

The Bayesian D-Optimal Design In Mixture Experimental Design

Uqwatul Alma Wizza¹, Utami Dyah Syafitri², Aji Hamim Wigena³
{uqwatulalma@gmail.com¹, utamids@gmail.com², ajiwigena@gmail.com³}

Department of Statistics, IPB University, Bogor, Indonesia^{1,2,3}

Abstract. Mixture design is known as experimental design which is often used. The total number of components in the mixture is 100% and the value of each component must be greater than or equal to 0%. The industry sector is usually used the mixture design. Then, the D-optimality criterion can help to determine the possible compositions of the mixture to conduct some trial and error composition of the product. However, this criterion very depend on the assumption of the model. To reduce its dependence, the Bayesian approximation is used. The Bayesian D-optimal algorithm applied to a mixture consisting of two components with constraint functions. Ten design points formed from eleven candidate points. By applying the Bayesian D-optimal algorithm on two components of the mixture, the design has no convergent design as the result. So, to find the result, the classical D-optimal was used and three different points was formed.

Keywords: Bayesian, D-optimal design, mixture design

1 Introduction

Mixture design is a design that involves two or more components (factors) with a value equal to or more than 0% and 100% total of the mixture. The response depends only on the proportion of components in the mixture, not depends on its total amount. If the proportion of a factor increase than at least one other will decrease, and vice versa [1].

Mixture experimental designs are necessary for the industrial sector. A company needed to add a variety of products on the market by doing an innovation. Based on the OSLO Manual "Innovation is an implementation of something new or develop that is significant both on products (goods and services), processes, marketing methods or organizational methods" [2]. Grace et al [2] revealed that from 1,223 respondents surveyed 61% claimed to provide innovation in their industry.

To make an innovation of the product, practitioners often do not have the right information about the design of the products. Practitioners usually determine the composition by trial and error that will give a large combination of experiments to try one by one. The more ingredients (components) involved, the more combinations will be formed. The consequences are more resources, time, energy, and costs will spend. The presence of experimental design provides a solution for the industry to achieve better efficiency. An experimental design will help in obtaining an optimal design. This procedure can reduce the number of possible combinations formed by trial and error.

One of the most often used of the optimal design criteria is the D-optimal design criterion [3]. The D-optimal design is used the maximum determinant of the information matrix as

criteria to find the optimal design. This criterion reduces uncertainty in estimating the model coefficient. Mitchell [4] conducted a D-optimal design with linear models that provided good design results for five sample cases. However, there is a lack of the D-optimal design criterion, which is a large dependence on the assumption of the model. Furthermore, DuMouchel and Jones [5] modified the Bayesian D-optimal design to reduce dependence on the model that assumed. By this modification, it was found the algorithm was more resistant to bias due to the model assumption error. Andere-Rendon et al. [6] were modified Bayesian D-optimal on mixtures of three and four components with and without constraint functions. Based on these modifications, it was found that the Bayesian D-optimal criterion provides better results, smaller error bias, allow for a larger order in the model, increase the scope of factor region, and better variety.

Research with a mixture design that consists of two and three components with constraint functions has been carried out by Andani [7] and Rahayu [8]. Andani [7] conducted a studying of optimal design using a linear model for numerical responses and nonlinear models for ordinal responses. Simplex Lattice, Simplex centroid, D-optimal design, and I-optimal design were generated using the JMP device. It was found that the design uses the assumption of a linear model with the I-optimal criterion and the assumption of a nonlinear model with the D-optimal criterion gave the best result for all cases. Furthermore, Rahayu [8] built the D-optimal algorithm with R device to determine the optimal design by assuming a nonlinear model for the ordinal responses. Based on the algorithm, the design by D-optimal was more efficient than using classical design.

Continuing research by Andani [7] and Rahayu [8], this study will use the Bayesian D-optimal design criterion to construct the optimal mixture design for two components of a mixture. The limited development of research in the Bayesian approach to D-optimal design, this study is assuming to use linear models for numerical responses.

2 Material

This mixture composition consists of two components, the proportion of component 1 (x_1) and the proportion of component 2 (x_2), with constraint functions $x_1 \geq 0.8$ and $x_2 \geq 0.1$. The constraint functions provided in each component can be visualized as a design region.

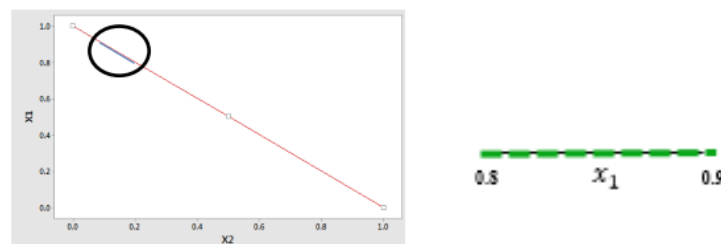


Fig.1. Design region of the constraint functions

Each point along the line is the possible point of the design. The lower and upper ends in the design region are the minimum and maximum values of the first component (x_1).

3Methods

The procedures of the Bayesian D-optimal design are as followed.

1. Determine the model assumption. The model used a quadratic linear model, following the model for the mixture design of the canonical form of the response function by Shefee[9].

$$y_i = \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{1i} x_{2i} \quad (1)$$

y_i = response for $i = 1, 2, \dots, n$.

β_j = parameter of the model for $j = 1, 2, 3$.

x_{ji} = value of dependent variables for $i = 1, 2, \dots, n, j = 1, 2, 3$.

2. Determine the design matrix with n is the number of design points. The design matrix will be used for the information matrix.

$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} & x_{11}x_{12} \\ x_{21} & x_{22} & x_{21}x_{22} \\ \vdots & \vdots & \vdots \\ x_{n1} & x_{n2} & x_{n1}x_{n2} \end{bmatrix} \quad (2)$$

3. Determine primary terms ($p = 2$) and potential terms ($q = 1$) from the model. Primary terms are the important term in the model and potential terms are the term that potential to the model. First-order in the model is decide as primary terms and the second-order in the model is decide as potential terms.
4. Determine the matrix \mathbf{K} whichelements is 0 and 1, p diagonal is 0 and q diagonal is 1, and 0 for others.

$$\mathbf{K} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (3)$$

5. Determine the number of design points that will be conducted as a result of optimal design, $n = 10$ design points.
6. Determine the candidate set, which is a collection of design points that allow optimal design candidates. The candidate points conducted by selecting the possible point from the constraint function of components.
7. Change the values of the candidate point to the scaling convention[5]. This procedure used to minimize the correlation between primary terms (\mathbf{X}_{pri}) and potential terms (\mathbf{X}_{pot}).
 - a. Suppose $\mathbf{X} = [\mathbf{X}_{pri} | \mathbf{X}_{pot}]$ is a set of points in the matrix design. Potential terms are regressed towards primary terms, the least square of \mathbf{X}_{pot} by \mathbf{X}_{pri} is

$$\boldsymbol{\alpha} = (\mathbf{X}_{pri}^T \mathbf{X}_{pri})^{-1} \mathbf{X}_{pri}^T \mathbf{X}_{pot} \quad (4)$$

- b. Defined

$$\mathbf{R} = \mathbf{X}_{pot} - \mathbf{X}_{pri}\boldsymbol{\alpha} \quad (5)$$

and

$$\mathbf{Z} = \frac{\mathbf{R}}{\max(\mathbf{R}) - \min(\mathbf{R})} \quad (6)$$

\mathbf{R} contains residuals from a regression between potential and primary terms. Definition \mathbf{X} is changed to $\mathbf{X} = [\mathbf{X}_{pri}|\mathbf{Z}]$. From this stage, the primary and potential terms are not correlated.

8. Choosing randomly n design points from the candidates set after scaled, then formed into a design matrix (equation 2).
9. Calculate the determinant of the information matrix from n design points using Bayesian D-optimal criterion [5,10,11].

$$\mathbf{M} = (\mathbf{X}^T \mathbf{X} + \frac{\mathbf{K}}{\tau^2}) \quad (7)$$

With τ^2 is a constant which is the ratio of variety on the potential terms with the error rate. Try for several τ to get convergent in determinant.

10. Exchange one by one the set candidates that have been selected in step 8 with one other set of candidates and calculate the matrix \mathbf{M} (equation 7) for every change.
11. Repeat steps 8-10 to a certain iteration. The design that produces the largest and convergent determinant \mathbf{M} (information matrix) of all replications was chosen as the optimal design.

4 Result and Discussion

Eleven candidates were constructed based on the constraint functions of components that represent all the possible design points. The candidates produced by choosing the possible points to represent the overall design region. Candidate points can be written in Table 1 and present in Figure 2. In the next step, ten from eleven candidates chose as the optimal design of the mixture. Ten design points were selected using the Bayesian D-optimal criterion. Optimal design using the Bayesian D-optimal is highly depended on the selection of τ^2 . Theoretically, the large values of τ^2 indicate that some of the potential terms are proper on the model. Conversely, small values of τ^2 indicate all potential terms should not be included in the model [10].

Table 1. Candidates points of mixture

No	Proportion of components	
	x_1	x_2
1	0.8000	0.2000
2	0.8307	0.1693
3	0.8607	0.1393
4	0.8907	0.1093
5	0.9000	0.1000
6	0.8693	0.1307
7	0.8386	0.1614
8	0.8079	0.1921
9	0.8500	0.1500
10	0.8193	0.1807
11	0.8807	0.1193

Each candidate points can be figured out by Fig.2, it shows that eleven points spread along the region of design.

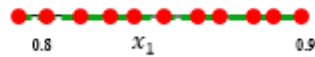


Fig. 2. Candidate points in design region

First, the design evaluated based on the choice of several values of τ . The results of the determinant of design points as in Figure 3.

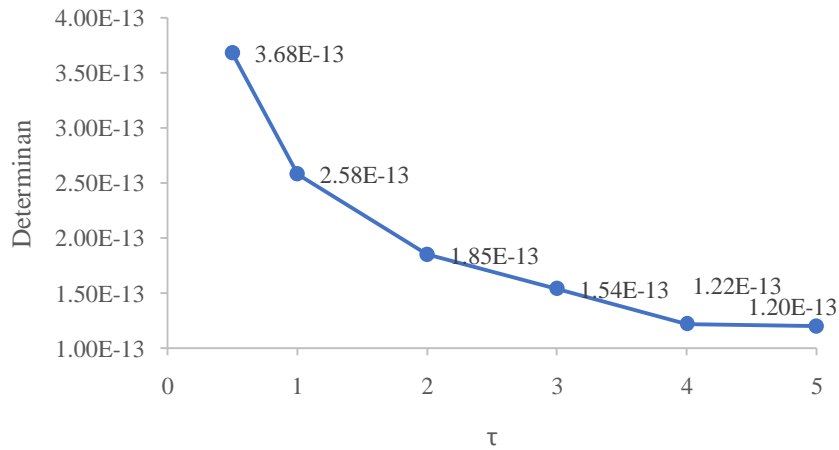


Fig. 3. Determinant for several τ

In Figure 3 the determinant of design point in $\tau = 0.5$ about $3.68E-13$. Then, determinant will be decreased for every increase of τ . Determinant highly decrease in $\tau = 1$ about $2.5E-13$ also

in $\tau = 2$ about $1.8E-13$. But in $\tau = 3, 4,$ and 5 the determinant slowly decrease about $1.54E-13$; $1.22E-13$; and $1.20E-13$ consecutively. Design points of each value of τ illustrated by the design region in Figure 4.

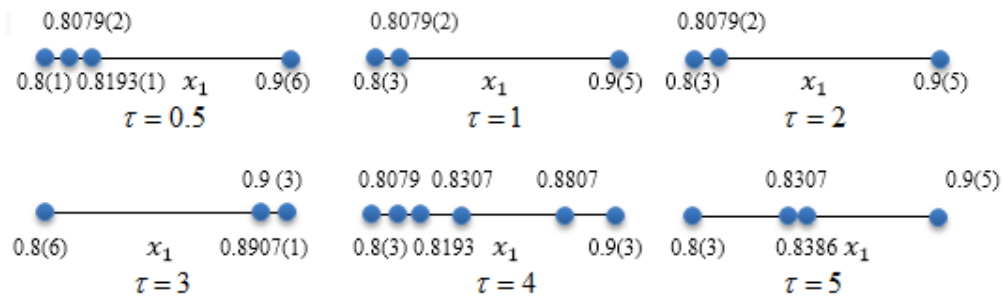


Fig. 4. Design region for several τ

Each point in the design region represents the first component (x_1), then $1-x_1$ represents the second component (x_2), and the number in the bracket shows the number of replication at that point. The design point was not convergent from $\tau = 0.5 \sim 5$. All of the design was different. This condition leads to two possibilities, firstly the model assumed was not fit to the case or secondly the potential term should not include the model, so the model only consists of the primary term. By this condition, if the potential terms changed to a primary term, then the optimal design constructed by D-optimal classic (Figure 5).

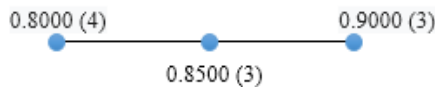


Fig. 5. Optimal design by D-optimal criterion

Figure 5 shows the optimal design using the D-optimal classic. The result is better than the D-optimal Bayesian. Three different points from 10 points formed with 3 and 4 replications. It placed in the end and middle of the design region. This design is closely formed with design in $\tau = 5$ by using D-optimal Bayesian. But, by D-optimal Bayesian, the design can not be convergent even in a high number of τ .

5 Conclusion

In this example of a mixture with two constraint components, it found that the optimality method using D-optimal Bayesian was not given a good result. The design was not convergent in several numbers of τ from 0.5 until 5. It needs more research to find what is the problem that caused the design can not be convergent. So, for the result of this case, the D-optimal classic was used with three different points as a result of optimal design.

6 Acknowledgements

We would like to thank Ministry of Research, Technology and Higher Education of the Republic Indonesia (Kemenristekdikti) for funding this research and the partnership company for collaboration.

References

- [1] Atkinson AC, Donev AN, Tobias RD: Optimum Experimental Designs, with SAS. United States (US): Oxford University Press (2007)
- [2] Grace N, Wijayanti R, Mardiana E, Handayani T, Rahmaida R, & Nadhiroh IM: Indikator Iptek Indonesia 2011. Aminullah E, editor. Jakarta(ID): LIPI (2011)
- [3] Kiefer J, Wolfowitz J: Optimum Design in Regression Problems. *Annals of Mathematical Statistics*, Vol. 30, pp. 271-294 (1959)
- [4] Mitchell, TJ: An Algorithm for the Construction of “D-Optimal” Experimental Designs. *Technometrics*, Vol: 16(2), pp. 203-210 (1974)
- [5] DuMouchel W, Jones B: A Simple Bayesian Modification of D-optimal Design to Reduce Dependence on an Assumed Model. *Technometrics*, Vol. 36(1), pp. 37-47 (1994)
- [6] Andere-Rendon J, Montgomery, DC & Rollier, DA: Design of Mixture Experiments Using Bayesian D-Optimality. *Journal of Quality Technology*, Vol. 29(4), pp. 451-463 (1997)
- [7] Andani W: Kajian Penggunaan Model Linear Untuk Respon Ordinal Pada Rancangan Campuran [Tesis]. Bogor (ID): Institut Pertanian Bogor (2019)
- [8] Rahayu WD. Rancangan D-optimal Untuk Respon Ordinal Pada Rancangan Campuran [Tesis]. Bogor (ID): Institut Pertanian Bogor (2019)
- [9] Scheffé H: Experiments with Mixtures. *Journal of the Royal Statistical Society*. Vol. 20, pp: 344-360 (1958)
- [10] Hefang L, Myers RH & Keying Y: Bayesian Two-Stage Optimal Design for Mixture Models. *Journal of Statistical Computation and Simulation*, Vol. 66(3), pp: 209-231 (2000)
- [11] Pilz J: Bayesian Estimation and Experimental Design in Linear Regression Models. New York (US): John Wiley & Sons, Inc (1991)