Design Workshop Lecture Notes
ISI, Kolkata, November, 25-29, 2002, pp. 51-60

Optimality Status of Incomplete Layout Three-Way Balanced Designs and Optimal Designs Under Heteroscedastic Errors in Linear Regression

Bikas K. Sinha
Theoretical Statistics and Mathematics Unit, Indian Statistical Institute
Kolkata, India

TOPIC I : Optimality Status of Incomplete Layout Three-Way Balanced Designs

The key references to this section are Agrawal (1966), Shah and Sinha (1990), Heiligers and Sinha (1995) and Saharay (1996). Other related references are Bagchi and Shah (1989), Hedayat and Raghavarao (1975) and Shah and Sinha (1989, 1996).

The set-up is that of a row-column design (also called a three-way design) and the model is that of fixed effects additive model with homoscedastic errors. We say that the row and column classifications represent known heterogeneity directions and the treatment effects correspond to the third component of variation. Our primary task is to provide efficient comparisons of the treatment effects. Further to this, we have two sets of differential effects, each within a specific heterogeneity direction, to be compared (or at least eliminated). Technically, these are referred to as row effects and column effects.

A three-way design involving R rows, C columns and v treatments is said to possess a complete (incomplete) layout if the number of experimental units (eu's) in the design is equal to (less than, respectively) RC. The row-column layout is generally made available to the experimenter who has to design treatment allocation over the eu's. The statistical analysis of data arising out of such a design is fairly straightforward and we refer to Shah and Sinha (1996) for this. There are three (assignable) sources of variation: rows, columns and treatments. These correspond to three classifications and the resulting data is often referred to as three-way classified data in the literature.

A three-way design is said to be *three-way balanced* (or, *totally balanced*) whenever the variances of estimates of normalized effects contrasts are the same for each classification. We readily verify that the simplest example of a three-way balanced design is a Latin square which is known to possess a complete layout involving the same number of rows, columns and treatments.

Incomplete three-way balanced designs are not easy to come across. Agrawal (1966) was the first to provide some series of such designs. Afterwards, Hedayat and Raghavarao (1975) also provided one series of such designs.

Note that a three-way balanced design in an incomplete three-way layout provides complete symmetry (c.s.) of the $\mathbf{C}$-matrices involving comparisons of row effects, or the column effects, or the treatment effects. In view of this kind of strong symmetry possessed by these designs, it is reasonable to expect that such designs will be optimal for inference on contrasts of the parameters for each of the three sources of variation: rows, columns and treatments !!

However, subsequent studies by Shah and Sinha (1990), Heiligers and Sinha (1995) and Saharay (1996) proved otherwise for most of the cases. Below we present salient features of these studies. Once for all, we use the notations $\mathbf{N}_{rc}$, $\mathbf{N}_{rt}$ and $\mathbf{N}_{ct}$ to denote respectively the row-column, row-treatment and column-treatment incidence matrices of appropriate orders in a given context. Likewise, $\mathbf{C}_r$, $\mathbf{C}_c$ and $\mathbf{C}_t$ will denote the $\mathbf{C}$-matrices for row effects comparisons, column effects comparisons and treatment effects comparisons, respectively.

It must be emphasized that in a given situation, we have the eu's available only at certain positions in the row-column setting as dictated by the row-column incidence matrix $\mathbf{N}_{rc}$. For given $\mathbf{N}_{rc}$, the allocation of the treatments to the eu's is to be so chosen that the corresponding incidence matrices $\mathbf{N}_{rt}$ and $\mathbf{N}_{ct}$ are feasible and consistent with the layout suggested by $\mathbf{N}_{rc}$.

Three-Way Balanced Designs Based on Agrawal's Method 1

We will first present results related to three-way balanced designs with the parameters $R = C = v = 4t + 3$, v a prime, for which the row-column incidence pattern is given by

$$\mathbf{N} = \mathbf{N}_{rc} = (0, 1) \text{ with } \mathbf{N} + \mathbf{N}' = \mathbf{J} - \mathbf{I} \tag{0.1}$$

and

$$\mathbf{N}\mathbf{N}' = (rt)\mathbf{I} + t\mathbf{J},\ r = 2t + 1. \tag{0.2}$$

Method 1 of Agrawal provides a design for which $\mathbf{N}_{rc} = \mathbf{N}_{rt} = \mathbf{N}_{ct} = \mathbf{N}$ as in (0.1)

It follows that for Agrawal's design

$$\mathbf{C}_A = \mathbf{C}_r = \mathbf{C}_c = \mathbf{C}_t = [(2t^2 - 1)/t][\mathbf{I} - \mathbf{J}/v]. \tag{0.3}$$

For treatment comparisons, however, the $\mathbf{C}_t$-matrix based on specified $\mathbf{N}_{rc}$ as above and arbitrary incidence matrices $\mathbf{N}_{rt}$ and $\mathbf{N}_{ct}$ is given by

$$\begin{aligned}\mathbf{C}_t &= (2t+1)[\mathbf{I} - \mathbf{J}/(4t+3)] - [(2t+1)/t(4t+3)][\mathbf{N}'_{rt}\mathbf{N}_{rt} + \mathbf{N}'_{ct}\mathbf{N}_{ct}] \\ &\quad +[1/t(4t+3)][\mathbf{N}'_{rt}\mathbf{N}\mathbf{N}_{ct} + \mathbf{N}'_{ct}\mathbf{N}'\mathbf{N}_{rt}]\end{aligned} \tag{0.4}$$

At this stage, for the choice $\mathbf{N}_{rt} = \mathbf{N}, \mathbf{C}_t$ simplifies to

$$\begin{aligned}\mathbf{C}_t &= [(2t+1)(4t^2+2t+1)/t(4t+3)]\mathbf{I} - [(2t+1)/t(4t+3)](\mathbf{N}'_{ct}\mathbf{N}_{ct}) \\ &\quad +[(t+1)/t(4t+3)](\mathbf{N}'_{ct} + \mathbf{N}_{ct})\end{aligned} \tag{0.5}$$

Now, we note that the last term in (0.4) has a positive coefficient. Agrawal's choice viz., $\mathbf{N}_{ct} = \mathbf{N}$ implies $\mathbf{N}'_{ct} + N_{ct} = \mathbf{J} - \mathbf{I}$ [vide (0.1))]. This was the point made by Shah and Sinha(1990). They argued that it is possible to make a choice of $\mathbf{N}_{ct}$ in such a way that $\mathbf{N}'_{ct} + \mathbf{N}_{ct}$ has 2 along the diagonal so that it contributes $2(t+1)/t$ to $trace(\mathbf{C}_t)$ in excess of what is obtained from Agrawal's choice. Of course, for large values of t, this increase may not be substantial. We now specialize to $R = C = v = 7$ and $\mathbf{N}_{rc} = (0\ 1\ 1\ 0\ 1\ 0\ 0)$, a circulant of order 7 with elements $[0, 1, 1, 0, 1, 0, 0]$ in the first row.

Shah and Sinha (1990) succeeded in finding an alternative to Agrawal's design for which the $\mathbf{C}_t$-matrix $\mathbf{C}_{S-S}$ say, satisfies the relation $\mathbf{C}_{S-S} > \mathbf{C}_A$ in the Loewner Ordering sense! However, $\mathbf{C}_{S-S}$ turned out *not* to be completely symmetric (c.s.). Their choice is $\mathbf{N}_{ct} = (1\ 0\ 1\ 0\ 0\ 1\ 0)$. Carrying this investigation further, Heiligers and Sinha (1995)

re-established superiority of another c.s. $\mathbf{C}_t$-matrix in the same set-up. Below we display all the three competing designs in terms of the row versus treatment incidence matrices (with rows, columns and treatments denoted by the numbers 0, 1, ..., 6 (mod 7):

Agrawal's Design: Initial Row = $\mathbf{N}_{rt}$ = (- 2 4 - 1 - -) and other rows are to be obtained by *developing* the initial row.

Shah and Sinha Design: Initial Row = $\mathbf{N}_{rt}$ = (- 1 4 - 2 - -) and other rows are to be obtained by *developing* the initial row.

Heiligers and Sinha Design: Initial Row = $\mathbf{N}_{rt}$ = (- 1 6 - 2 - -) and other rows are to be obtained by *developing* the initial row.

Computations yield

$$\begin{aligned} \mathbf{C}_A &= (6-1-1-1-1-1-1)/7; \\ \mathbf{C}_{S-S} &= (100-2-3-3-20)/7 \end{aligned}$$

and

$$\mathbf{C}_{H-S} = (12-2-2-2-2-2-2)/7 = 2\mathbf{C}_A.$$

It is interesting to note that Shah and Sinha design also dominates over Agrawal's design with respect to row effects comparisons and column effects comparisons as well.

Three-Way Balanced Designs Based on Agrawal's Method 3

In Method 3, Agrawal has constructed three-way balanced designs with the parameters $R = C = v$ and $\mathbf{N}_{rc} = \mathbf{N}_{rt} = \mathbf{N}_{ct} = \mathbf{J} - \mathbf{I}$. A simpler method to construct such designs would be to construct Latin squares of order $v \times v$ with diagonal elements all different and then to delete the diagonal. Anyway, such designs have so much symmetry that it seems impossible to beat them for any of the three sources of variation.

Computations yield:

$$\mathbf{C}_A = \mathbf{C}_r = \mathbf{C}_c = \mathbf{C}_t = [v(v-3)/(v-2)][\mathbf{I} - \mathbf{J}/v]. \tag{0.6}$$

Further, for arbitrary choices of $\mathbf{N}_{rt}$ and $\mathbf{N}_{ct}$ subject to $\mathbf{N}_{rc} = \mathbf{J} - \mathbf{I}$, it follows that

$$\begin{aligned} \mathbf{C}_t &= \mathbf{r}^d + [1/(v-1)(v-2)]\mathbf{r}\mathbf{r}' \\ &- \frac{1}{v(v-2)}\left[(v-1)(\mathbf{N}'_{ct}\mathbf{N}_{ct} + \mathbf{N}'_{rt}\mathbf{N}_{rt}) + (\mathbf{N}'_{ct}\mathbf{N}_{rt} + \mathbf{N}'_{rt}\mathbf{N}_{ct})\right]. \end{aligned} \tag{0.7}$$

At this stage, Shah and Sinha (1990) argued that Agrawal's design has minimum trace($\mathbf{C}_t$) in the binary class and that it is possible to increase trace($\mathbf{C}_t$) by allowing for unequal replications.

Carrying this investigation further, Saharay (1996) made the choice: $\mathbf{N}_{rt} = \mathbf{J} - \mathbf{I}$; $\mathbf{N}_{ct} = \mathbf{J} - \mathbf{P}$, $\mathbf{P}$ being a cyclic permutation matrix of order v, different from the identity permutation. For this choice

$$\mathbf{C}_S = \mathbf{C}_A + [1/v(v-2)][2\mathbf{I} - \mathbf{P} - \mathbf{P}']. \tag{0.8}$$

Since $2\mathbf{I} - \mathbf{P} - \mathbf{P}'$ is an *nnd* matrix, we have Loewner domination (LOD) of $\mathbf{C}_A$ by $\mathbf{C}_S$. Next, she established that designs with $\mathbf{C}$-matrix of the form (0.8) are E-optimal. Also, by extensive computer search, she verified that such designs are A- and D- optimal at least for values of v up to and including 20 within the class of binary equireplicate designs. Finally, she developed a systematic method of constructing such designs.

We now display below the designs by Agrawal, Shah and Sinha and Saharay for the case of $v = 7$. The rows, columns and treatments are denoted by the numbers 1, 2, ..., 7.

Agrawal's Design:

–	2	3	4	5	6	7
6	–	1	2	3	4	5
4	5	–	7	1	2	3
2	3	4	–	6	7	1
7	1	2	3	–	5	6
5	6	7	1	2	–	4
3	4	5	6	7	1	–

Shah and Sinha Design:

–	7	6	2	1	3	5
4	–	7	6	2	1	3
5	4	–	7	6	2	1
3	5	4	–	7	6	2
1	3	5	4	–	7	6
2	1	3	5	4	–	7
7	2	1	3	5	4	–

Saharay Design:

–	1	4	3	2	5	6
2	–	3	4	5	6	7
3	4	–	5	6	7	1
4	5	6	–	7	1	2
5	6	7	1	–	2	3
7	3	2	6	1	–	4
1	2	5	7	4	3	–

It is also interesting to note that in the above example, Shah and Sinha design improves over Agrawal's design with respect to D-optimality criterion only for treatment effects comparisons while for row effects and column effects comparisons, it does improve over Agrawal's with respect to both A- and D-optimality criteria. There is no Loewner domination (LOD) of Shah and Sinha designs over Agrawal's.

Three-Way Balanced Designs Based on Agrawal's Method 4

In his Method 4, Agrawal (1966) constructed designs with the parameters

$$\begin{aligned}
&R = C = 4t+3 \text{ (a prime)},\ v = 2(4t+3),\\
&\mathbf{N}_{rc} = \mathbf{J} - \mathbf{I},\ \mathbf{N}_{rt} = (\mathbf{N}, \mathbf{N}) \text{ and } \mathbf{N}_{ct} = (\mathbf{N}, \mathbf{N})\\
&\text{with } \mathbf{N} + \mathbf{N}' = \mathbf{J} - \mathbf{I} \text{ and } \mathbf{N}'\mathbf{N} = (t+1)\mathbf{I} + t\mathbf{J}.
\end{aligned} \tag{0.9}$$

It can be seen that this solution is obtained by duplicating the solution in Agrawal's Method 1. We have already observed that designs based on Method 1 can be improved. This prompted Shah and Sinha (1990) to suggest improved designs in this situation as well. They started with the choice $\mathbf{N}_{rt} = [\mathbf{N}\ \mathbf{P}]$, $\mathbf{N}_{ct} = [\mathbf{P}\ \mathbf{N}]$ for some $\mathbf{P}$ where $\mathbf{N}$ is as in (0.9). Then the $\mathbf{C}_t$ matrix assumes the form (apart from a term involving $\mathbf{J}$),

$$\mathbf{C}_t = \begin{pmatrix} (2t+1)\mathbf{I} - \frac{\mathbf{N}'\mathbf{N}}{(4t+2)} & -\frac{\mathbf{N}'\mathbf{P}}{(4t+2)} \\ & (2t+1)\mathbf{I} - \frac{\mathbf{P}'\mathbf{P}}{(4t+2)} \end{pmatrix} - \frac{(4t+2)}{(4t+1)(4t+3)}\mathbf{H}, \tag{0.10}$$

where

$$\mathbf{H} = \begin{pmatrix} \mathbf{P'P} + \frac{(\mathbf{P'N}+\mathbf{N'P})}{(4t+2)} + \frac{\mathbf{N'N}}{(4t+2)^2} & \mathbf{P'N} + \frac{(\mathbf{P'P}+\mathbf{N'N})}{(4t+2)} + \frac{\mathbf{N'P}}{(4t+2)^2} \\ & \mathbf{N'N} + \frac{(\mathbf{N'P}+\mathbf{P'N})}{(4t+2)} + \frac{\mathbf{P'P}}{(4t+2)^2} \end{pmatrix}. \tag{0.11}$$

If we restrict to a binary equireplicate design, then $trace(\mathbf{P'P})$ coincides with $trace(\mathbf{N'N})$. From (0.9) it is evident that in order to maximize $trace(\mathbf{C}_t)$, we have to make a choice of $\mathbf{P}$ so as to minimize $trace(\mathbf{P'N} + \mathbf{N'P})$. For Agarwal's choice $\mathbf{P} = \mathbf{N}$, we have $trace(\mathbf{P'N} + \mathbf{N'P}) = (2t+1)v$.

At this stage Shah and Sinha (1990) argued that it is indeed possible to make a choice of $\mathbf{P}$ so that $trace(\mathbf{P'N} + \mathbf{N'P}) << v(2t+1)$. It is true for $v = 7$ as was illustrated by them by taking $\mathbf{P} = (1\ 0\ 1\ 0\ 0\ 1\ 0)$, a circulant of order 7. It follows that the resulting design does better than Agrawal's design with respect to A-, D- and E-optimality criteria for treatment effects comparisons. Further, it shows Loewner dominance over Agrawal's for both row effects and column effects comparisons.

Concluding Remarks

For Method 2 of Agrawal (1966), it has been conjectured in Shah and Sinha (1990) that Agrawal's designs are at least A- , D- and E-optimal. Optimality of Agrawal's designs given by Method 5 was discussed in Bagchi and Shah (1989) and also in Shah and Sinha (1989). These designs are found to be Ψ_f-optimal in the class of equireplicate designs and E- optimal in the unrestricted class. These designs are based on $\mathbf{N}_{rc} = \mathbf{J}$ so that the row-column classification to start with is orthogonal.

TOPIC II : Optimal Designs Under Heteroscedastic Errors in Linear Regression

Most optimality studies in the context of linear, quadratic and polynomial regression models deal with homoscedastic error structure. However, some recent studies also reflect situations wherein the use of heteroscedastic error structureheteroscedastic error structure in the context of linear regression is called for. See, for example, Abdelbasit and Plackett (1983), Minkin (1987, 1993), Wu (1988), Sitter and Wu (1993), Khan and Yazdi (1988), Ford *et al.* (1992), Atkinson and Cook (1995), Hedayat *et al.* (1997), Sebastiani and Settimi (1997) and Mathew and Sinha (2001).

The key reference to this section is Liski et al (2002) and Minkin (1993) and we intend to re-visit an interesting optimality result reported therein. It turns out that there is a unified approach to the problem studied there. We exploit de la Garza phenomenon and Loewner order domination (LOD) for asymmetric experimental domains and succeed in characterizing a complete class of experiments, using continuous design theory. This study enables us extend optimality results in Minkin (1993) in a natural manner.

A Model with Heteroscedastic Errors

Following Minkin (1993), we consider a Poisson count model

$$\begin{aligned} &Y_x \sim \text{ Poisson}[\mu(x)] \\ &\text{with } \mu(x) = c(x)e^{\theta(x)},\ \theta(x) = \alpha + \beta x \text{ and } \beta = -\beta^* < 0, \end{aligned} \tag{0.12}$$

where α and β are unknown parameters and x is a non-stochastic covariate over the domain $\mathcal{T} = [0, \infty)$. Further, $c(x)$ is a known positive quantity, independent of the unknown parameters.

Since we are primarily interested in the regression parameters, we assume,for simplicity, that we are in a situation where we can use approximate theory and accordingly, confine to the n-point design (vide Minkin 1993):

$$d_n = [(0 \leq x_1 < \cdots < x_n < \infty;\ 0 < p_1, p_2, \ldots, p_n;\ \sum c(x_i)p_i = 1]. \tag{0.13}$$

Here $c(x)$'s are known positive (finite) real numbers.

Next we write down the log-likelihood function and hence deduce the form of the asymptotic information matrix for the parameters as

$$\mathbf{I}(\alpha, \beta) = \sum c(x_i)p_i e^{\theta(x_i)}(1, x_i)'(1, x_i). \tag{0.14}$$

Note that the information matrix has an equivalent representation given by

$$\begin{aligned} &\mathbf{I}(\alpha, \beta) = \sum P_i e^{\theta(x_i)}(1, x_i)'(1, x_i); \\ &P_i = c(x_i)p_i,\ 1 \leq i \leq n; \quad \sum P_i = 1. \end{aligned} \tag{0.15}$$

With reference to $\mathbf{I}$, optimality problem refers to optimal choice of the x_i's and the corresponding P_i's subject to $\sum P_i = 1$. Note that Minkin (1993) characterized the nature of optimum design for minimizing $V(\hat{\beta}^{-1})$. As is typical in nonlinear settings, the optimum design depends on the unknown parameters α and β.

The components of I are given by

$$I_{11} = \sum_{i=1}^{n} P_i e^{\alpha - \beta^* x_i},\ I_{12} = \sum_{i=1}^{n} P_i x_i e^{\alpha - \beta^* x_i},\ I_{22} = \sum_{i=1}^{n} P_i x_i^2 e^{\alpha - \beta^* x_i}. \tag{0.16}$$

It now follows that without any loss of generality, we can ignore the factor e^{α} in the expression for I. Further, let us write $\beta^* x_i = x_i^*$ for each i. For convenience, we re-name the information matrix as I^* having elements

$$I_{11}^* = \sum_{i=1}^{n} P_i e^{-x_i^*},\ I_{12}^* = \sum_{i=1}^{n} P_i x_i e^{-x_i^*},\ I_{22}^* = \sum_{i=1}^{n} P_i x_i^2 e^{-x_i^*}. \tag{0.17}$$

Remark : We note in passing that I^* can be written as a convex combination of component I^*'s based on subsets of points in $\mathcal{T}$. In other words,

$$I^* = \sum_{r=1}^{k} \pi_r I^*(r) \tag{0.18}$$

where $\pi_1, \pi_2, \ldots, \pi_k$ are sums over mutually exclusive and exhaustive subsets of the set of P_i's and the $I^*(r)$'s are the component information matrices based on the corresponding mutually exclusive and exhaustive subsets of the set of x_i's with the revised P-values as $P_i^* = \frac{P_i}{\pi_r}$ whenever P_i is in the rth subset of the P-values. Minkin (1993) established that for optimal estimation of the slope parameter, an optimal design is a 2-point design including the point 0.

Complete Class of Designs in Minkin's Set-Up

We provide here a unified approach to arrive at a very general result to the effect that for inference on the regression parameters, a complete class of continuous designs must necessarily comprise of 2-point designs including the point 0.

Towards establishing the complete class result stated above, first we start with a 2-point design viz.,

$$\begin{aligned} &d_2 = [a, b; P, Q]; \\ &\text{with } 0 < a < b < \infty; \quad 0 < P, Q < 1; \quad P + Q = 1. \end{aligned} \tag{0.19}$$

Then, writing $a^* = a\beta^*$ and $b^* = b\beta^*$,

$$I_{11}^* = Pe^{-a^*} + Qe^{-b^*}; \tag{0.20}$$

$$I_{12}^* = Pae^{-a^*} + Qbe^{-b^*}; \tag{0.21}$$

$$I_{22}^* = Pa^2e^{-a^*} + Qb^2e^{-b^*}. \tag{0.22}$$

The following is the main result in this study.

Given d_2 in (0.19), there exists another 2-point design $d_2^* = [0, c; \lambda, 1-\lambda]$ that dominates d_2 in the Loewner domination (LOD) sense i.e.

$$d_2^* = [0, c; \lambda, 1 - \lambda] \succ d_2 = [a, b; P, Q] \tag{0.23}$$

such that $\mathcal{I}^*(d_2^*) - \mathcal{I}^*(d_2)$ is a nnd matrix.

Proof. Utilization of (0.20) - (0.22) yields

$$I_{11}^*(d_2^*) - I_{11}^*(d_2) = \lambda + (1 - \lambda)e^{-c^*} - Pe^{-a^*} - Qe^{-b^*}; \tag{0.24}$$

$$I_{12}^*(d_2^*) - I_{12}^*(d_2) = (1 - \lambda)ce^{-c^*} - Pae^{-a^*} - Qbe^{-b^*}; \tag{0.25}$$

$$I_{22}^*(d_2^*) - I_{22}^*(d_2) = (1 - \lambda)c^2e^{-c^*} - Pa^2e^{-a^*} - Qb^2e^{-b^*}. \tag{0.26}$$

We now proceed as in Liski *et al.* (2002), taking the clue from Pukelsheim (1993). In other words, in the expression for the difference of the two information matrices, we equate both the terms in $(1, 1)$th and $(1, 2)$th positions to 0 and solve for c and λ. Then we show that the term in the $(2, 2)$th position is strictly positive.

The equations in terms of $c^* = c\beta^*$ and λ are given by

$$\lambda + (1 - \lambda)e^{-c^*} = Pe^{-a^*} + Qe^{-b^*}; \tag{0.27}$$

$$(1 - \lambda)ce^{-c^*} = Pae^{-a^*} + Qbe^{-b^*} \tag{0.28}$$

whence, eliminating λ, we obtain the following equation involving c^*:

$$\begin{aligned} &\phi(c^*) = w\phi(a^*) + (1 - w)\phi(b^*), \\ \text{where} \qquad &\phi(x^*) = \frac{e^{x^*} - 1}{x^*} \text{ and } w = \frac{Pa^*e^{-a^*}}{Pa^*e^{-a^*} + Qb^*e^{-b^*}}. \end{aligned} \tag{0.29}$$

It is readily seen that the function $\phi(x^*)$ is convex and increasing in x^* over $[0, \infty)$ so that we have a unique solution for c^* above. Moreover, $a^* < c^* < b^*$ i.e., $a < c < b$. Once c^* is known, λ is obtained from the relation

$$1-\lambda = \frac{P(1-e^{-a^*})+Q(1-e^{-b^*})}{(1-e^{-c^*})}. \tag{0.30}$$

It now remains to verify that $0<\lambda<1$ which is equivalent to verifying that

$$e^{-c^*} < Pe^{-a^*} + Qe^{-b^*}. \tag{0.31}$$

The proof of this claim is given in the Appendix.

Finally, to show strict positivity of the $(2,2)$th term in (0.26), upon simplification, we find that we have to establish the inequality: $c^* > wa^* + (1-w)b^*$. This follows readily from the strict convexity of the function $\phi(x^*)$ and the defining equation for c^*. This establishes the result.

Corollary. Given d_n in (0.13), there exists a 2-point design $d_2 = [0, c; \lambda, 1-\lambda]$ that dominates d_n in the Loewner domination sense.

Proof. In view of Remark , the proof follows by induction on the number n of support points of d_n.

We have thus arrived at a characterization of a complete class of continuous designs for inference on the regression parameters in the set-up described in (0.12) and (0.13). It is now a routine task to determine specific optimal designs for one or both the parameters with respect to various optimality criteria. This is studied in the next section. We note that the asymptotic variance-covariance matrix of the maximum likelihood estimators of α and β is obtained by inverting the information matrix $\mathbf{I}$.

Specific Optimal Designs

For inference on the parameter β or on β^{-1} (as discussed in Minkin (1993)) with minimum asymptotic variance i.e., maximum information, we are supposed to maximize $I^*_{22.1}$ for proper choices of c^* and λ.

Algebraically, we have to maximize

$$f(s,c) = sc^2e^{-c^*} - \frac{[sce^{-c^*}]^2}{[1-s(1-e^{-c^*})]},$$

where $s=(1-\lambda)$. Note that this is the same as maximizing

$$f(s,c^*) = sc^{*2}e^{-c^*} - \frac{[sc^*e^{-c^*}]^2}{[1-s(1-e^{-c^*})]}.$$

For fixed c^*, we can maximize the above expression in terms of s and obtain $s_{opt}(c^*) = \frac{1}{[1+e^{-c^*/2}]}$ whence, upon substituting the expression for s, we need to maximize $f(c^*) = [\frac{c^*}{1+e^{c^*/2}}]^2$. It follows readily that $c^*_{opt} = 2.557$; $\lambda_{opt} = 0.218$. Minkin (1993) deduced this result through a different approach.

For inference on α, it follows that an optimal design is singular with $P_{[x=0]} = 1$. For D-optimality involving both the parameters α and β, we have to maximize $f(s,c^*) = s(1-s)c^{*2}e^{-c^*}$ and this yields $c^*_{opt} = 2.0$, $\lambda_{opt} = 0.5$.

For A-optimality, we have to minimize $f(s,c^*) = \frac{A}{s} + \frac{B}{(1-s)}$, where $A = \frac{e^{c^*}}{c^{*2}}$; $B = 1 + \frac{1}{c^{*2}}$. This yields $c^*_{opt} = 2.261$, $\lambda_{opt} = 0.444$. Again, for E-optimality, we have to maximize

$$f(s,c^*) = [1-L_1s] - [(1-L_2s)^2 + Qs^2]^{1/2},$$

where $L_1 = 1 - e^{-c^*} - c^{*2}e^{-c^*}$, $L_2 = 1 - e^{-c^*} + c^{*2}e^{-c^*}$ and $L = 4c^{*2}e^{-2c^*}$.

Numerical calculations yield: $c^*_{opt} = 2.565$; $\lambda_{opt} = 0.4002$. Finally, for MV-optimality, we need to minimize $f(s, c^*)$ which corresponds to the larger of the two variance expressions. Routine calculations yield: $c^*_{opt} = 1 + \sqrt{2} = 2.4142$ and $s_{opt} = [1 - e^{-c^*} + c^{*2}e^{-c^*}]^{-1} = 0.6984$ and, hence, $\lambda_{opt} = 0.3016$.

Appendix

Proof of (0.31). Note first the following chain of equivalent inequalities:

$$e^{-c^*} < Pe^{-a^*} + Qe^{-b^*} \tag{A.1}$$

$$\Longleftrightarrow c^* > -\log(Pe^{-a^*} + Qe^{-b^*}) \tag{A.2}$$

$$\Longleftrightarrow \phi(c^*) > \phi(-\log(Pe^{-a^*} + Qe^{-b^*})) \tag{A.3}$$

$$\Longleftrightarrow w\phi(a^*) + (1-w)\phi(b^*) > \frac{e^{-\log(Pe^{-a^*}+Qe^{-b^*})} - 1}{-\log(Pe^{-a^*} + Qe^{-b^*})} \tag{A.4}$$

$$\Longleftrightarrow \frac{P(1-e^{-a^*}) + Q(1-e^{-b^*})}{Pa^*e^{-a^*} + Qb^*e^{-b^*}} > \frac{(1 - Pe^{-a^*} - Qe^{-b^*})}{[(Pe^{-a^*} + Qe^{-b^*})(-\log(Pe^{-a^*} + Qe^{-b^*}))]} \tag{A.5}$$

$$\Longleftrightarrow [(Pe^{-a^*} + Qe^{-b^*})(-\log(Pe^{-a^*} + Qe^{-b^*}))] > Pa^*e^{-a^*} + Qb^*e^{-b^*} \tag{A.6}$$

$$\Longleftrightarrow [1 + \frac{P}{Q}e^{b^*-a^*}][b^* - \log Q - \log(1 + \frac{P}{Q}e^{b^*-a^*})] > b^* + \frac{Pa^*}{Q}e^{b^*-a^*} \tag{A.7}$$

$$\Longleftrightarrow (Q + Pe^{b^*-a^*})\log(Q + Pe^{b^*-a^*}) < P(b^* - a^*)e^{b^*-a^*} \tag{A.8}$$

$$\Longleftrightarrow (Q + Pe^{x^*})\log(Q + Pe^{x^*}) < Px^*e^{x^*}; x^* = b^* - a^* > 0. \tag{A.9}$$

Define $\delta(x^*) = Px^*e^{x^*} - (Q + Pe^{x^*})\log(Q + Pe^{x^*})$, $x^* > 0$. It is easy to verify that $\delta(x^*)$ is increasing in x^* since $x^* = \beta^* x$ and β^* are positive. Further, $\delta(0) = 0$. Hence, the result follows.

References

Abdelbasit, K. M. and Plackett, R. L. (1983). Experimental design for binary data. *Journal of the American Statistical Association* **78**, 90–98.

Agrawal, H. L. (1966). Some systematic methods of construction of designs for two-way elimination of heterogeneity. *Calcutta Statistical Association Bulletin* **15**, 93–108.

Atkinson, A. C. and Cook, R. D. (1995). D-optimum designs for heteroscedastic linear models. *Journal of the American Statistical Association* **90**, 204–212.

Bagchi, S. and Shah, K. R. (1989). On the optimality of a class of row-column designs. *Journal of Statistical Planning and Inference* **23**, 397–402.

Ford, I., Torsney, B. and Wu, C. F. J. (1992). The use of a canonical form in the construction of locally optimal designs for non-linear problems. *Journal of the Royal Statistical Society,Series* **54**, 569–583.

Heiligers, B. and Sinha, Bikas K. (1995). Optimality aspects of Agrawal designs - Part II. *Statistica Sinica* **5**, 599–604.

Hedayat, A. S. and Raghavarao, D. (1975). 3-way BIB designs. *Journal of Combinatorial Theory, Series A* **18**, 207–209.

Hedayat, A. S., Yan, B. and Pezzuto, J. M. (1997). Modeling and identifying optimum designs for fitting dose-response curves based on raw optical density data. *Journal of the American Statistical Association* **92**, 1132–1140.

Khan, M. K. and Yazdi, A. A. (1988). On D-optimal designs for binary data. *Journal of Statistical Planning and Inference* **18**, 83–91.

Liski, E. P. Mandal, N. K. Shah, K. R. and Sinha, B. K. (2002) : Topics in Optimal Design. Springer-Verlag Lecture Notes in Statistics.

Mathew, T. and Sinha, Bikas K. (2001). Optimal designs for binary data under logistic regression. *Journal of Statistical Planning and Inference* **93**, 295–307

Minkin, S. (1987). Optimal design for binary data. *Journal of the American Statistical Association* **82**, 1098–1103.

Minkin, S. (1993). Experimental design for clonogenic assessment in chemotherapy. *Journal of American Statistical Association* **88**, 410–420.

Pukelsheim, F. (1993). *Optimal design of experiments.* Wiley, New York.

Saharay, R. (1996). A class of optimal row-column designs with some empty cells. *Statistica Sinica* **6**, 989–996.

Sebastiani, P. and Settimmi, R. (1997). A note on D-optimal designs for a logistic regression model. *Journal of Statistical Planning and Inference* **59**, 359–368.

Shah, K. R. and Sinha, Bikas K. (1990). Optimality aspects of Agrawal designs. *Gujarat Statistical Review: Professor C. G. Khatri Memorial Volume* **7**, 214–222.

Shah, K. R. and Sinha, B. K. (1989). Theory of Optimal Designs. Springer-Verlag Lecture Notes in Statistics, Vol. 53.

Shah, K. R. and Sinha, Bikas K. (1996). Row-column designs. In *Handbook of Statistics* **13** : Design and Analysis of Experiments. Ed. S. Ghosh and C. R. Rao. North Holland, 903–938.

Sitter, R. R. and Wu, C. F. J. (1993). Optimal designs for binary response experiments: Fieller-, D, and A criteria. *Scandinavian Journal of Statistics* **20**, 329–342.

Wu, C. F. J. (1988). Optimal design for percentile estimation of a quantal response curve. In *Optimal Design and Analysis of Experiments*, Eds. Y. Dodge, V. Federov and H. P. Wynn. Elsevier, Amsterdam, 213–223.