. Biometrics, 30, 11-31.

FINNEY, D.J. (1971). Probit Analysis (3rd ed.). London: Cambridge University Press.

LOEWE, S. (1957). Antagonisms and Antagonists. Pharmac Rev. 9, 237-242.

MOODY, A.J., STAN, M.A., STAN, M., and GLIEMANN, J. (1974). A simple free fat cell bioassay for insulin. Horm. Metab. Res., 6, 12-16.

TABLE 1. RESULTS OF ASSAY

Mixture	Ratio of Insulin to A1-B29 Suberoyl Insulin	Total Dose (pmol l^{-1})	Responses for 4 replicates			
1	1:0	20.9	14.0	14.4	14.3	15.2
		41.9	24.6	22.4	22.4	26.7
2	1:1.85	52.9	11.7	15.0	12.9	8.3
		106.	20.6	18.0	19.6	20.5
3	1:5.56	101.	10.6	13.9	11.5	15.5
		202.	23.4	19.6	20.0	17.8
4	1:16.7	181.	13.8	12.6	12.3	14.0
		362.	15.8	17.4	18.0	17.0
5	1:50.0	261.	8.5	9.0	13.4	13.5
		522.	20.6	17.5	17.9	16.8
6	1:150	309.	12.7	9.5	12.1	8.9
		617.	18.6	20.0	19.0	21.1
7	0:1	340.	12.3	15.0	10.1	8.8
		681.	20.9	17.1	17.2	17.4

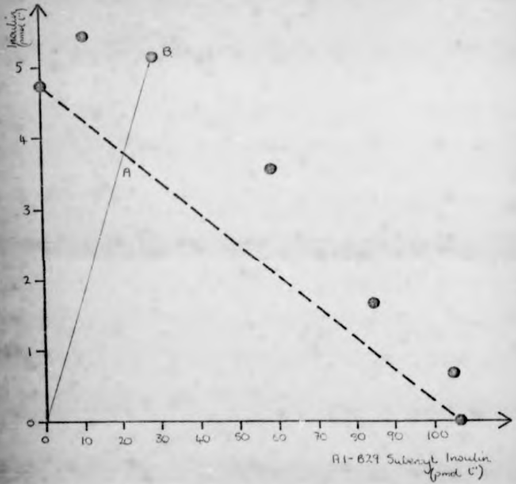


FIG 1. Isobol of assay data. The points are the estimated doses required to produce zero response under the alternative hypothesis. The dotted line represents the theoretical result for additive drugs. OA/OB is a point estimate of $f(\pi_3)$.

References

Armitage, P., Barnett, B.M. and Finney, D.J. (1976). Point and interval estimation in the combination of bioassay results. Journal of Hygiene, 76, 147-162.

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

de Finetti, B. (1975). Theory of Probability. Vols 1 & 2. London, Wiley.

Draper, N.R. and Hunter, W.G. (1967). The use of prior distributions in the design of experiments for parameter estimation in non-linear situations. Biometrika, 54, 147-153.

Finney, D.J. (1964). Statistical Method in Biological Assay, 2nd ed. London: Griffin & Co. Ltd.

Froberg, C.E. (1965). Introduction to Numerical Analysis. Addison-Wesley.

Lindley, D.V. (1961). The use of prior probability distributions in statistical inference and decisions. Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, University of California Press, 1, 453-468.

Lindley, D.V. (1971 a). Bayesian Statistics, A Review. Philadelphia: Society for Industrial and Applied Mathematics.

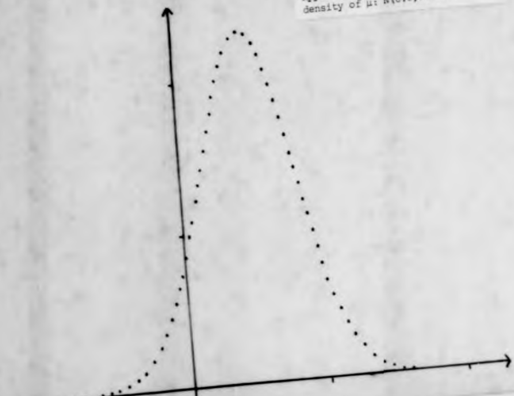
Lindley, D.V. (1971 b). The estimation of many parameters. In 'Foundations of Statistical Inference' (V.P. Godambe and D.A. Sprott, eds.) pp. 435-455. Toronto: Holt, Rinehart and Winston.

Lindley, D.V. and Smith, A.F.M. (1972). Bayes estimates for the linear model. J.R. Statist. Soc., B, 34, 1-41.

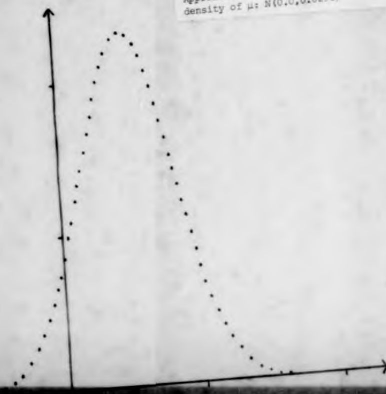
References / cont.

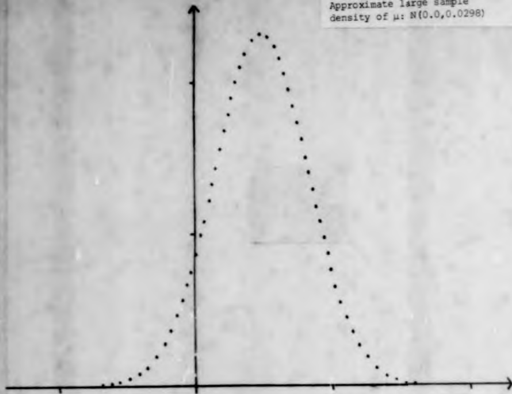
Smith, A.F.M. (1973). Bayes estimates in one-way and two-way models. Biometrika, 60, 319-329.

Transparency 1.
Approximate large sample
density of μ: N(0.0,0.0298)

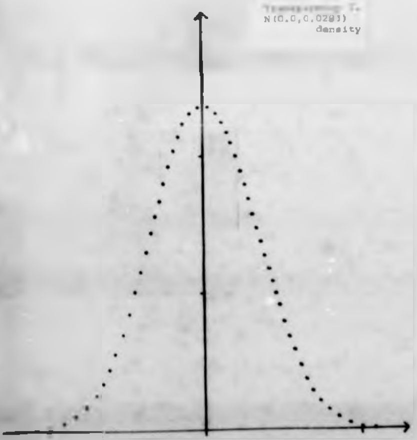


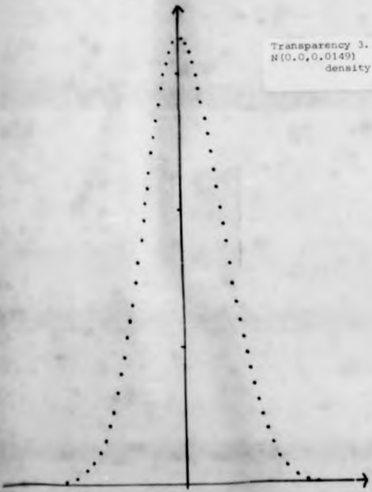
Transparency 1.
Approximate large sample
density of μ: N(0.0,0.0298)



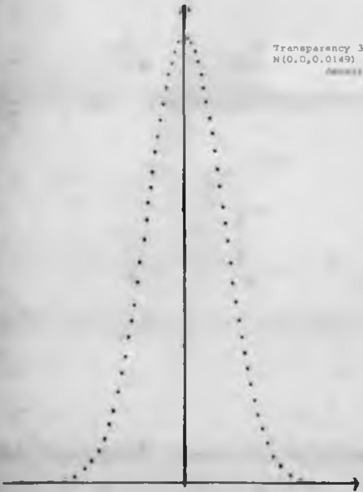


Transparency 1.
Approximate large sample
density of μ: N(0.0,0.0298)

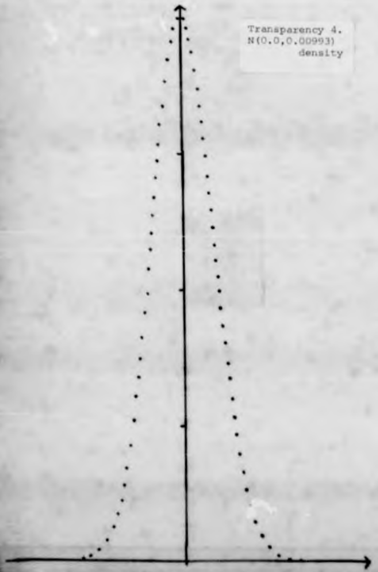

density



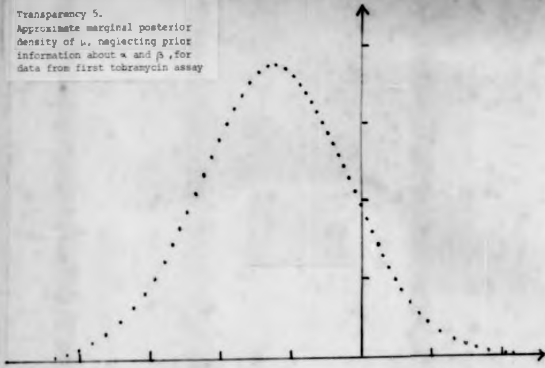
Transparency 3.
N(0.0,0.0149)
density



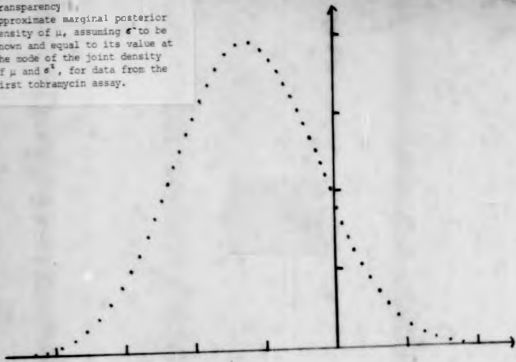
Transparency 3.
N(0.0,0.0149)

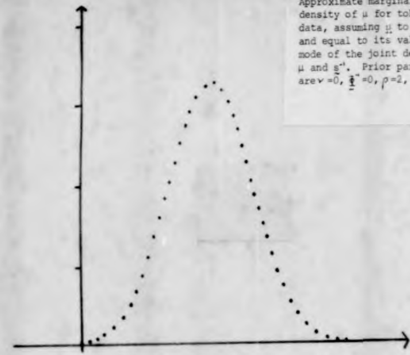

Transparency 4.
N(0.0,0.00993)
density

Transparency 5.
Approximate marginal posterior density of μ, neglecting prior information about α and β, for data from first tobramycin assay

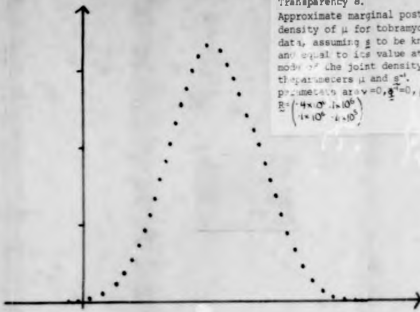


Transparency [illegible]
Approximate marginal posterior density of μ, assuming σ^2 to be known and equal to its value at the mode of the joint density of μ and σ^2, for data from the first tobramycin assay.





Transparency 7.
Approximate marginal posterior density of μ for tobramycin data, assuming $\underset{\sim}{\mu}$ to be known and equal to its value at the mode of the joint density of μ and $\underset{\sim}{s}^{-1}$. Prior parameters are $\nu = \tilde{0}$, $\underset{\sim}{\Phi}^{-1} = 0$, $\rho = 2$, $R = \begin{pmatrix} \cdot 4\times10^{4} & \cdot 1\times10^{4} \\ \cdot 1\times10^{4} & \cdot 4\times10^{3} \end{pmatrix}$



Transparency 8.
Approximate marginal posterior density of μ for tobramycin data, assuming $\underline{s}$ to be known and equal to its value at the mode of the joint density of the parameters μ and $\underline{s}^{-1}$. Prior parameters are $v=0$, $\underline{s}^{-1}=0$, $\rho=2$,
$\underline{R}=\begin{pmatrix} \cdot 4\times 10^{6} & \cdot 1\times 10^{6} \\ \cdot 1\times 10^{6} & \cdot 4\times 10^{5} \end{pmatrix}$