Available online at www.isas.org.in/jisas
JOURNAL OF THE INDIAN SOCIETY OF AGRICULTURAL STATISTICS 75(1) 2021 19–23

Augmented Simplex-centroid Designs for Mixture Experiments

Debopam Rakshit, B.N. Mandal, Rajender Parsad and Sukanta Dash
ICAR-Indian Agricultural Statistics Research Institute, New Delhi

Received 29 May 2020; Accepted 30 December 2020

SUMMARY

For an experiment with mixtures, it is presumed that the response is dependent on the relative proportions of the components present in the mixture and it is invariant of the total amount of the mixture used in the experiment. More number of the design points from the interior of the simplex space can help for better exploration of the entire simplex space for the standard simplex-centroid designs. In this article, a new method for obtaining augmented points from the interior of the simplex space for three-component simplex-centroid designs has been proposed. A special property of equilateral triangles has been used for this purpose. D-efficiency and G-efficiency of these augmented designs are evaluated to measure the efficiency of these obtained designs.

Keywords: Mixture experiments, Simplex-centroid design, Augmented, Component.

1. INTRODUCTION

For an experiment with mixtures, the proportions of constituent components are expressed as fractions of a mixture and they are consistently non-negative and sum to be one always, i.e. $0 \leq x_i \leq 1.0$ and $\sum_{i=1}^{q} x_i = 1; i = 1, 2, \ldots, q$, where q is the number of components in the mixture and x_i, $i = 1, 2, \ldots, q$ represents the proportion of component i forming the mixture. In mixture experiments, it is assumed that the observed response depends on the proportions of the ingredients present in the mixture and it does not depend on the amount of the mixture (Cornell, 2002). More details about mixture experiments can be found in Cornell (1973, 1979, 2002 and 2011). Simplex-centroid designs (Scheffé, 1963) are among the most widely used designs of mixture experiments.

A simplex-centroid design for q components mixture experiment consists of all possible subsets of these q components, which are present in equal proportions. The design consist of only one full mixture blend and that point is represented by the overall centroid. The total number of design points of standard simplex-centroid design is $(2^q - 1)$. These design points are obtained as ${}^qC_1$ permutations of (1,0,…,0), ${}^qC_2$ permutations of (1/2,1/2,0,…,0), ${}^qC_3$ permutations of (1/3,1/3,1/3,0,…,0), and so on and so forth, and finally the overall centroid of the triangle $(1/q, 1/q, \ldots, 1/q)$. For example, a three-component simplex-centroid design consists of $2^3 - 1 = 7$ points (Fig. 1(a)). The design points of this standard simplex-centroid are given in Fig. 1(b).

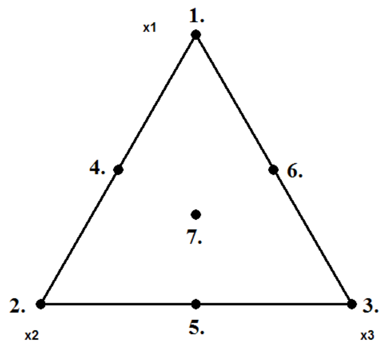
Let the functional relationship between observed response (η) and the proportions of components $x_1, x_2, \ldots, x_q$ be

$$\eta = \varphi(x_1, x_2, \ldots, x_q) + \varepsilon, \ \varepsilon \sim i.i.d\ \mathrm{N}(0, \sigma^2) \qquad (1)$$

A very basic assumption is made that the response surface, denoted by the function φ, is a continuous function of the x_i, $i = 1, 2, \ldots, q$.

The problem of determining the shape of the response surface with the component compositions depends on finding the mathematical equation that satisfactorily describe the function φ in (1). Due to the constraint $\sum_{i=1}^{q} x_i = 1$, these polynomial functions are different from the usual regression equations.

Corresponding author: B.N. Mandal
E-mail address: bn.mandal@icar.gov.in



(a)

Points	x1	x2	x3
1	1	0	0
2	0	1	0
3	0	0	1
4	1/2	1/2	0
5	0	1/2	1/2
6	1/2	0	1/2
7	1/3	1/3	1/3

(b)

Fig. 1. Three-component 7-point standard simplex-centroid design

To deal with data obtained from mixture experiments, Scheffé (1958, 1963) derived the canonical polynomial by applying the constraint that sum of mixture components proportions adds to one in standard polynomial. These polynomials are devoid of any intercepts.

Linear: $$\eta = \sum_{i=1}^{q} \beta_i x_i + \varepsilon \quad (2)$$

Quadratic: $$\eta = \sum_{i=1}^{q} \beta_i x_i + \sum_{i}^{q} \sum_{j>i}^{q} \beta_{ij} x_i x_j + \varepsilon \quad (3)$$

Special cubic:

$$\eta = \sum_{i=1}^{q} \beta_i x_i + \sum_{i}^{q} \sum_{j>i}^{q} \beta_{ij} x_i x_j + \sum_{i}^{q} \sum_{j}^{q} \sum_{k}^{q} \beta_{ijk} x_i x_j x_k + \varepsilon \quad (4)$$

where $\varepsilon \sim i.i.d\ N(0, \sigma^2)$

The terms present in canonical polynomials can be interpreted as usual regression coefficients. The coefficients of x_i, $i = 1, 2, \ldots, q$ denote the effects due to the i^{th} component, whereas the coefficients of x_{ij} denote the joint effect of the components *i* and *j*.

When the mixtures of components are strictly additive in nature, linear canonical polynomial (first degree) is best suitable to represent the surface, whereas canonical polynomials of higher degree are being used to represent the surface with curvature due to non-linear blending between the pairs of components. In this study, we confine ourselves to only quadratic and special cubic canonical polynomial.

Augmentation of simplex-centroid designs was first introduced by Cornell (1986). An application of 10-point augmented simplex centroid design can be found in Sifaoui (2016). Augmentation is suggested in those situations where researcher is fitting a low degree (first or second) canonical polynomial but is uncertain about the shape of the surface above the simplex region. If the number of design points is equal or less than the number of parameters in the fitted model, test for the significance of the parameters of the fitted model cannot be conducted. If the design is augmented with extra design points then only it has more number of design points than the number of parameters in the fitted model. With these more number of design points when an experiment is carried out, then test of significance of the parameters can be conducted. Classical designs position the points at the periphery of the experimental space and are potentially ill-adapted to deal with irregularities inside the space [Gomes *et al.* (2018)]. To limit these risks, augmentation of the design may be done in the interior of the simplex space. Further, more number of the design points from the interior of the simplex space can help for better exploration of the entire simplex space for the standard simplex-centroid designs.

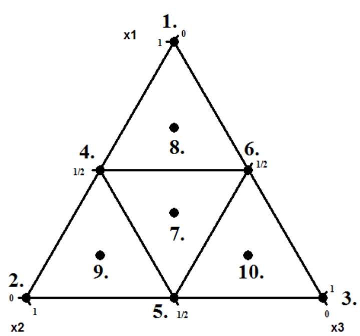
Thus, efforts are needed to be made for obtaining suitable ways to augment a standard simplex-centroid design so that the testing of significance of the parameters can be performed and testing of hypothesis that there is no lack of fit of the model being used with replicated points can be done.

In this paper, we propose a method of construction of simplex-centroid design with augmented points. These additional design points help to investigate the entire simplex area in a more precise way. D-efficiency and G-efficiency per point of the constructed designs are also computed to measure their efficacy.

2. METHOD OF CONSTRUCTION OF AUGMENTED SIMPLEX-CENTROID DESIGNS

A simplex-centroid design for three components is characterized by an equilateral triangle. From the geometry of triangles, it is known that any equilateral triangle can be partitioned into t^2 equilateral triangles, where t is a positive integer. So the simplex design space of a simplex-centroid design for three components can be partitioned into t^2 equilateral triangles. Centroids of these t^2 equilateral triangles can be used as augmented points of the design. This principle may be used for augmentation of standard designs for three-component mixture experiments. We illustrate the method of construction using two examples.

Example 2.1: Let $t = 2$. In this situation the whole simplex space is divided into $t^2 = 4$ small equilateral triangles as shown in Fig. 2. From each of these four small triangles, their centroids are obtained to include them as augmented design points. One of the centroid point is identical with the overall centroid of the triangle. Hence, we get three additional points to include as additional design points. Now the design becomes a ten-point augmented simplex-centroid design. The design is depicted by Fig. 2(a) and the design points are tabulated in Fig. 2(b).



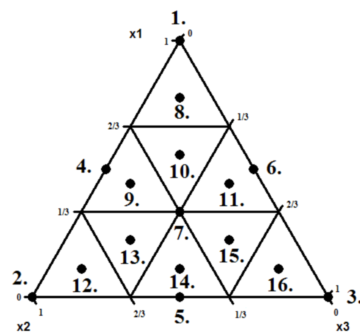
(a)

POINTS	x_1	x_2	x_3
1	1	0	0
2	0	1	0
3	0	0	1
4	1/2	1/2	0
5	0	1/2	1/2
6	1/2	0	1/2
7	1/3	1/3	1/3
8	2/3	1/6	1/6
9	1/6	2/3	1/6
10	1/6	1/6	2/3

(b)

Fig. 2. Three-component 10-point design, a simplex-centroid design with 3 components

Example 2.2: Consider $t = 3$. Then the triangle is divided into $t^2 = 9$ smaller equilateral triangles. From each of these nine triangles, nine centroid points are obtained and they are included in the design as augmented points. As none of the centroid is identical with the overall centroid of the triangle, all of them are included as additional points. Hence, we obtain a sixteen-point augmented simplex-centroid design. The resultant design is depicted by Fig. 3(a) and the design points are tabulated in Fig. 3(b).



(a)

Points	x_1	x_2	x_3
1	1	0	0
2	0	1	0
3	0	0	1
4	1/2	1/2	0
5	0	1/2	1/2
6	1/2	0	1/2
7	1/3	1/3	1/3
8	7/9	1/9	1/9
9	4/9	4/9	1/9
10	5/9	2/9	2/9
11	4/9	1/9	4/9
12	1/9	7/9	1/9
13	2/9	5/9	2/9
14	1/9	4/9	4/9
15	2/9	2/9	5/9
16	1/9	1/9	7/9

(b)

Fig. 3. Three-component 16-point design, a simplex-centroid design with 3 components

Augmented simplex-centroid designs for $t = 4$ and 5 are given in the Appendix.

3. EFFICIENCIES OF THE AUGMENTED DESIGNS

To compare the augmented simplex centroid designs with their non-augmented counterpart, D-efficiency and G-efficiency are computed. For this purpose, model matrix **X** is obtained for each augmented design for the study. Two efficiency measures such as D-efficiency per point and G-efficiency per point are calculated for each of the augmented and non-augmented designs. These measures are given by

$$\text{D-efficiency} = \left(\frac{|\mathbf{X'X}|^{1/p}}{n}\right) \times 100 \text{ (Hasan } et\ al.\text{, 2018)}$$

$$\text{G-efficiency} = \left(\frac{p}{nd}\right) \times 100$$

where n denotes the number of points in the design, p denotes the number of parameters in the model and d denotes the maximum value of var[$\hat{y}(\mathbf{X})$] $/\sigma^2$ over all candidate points.

Table 1 gives the D-efficiency and G-efficiency of augmented simplex-centroid designs for Scheffé quadratic model.

Table 1. D-efficiency and G-efficiency of augmented simplex-centroid designs for Scheffé quadratic model

	Total design points (n)	$\lvert X'X \rvert^{1/p}$	D-efficiency per point(%)	G-efficiency per point(%)
Non-augmented	7 + 0 = 7	0.27	3.87	86.36
t = 1	7 + 0 = 7	0.27	3.87	86.36
t = 2	7 + 3 = 10	0.31	3.15	64.51
t = 3	7 + 9 = 16	0.42	2.60	45.90
t = 4	7 + 15 = 22	0.53	2.40	38.61
t = 5	7 + 24 = 31	0.68	2.20	32.32

Both D-efficiency per point and G-efficiency per point of augmented designs are decreasing with increasing number of design points when Scheffé quadratic model is fitted. It seems to be because of the term "n" (number of design points) in the denominator. But it is also seen that the quantity $\lvert X'X \rvert^{\frac{1}{p}}$ increases with the increasing number of design points. This quantity reflects the amount of information being extracted from the design. Hence, with additional design points, the augmented designs are extracting more information.

In Table 2, we present the D-efficiency and G-efficiency of augmented simplex-centroid designs for Scheffé special cubic model.

Table 2. D-efficiency and G-efficiency of augmented simplex-centroid designs for Scheffé special cubic model

	Total design points (n)	$\lvert X'X \rvert^{1/p}$	D-efficiency per point(%)	G-efficiency per point(%)
Non-augmented	7 + 0 = 7	0.12	1.70	100
t = 1	7 + 0 = 7	0.12	1.70	100
t = 2	7 + 3 = 10	0.14	1.38	74.91
t = 3	7 + 9 = 16	0.18	1.13	52.82
t = 4	7 + 15 = 22	0.22	1.01	44.18
t = 5	7 + 24 = 31	0.29	0.93	36.40

While fitting Scheffé special cubic model, results similar to fitting Scheffé quadratic model are obtained. The quantity $\lvert X'X \rvert^{1/p}$ increases with increasing number of design points but it is comparatively lower as compared to quadratic model fitting. D-efficiency values are smaller for fitting Scheffé special cubic model as compared to fitting Scheffé quadratic model for the same design. In contradiction, G-efficiency is more for fitting Scheffé special cubic model than fitting quadratic model for the same design.

4. CONCLUSION

For the standard designs of three-component mixture experiments most of the design points are spread on the edge of the simplex design space. Augmentation of design points helps to explore the whole design space more effectively when these augmented points are from the interior portion of the simplex space. In this article, an easy and systematic method for obtaining augmented design points has been discussed.

It is seen that the quantitative value of $\lvert X'X \rvert^{1/p}$ increases as the number of design point increases by augmenting the standard design with more number of points. It implies that more amount information are being extracted by augmenting the standard simplex-centroid design. Again, it is also seen that G-efficiency are being reduced with the increasing number of augmented design points. This is due to bigger value of n in the denominator in the equation. Wheeler (1972) recommended that a design with G-efficiency more than or equal to 50% may be considered as "good" for practical purposes. It is seen that for augmented simplex-centroid design for Scheffé quadratic model considering up to t = 2 can be called as "good" and for cubic model it is up to t = 3.

ACKNOWLEDGEMENTS

The authors are thankful to the anonymous referee for their valuable comments.

REFERENCES

Cornell, J.A. (1973). Experiments with mixtures: A review. *Technometrics*, **15(3)**, 437-455.

Cornell, J.A. (1979). Experiments with mixtures: An update and bibliography. *Technometrics*, **21**, 95-106.

Cornell, J.A. (1986). A comparison between two ten-point designs for studying three component mixture systems. *Journal of Quality Technology*, **18**, 1-15.

Cornell, J.A. (2002). *Experiments with Mixtures: Designs, Models, and the Analysis of Mixture Data*, 3rd ed., John Wiley & Sons, Inc., New Jersey.

Gomes, C., Claeys-Bruno, M. and Sergent, M.(2018). Space-filling designs for mixtures. *Chemometrics and Intelligent Laboratory Systems*, **174**, pp.111-127.

Hasan, T., Ali, S. and Ahmed, M. (2018). Shrinkage simplex-centroid designs for a quadratic mixture model. *Journal of Industrial Engineering International*, **14**, 87-93.

Cornell, J.A. (2011). *A Primer on Experiments with Mixtures*, John Wiley & Sons, Inc., New Jersey.

Scheffé, H. (1958). Experiments with mixtures. *J. Roy. Statist. Soc.*, **B 20**, 344-360.

Scheffé, H. (1963). Simplex-centroid designs for experiments with mixtures. *J. Roy. Statist. Soc.*, **B 25**, 235-263.

Sifaoui, I., Mecha, E., Silva, A., Chammem, N., Mejri, M., Abderabba, M., and Bronze, M.R. (2016). Optimized extraction of antioxidants from olive leaves using augmented simplex centroid design. *Analytical Letters*, **49(9)**, 1323-1333.

Wheeler, R.E. (1972). Efficient experimental design, Montreal, Canada: American Statistical Association.

APPENDIX

Augmented simplex-centroid designs

(a) $t = 4$

Points	x_1	x_2	x_3
1	1	0	0
2	0	1	0
3	0	0	1
4	1/2	1/2	0
5	0	1/2	1/2
6	1/2	0	1/2
7	1/3	1/3	1/3
8	5/6	1/12	1/12
9	7/12	1/3	1/12
10	2/3	1/6	1/6
11	7/12	1/12	1/3
12	1/3	7/12	1/12
13	5/12	5/12	1/6
14	5/12	1/6	5/12
15	1/3	1/12	7/12
16	1/12	5/6	1/12
17	1/6	2/3	1/6
18	1/12	7/12	1/3
19	1/6	5/12	5/12
20	1/12	1/3	7/12
21	1/6	1/6	2/3
22	1/12	1/12	5/6

(b) $t = 5$

Points	x_1	x_2	x_3
1	1	0	0
2	0	1	0
3	0	0	1
4	1/2	1/2	0
5	0	1/2	1/2
6	1/2	0	1/2
7	1/3	1/3	1/3
8	13/15	1/15	1/15
9	2/3	4/15	1/15
10	11/15	2/15	2/15
11	2/3	1/15	4/15
12	7/15	7/15	1/15
13	8/15	1/3	2/15
14	7/15	4/15	4/15
15	8/15	2/15	1/3
16	7/15	1/15	7/15
17	4/15	2/3	1/15
18	1/3	8/15	2/15
19	4/15	7/15	4/15
20	4/15	4/15	7/15
21	1/3	2/15	8/15
22	4/15	1/15	2/3
23	1/15	13/15	1/15
24	2/15	11/15	2/15
25	1/15	2/3	4/15
26	2/15	8/15	1/3
27	1/15	7/15	7/15
28	2/15	1/3	8/15
29	1/15	4/15	2/3
30	2/15	2/15	11/15
31	1/15	1/15	13/15