Mathematics

http://www.aimspress.com/journal/Math

AIMS Mathematics, 7(3): 4540–4551.
DOI: 10.3934/math.2022253
Received: 29 September 2021
Revised: 15 December 2021
Accepted: 19 December 2021
Published: 23 December 2021

Research article

R-optimal designs for second-order Scheffé model with qualitative factors

Ling Ling[1], **Guanghui Li**[2], **Xiaoyuan Zhu**[1] **and Chongqi Zhang**[1,*]

[1] School of Economics and Statistics, Guangzhou University, Guangzhou 510006, China
[2] School of Science, Kaili University, Kaili 556011, China

* **Correspondence:** Email: cqzhang@gzhu.edu.cn.

Abstract: Considering a mixture model with qualitative factors, the *R*-optimal design problem is investigated when the levels of the qualitative factor interact with the mixture factors. In this paper, the conditions for *R*-optimality of designs with mixture and qualitative factors are presented. General analytical expressions are also derived for the decision function under the *R*-optimal designs, in order to verify that the resulting designs satisfy the general equivalence theorem. In addition, the relative efficiency of the *R*-optimal design is discussed.

Keywords: mixture experiments; qualitative factors; *R*-optimality; equivalence theorem
Mathematics Subject Classification: 62K05, 62K99

1. Introduction

In recent years, the mixture models have been widely used in sciences, engineering and business applications. Cornell [1] applied models for data from mixture experiments to chick feeding, paper coatings, and many other products. Zijlstra et al. [2] embed mixture models in a panel mixed logit framework to describe preferences and preference heterogeneity for the mobility budget. Goos and Hamidouche [3] applied the mixture models to a real-life cocktail experiment. In mixture experiments, the factors under study are proportions of the ingredients of a mixture. As the application involves the choice of cocktails, the taste of cocktails only depends on the ingredient proportions, and not on the total amount of cocktail.

To obtain the optimal design under certain optimality criteria has always been a hot topic in the research of the mixture design. Available optimality criteria include *D*-, *A*-, *R*-, and *I*-optimality, etc. (see [4, 5]). The *D*-optimality and *A*-optimality have been used most frequently. In practice, the efficiency of a specific optimality criterion depends on statistical problem sensitively, and in many cases, the *D*- and *A*-optimal designs are reasonable. However, there are also various situations where the application of the *D*- and *A*-optimality criteria is not appropriate. The efficiency loss may be considerable when the D-optimal design is used for the construction of rectangular confidence regions, as stated by Dette [6]. For this reason, Dette [6] proposed the R-optimality criterion, which minimizes the volume of the p-dimensional rectangular confidence region for the unknown parameters. Furthermore, the R-optimality criterion not only has an excellent statistical interpretation but also satisfies a beneficial invariance property. It allows an easy calculation of optimal designs on many linearly transformed design spaces. Hao et al. [7] investigated the R-optimal design for the second-order mixture model, while the qualitative factors are not considered.

When an experiment includes qualitative factors, the effects between the quantitative and qualitative factors should be taken into consideration. For example, Lee and Huang [8] demonstrated through an example on chemical study how the D-optimal design may help to design an experiment with both quantitative and qualitative factors more efficiently. Yue et al. [9] investigated the D-optimal designs for multiresponse polynomial regression models with both quantitative and qualitative factors. Kao and Hazar [10] investigated the optimal designs for mixed continuous and binary responses with quantitative and qualitative factors. Zhang et al. [11] proposed a new uniformity criterion for designs with both qualitative and quantitative factors. Although qualitative factors have received increasing attention in the literature, little is known concerning mixture designs with qualitative factors. Following the seminal work of Donev [12], the only published results concerning qualitative factors for mixture models are presented in [13], which generated the A-optimal designs for mixture central polynomial model with qualitative factors. This paper proposes a novel optimal design of mixture model with qualitative factors.

This paper aims at developing a R-optimal design for Scheffé models, when the response depends on the joint influence of both mixture and qualitative factors. The remainder of this paper is organized as follows. Section 2 and Section 3 briefly introduce the R-optimality, and some basic notations of the two kinds of mixture designs. The main results are presented in Section 4. Finally, the conclusion is drawn in Section 5.

2. Preliminaries

A mixture experiment design usually assumes that the response variable y is only related to the proportion of each component $x_1, \cdots, x_q$, but not to the total. The experimental region determined by the component proportions can be expressed as

$$\mathcal{X} = \left\{ \mathbf{x} = \left(x_1, x_2, \cdots, x_q\right) : \sum_{i=1}^{q} x_i = 1, x_i \geq 0, i = 1, 2, \cdots, q, C's \right\},$$

in which $C's$ are some other constraint conditions. When the model does not contain such constraints, it is called a $(q-1)$-dimensional simplex, denoted as S^{q-1}. The response at $\mathbf{x}$ can be written as

$$y = \beta^T f(\mathbf{x}) + \varepsilon, \tag{2.1}$$

where $\beta = (\beta_1, \beta_2, \cdots, \beta_m)^T$ is the m-vector of unknown parameters, $f(\mathbf{x}) = (f_1(\mathbf{x}), f_2(\mathbf{x}), \cdots, f_m(\mathbf{x}))^T$ is a given m-vector of regression functions of $\mathbf{x} = (x_1, x_2, \cdots, x_q)^T \in \mathcal{X} \subseteq S^{q-1}$. The error ε is assumed to be independent of $\mathbf{x}$ and independent identically distributed with a normal distributed with mean 0 and constant variance σ^2.

Based on the model (2.1), an approximate design can be expressed as

$$\xi = \begin{pmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_k \\ r_1 & r_2 & \cdots & r_k \end{pmatrix},$$

where $\mathbf{x}_i \in \mathcal{X}$ are support points and their weights $\xi(\mathbf{x}_i) = r_i$ satisfy $r_i > 0$ and $\sum_{i=1}^{k} r_i = 1, i = 1, 2, \cdots, k$. For a design ξ, the Fisher information matrix is given by

$$M(\xi) = \int_X f(x) f^T(x) \xi(dx).$$

Moreover, let Ξ denote the set of all designs with non-singular information matrix on $\mathcal{X}$.

Definition 1. *A design $\xi^* \in \Xi$ is called R-optimal for the model (2.1) if it minimizes*

$$\Phi_R(\xi) = \prod_{i=1}^{m} \left(M^{-1}(\xi)\right)_{ii}, \tag{2.2}$$

where m is the number of unknown parameters, $(M^{-1}(\xi))_{ii}$ is the element on the main diagonal of $M^{-1}(\xi)$.

For a given design, the equivalence theorem seeks to prove whether the design is the optimal design. Dette [6] investigated the equivalence theorem of R-optimal design.

Lemma 1. *For any $\mathbf{x} \in \mathcal{X}$, the design ξ^* is R-optimal if and only if*

$$\psi_R(\mathbf{x}, \xi) = \sum_{j=1}^{m} \frac{\left(e_j^T M^{-1}(\xi) f(\mathbf{x})\right)^2}{e_j^T M^{-1}(\xi) e_j} - m \leq 0, \tag{2.3}$$

where e_j is the unit vector whose j-th component is equal to 1 and all others are 0. Moreover, the equality is attained at the support points of the R-optimal design ξ^*.

3. Models and analysis

In the mixture experiments, a simplex-lattice design for q ingredients involves all possible mixture formulations, the set of q components n order lattice points can be expressed as

$$\mathcal{L}\{q, n\} = \left\{\left(\frac{\alpha_1}{n}, \frac{\alpha_2}{n}, \ldots, \frac{\alpha_q}{n}\right) : \sum_{i=1}^{q} \alpha_i = n, \alpha_i \in \mathbf{Z}^+, i = 1, 2, \ldots, q\right\},$$

and there are $\begin{pmatrix} q+n-1 \\ n \end{pmatrix}$ points in $\mathcal{L}\{q, n\}$.

The matrix $H(\mathbf{x}) = \left\{x_{i_j}\right\}_{i,j=1}^{N,q}$ generated by all permutations of the coordinates of $\mathbf{x} = \left(x_1, x_2, \cdots, x_q\right)^{\mathbf{T}} \in \mathcal{X}$ is called permutation matrix, where $i_1, i_2, \cdots, i_q$ is a permutation of the $1, 2, \cdots, q$. Then each row of $H(\mathbf{x})$ is a mixture point and $\mathcal{H}(\mathbf{x}) = \{\mathbf{x}, \mathbf{x}_1, \cdots, \mathbf{x}_{N-1}\}$ is called permutation point set about $\mathbf{x}$. The set of q components centroid points can be expressed as $C\{q\} = \bigcup_{i=1}^{q} \mathcal{H}(\mathbf{x}_i)$. A simplex-centroid design consists of $2^q - 1$ points: q permutations of $(1, 0, \cdots, 0)$, $\binom{q}{2}$ permutations of $(1/2, 1/2, 0, \cdots, 0)$, $\cdots$, and one permutation of $(1/q, 1/q, \cdots, 1/q)$. The corresponding model to be fitted to data at the points of the simplex-centroid design is

$$E(y) = \sum_{i=1}^{q} \beta_i x_i + \sum_{i=1}^{q-1} \sum_{i<j}^{q} \beta_{ij} x_i x_j + \cdots + \beta_{12\cdots q} x_1 x_2 \cdots x_q.$$

The Scheffé central polynomial model can reduce the number of trials without reducing the order of the model and maintaining prediction accuracy.

In many practical problems, however, the response is influenced not only by the components' proportion in the mixture, but also by other variables. We can consider the general model which was introduced by Lee and Huang [8], this type of model can be expressed as follows:

$$E[y(j, \mathbf{x})] = f_1^T(\mathbf{x})\beta_j + f_2^T(\mathbf{x})\gamma, \mathbf{x} \in \mathcal{X} \subseteq S^{q-1}, j = 1, 2, \cdots, s. \tag{3.1}$$

In the model, $f_1^T(\mathbf{x})\beta_j$ describes the part of the response that is influenced by a qualitative factor, which is generally referred to as the factorial effect, among them, $f_1^T(\mathbf{x}) = (f_{11}(\mathbf{x}), f_{12}(\mathbf{x}), \cdots, f_{1p_1}(\mathbf{x}))$, $\beta_j = (\beta_{j1}, \beta_{j2}, \cdots, \beta_{jp_1})^T$, s are the level of qualitative factor, $y(j, \mathbf{x})$ is the value of the response variable at design point $\mathbf{x} = (x_1, x_2, \cdots, x_q)^T \in \mathcal{X} \subseteq S^{q-1}$ at the jth level. $f_2^T(\mathbf{x})\gamma$ is the part of the response that is not affected by the qualitative factors, generally referred to as fixed effect, where $f_2^T(\mathbf{x}) = (f_{21}(\mathbf{x}), f_{22}(\mathbf{x}), \cdots, f_{2p_2}(\mathbf{x}))$, $\gamma = (\gamma_1, \gamma_2, \cdots, \gamma_{p_2})^T$.

Taking the second-order general Scheffé central polynomial model as an example, the model (3.1) can be expressed in the following three forms:

$$E\left[y(j, \mathbf{x})\right] = (f_{L_1}^T(\mathbf{x}), f_{L_2}^T(\mathbf{x}))\beta_j, j = 1, 2, \cdots, s; \tag{3.2}$$

$$E\left[y(j, \mathbf{x})\right] = f_{L_1}^T(\mathbf{x})\beta_j^{(L_1)} + f_{L_2}^T(\mathbf{x})\gamma^{(L_2)}, j = 1, 2, \cdots, s; \tag{3.3}$$

$$E\left[y(j, \mathbf{x})\right] = f_{L_2}^T(\mathbf{x})\beta_j^{(L_2)} + f_{L_1}^T(\mathbf{x})\gamma^{(L_1)}, j = 1, 2, \cdots, s. \tag{3.4}$$

Where $f_{L_1}(\mathbf{x}) = (x_1, x_2, \cdots, x_q)^T$ and $f_{L_2}(\mathbf{x}) = (x_1x_2, x_1x_3, \cdots, x_{q-1}x_q)^T$. Among them, both the primary and interaction terms in model (3.2) are affected by qualitative factors; while these terms in model (3.3) are factorial effects and fixed effects, respectively, and the case is quite the contrary in model (3.4), with the two terms being fixed and factorial effects, respectively.

For example, a 21-herb formula, referred to as the modified Qing Fei Pai Du Tang (mQFPD) for treating COVID-19, is now being tested in a clinical trial sponsored by the University of California, San Diego. The results obtained by these experiments define the efficacy of the drugs. In this case, the difference between the efficacies of the drugs on the gender of the patient can be considered as the first part of models (3.3) and (3.4), while the fixed efficacy of the drugs can be considered as the second part of the two models. Afterwards, the R-optimal design for models (3.3) and (3.4), is presented.

4. Main results

Denote the index set of the qualitative factor as $\mathcal{D}_s = \{1, 2, \cdots, s\}$, and denote $z = (j, \mathbf{x}) \in \Omega$, $\Omega = \mathcal{D}_s \times \mathcal{X}$ as the product set formed by the mixture experiment region and the qualitative factor.

Then the design on the experimental region Ω can be expressed as a product of two partial designs, i.e., $\zeta(j,\mathbf{x}) = \eta(j)\times\xi(\mathbf{x})$, where η is the design on $\mathcal{D}_s$ and $\xi(\mathbf{x})$ is the design on $\mathcal{X}$. Then ζ has the following form

$$\zeta = \begin{pmatrix} (1;\mathbf{x}_{11},\cdots,\mathbf{x}_{1n_1}) & (2;\mathbf{x}_{21},\cdots,\mathbf{x}_{2n_2}) & \cdots & (s;\mathbf{x}_{s1},\cdots,\mathbf{x}_{sn_s}) \\ (\eta(1);r_{11},\cdots,r_{1n_1}) & (\eta(2);r_{21},\cdots,r_{2n_2}) & \cdots & (\eta(s);r_{s1},\cdots,r_{sn_s}) \end{pmatrix}.$$

First note that the information matrix of the design ξ is

$$M_f(\xi) = \begin{bmatrix} M_{11}(\xi) & M_{12}(\xi) \\ M_{21}(\xi) & M_{22}(\xi) \end{bmatrix},$$

which is associated with the quantitative model $E[y(\mathbf{x})] = \left[f_1^T(\mathbf{x}), f_2^T(\mathbf{x})\right]\left(\beta^T,\gamma^T\right)^T$. Let the general model (3.1) be rewritten as

$$E\left[y(j,\mathbf{x})\right] = \left[e_j^T \otimes f_1^T(\mathbf{x}), f_2^T(\mathbf{x})\right]\left(\beta_1^T,\beta_2^T,\cdots,\beta_s^T,\gamma^T\right)^T = g^T(j,\mathbf{x})\theta. \tag{4.1}$$

By a result in Lee and Huang [8], the information matrix of the design ζ is

$$M_g(\zeta) = \begin{bmatrix} D\otimes M_{11}(\xi) & \eta\otimes M_{12}(\xi) \\ \eta^T\otimes M_{21}(\xi) & M_{22}(\xi) \end{bmatrix},$$

where $D = \text{diag}\left\{\eta(1),\eta(2),\cdots,\eta(s)\right\}, \eta = \left[\eta(1),\eta(2),\cdots,\eta(s)\right]^T$.

According to the R-optimality criterion, the minimization of $\prod_{i=1}^{m}(M_g^{-1}(\zeta))_{ii}$ amounts to minimization of $tr[log[\prod_{i=1}^{m}(M_g^{-1}(\zeta))_{ii}]]$ for all $\zeta\in\Xi$. The following theorem defines the conditions for R-optimality of designs and the proof is given in the Appendix.

Theorem 1. *For a design $\zeta(j,\mathbf{x}) = \eta(j)\times\xi(\mathbf{x})$, where η is the design on $\mathcal{D}_s = \{1,2,\cdots,s\}$ and ξ is the design on $\mathcal{X}\subseteq S^{q-1}$, the following equation holds for model (3.1):*

$$tr[log[\prod_{i=1}^{m}(M_g^{-1}(\zeta))_{ii}]] = \sum_{j=1}^{s} tr\left(\log\left(\frac{1}{\eta(j)}M_{11}^{-1}(\xi) + K_{(1)}\right)\right) + tr\left\{\log\left[D_{22}(\xi)\right]\right\},$$

where

$$K_{(1)} = M_{11}^{-1}(\xi)M_{12}(\xi)D_{22}(\xi)M_{21}(\xi)M_{11}^{-1}(\xi),$$
$$D_{22}(\xi) = \left[M_{22}(\xi) - M_{21}(\xi)M_{11}^{-1}(\xi)M_{12}(\xi)\right]^{-1}.$$

Moreover, it also follows that, for ζ to be optimal, all the elements of η must be equal, i.e. $\eta(1) = \eta(2) = \cdots = \eta(s) = 1/s$.

If the optimal marginal design η is the uniform design on $\mathcal{D}_s = \{1,2,\cdots,s\}$, then we have

$$tr[log[\prod_{i=1}^{m}(M_g^{-1}(\zeta))_{ii}]] = s\cdot tr\left(\log\left(s\cdot M_{11}^{-1}(\xi) + K_{(1)}\right)\right) + tr\left\{\log\left[D_{22}(\xi)\right]\right\}.$$

To verify the R-optimality of the designs by the equivalence theorem, the decision function presented in Eq (2.3) can be simplified to

$$\psi_R(\mathbf{x},\xi) = f^T(\mathbf{x})M^{-1}(\xi)BM^{-1}(\xi)f(\mathbf{x}) - m \leq 0,$$

where $B = \text{diag}(1/M_{11}, 1/M_{22}, \ldots, 1/M_{mm})$, $M_{ii} = (M^{-1}(\xi))_{ii}$.

Theorem 2. *For a design ζ on Ω, let b_{ii}, d_{jj} and c_{kk} are the element on the main diagonal of $M_{11}^{-1}(\xi)$, $D_{22}(\xi)$ and $K_{(1)}$, respectively, if η is a uniform design on $\mathcal{D}_s$, the decision function of the model (4.1) under the R-optimal design is*

$$\begin{aligned}\psi_{R-g}(j,\mathbf{x};\zeta) =& g^{\mathrm{T}}(j,\mathbf{x})M_g^{-1}(\zeta)B_gM_g^{-1}(\zeta)g(j,\mathbf{x})\\ =& s^2u(\mathbf{x};\xi)+s\cdot\sum_{i=1}^{4}v_i(\mathbf{x};\xi)+\sum_{i=1}^{4}\omega_i(\mathbf{x};\xi),\end{aligned}$$

where $B_g = \begin{bmatrix} I_s\otimes B(s,1) & 0\\ 0 & H\end{bmatrix}$, $B(u,v) = \operatorname{diag}\left(1/\left(ub_{11}+vc_{11}\right),1/\left(ub_{22}+vc_{22}\right),\cdots,1/\left(ub_{p_1p_1}+vc_{p_1p_1}\right)\right)$, $H=\operatorname{diag}\left(1/d_{11},1/d_{22},\cdots,1/d_{p_2p_2}\right)$.

In the following, the approximate R-optimal design on the model (3.3) and (3.4) are derived from the results of theorem 2. Suppose ζ^* is a R-optimal design on the region $\mathcal{D}_s\times S^{q-1}$ and denoted by $\eta^*\times\xi^*$, where η^* is the uniform design on $\mathcal{D}_s$. For the model (3.3), the design ξ over the mixture region is arranged by using the set of second-order generalized simplex-centroid points $C\{q,2\}$, denoted as

$$\xi=\begin{pmatrix}\mathcal{H}(\mathbf{x}_1) & \mathcal{H}(\mathbf{x}_2)\\ r_1 & r_2\end{pmatrix},$$

where $\mathbf{x}_1=(1,0,\cdots,0)^T$ and $\mathbf{x}_2=(1/2,1/2,\cdots,0)^T$ are the vertices and mid-points of the edges of the simplex, respectively. $\mathcal{H}(\mathbf{x}_i)=\{\mathbf{x}_{i1},\mathbf{x}_{i2},\cdots,\mathbf{x}_{in_i}\}$ is permutation point set about $\mathbf{x}_i$, $i=1,2$, and the weights of corresponding point satisfy $n_1r_1+n_2r_2=1$. Then the function $\psi_{R-g}(j,\mathbf{x};\zeta)$ can be expressed as

$$\begin{aligned}\psi_{R-g}(j,\mathbf{x};\zeta) =& \left(\frac{a_0+a_1}{b_0}+\frac{(q-1)\,r_2}{2r_1\,(2r_1+r_2)}\right)\sum_{i=1}^{q}x_i^2+\left(\frac{a_2}{b_0}+\frac{r_2}{r_1\,(2r_1+r_2)}\right)\sum_{i<j}x_ix_j\\ &-\left(\frac{a_3}{b_1}+2qr_2^2+8r_1r_2\right)\sum_{i\neq j}x_i^2x_j-\frac{12r_2}{r_1\,(2r_1+r_2)}\sum_{i<j<k}x_ix_jx_k\\ &+\frac{32r_1^2+32r_1r_2+4qr_2^2}{r_1r_2\,(2r_1+r_2)}\sum_{i<j}x_i^2x_j^2+\frac{32r_1+(4q+8)\,r_2}{r_1\,(2r_1+r_2)}\sum_{i\neq j\neq k}x_i^2x_jx_k,\end{aligned}$$

where

$$\begin{aligned}a_0 =& 16s^2r_1^4+8\,(2q-3)\,s^2r_1^3r_2+\left(4q^2-5q+8\right)s^2r_1^2r_2^2,\\ a_1 =& (q-1)\,s\left[8r_1^3r_2+(9q-16)\,r_1^2r_2^2+(q-2)\,(3q-4)\,r_1r_2^3+\frac{(q-2)^2\,(q-1)}{4}r_2^4\right],\\ a_2 =& s\,(1-s)\left[16r_1^3r_2+(6q-8)\,r_1^2r_2^2\right],\\ a_3 =& 8sr_1^2+(6q-8)\,sr_1r_2+(q-1)\,(q-2)\,sr_2^2,\\ b_0 =& b_1\left[4r_1^2+(3q-4)\,r_1r_2+\frac{(q-1)\,(q-2)}{2}r_2^2\right],\\ b_1 =& 4sr_1^3+(2sq+q-3s-1)\,r_1^2r_2+\frac{(q-1)\,(q-2)}{2}r_1r_2^2.\end{aligned}$$

Let $q = 3$, $s = 2$, $\Phi_R(\zeta, r_1) = \prod_{i=1}^{m}(M_g^{-1}(\zeta))_{ii}$, $h_i(r_1) = \psi_{R-g}(j, \mathbf{x}; \zeta), \mathbf{x} \in \mathcal{H}(\mathbf{x}_i), i = 1, 2$ is adopted to denote the values of the decision function at the two types of design points, respectively. If ζ^* is a R-optimal design, on the one hand there should be

$$r_1^* = \arg \min_{r_1 \in (0, 1/n_1)} \prod_{i=1}^{m}(M_g^{-1}(\zeta))_{ii},$$

on the other hand

$$h_i(r_1^*) = \psi_{R-g}(j, \mathbf{x}; \zeta^*) = sn_1 + n_2, \mathbf{x} \in \mathcal{H}\left(\mathbf{x}_i\right), i = 1, 2.$$

As r_1 transforms, $\Phi_R(\zeta, r_1)$ changes as shown in Figure 1(a), we can find the optimal design

$$\xi^* = \begin{pmatrix} \mathcal{H}(\mathbf{x}_1) & \mathcal{H}(\mathbf{x}_2) \\ 0.2269 & 0.1065 \end{pmatrix}.$$

Figure 1(b) shows the changes of $h_1(r_1)$ and $h_2(r_1)$, which can be seen when $r_1 = r_1^*$, $h_1(r_1)$ and $h_2(r_1)$ intersect at a point, whose coordinates are $\left(r_1^*, sn_1 + n_2\right) = (0.2269, 9)$. The contour map of the function $\psi_{R-g}(j, \mathbf{x}; \zeta^*)$ on the region Ω is shown in Figure 2. It is obvious to see that the maximum of the decision function is no greater than the number of model parameters and the maximum is attained at the support points of ξ^*. According to the Eq (2.3), the design ξ^* is R-optimal for model (3.3) and satisfies the equivalence theorem.

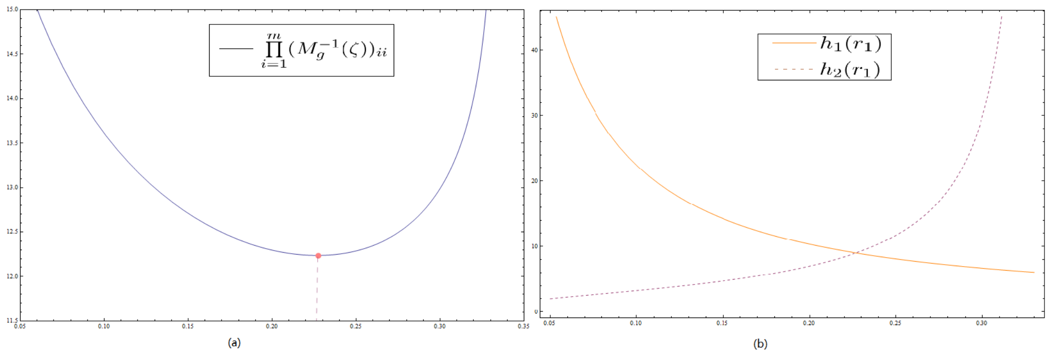

Figure 1. R-optimal design on model (3.3) with $q = 3$ and $s = 2$.

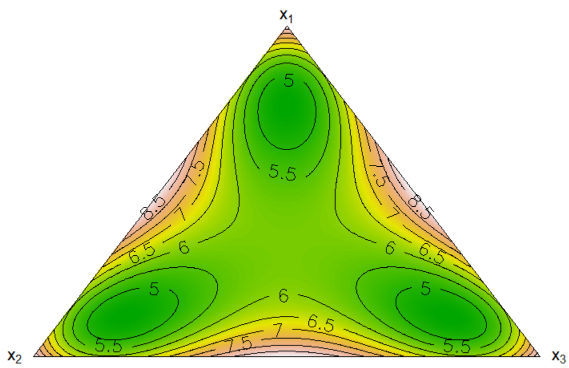

Figure 2. The contour map of decision function of ζ^*.

Table 1 lists the weights of R-optimal designs for model (3.3) and model (3.4) with $3 \leq q \leq 6$ and $2 \leq s \leq 6$. The corresponding values of the common logarithm of decision function, $lg[\Phi_R(M(\zeta))]$, are also listed in the fifth and eighth columns of the table. In model (3.3), for any q, the weights of vertices increase with the total number of level combinations. On the contrary, in model (3.4), the weights of vertices decrease with the increase of the total number of level combinations.

Table 1. The weights of R-optimal designs for $3 \leq q \leq 6$ and $2 \leq s \leq 6$.

		Model(3.3)			Model(3.4)		
q	s	$r_1(\mathbf{x}_1)$	$r_2(\mathbf{x}_2)$	$lg[\Phi_R(M(\zeta))]$	$r_1(\mathbf{x}_1)$	$r_2(\mathbf{x}_2)$	$lg[\Phi_R(M(\zeta))]$
3	2	0.2269	0.1065	12.2362	0.1589	0.1744	16.6097
3	3	0.2412	0.0921	16.3870	0.1333	0.2000	24.9193
3	4	0.2521	0.0812	20.9449	0.1168	0.2165	33.5317
3	5	0.2607	0.0726	25.8066	0.1051	0.2282	42.3861
3	6	0.2677	0.0656	30.9111	0.0963	0.2371	51.4420
4	2	0.1486	0.0676	23.3663	0.0868	0.1088	33.1658
4	3	0.1595	0.0603	29.4999	0.0625	0.1250	48.7669
4	4	0.1684	0.0544	36.1694	0.0560	0.1293	61.7016
4	5	0.1758	0.0495	43.2376	0.0513	0.1325	74.8992
4	6	0.1820	0.0453	50.6236	0.0477	0.1349	88.3275
5	2	0.1064	0.0468	38.2574	0.0522	0.0739	55.3620
5	3	0.1146	0.0427	46.5072	0.0448	0.0776	72.5382
5	4	0.1216	0.0392	55.4224	0.0400	0.0800	90.0000
5	5	0.1277	0.0362	64.8325	0.0365	0.0817	107.7430
5	6	0.1330	0.0335	74.6359	0.0339	0.0830	125.7450
6	2	0.0807	0.0344	57.0488	0.0400	0.0506	78.8374
6	3	0.0871	0.0318	67.5197	0.0342	0.0530	100.8750
6	4	0.0926	0.0296	78.7860	0.0305	0.0545	123.1450
6	5	0.0975	0.0276	90.6445	0.0278	0.0556	145.6890
6	6	0.1019	0.0258	102.973	0.0257	0.0564	168.5050

For any design ζ of the mixture model with qualitative factors, we measure its quality by efficiency, and the efficiency under A-optimality criterion can be expressed as:

$$A_{eff}(\zeta) = \frac{trM_g^{-1}(\zeta_A)}{trM_g^{-1}(\zeta)}.$$

The relative A-efficiencies of different R-optimal designs in model (3.3) are displayed in Table 2. In model (3.3), for any q, whereas the efficiencies decrease with the increase of the total number of qualitative levels, the R-optimal designs in most cases have relatively high efficiencies. It should be noticed that the R-optimal designs for $q = 6$ and $s = 2$ have an A-efficiency of 98.89 % relative to the A-optimal designs. In other words, the R-optimal designs in model (3.3) perform excellently in terms of the A-optimality criterion.

Table 2. The relative A-efficiencies of R-optimal designs for $3 \leq q \leq 6$ and $2 \leq s \leq 6$.

q	s	$A_{eff}(\zeta_R^*)$	q	s	$A_{eff}(\zeta_R^*)$
3	2	0.9666	5	2	0.9852
3	3	0.9616	5	3	0.9807
3	4	0.9601	5	4	0.9773
3	5	0.9607	5	5	0.9749
3	6	0.9623	5	6	0.9733
4	2	0.9788	6	2	0.9889
4	3	0.9735	6	3	0.9852
4	4	0.9702	6	4	0.9821
4	5	0.9685	6	5	0.9797
4	6	0.9679	6	6	0.9778

5. Conclusions

This paper tackles the R-optimal design under the interaction of qualitative factors and mixture components. It considers the continuous R-optimal designs for the second-order Scheffé model with qualitative factors. The obtained results show that the optimal marginal design η is the uniform design under R-optimality criterion. For model (3.3), general analytical expressions for the decision function under the R-optimal designs are derived, and the R-optimality is confirmed by the equivalence theorem. Finally, we remark further that the performance of R-optimal designs in terms of the A-optimality criterion is excellent.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (Grant No. 12071096, 11901260), Science and Technology Fund for basic research of Guizhou Province ([2020]1Y010).

Conflict of interest

The authors declare that there is no conflict of interest.

References

1. J. A. Cornell, *Experiments with mixtures: designs, models, and the analysis of mixture data*, New York: John Wiley & Sons, 2002.

2. T. Zijlstra, P. Goos, A. Verhetsel, A mixture-amount stated preference study on the mobility budget, *Transport. Res. A-Pol.*, **126** (2019), 230–246. https://doi.org/10.1016/j.tra.2019.06.009

3. P. Goos, H. Hamidouche, Choice models with mixtures: An application to a cocktail experiment, *Food Qual. Prefer.*, **77** (2019), 135–146. https://doi.org/10.1016/j.foodqual.2019.04.006

4. P. Goos, B. Jones, U. Syafitri, I-optimal design of mixture experiments, *J. Am. Stat. Assoc.*, **111** (2016), 899–911. https://doi.org/10.1080/01621459.2015.1136632

5. X. F. Zhang, Z. B. Zhu, C. Q. Zhang, Multi-stage differential evolution algorithm for constrained D-optimal design, *AIMS Mathematics*, **6** (2021), 2956–2969. https://doi.org/10.3934/math.2021179

6. H. Dette, Designing experiments with respect to standardized optimality criteria, *J. R. Stat. Soc. B*, **59** (1997), 97–110. https://doi.org/10.1111/1467-9868.00056

7. H. H. Hao, X. Y. Zhu, X. F. Zhang, C. Q. Zhang, R-optimal design of the second-order scheffé mixture model, *Stat. Probabil. Lett.*, **173** (2021), 109069. https://doi.org/10.1016/j.spl.2021.109069

8. C. P. Lee, M. N. L. Huang, D-optimal designs for second-order response surface models with qualitative factors, *J. Data Sci.*, **9** (2011), 139–153. https://doi.org/10.6339/JDS.201104_09(2).0001

9. R. X. Yue, X. Liu, K. Chatterjee, D-optimal designs for multiresponse linear models with a qualitative factor, *J. Multivariate Anal.*, **124** (2014), 57–69. https://doi.org/10.1016/j.jmva.2013.10.011

10. M. H. Kao, K. Hazar, Optimal designs for mixed continuous and binary responses with quantitative and qualitative factors, *J. Multivariate Anal.*, **182** (2021), 104712. https://doi.org/10.1016/j.jmva.2020.104712

11. M. Zhang, F. Yang, Y. D. Zhou, Uniformity criterion for designs with both qualitative and quantitative factors, *Statistics*, **55** (2021), 90–109. https://doi.org/10.1080/02331888.2021.1873993

12. A. N. Donev, Design of experiments with both mixture and qualitative factors, *J. R. Stat. Soc. B*, **51** (1989), 297–302. https://doi.org/10.1111/j.2517-6161.1989.tb01767.x

13. Z. B. Zhu, G. H. Li, C. Q. Zhang, A-optimal designs for mixture central polynomial model with qualitative factors, *Commun. Stat. Theor. M.*, **48** (2019), 2345–2355. https://doi.org/10.1080/03610926.2018.1472783

Appendix

A.1 Proof of Theorem 1

The inverse matrices of $M_f(\xi)$ is

$$M_f^{-1}(\xi) = \begin{bmatrix} M_{11}^{-1}(\xi) + K_{(1)} & -M_{11}^{-1}(\xi)M_{12}(\xi)D_{22}(\xi) \\ -D_{22}(\xi)M_{21}(\xi)M_{11}^{-1}(\xi) & D_{22}(\xi) \end{bmatrix},$$

we have

$$M_g^{-1}(\zeta) = \begin{bmatrix} F_{11}(\zeta) & F_{12}(\zeta) \\ F_{21}(\zeta) & F_{22}(\zeta) \end{bmatrix},$$

where $\zeta = \eta \times \xi$,

$$\begin{aligned}
F_{11}(\zeta) =& (D \otimes M_{11}(\xi))^{-1} + (D \otimes M_{11}(\xi))^{-1}\, \eta \otimes M_{12}(\xi) \\
& \left[M_{22}(\xi) - \eta^T \otimes M_{21}(\xi)\, (D \otimes M_{11}(\xi))^{-1}\, \eta \otimes M_{12}(\xi)\right]^{-1} \\
& \eta^T \otimes M_{21}(\xi) \left[D^{-1} \otimes M_{11}^{-1}(\xi)\right]^{-1} \\
=& \left(D^{-1} \otimes M_{11}^{-1}(\xi)\right) + \left(D^{-1} \otimes M_{11}^{-1}(\xi)\right) \eta \otimes M_{12}(\xi) \\
& \left[M_{22}(\xi) - \eta^T D^{-1} \eta \otimes M_{21}(\xi) M_{11}^{-1}(\xi) M_{12}(\xi)\right]^{-1} \eta^T D^{-1} \otimes M_{21}(\xi) M_{11}^{-1}(\xi) \\
=& D^{-1} \otimes M_{11}^{-1}(\xi) + \left(D^{-1}\eta \otimes M_{11}^{-1}(\xi) M_{12}(\xi)\right) D_{22} \eta^T D^{-1} \otimes M_{21}(\xi) M_{11}^{-1}(\xi) \\
=& D^{-1} \otimes M_{11}^{-1}(\xi) + D^{-1}\eta\eta^{\mathrm{T}} D^{-1} \otimes M_{11}^{-1}(\xi) M_{12}(\xi) M_{21}(\xi) M_{11}^{-1}(\xi) \\
=& D^{-1} \otimes M_{11}^{-1}(\xi) + J_s \otimes K_{(1)}, \\
F_{12}(\zeta) =& -(D \otimes M_{11}(\xi))^{-1} (\eta \otimes M_{12}(\xi)) \\
& \left[M_{22}(\xi) - \eta^T \otimes M_{21}(\xi)\, (D \otimes M_{11}(\xi))^{-1}\, \eta \otimes M_{12}(\xi)\right]^{-1} \\
=& - D^{-1} \otimes M_{11}^{-1}(\xi) \eta \otimes M_{12}(\xi) D_{22}(\xi) \\
=& - D^{-1}\eta \otimes M_{11}^{-1}(\xi) M_{12}(\xi) D_{22}(\xi) \\
=& - 1_s \otimes M_{11}^{-1}(\xi) M_{12}(\xi) D_{22}(\xi) \\
=& F_{21}^{\mathrm{T}}(\zeta), \\
F_{22}(\zeta) =& \left[M_{22}(\xi) - M_{21}(\xi) M_{11}^{-1}(\xi) M_{12}(\xi)\right]^{-1} \\
=& D_{22}(\xi),
\end{aligned}$$

and $\mathbf{1}_s$ is a $s \times 1$ vector of all ones, j_s is the $s \times s$ matrix with all elements equal to unity.

As a result, we have

$$\begin{aligned}
tr[log[\prod_{i=1}^{m} (M_g^{-1}(\zeta))_{ii}]] &= \mathrm{tr}\left\{\log\left[D^{-1} \otimes M_{11}^{-1}(\xi) + J_s \otimes K_{(1)}\right]\right\} + tr\left\{\log\left[D_{22}(\xi)\right]\right\} \\
&= \sum_{j=1}^{s} \mathrm{tr}\left(\log\left(\frac{1}{\eta(j)} M_{11}^{-1}(\xi) + K_{(1)}\right)\right) + tr\left\{\log\left[D_{22}(\xi)\right]\right\} \\
&\geq \prod_{j=1}^{s} \mathrm{tr}\left(\log\left(\frac{1}{\eta(j)} M_{11}^{-1}(\xi) + K_{(1)}\right)\right) + tr\left\{\log\left[D_{22}(\xi)\right]\right\},
\end{aligned}$$

the equality is attained if and only if the elements of η is equal, i.e. $\eta(1) = \eta(2) = \cdots = \eta(s) = 1/s$.

A.2 Proof of Theorem 2

$$\begin{aligned}
\psi_{R-g}(j, \mathrm{x}; \zeta) =& g^{\mathrm{T}}(j, \mathrm{x}) M_g^{-1}(\zeta) B_g M_g^{-1}(\zeta) g(j, \mathrm{x}) \\
=& s^2 u(\mathrm{x}; \xi) + s \cdot \sum_{i=1}^{4} v_i(\mathrm{x}; \xi) + \sum_{i=1}^{4} \omega_i(\mathrm{x}; \xi)
\end{aligned}$$

where

$$
\begin{aligned}
u(\mathrm{x};\xi) &= f_1^T(\mathrm{x})M_{11}^{-1}(\xi)B(s,1)M_{11}^{-1}(\xi)f_1(\mathrm{x}),\\
v_1(\mathrm{x};\xi) &= 2f_1^T(\mathrm{x})M_{11}^{-1}(\xi)B(s,1)K_{(1)}f_1(\mathrm{x}),\\
v_2(\mathrm{x};\xi) &= f_1^T(\mathrm{x})K_{(1)}B(s,1)K_{(1)}f_1(\mathrm{x}),\\
v_3(\mathrm{x};\xi) &= -f_1^T(\mathrm{x})M_{11}^{-1}(\xi)B(s,1)M_{11}^{-1}(\xi)M_{12}(\xi)D_{22}(\xi)f_2(\mathrm{x}),\\
v_4(\mathrm{x};\xi) &= -f_1^T(\mathrm{x})K(1)B(s,1)M_{11}^{-1}(\xi)M_{12}(\xi)D_{22}(\xi)f_2(\mathrm{x}),\\
\omega_1(\mathrm{x};\xi) &= f_1^T(\mathrm{x})M_{11}^{-1}(\xi)M_{12}(\xi)D_{22}(\xi)HD_{22}(\xi)M_{21}(\xi)M_{11}^{-1}(\xi)f_1(\mathrm{x}),\\
\omega_2(\mathrm{x};\xi) &= -f_1^T(\mathrm{x})M_{11}^{-1}(\xi)M_{12}(\xi)D_{22}(\xi)HD_{22}(\xi)f_2(\mathrm{x}),\\
\omega_3(\mathrm{x};\xi) &= -f_2^T(\mathrm{x})D_{22}(\xi)HD_{22}(\xi)M_{21}(\xi)M_{11}^{-1}(\xi)f_1(\mathrm{x}),\\
\omega_4(\mathrm{x};\xi) &= f_2^T(\mathrm{x})D_{22}(\xi)HD_{22}(\xi)f_2(\mathrm{x}).
\end{aligned}
$$