Bayesian Approach to The D-Optimal for Mixture Experimental Design

U A Wizsa[1], U D Syafitri[1*], A H Wigena[1]

[1] Department of Statictics, IPB University, INA
*correspondence author: utamids@apps.ipb.ac.id

Abstract. A mixture experiment is a special case of response surface methodology in which the value of the components are proportions. In case there are constraints on the proportions, the experimental region can be not a simplex. The classical designs such as a simplex-lattice design or a simplex-centroid design, in some cases, cannot fit to the problem. In this case, optimal design come up as a solution. A D-optimal design is seeking a design in which minimizing the covariance of the model parameter. Some model parameters are important and some of them are less important. As the priority of the parameters, the prior information of parameters is needed in advance. This brings to a Bayesian D-optimal design. This research was focus on a baking experiment in which consisted of three ingredients with lower bounds on the proportion of the ingredients. The assumption model was a quadratic model. Due to the priority of the model parameters, the Bayesian D-optimal design was used to solve the problem. A point-exchange algorithm was developed in order to find the optimal design. The results show that the Bayesian D-optimal design was better than the D-optimal design in terms of parameter estimation for the baking experiment. Nineteen candidates is used to choose twelve design points. It found that the potential term is feasible to the actual model. The design points also represent overall points in the design area.

1. Introduction

Mixture experimental design used especially in industrial sector. A Mixture experiment is a design in which the components are proportions. The proportions lie between 0 and 1 and the sum of the proportions among the components is unity [11]. Due to the restriction of the mixture experiment, the proportions of the components are dependence [2].

Unlike other designs, the compositions of a mixture design depends on the assumption model. The model parameters are some important but others are not. The prior information of the parameters is needed. In addition, there are some constraints on the proportions. The constraints effects to the experimental region. In some cases, the classical mixture designs cannot suite to the problem. Hence, it needs an optimal design approach. The optimal design is seeking a design based on a certain criterion. A D-optimality criterion is a widely criterion used in mixture experiments. It can reduce the uncertainty of estimating parameters model [7]. This criterion is useful for the performance of design under the assumption of model [4]. The D stands for determinant. The D-optimality criterion is a criterion which minimizing the determinant of the invers of the information matrix or it is equivalence with maximizing the determinant of the information matrix [3].

The optimal design based on the matrix of design is very depend on the assumtion of model [6]. Inaccuracies in model assumption impact to inaccuracies in design result. The used of D-optimality criterion with the Bayesian approach can retrieve basic information from the distribution of parameter model, so the dependence of D-optimal to model assumption can be reduced.

2. Bayesian D-Optimal Design

Assumption model used the quadratic canonical Shefee model [12]. The general form of quadratic canonical model is

$$\eta(x) = \sum_{i} \beta_i x_i + \sum_{i<j} \beta_{ij} x_i x_j \tag{1}$$

Total proportion of components which amounted to 100% in the mixture caused the form of a linear model do not have intercept. Then, the cross product $\boldsymbol{x_i x_j}$ and $\boldsymbol{x_i^2}$ cannot included together to the model, because it caused perfect collinierity [10]. Estimating model parameters using least squares estimation can be calculated as

$$\boldsymbol{b} = (\boldsymbol{X^T X})^{-\mathbf{1}} \boldsymbol{X^T Y} \tag{2}$$

Variance of $\boldsymbol{\beta}$ is $\sigma^2(\boldsymbol{X^T X})^{-1}$. Minimizing the variety of $\boldsymbol{\beta}$ can be calculate as maximize of determinant $(\boldsymbol{X^T X})$.

In the Bayesian approach of D-optimal design, matrix $\boldsymbol{X}$ divided into two conditions, namely primary and potential terms. The most important term that really want to fit in model lead to primary terms and the term that potential (possibly important) to the model called as potential terms [5]. Bayesian used initial information to predict the condition. Thereafter, the initial information is taken from estimating parameters of primary and potential terms. The initial information can be formed into prior distribution. Prior distribution for primary term is $\boldsymbol{\beta_{pri}} \sim N(\mu, \sigma^2)$ and prior distribution for potential term is $\boldsymbol{\beta_{pot}} \sim N(0, \tau^2\sigma^2 I)$. The value of τ is the ratio of the variety on the potential terms with the error rate. A large value of τ means that some of the potential terms are feasible on the model. Conservely, a small value of τ means all potential terms should not be included in the model [6].

Based on *prior distributions* above, *posterior distribution* to parameters of model [5,8] is

$$p(\boldsymbol{\beta}|\boldsymbol{y}, \sigma) \sim N\left[(\boldsymbol{X^T X} + \tfrac{\boldsymbol{K}}{\tau^2})^{-1} \boldsymbol{X^T Y}\,, \sigma^2(\boldsymbol{X^T X} + \tfrac{\boldsymbol{K}}{\tau^2})^{-1}\right] \tag{3}$$

$\boldsymbol{K}$ is $(p+q) \times (p+q)$ identity matrix with p diagonal is zero and q others diagonal is one. Variance of posterior distribution is $\sigma^2(\boldsymbol{X^T X} + \boldsymbol{K}/\tau^2)^{-1}$. So, minimize the variance similar with maximize the value of determinan information matrix below [1]

$$(\boldsymbol{X^T X} + \boldsymbol{K}/\boldsymbol{\tau^2}) \tag{4}$$

DuMouchel and Jones also introduce the procedure called scalling convention [5]. This procedure is to make the effect of potential term can be interpret. Then, scalling convention also minimize correlation between primary and potential terms. Let $\boldsymbol{X} = [\boldsymbol{X_{pri}} | \boldsymbol{X_{pot}}]$ is design matrix that divided to primary and potential terms. Each unconstant primary terms transformed by interval -1 to 1. Then potential terms transformed such that max($\boldsymbol{X_{pot}}$)-min($\boldsymbol{X_{pot}}$)=1 [5].

For each primary term the midpoint value is $M = (L+U)/2$ and the half of the range is $\Delta = (U-L)/2$, L and U is lower and upper value of primary term. Then the scaled, l_k, from the primary term x_k is

$$x_k = (l_k - M)/\Delta \tag{5}$$

Potential terms regressed to primary terms, by least square estimation the coefficients regression ($\boldsymbol{\alpha}$) of $\boldsymbol{X_{pot}}$ to $\boldsymbol{X_{pri}}$ is

$$\boldsymbol{\alpha} = (\boldsymbol{X}_{pri}^{T}\boldsymbol{X}_{pri})^{-1}\boldsymbol{X}_{pri}^{T}\boldsymbol{X}_{pot} \tag{6}$$

Defined $\boldsymbol{R}$ residual of regression of $\boldsymbol{X}_{pot}$ to $\boldsymbol{X}_{pri}$, $\boldsymbol{R} = \boldsymbol{X}_{pot} - \boldsymbol{X}_{pri}\boldsymbol{\alpha}$ and $\boldsymbol{Z}$ is transformation of $\boldsymbol{R}$.

$$\boldsymbol{Z} = \boldsymbol{R}/\left(max(\boldsymbol{R}) - min(\boldsymbol{R})\right) \tag{7}$$

The result, definition of $\boldsymbol{X} = \left[\boldsymbol{X}_{pri}|\boldsymbol{X}_{pot}\right]$ become $\boldsymbol{X} = \left[\boldsymbol{X}_{pri}|\boldsymbol{Z}\right]$.

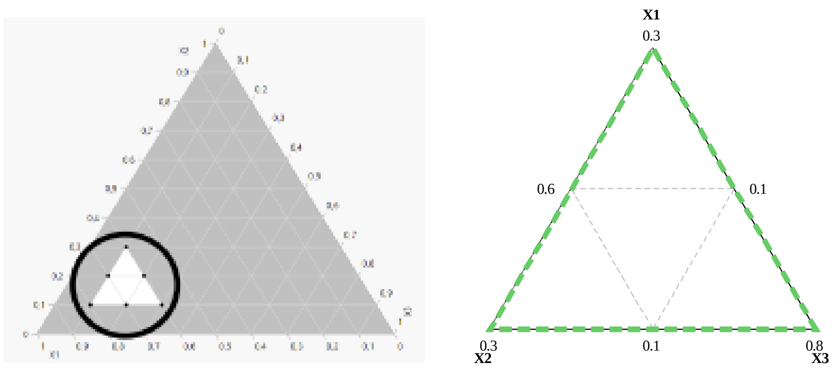
3. Point Exchange Algorithm

The process of finding optimal design points from candidate points is carried out with an algorithm. One simple algorithm that can be use is point exchange algorithm. This algorithm starts by randomly selecting the design points on the set of candidate points as many as *n* points as the initial design. Next, one by one the points in the initial design are exchanged with one other point from the set of candidate points in sequence. This process is carried out in an effort to improve the initial draft criteria that have been selected.

Randomization of points for the initial design does not guarantee that the design obtained is the optimal design. The point replacement process also does not find much permutation from all of the candidate points owned. Therefore, some initial designs were chosen randomly and carried out iteratively to overcome the limitations of point changes.

4. A Practical Example

As a practical used of the D-optimal design using Bayesian approach, an example of mixture experimental design with three components and constraints function was presented. Defined the first component as x_1, the second component as x_2, and the third component as x_1. The constraint function of the first and second component greater than or equal to 0.1 and third component greater than or equal to 0.6.



(a) (b)

Figure 1 Design region in full area (a) Design region by the constraint function (b)

The constraint function of each components can be figure out as design region. The region with constraints function is a small part of the full design area (Figure 1a). If the region was magnified, it can be seen as Figure 1b and points along the design can be choose as candidate point. Points in the edge of region called as extreme vertices. It represent the maximum proportion of each components. The other points can be in the middle of line, within, or in center of region.

Nineteen candidates generated using Cox-direction. Define first component as x_i and other component as x_j, so changes for every component can be calculate $\tilde{x}_i = x_i + \delta_i$ dan $\tilde{x}_j = x_j - \delta_i x_j/(1 - x_i)$ for $i \neq j$ and $i = 1,2,\ldots,p$, p is the number of candidates can be generate [9]. List of candidates of this case can be show in table 1.

Table 1. Candidates Set of Mixture

No	Proportion		
	x_1	x_2	x_3
1	0.1000	0.1000	0.8000
2	0.1000	0.1500	0.7500
3	0.1000	0.2000	0.7000
4	0.1000	0.2500	0.6500
5	0.1000	0.3000	0.6000
6	0.1333	0.1333	0.7334
7	0.1333	0.2333	0.6334
8	0.1500	0.1000	0.7500
9	0.1500	0.1500	0.7000
10	0.1500	0.2000	0.6500
11	0.1500	0.2500	0.6000
12	0.1666	0.1667	0.6667
13	0.2000	0.1000	0.7000
14	0.2000	0.1500	0.6500
15	0.2000	0.2000	0.6000
16	0.2333	0.1333	0.6334
17	0.2500	0.1000	0.6500
18	0.2500	0.1500	0.6000
19	0.3000	0.1000	0.6000

The model used canonical quadratic model, it written as

$$y_i = \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \beta_4 x_{1i} x_{2i} + \beta_5 x_{1i} x_{3i} + \beta_6 x_{2i} x_{3i} \tag{8}$$

which

y_i= respose, $i = 1,2,...,n$.

β_j= model parameters, $j = 1,2,...,6$.

x_{ji}= value of independence variables $i = 1,2,...,n$, $j = 1,2,...,6$.

From the model above, the design matrix can be constructed as

$$\boldsymbol{X} = \begin{bmatrix} x_{11} & x_{12} & x_{13} & x_{11}x_{12} & x_{11}x_{13} & x_{12}x_{13} \\ x_{21} & x_{22} & x_{23} & x_{21}x_{22} & x_{21}x_{23} & x_{22}x_{23} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{n1} & x_{n1} & x_{n3} & x_{n1}x_{n2} & x_{n1}x_{n3} & x_{n2}x_{n3} \end{bmatrix} \tag{9}$$

The first order of model in this case assumed as primary terms, and the second order assumed as potential terms. Then to calculate the determinant of matrix information K matrix can be used as

$$\boldsymbol{K} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{10}$$

From design matrix, the matrix information can be construct to get optimal design from several value of τ by equation 4. The result of design points can form as design region below.

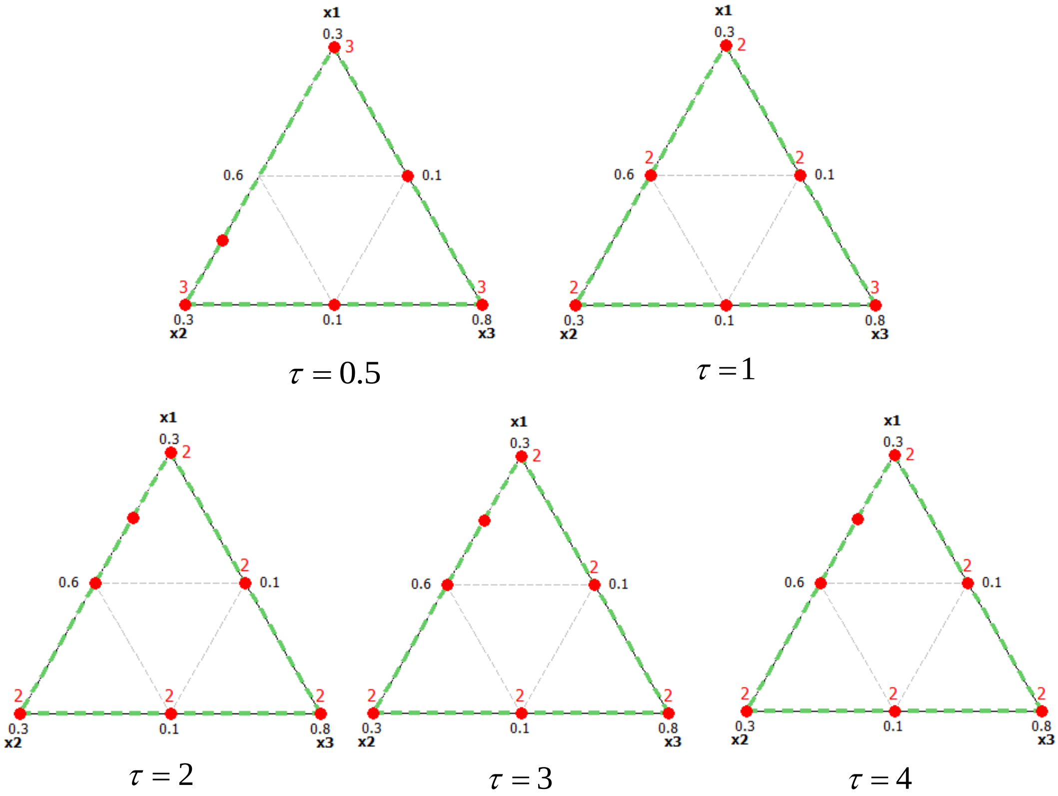


Figure 2. Design region for several value of τ

The design region in $\tau = 0.5$ have lack points as design, it can be seen from some points was not detect in the region. Furthermore, in $\tau = 1$ the design points better then design in $\tau = 0.5$. The design point detect in every extreme points and half of design region. Even the design was better, if the value of τ increase, when $\tau = 2$ the design points was similiar with $\tau = 3$ and $\tau = 4$. The design have addition one point from the design in $\tau = 1$. So, from this result the design constant strart from $\tau = 2$, and this design choosed as optimal design points. The result of optimal design points using Bayesian D-optimal design in two components of mixture can be seen in table 2. Determinant for this design about 1464.09 with seven different points of design.

Table 2. Design Points of Optimal Design

No	Proportion		
	x_1	x_2	x_3
1	0.20	0.10	0.70
2	0.10	0.20	0.70
3	0.10	0.20	0.70
4	0.10	0.30	0.60
5	0.30	0.10	0.60
6	0.25	0.15	0.60
7	0.10	0.10	0.80
8	0.20	0.20	0.60
9	0.10	0.30	0.60
10	0.30	0.10	0.60
11	0.20	0.10	0.70
12	0.10	0.10	0.80
Determinant	1464.09		

5. References

[1] Andere-Rendon J Montgomery DC and Rollier DA 1997 *Design of Mixture Experiments Using Bayesian D-Optimality* (Journal of Quality Technology vol 29) no 4 p 451-463.

[2] Cornell JA 2002 *Experiments with Mixtures: Designs, Models, and the Analysis of Mixture Data, Third Edition* (New York (US): John Wiley & Sons, Inc)

[3] Chaloner K and Verdinelli I 1995 *Bayesian Experimental Design: A Review* (Statist. Sci 10) no 3 p 237-304

[4] DasGupta A 1996 *Review of Optimal Bayes Designs* (Handbook of Statistics vol 13: Elsevier) p 1090-1147

[5] DuMouchel W and Jones B 1994 *A Simple Bayesian Modification of D-Optimal Design to Reduce Dependence on an Assumed Model* (Technometrics vol 36) no 1 p 37-47.

[6] Hefang L Myers RH and Keying Y 2000 *Bayesian Two-Stage Optimal Design for Mixture Models* (Journal of Statistical Computation and Simulation vol 66) no 3 p 209-231.

[7] Mitchell TJ 1974 *An Algorithm for the Construction of "D-Optimal" Experimental Designs*. (Technometrics vol 16) no 2 p 203-210.

[8] Pilz J 1991 *Bayesian Estimation and Experimental Design in Linear Regression Models* (New York John Wiley & Sons, Inc).

[9] Rahayu WD 2019 *Rancangan D-Optimal untuk Respon Ordinal Pada Rancangan Campuran* (Bogor : Institut Pertanian Bogor).

[10] Ruseckaite A Goos P Fok D 2016 *Bayesian D-optimal Choice Design for Mixtures* (Rotterdam: Journal of the Royal Statistical Society: Series C Applied Statistics) vol 66

[11] Smith WF 2005 *Experimental Design for Formulation* (Virginia American Statistical Assosiation and the Society for Industrial and Applied Mathematics).

[12] Schefee H 1958 *Experimentas with Mixtures* (Journal of the Royal Statistical Society, B 20) p 344-360.