Statistics and Applications (ISSN 2452-7395 (online))
Volume 18, No. 2, 2020 (New Series), pp 207 – 221

Bayesian Mixture Designs for Hypothesis Testing

N.K. Mandal[1] and Manisha Pal[2]
[1]*Former Professor, Department of Statistics, University of Calcutta*
[2]*Department of Statistics, University of Calcutta*

Received: 07 June 2020; Revised: 22 July 2020; Accepted: 28 July 2020

Abstract

Scheffé (1958) first introduced models of different degrees to represent the response function in a mixture experiment. He also introduced corresponding designs for the estimation of the model parameters. Later, several authors studied the problem of finding optimum designs, especially for quadratic and cubic models. In this paper, we consider the problem of designing an optimal experiment for the purpose of performing one or more hypothesis tests in a first-degree mixture model. The Bayesian decision theoretic approach is used for this purpose.

Key words: Mixture experiments; Hypothesis testing; Bayes optimality; Normal prior; Optimum designs.

0. Tribute to Professors Sinha

We feel privileged to be able to contribute to this special issue of Statistics and Applications. We are fortunate to have come in contact with the highly acclaimed statisticians, the Sinha brothers, especially Professor Bikas K. Sinha, with whom we share a very close bond. We have been working with Professor Bikas K. Sinha, whom we fondly call Bikas Da, for more than a decade. His advice, support, positive attitude and, above all, his unbounded energy have been highly inspirational to us.

The first author had the good fortune of being taught by Bikas Da in the Statistics post-graduate course of Calcutta University in the mid seventy's, and, over time, has developed a close brotherly bond with him. The second author feels grateful for the constant encouragement that she has received from him, and for her enhanced knowledge of DoE through her association with him.

1. Introduction

Literature on Bayesian optimal design is generally based on linear models and the loss functions are chosen so as to be appropriate for estimation of the unknown parameters, and also for prediction purpose. See, for example, Chaloner (1982, 1984), El-Krunz and Studden (1991), Chaloner and Verdinelli (1995), Dasgupta (1996). Keeping in mind that the data analyst may also be interested in testing hypotheses regarding the parameters, Toman (1996)

Corresponding Author: Manisha Pal
E-mail: manishapal2@gmail.com

attempted to find optimum designs for hypotheses testing, using a suitable loss function in the discrete set-up. He considered both the cases of single hypothesis and multiple hypotheses testing. In the former case, the optimal design minimizes the Bayes risk, while in the latter case, where more than one decision is to be made, Toman (1996) suggested two approaches – one to minimize a Bayes risk while the other risks are constrained to be less than specified values, and the other to minimize the weighted sum of the Bayes risks, the weights being suitably selected by the experimenter. No further studies along this line have come to our notice. It is also noteworthy that the problem of determining the optimum Bayesian designs for hypotheses tests in the mixture set-up has not been addressed so far.

In this paper, we have considered the first-degree homogeneous mixture model due to Scheffé (1958). With suitably defined loss function, we have determined the Bayes optimal designs for testing both single and multiple hypotheses. The paper has been organized as follows. In Section 2 we have discussed the loss function and the Bayes risk for a linear model. In Section 3 we have obtained Bayes optimal designs for single hypotheses regarding the parameters of the model. Bayes designs for multiple hypotheses are discussed in Section 4. In Section 5, examples have been cited for multiple tests, and a discussion on the article has been given in Section 6.

2. The Loss Function and the Bayes Risk

Consider the linear model $\boldsymbol{Y} \sim N_p(\boldsymbol{X\theta}, \sigma^2 I_p)$, where $\boldsymbol{Y}$ is the vector of observations, $\boldsymbol{X}$ is the design matrixof order $p\times q$ and $\boldsymbol{\theta}$ is the parameter vector of order $q\times 1$. Suppose the prior distribution of $\boldsymbol{\theta}$ is $N_q(\boldsymbol{\tau}, \sigma^2 R)$, where R is a positive definite matrix.

Then, the posterior distribution of $\boldsymbol{\theta}$ is $N_q(\mu, V)$, where

$$\begin{aligned}
\mu &= \hat{\boldsymbol{\theta}} - (M + R^{-1})^{-1} R^{-1}(\hat{\boldsymbol{\theta}} - \boldsymbol{\tau}) \\
V &= \sigma^2 (M + R^{-1})^{-1} \\
M &= \boldsymbol{X'X},
\end{aligned}$$

and $\widehat{\boldsymbol{\theta}}$ is the least squared estimator of $\boldsymbol{\theta}$.

Suppose that one is interested to test the following k hypotheses:

H_{0i}: $\boldsymbol{c}_i'\boldsymbol{\theta} \geq v_i$ versus H_{1i}: $\boldsymbol{c}_i'\boldsymbol{\theta} < v_i$; $i = 1, 2, \ldots, k$,

where $\boldsymbol{c}_i$ and v_i, $i = 1, 2, \ldots, k$ are specified.

Let $\boldsymbol{d} = (d_1, d_2, \ldots, d_k)$ be the decision vector corresponding to a given $\boldsymbol{\theta}$, where d_i denotes the decision for the i-th testing problem $i = 1, 2, \ldots, k$. Let us denote by a_{i1} the action favoring $\boldsymbol{c}_i'\boldsymbol{\theta} \geq v_i$ and by a_{i2} that favoring $\boldsymbol{c}_i'\boldsymbol{\theta} < v_i$ in the i-th problem.

DeGroot (1970) showed the following loss function to be appropriate for the i-th problem in the general set-up:

$L_i(\boldsymbol{\theta}, a_{i1}) = 0, \quad$ if $\boldsymbol{c}_i'\boldsymbol{\theta} \geq v_i$

$= v_i - \boldsymbol{c}_i'\boldsymbol{\theta}$, if $\boldsymbol{c}_i'\boldsymbol{\theta} < v_i$,

and

$L_i(\boldsymbol{\theta}, a_{i2}) = \boldsymbol{c}_i'\boldsymbol{\theta} - v_i$, if $\boldsymbol{c}_i'\boldsymbol{\theta} \geq v_i$

$= 0$, if $\boldsymbol{c}_i'\boldsymbol{\theta} < v_i$, for $i = 1\ (1)\ k$.

Let $\delta^i(\boldsymbol{y})$ be the Bayes decision rule for the i-th problem. The Bayes risk is obtained by averaging the losses over both $\boldsymbol{Y}$ and $\boldsymbol{\theta}$.

Given a design with information matrix M, the Bayes risk $r_i(\delta^i, M)$ for the i^{th} testing problem is given as (*cf.* DeGroot, 1970):

$$r_i(\delta^i, M) = \sigma(\boldsymbol{c}_i'R\,\boldsymbol{c}_i)^{1/2}\,\{\Psi(s_i) - (1-p_i)^{1/2}\,\Psi(s_i/(1-p_i)^{1/2})\}, \tag{1}$$

where $\psi(s) = \varphi(s) - s[1-\Phi(s)]$, $\varphi(s)$ and $\Phi(s)$ are the density and cumulative distribution functions respectively of a standard normal variate, $p_i = \boldsymbol{c}_i'(R^{-1}+M)^{-1}\boldsymbol{c}_i/(\boldsymbol{c}_i'R\,\boldsymbol{c}_i)$ is the ratio of the posterior variance to the prior variance of $\boldsymbol{c}_i'\boldsymbol{\theta}$, and $s_i = (v_i - \boldsymbol{c}_i'\boldsymbol{\tau})/\sigma(\boldsymbol{c}_i'R\,\boldsymbol{c}_i)^{1/2}$, the standardized difference between the constant v_i and the prior mean of $\boldsymbol{c}_i'\boldsymbol{\theta}$. For the single hypothesis case, where the hypotheses are H_0: $\boldsymbol{c}'\boldsymbol{\theta} \geq v$ against H_1: $\boldsymbol{c}'\boldsymbol{\theta} < v$, the optimum design is selected so as to minimize the Bayes risk.

Now, $\frac{\partial}{\partial p} r(\delta, M) = 2\sigma(\boldsymbol{c}'R\boldsymbol{c})^{1/2}(1-p)^{1/2}\varphi(s/(1-p)^{1/2})$, which is positive, whatever be s. Hence, for every s, the risk function $r(\delta, M)$ is increasing in p. So, minimization of $r(\delta, M)$ can be achieved through minimization of p, which does not involve s, and hence the prior mean. Again, since $p = \boldsymbol{c}_i'(R^{-1}+M)^{-1}\boldsymbol{c}_i/(\boldsymbol{c}_i'R\,\boldsymbol{c}_i)$, and the denominator is free from the design, the Bayes ψ-optimal design will be obtained by minimizing

$$\phi(M) = \text{Trace}\,[\boldsymbol{cc}'(R^{-1} + M)^{-1}]. \tag{2}$$

We consider Scheffé's first order mixture model and work in a continuous design setting.

3. Optimal Mixture Design for a Single Test

Consider the mixture model given below:

$$\eta_{\boldsymbol{x}} = E(Y \mid \boldsymbol{x}) = \sum_{i=1}^{q} \beta_i x_i, \tag{3}$$

where Y denotes the response and $\boldsymbol{x} = (x_1, x_2, \ldots, x_q)$ the mixing proportions of the ingredients. The experimental region is $\Xi = \{(x_1, x_2, \ldots, x_q) : x_i \geq 0, i = 1(1)q, \sum_{i=1}^{q} x_i = 1\}$.

Here, one may be interested in single tests of the following form:

(I) H_0: $\beta_i \geq 0$ versus H_1: $\beta_i < 0$, for some i, $1 \leq i \leq q$

(II) H_0: $\beta_i - \beta_j \geq 0$ versus H_1: $\beta_i - \beta_j < 0$, for some i, j, $1 \leq i < j \leq q$

(III) H_0: $c'\boldsymbol{\beta} \geq 0$ versus H_1: $c'\boldsymbol{\beta} < 0$, where c is a $q \times 1$ real vector.

Before proceeding further, we note an important property of $\phi(M)$, defined in (2).

Property 1: $\phi(M)$, given by (2), is convex in M.

Proof: Let us write

$$\phi(M) = \text{Trace}\,[\boldsymbol{cc}'(R^{-1}+M)^{-1}] = [\boldsymbol{c}'(R^{-1}+M)^{-1}\boldsymbol{c}].$$

We have to show that $\phi(\lambda M_1 + (1-\lambda)M_2) \le \lambda\phi(M_1) + (1-\lambda)\phi(M_2)$.
Now, $\phi(\lambda M_1 + (1-\lambda)M_2) = [\boldsymbol{c}'(R^{-1}+(\lambda M_1+(1-\lambda)M_2))^{-1}\boldsymbol{c}]$

$$= [\boldsymbol{c}'(\lambda(R^{-1}+M_1) + (1-\lambda)(R^{-1}+M_2))^{-1}\boldsymbol{c}]. \qquad (4)$$

But it is known that

$$[\lambda(R^{-1}+M_1) + (1-\lambda)(R^{-1}+M_2)]^{-1} \le \lambda(R^{-1}+M_1)^{-1} + (1-\lambda)(R^{-1}+M_2)^{-1} \qquad (5)$$

(*cf.* Fedorov (1972)).

Hence from (4) and (5), we have

$$\phi(\lambda M_1 + (1-\lambda)M_2) \le \lambda\boldsymbol{c}'((R^{-1}+M_1)^{-1}\boldsymbol{c} + (1-\lambda)\,\boldsymbol{c}'(R^{-1}+M_2)^{-1}\boldsymbol{c})$$
$$= \lambda\phi(M_1) + (1-\lambda)\,\phi(M_2),$$

which establishes the convexity of $\phi(M)$. □

For the problem of minimizing $\phi(M)$, given by (2), a necessary and sufficient condition for a design to be Bayes optimal is obtained by applying the generalized equivalence theorem (*cf.* Whittle (1973); Silvey (1980)), which gives:

Theorem1: Any one of the following 3 conditions is necessary and sufficient for a design ξ_0 with information matrix M_0, to be optimal:

(i) $F_\varphi(R^{-1}+M_0, R^{-1}+M) \ge 0$ for all $R^{-1} + M$, for all $M \in \mathsf{M}$
(ii) $F_\varphi(R^{-1}+M_0, R^{-1}+\boldsymbol{xx}' \ge 0$ for all $\boldsymbol{x}\in\Xi$
(iii) $\min\limits_{\boldsymbol{x}\in\Xi}[F_\varphi(R^{-1}+M_0, R^{-1}+\boldsymbol{xx}')] = \max\limits_{R^{-1}+\boldsymbol{xx}'\in \boldsymbol{S}}\ \min\limits_{\boldsymbol{x}\in\Xi}[F_\varphi(R^{-1}+M, R^{-1}+\boldsymbol{xx}')]$,

where $F_\varphi(M_1, M_2)$ is the Fréchet derivative of M_1 in the direction M_2, M is the class of all information matrices, Ξ is the domain of $\boldsymbol{x}$ and $\boldsymbol{S}$ is the class of all non-singular matrices. Further, if $M_0 = \sum v_i\, \boldsymbol{x}_i\boldsymbol{x}_i'$, where $\boldsymbol{x}_i \in \Xi$ is the i-th design point of M_0, with mass $v_i > 0$, $i = 1,\ldots,m$, such that $\sum v_i = 1$, then for each i,

$$F_\varphi(R^{-1}+M_0, R^{-1}+\boldsymbol{x}_i\boldsymbol{x}_i') = 0.$$

The proof of Theorem 1 follows along the lines of the proof in Silvey (1980), pages 19- 23. Condition (ii) of Theorem 1 reduces to condition (ii)′ given below:

(ii)′ If ξ_0 be the optimal design with information matrix $M_0 = \sum_{i=1}^{m} v_i\,\boldsymbol{x}_i\boldsymbol{x}_i'$, where $\boldsymbol{x}_i\in\Xi$ is the i-th support point of ξ_0 with mass $v_i \ge 0$, $i = 1, \ldots, m$, and $\sum v_i=1$, then, for any $\boldsymbol{x}\in\Xi$,

$$\boldsymbol{x}'(R^{-1}+M_0)^{-1}\,\psi(R^{-1}+M_0)^{-1}\boldsymbol{x} \;\le \boldsymbol{x}_i'(R^{-1}+M_0)^{-1}\,\psi\,(R^{-1}+M_0)^{-1}\boldsymbol{x}_i\,, \tag{6}$$

where $\psi = \boldsymbol{cc}'$. It is obvious that equality in (6) holds at the support points of ξ_0.

Inequality (6) helps to identify the nature of the support points of the optimum design. For single hypothesis testing, the left-hand side of (6) is pseudo-convex in $\boldsymbol{x}$, that is, it behaves like a convex function with respect to finding its local minima, but may not actually be convex. As such, the maximum may be attained at the boundary points of Ξ and also at some non-boundary points. This is true irrespective of the form of R.

3.1. Consider problem I

For simplicity sake, let us take $i = 1$ in the above hypothesis testing, so that $\boldsymbol{c}' = (1, 0, 0,\ldots,0)$. Then a Bayes optimal design minimizes $\phi(M) = \boldsymbol{c}'\,(R^{-1} + M)^{-1}\boldsymbol{c}$, for the given $\boldsymbol{c}$.

To find a closed form solution, let us assume that the prior matrix R is invariant with respect to $\beta_2, \beta_3, \ldots,\beta_q$, so that R^{-1} can be written

$$R^{-1} = \begin{bmatrix} r^{11} & r^{12}\mathbf{1}'_{q-1} \\ & R^{22} \end{bmatrix},$$

where $R^{22} = u_1 I_{q-1} + u_2 J_{q-1}$, for some scalars u_1 and u_2, $\mathbf{1}_{q\text{-}1}$ is a $(q\text{-}1)$-vector with all elements 1, and $J_{q-1} = \mathbf{1}_{q\text{-}1}\mathbf{1}'_{q\text{-}1}$. Then, from Property 2 below, the criterion function ϕ will be invariant with respect to $x_2,\ldots,x_q$.

Property2: If M be invariant with respect to the components of $\boldsymbol{x}_{(1)} = (x_1,\ldots,x_{q_1})$ and $\boldsymbol{x}_{(2)} = (x_{q_1+1},\ldots,x_q)$, and R be invariant with respect to the components of $\boldsymbol{\beta}_{(1)} = (\beta_1,\ldots,\beta_{q_1})$ and $\boldsymbol{\beta}_{(2)} = (\beta_{q_1+1},\ldots,\beta_q)$, then ϕ is invariant with respect to the components of $\boldsymbol{x}_{(1)}$ and $\boldsymbol{x}_{(2)}$.

Proof: Consider P_1 and P_2 to be two permutation matrices of orders q_1 and q_2, respectively $(q_1 + q_2 = q)$. Let,

$$P = \begin{bmatrix} P_1 & \mathrm{O} \\ \mathrm{O} & P_2 \end{bmatrix}.$$

Then, P is a permutation matrix of order q.

Now, if we take the same permutations of the components of both $\boldsymbol{x}_{(1)}$ and $\boldsymbol{\beta}_{(1)}$ and similarly of the components of $\boldsymbol{x}_{(2)}$ and $\boldsymbol{\beta}_{(2)}$, then it is clear that, for the new set, the M and R matrices will reduce to PMP' and PRP'. Therefore, the criterion function will reduce to $\boldsymbol{c}'P'P(R^{-1}+M)^{-1}P'P\boldsymbol{c}$, which is same as $\boldsymbol{c}'(R^{-1}+M)^{-1}\boldsymbol{c}$. This establishes Property 2.

Using properties 1 and 2, we get the following theorem:

Theorem 2: A Bayes-optimal design is invariant with respect to $x_2,\ldots,x_q$.

Let us denote the class of all designs invariant with respect to $(x_2,\ldots,x_q)$ by D_1. For a design $\xi \in D_1$, $M(\xi)$ is of the form

$$M = \begin{bmatrix} m_{11} & m_{12}\mathbf{1}' \\ & M_{22} \end{bmatrix}, \tag{7}$$

where $M_{22} = g\, I_{q\text{-}1} + h\, J_{q\text{-}1}$, for some scalars g and h. Then,

$$[\boldsymbol{c}'\,(R^{-1} + M)^{-1}\boldsymbol{c}]^{-1} = (m_{11} + r^{11}) - (m_{12} + r^{12})^2\{\mathbf{1}'(M_{22} + R^{22})^{-1}\mathbf{1}\}. \tag{8}$$

From the structure of R^{22}, it is clear that $(R^{22})^{-1}$ will also be of the form $aI_{q-1} + bJ_{q-1}$, where $J_{q-1} = \mathbf{1}_{q\text{-}1}\mathbf{1}'_{q\text{-}1}$.

The following theorem indicates the Bayes optimal design under certain restriction.

Theorem 3: The Bayes optimal design is a singular design with only one support point at (1,0,0,…,0) *for* $0 \le |r^{12}|[a + (q-1)b] \le 1$.

Proof: Let ξ_0 be the singular design with one support point at (1,0,0,…,0). Then,

$$M(\xi_0) = \begin{bmatrix} 1 & \mathbf{0}' \\ \mathbf{0} & \mathbf{O} \end{bmatrix},\ [\boldsymbol{c}'\,(R^{-1} + M(\xi_0))^{-1}\boldsymbol{c}]^{-1} = (1 + r^{11}) - (r^{12})^2[\mathbf{1}'(R^{22})^{-1}\mathbf{1}],$$

$$\boldsymbol{x}'(R^{-1} + M(\xi_0))^{-1}\boldsymbol{c} = A[x_1 - \boldsymbol{x}'_{(2)}r^{12}(R^{22})^{-1}\mathbf{1}],$$

where $A = [(1 + r^{11}) - (r^{12})^2\{\mathbf{1}'(R^{22})^{-1}\mathbf{1}\}]^{-1}$, $\boldsymbol{x} = (x_1, x_2, \ldots, x_q)' = (x_1, \boldsymbol{x}'_{(2)})' \in \Xi$.

For ξ_0 to be Bayes optimal design, it must satisfy (3.2) for all $\boldsymbol{x} \in \Xi$, with equality holding at the support point of ξ_0. This is equivalent to satisfying

$$[x_1 - \boldsymbol{x}'_{(2)}r^{12}(R^{22})^{-1}\mathbf{1}]^2 \le 1, \text{ for all } \boldsymbol{x} \in \Xi, \tag{9}$$

with equality holding at (1,0,0,…,0).

For $r^{12} = 0$, (9) holds trivially. For $r^{12} > 0$, by the condition of the theorem, we have

$$\boldsymbol{x}'_{(2)}r^{12}(R^{22})^{-1}\mathbf{1} = r^{12}[a + (q-1)b]\boldsymbol{x}'_{(2)}\mathbf{1} \le 1,$$

since $\boldsymbol{x}'_{(2)}\mathbf{1} \le 1$. Hence, (9) is satisfied.

For $r^{12} < 0$, we can write

$$\text{l.h.s. of (9)} = [x_1 + \boldsymbol{x}'_{(2)}\,|\,r^{12}\,|\,(R^{22})^{-1}\mathbf{1}]^2$$
$$= [1 - \boldsymbol{x}'_{(2)}\{I - |\,r^{12}\,|\,(R^{22})^{-1}\}\,\mathbf{1}]^2,$$

Now,

$$1-\boldsymbol{x}'_{(2)}[I-|r^{12}|(R^{22})^{-1}]\mathbf{1} = 1\text{-}[1-|r^{12}|\{a+(q-1)b\}]\boldsymbol{x}'_{(2)}\mathbf{1} \leq 1,$$

as $\boldsymbol{x}'_{(2)}\mathbf{1} \leq 1$ and from the condition of the theorem. Hence, (9) is satisfied.

Remark: For $|r^{12}|[a+(q-1)b] > 1$, condition (9) is violated for $\boldsymbol{x}$ with $0 \leq x_1 \leq \frac{|r^{12}|[a+(q-1)b]}{1+|r^{12}|[a+(q-1)b]}$, when $r^{12} > 0$, and for all $\boldsymbol{x}$, when $r^{12} < 0$. Hence, the singular design with support point (1,0,…,0) will not be optimal.

It is difficult to analytically find the support points of the Bayes optimal design when $|r^{12}|[a+(q-1)b] > 1$. We, therefore, obtain the same through computation in the following example with $q = 3$:

Example 1: Suppose $\sigma = 1$ and the prior covariance matrix of the regression coefficients in Scheffé's first order model for $q = 3$ is

$$\text{R} = \begin{bmatrix} 0.70 & 0.75 & 0.75 \\ 0.75 & 3.00 & 1.00 \\ 0.75 & 1.00 & 3.00 \end{bmatrix}.$$

Consider testing of the hypothesis H0: $\beta_1 \geq 0$ versus H1: $\beta_1 < 0$. We have

$$R^{-1} = \begin{bmatrix} 2.38806 & -0.44776 & -0.44776 \\ -0.44776 & 0.458955 & -0.04104 \\ -0.44776 & -0.04104 & 0.458955 \end{bmatrix}$$

and

$$R_{22}^{-1} = \begin{bmatrix} 2.196429 & 0.196429 \\ 0.196429 & 2.196429 \end{bmatrix}$$

Hence, $|r^{12}|[a+(q-1)b] = 1.07 > 1$.

Using MATLAB, we obtain the Bayes optimal design as having support points (1,0,0), (0,1,0) and (0,0,1) with masses 0.94, 0.03 and 0.03, respectively. Thus, even when $|r^{12}|[a+(q-1)b] > 1$, the example gives the optimum support points at (1,0,0), (0,1,0) and (0,0,1).

Remark: It is interesting to note that the design with support points (1,0,0) and (0,½,½) gives minimum risk for masses 0.94 and 0.06 at (1,0,0) and (0,½,½) respectively, and the risk is same as that obtained for the optimum design with support points at (1,0,0), (0,1,0) and (0,0,1).

3.2. Consider testing of hypothesis II

For simplicity sake, let us take i = 1 and j = 2in the above hypothesis testing, so that $\boldsymbol{c}'$ = (1,−1,0,...,0).

We assume that the prior parameter matrix R is invariant with respect to β_1 and β_2, and with respect to $\beta_3, \beta_4, \ldots, \beta_q$, so that R^{-1} can be written as

$$R^{-1} = \begin{bmatrix} R^{11} & R^{12} \\ R^{12'} & R^{22} \end{bmatrix},$$

where $R^{ii} = s_i I_2 + t_i J_2$, for some scalars s_i and t_i, i = 1, 2, are positive definite matrices, R^{11} and R^{22} are of orders 2×2 and $(q-2)\times(q-2)$ respectively, and $R^{12} = r_0 \begin{pmatrix} \mathbf{1}'_{q-2} \\ \mathbf{1}'_{q-2} \end{pmatrix}$, where r_0 is a scalar.

From property 2, we have the following:

Property 3: ϕ is invariant with respect to (x_1, x_2) and with respect to $(x_3, x_4, \ldots, x_q)$.

Using properties 1 and 3, we get the following theorem:

Theorem 4: A Bayes-optimal design is invariant with respect to (x_1, x_2), and with respect to $(x_3, \ldots, x_q)$.

Let D_2 denote the class of all designs invariant with respect to (x_1, x_2), and with respect to $(x_3, \ldots, x_q)$.

For a design $\xi \in \mathsf{D}_2$, $M(\xi)$ is of the form

$$M(\xi) = \begin{bmatrix} M_{11} & m_0 \begin{pmatrix} \mathbf{1}'_{q\text{-}2} \\ \mathbf{1}'_{q-2} \end{pmatrix} \\ m_0 \left(\mathbf{1}_{q\text{-}2}\ \mathbf{1}_{q-2} \right) & M_{22} \end{bmatrix}, \tag{10}$$

where $M_{11} = g_1 I_2 + h_1 J_2$ and $M_{22} = g_2 I_{q\text{-}2} + h_2 J_{q\text{-}2}$, for some scalars g_1, g_2, h_1 and h_2.

Then,

$$R^{-1} + M = \begin{bmatrix} (u_1 + g_1) I_2 + (v_1 + h_1) J_2 & [m_0 + r_0] \begin{pmatrix} \mathbf{1}'_{q\text{-}2} \\ \mathbf{1}'_{q-2} \end{pmatrix} \\ & (u_2 + g_2) I_2 + (v_2 + h_2) J_2 \end{bmatrix}. \tag{11}$$

Theorem 5: The Bayes optimal design is a singular design with two support points at (1,0,0,…,0) and (0,1,0,…,0), each with mass ½.

Proof: Let ξ_0 be the singular design with two support points at (1,0,0,…,0) and (0,1,0,…,0), each with mass ½. Then,

$$M(\xi_0) = \begin{bmatrix} \frac{1}{2}I_2 & \mathbf{0}' \\ \mathbf{0} & \mathbf{O}_{q\text{-}2} \end{bmatrix},$$

$$\boldsymbol{x}'(R^{-1} + M(\xi_0))^{-1}\boldsymbol{c} = \boldsymbol{x}'_{(1)}A\boldsymbol{c} + \boldsymbol{x}'_{(2)}B\boldsymbol{c},$$

where $\boldsymbol{x} = (\boldsymbol{x}'_{(1)}, \boldsymbol{x}'_{(2)})'$, $\boldsymbol{x}'_{(1)} = (x_1, x_2)$, $\boldsymbol{x}'_{(2)} = (x_3, \ldots, x_q)$, and A and B are 2×2 and (q-2)×2 matrices, respectively given by

$$A = [(R^{11} + \frac{1}{2})I_2) - R^{12}(R^{22})^{-1}R^{21}]^{-1} = [(u_1 + \frac{1}{2})I_2 + v_1 J_2 - r_0^2 \begin{pmatrix} 1'_{q-2} \\ 1'_{q-2} \end{pmatrix} (R^{22})^{-1}(1_{q-2}\, 1_{q-2})]^{-1}$$

$$= \frac{1}{(u_1 + 1/2)}\left[I_2 - \frac{v_1 - r_0^2 1'_{q-2}\{R^{22}\}^{-1}1_{q-2}}{u_1 + 1/2 + 2[v_1 - r_0^2 1'_{q-2}\{R^{22}\}^{-1}1_{q-2}]} J_2 \right]$$

$$B = -[R^{22} - R^{21}(R^{11})^{-1}R^{12}]^{-1}R^{21}(R^{11})^{-1} = -\frac{1}{u_2}\left[I_{q-2} - \frac{v_2 - r_0^2 1'_2\{R^{11}\}^{-1}1_2}{u_2 + (q-2)[v_2 - r_0^2 1'_2\{R^{11}\}^{-1}1_2]} J_{q-2} \right] \times R^{21}(R^{11})^{-1}.$$

For ξ_0 to be Bayes optimal design, it must satisfy (6) for all $\boldsymbol{x} \in \Xi$, with equality holding at the support point of ξ_0. Writing $\boldsymbol{c} = (\boldsymbol{c}'_{(1)}\ \ \mathbf{0}')'$, where $\boldsymbol{c}_{(1)} = (1\text{-}1)'$, and noting that $\mathbf{1}_2'\boldsymbol{c}_{(1)} = 0$, we have, after a little algebraic manipulation, that (6) is equivalent to

$$(x_1 - x_2)^2 \le 1, \text{ for all } \boldsymbol{x} \in \Xi, \tag{12}$$

with equality holding at (1,0,0,…,0) and (0,1,0,…,0).

Clearly, (12) holds for all $\boldsymbol{x} \in \Xi$, with equality at the support points of ξ_0. Thus, ξ_0 is the Bayes optimal design with

$$\phi(\xi_0) = \boldsymbol{c}'(R^{-1} + M(\xi_0))^{-1}\boldsymbol{c} = \boldsymbol{c}'_{(1)}A\boldsymbol{c}_{(1)} = \frac{2}{u_1 + 1/2}.$$

3.3. Consider testing of hypothesis III

The problem of finding optimum design for a general '$\boldsymbol{c}$' in a closed form seems difficult. For this, to find the optimum design, we have considered two specific choices of $\boldsymbol{c}$, namely, (i) $\boldsymbol{c}'$ = (1, 1, …, 1), and (ii) $\boldsymbol{c}' = (c_1, c_2, \ldots, c_q)$, where c_i = –1 or +1, such that $\sum_{i=1}^{q} c_i = 0$, that is, q is even with q/2 of the c_i's equal to +1 and the rest –1. .

3.3.1. Let $\boldsymbol{c}' = (1, 1, \ldots, 1)$.

Since the hypothesis is invariant with respect to the q components, to start with, let us assume that R also has the same invariance property, that is, $R = a\,I + b\,J$, for some scalars a, b, where I and J stand for an identity matrix and a matrix of 1's of appropriate order, respectively. This means that the prior dispersion matrix of $\boldsymbol{\beta}$ is invariant with respect to the coefficients β_i's. Then it is easy to check the following invariance property of the criterion function $\phi\{M(\xi)\}$.

Property 4: $\phi\{M(\xi)\}$ is invariant with respect to the permutation of the components of the mixture.

Because of the properties (1) and (4), we get the following Theorem:

Theorem 6: A Bayes-optimal design is necessarily invariant.

Thus, in view of Theorem 6, we can confine our search for the Bayes-optimal design within the class of ***invariant designs***.

There are three ways out to find the desired design:

(i) Express M in terms of the two moments $\mu_2 = E(x_1^{\,2})$ and $\mu_{11} = E(x_1 x_2)$ of the design, and then show that ϕ is decreasing in μ_{11}.
(ii) Use Lowener Order dominance to find ξ^* such that $M(\xi^*) \geq M(\xi)$ for every invariant design ξ [*cf.* Draper and Pukelsheim (1999)].
(iii) Use Equivalence Theorem to find the optimal design.

Approaches (i) and (ii) fail as soon as the complete symmetry property of the matrix R is violated. In general, it is difficult to find a closed form solution to the problem. However, it may be possible to indicate the nature of the support points of the optimal design using (iii). The following theorem identifies an optimal design satisfying the Equivalence Theorem.

Theorem 7: The Bayes optimal design is a saturated design with support points at (1,0,0,…,0) and its permutations, each with mass $1/q$.

Proof: Let ξ_0 be the saturated design with support points at (1,0,0,…,0), (0,1,0,…,0), …, (0,0,…,1), each with mass $1/q$. Then, $M(\xi_0) = \frac{1}{q} I_q$.

Since both R and M have the complete symmetry property, $(R^{-1} + M(\xi_0))^{-1}$ is also completely symmetric and is of the form $eI_q + fJ_q$, with $e + qf > 0$. Then,

$$\boldsymbol{x}'(R^{-1} + M(\xi_0))^{-1}\boldsymbol{c} = \boldsymbol{x}'[eI_q + fJ_q]\mathbf{1}_q = e + qf,$$

since $J_q = \mathbf{1}_q\mathbf{1}_q'$ and $\boldsymbol{x}'\mathbf{1}_q = 1$.

Thus, $\boldsymbol{x}'(R^{-1} + M(\xi_0))^{-1}\boldsymbol{c}$ is a constant, independent of $\boldsymbol{x}$, and therefore (6) is satisfied for all $\boldsymbol{x} \in \Xi$.

Hence, ξ_0 satisfies the Equivalence Theorem, and is, therefore, a Bayes optimal design.

3.1.2. Consider $q = 2k$, k a positive integer, and $\boldsymbol{c} = (c_1, c_2, \ldots, c_q)'$, with k of the $c_i = -1$ and remaining 1, so that $\sum_{i=1}^{q} c_i = 0$.

Let R be complete symmetric. Intuitionally, we feel that M will also be complete symmetric.

We start with a saturated design ξ_0 which has support points at the extreme points of **X,** each having mass $1/q$. Then, as before, $M(\xi_0) = \frac{1}{q} I_q$.

As $R^{-1}$is a complete symmetric and positive definite matrix, $(R^{-1} + M(\xi_0))^{-1}$will have the form $eI_q + fJ_q$, with $e > 0$, $e + qf > 0$. Then,

$$\boldsymbol{x}'(R^{-1} + M(\xi_0))^{-1}\boldsymbol{c} = \boldsymbol{x}'[eI_q + fJ_q]\boldsymbol{c} = e\,\boldsymbol{x}'\boldsymbol{c},$$

since $\boldsymbol{x}'\mathbf{1}_q = 1$ and $\mathbf{1}_q'\boldsymbol{c} = 0$.

Hence, for each support point of ξ_0, r.h.s. of (6) $= e^2$. The l.h.s. of (6) is $e\,\boldsymbol{x}'\boldsymbol{c}$ for all $\boldsymbol{x} \in \Xi$**,** which is clearly $\leq e^2$, and equality holds at the support points of ξ_0.

Hence, ξ_0 is Bayes optimal.

4. Optimal Mixture Design for Multiple Tests

For the first-degree mixture model (3), one may be interested in multiple tests of the form

H_{0i}: $\beta_i \geq 0$ versus H_{1i}: $\beta_i < 0$; $i = 1,2, \ldots, k$, $k \leq q$.

For a given design ξ or the corresponding moment matrix $M(\xi)$, let $r_i(\delta^i, M)$denote the Bayes risk for the i-th hypothesis, $i = 1, 2, \ldots, k$. So, now we have a vector of Bayes risks. In order to define a partial ordering of the designs in terms of the moment matrices (*cf.* Kiefer, 1959), we proceed as in Toman (1996), who uses an idea analogous to the classical decision theoretic concept of admissibility. Admissibility is defined with respect of the risk function in the classical decision theory, and not the Bayes risk. In the present case, the index i of the Bayes risk is treated as the parameter in classical risk.

The following definitions are due to Toman (1996).

Definition 1: A design ξ_1 is said to be r-better than design ξ_2 if $r_i(\delta^i, M(\xi_1)) \leq r_i(\delta^i, M(\xi_2))$ for $i = 1,2, \ldots, k$, with strict inequality for at least one i.

Definition 2: A design ξ is said to be r-admissible if there exists no r-better design.

From the above it is quite clear that r-admissibility is a desired property of any design.

Toman (1996) indicated two methods of determining r-admissible design under multiple optimality criteria. In the present set-up, they are as follows:

I. Minimize $r_k(\delta^k, M)$ subject to $r_i(\delta^i, M) \le a_i$, i = 1, 2, …, k – 1, where the index k and the scalars a_i, i = 1,2,.., k–1 are determined by the experimenter.

II. Define a single risk function by combining the k Bayes risks $r_i(\delta^i, M)$, i = 1, 2, …, k as follows:

$$r(\delta, M) = \sum_{i=1}^{k} w_i r_i(\delta^i, M), \tag{13}$$

where the weights $\{w_i\}$ represent a discrete probability measure on the index i so that $w_i \ge 0$, for i = 1, 2, …, k and $\sum_{i=1}^{k} w_i = 1$. The weights represent the relative importance of the corresponding decision problems. A design ξ with moment matrix $M(\xi)$ which minimizes (13) will be the r-admissible design (*cf.* Toman, 1999).

Method I can be equivalently written as:

Minimize $\boldsymbol{c}_k'(R^{-1}+M)^{-1}\boldsymbol{c}_k$, subject to $\boldsymbol{c}_i'(R^{-1}+M)^{-1}\boldsymbol{c}_i \le b_i$, i = 1, 2, …, k–1,

where b_i is some function of a_i. This is because, for each i, $r_i(\delta^i, M)$, given in (1), is an increasing function of $p_i = \dfrac{c_i'(R^{-1}+M)^{-1}c_i}{c_i' R c_i}$, and hence of $\boldsymbol{c}_i'(R^{-1}+M)^{-1}\boldsymbol{c}_i$. Thus, we have a constrained optimization problem, which yields an r-optimal design (*cf.* Theorem 5 in Toman, 1996).

Lemma 1below shows the equivalence of Methods I and II for some set of weights $\{w_i\}$:

Lemma 1: If M_0 minimizes the combined risk $r(\delta, M)$, given in (13), then it also minimizes $r_k(\delta^k, M)$, subject to the restrictions $r_i(\delta^i, M) \le a_{i0}$, for i = 1, 2, …, k–1, where $a_{i0} = r_i(\delta^i, M_0), i = 1,2,\ldots,k-1$.

The lemma is a consequence of the following lemma of Cook and Wong (1994):

Lemma 2: (Cook and Wong, 1994): For $\lambda \in (0,1)$, let ξ_λ maximize the functional $\phi(\xi \mid \lambda) = \lambda\phi_1(\xi) + (1-\lambda)\phi_2(\xi)$, and let $c_\lambda = \phi_1(\xi_\lambda)$, the primary design criterion evaluated at ξ_λ. Then ξ_λ maximizes $\phi_2(\xi)$ subject to the constraint $\phi_1(\xi) \ge c_\lambda$.

The Bayes risk $r(\delta, M)$ depends on the design only through the p_is, which give the ratios of the posterior and prior variances. Further, for any given $\boldsymbol{c}_i$, $p_i \to 0$ as the prior information matrix $R^{-1} \to 0$, provided M is nonsingular.

Approximating the Bayes risks $r_i(\delta^i, M_0), i = 1,2,\ldots,k,$ by a first-order Taylor series expansion around $p_i = 0$, we get

$$r_i(\delta^i, M_0) \cong L_i \boldsymbol{c}_i{}' (M + R^{-1})^{-1} \boldsymbol{c}_i,$$

where φ (.) denotes the standard normal density, and

$$s_i = \frac{v_i - \boldsymbol{c}_i{}' \boldsymbol{\tau}}{\sigma \sqrt{\boldsymbol{c}_i{}' R \boldsymbol{c}_i}}, \quad L_i = \frac{\sigma}{2\boldsymbol{c}_i{}' R \boldsymbol{c}_i} \phi(s_i), i = 1,2,\ldots,k.$$

Then, $r(\delta, M) = \sum_{i=1}^{k} w_i L_i \mathrm{c}_i{}' (M + R^{-1})^{-1} \mathrm{c}_i = \mathrm{Trace}[B(M + R^{-1})^{-1}],$ (14)

where $B = \sum_{i=1}^{k} w_i L_i \boldsymbol{c}_i \boldsymbol{c}_i{}'$.

It seems easier to study the problem of multiple hypotheses testing using the second criterion. Some examples are worked out in Section 5.

5. Examples of Multiple Hypotheses Testing

In this section we find Bayes optimal designs in two examples on multiple hypotheses testing.

Example 2: Consider Scheffé's homogeneous mixture model of first degree for a three-component mixture:

$$\eta_{\mathrm{x}} = E(Y \mid \mathrm{x}) = \sum_{i=1}^{3} \beta_i x_i,$$

Suppose σ = 1, and the prior mean and covariance matrix of the regression coefficients are

$$\tau = (5,5,0.1)', \; R = \begin{bmatrix} 5.00 & 0.25 & 0.25 \\ 0.25 & 5.00 & 0.25 \\ 0.25 & 0.25 & 5.00 \end{bmatrix}.$$

(a) Consider testing of the hypotheses H_{0i}: $\beta_i \geq 0$ against H_{Ai}: $\beta_i < 0$, for i = 1, 2, 3.

Using Method 2, we minimize Trace $B(M(\xi) + R^{-1})^{-1}$, given by (14). In the absence of any knowledge about the relative importance of the components, it may be assumed that component problems are equally important. We, therefore, take, $w_i = \frac{1}{3}$, for i = 1, 2, 3.

From the given data, we obtain, $s_1 = s_2 = -2.24$ and $s_3 = -0.045$. Hence, $L_1 = L_2 = 0.007, L_3 = 0.089,$ and $B = \mathrm{Diag}(0.007, 0.007, 0.089)$.

We restrict to the class of saturated designs. Within this class, we get min[Trace $B(M(\xi) + R^{-1})^{-1}$] = 0.045135, which is obtained for a design with mass 0.0874 at each of the support points (1,0,0) and (0,1,0), and mass 0.8252 at (0,0,1). Comparing this design with an alternative one, say a design which puts equal masses at its three design points, we get the Bayes risk as 0.064301, which is 14.2% more than the optimum Bayes risk.

(b) Now, suppose we are interested to test the hypotheses

$$H_{01}: \beta_1 - \beta_3 \geq 0 \text{ against } H_{A1}: \beta_1 - \beta_3 < 0$$
$$H_{02}: \beta_2 - \beta_3 \geq 0 \text{ against } H_{A2}: \beta_2 - \beta_3 < 0.$$

Here, $s_1 = s_2 = -1.58977$, $L_1 = L_2 = 0.005934$. Then,

$$B = 0.002867 \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ -1 & -1 & 2 \end{bmatrix}.$$

Restricting to the class of saturated designs, the optimum design puts mass 0.272 at each of the points (1,0,0) and (0,1,0), and 0.456 at the point (0,0,1), and the Bayes risk is 0.0205.

6. Discussion

This paper attempts to find Bayes optimal designs for testing of single and multiple hypotheses in Scheffé's homogeneous first-degree mixture model. Interestingly, under the hypotheses considered, the support points of the optimal designs are found to be at one or more of the extreme points of the experimental region. The study can be extended to other testing situations, and also to the cases of quadratic and cubic mixture models.

Acknowledgement

The authors thank the anonymous reviewer for reading the manuscript with great care and making fruitful suggestions, which immensely helped to improve the presentation.

References

Chaloner, K. (1982). *Optimal Bayesian Experimental Design for Linear Models*. Technical Report 238, Carnegie-Mellon University, Department of Statistics.

Chaloner, K.(1984). Optimal Bayesian experimental design for linear models. *Annals of Statistics*, **12(1)**, 283-300.

Chaloner, K. and Verdinelli, I. (1995). Bayesian experimental design: A review. *Statistical Science*, **10**, 273-304.

DasGupta, A. (1996). Review of optimal Bayes designs. *Handbook of Statistics*, **13**, 1099-1147.

Draper, N. R. and Pukelsheim, F. (1999). Kiefer ordering of simplex designs for first- and second-degree mixture models. *Journal of Statistical Planning and Inference*, **78**, 325-348.

El-Krunz, S. and Studden, W. J. (1991). Bayesian optimal designs for linear regression models. *The Annals of Statistics*, **19**, 2183-2208.

Fedorov, V. V.(1972). *Theory of Optimal Experiments*. Academic Press. New York.

Scheffé, H. (1958). Experiments with mixtures. *Journal of the Royal Statistical Society*, **B20**, 344-360.

Silvey, S. D. (1980). *Optimal Design*. Chapman and Hall, London.

Toman, B. (1996). Bayesian experimental design for multiple hypothesis testing. *Journal of the American Statistical Association*, **91**, 185-190.

Whittle, P. (1973). Some general points in the theory of optimal experimental design. *Journal of the Royal Statistical Society*, **B35**(**1**), 123-130.