Journal of Community Mobilization and Sustainable Development
Vol. 17(1), January-March 2022, 129-133

Cost friendly Experimental Designs for Product Mixtures in Agricultural Research

Rahul Banerjee[1], Seema Jaggi[2], Arpan Bhowmik[1]*, Eldho Varghese[3], Cini Varghese[1] and Anindita Datta[1]
[1]ICAR-Indian Agricultural Statistics Research Institute, New Delhi-110012
[2]ICAR Head Quarter, New Delhi-110012
[3]ICAR-Central Marine Fisheries Research Institute, Kochi-682018, Kerala

ABSTRACT

Sustainable agriculture practices are the one that caters for the present without compromising for the future generations and at the same time maintains and enhances environmental quality. With the depletion in the quality environmental conditions and an increase in the global population there is an immense need for ensuring sustainability in agricultural practices. Agricultural research can be thought to be as the backbone for bringing out alternative sustainable agricultural methodologies. Precise formulation of agricultural experiments for research requires the application of appropriate statistical tools particularly accurate situation specific designing and analysis of the experiments. There are several experiments in agricultural research where the response depends only on the proportions of the factors in the experiment. The theory of mixture experiments plays a crucial role in such situations and it has wide applicability in agricultural research. One of the difficulties in Mixture Experiments is the generation of design points specific to situation and model chosen. In this study we have attempted to develop an algorithm to obtain designs for mixture experiments specific to the situation. The algorithm is versatile in terms of the situation, number of runs in the design and other parameters and hence has wider applicability. The application of the algorithmic approach of design generation will also lead to minimization in the computational cost involved in the experiments.

Keywords: Algorithmic approach, Design generation, Mixture design, Mixture experiments, Nonlinear models

INTRODUCTION

Agriculture has predominantly been the most crucial source of subsistence for the mankind since times immemorial. It is the lifeblood to half of the global population even in the present scenario. As per, Food and Agriculture Organisation (FAO), Rome citizens of developing countries with an increasing trend in population size will face acute shortage of food security with a rise in hunger if the global food production does not increase on an average by 50-60% (Harender *et al.*, 2019). Thus, one of the most important contemporary requisites is to sustain natural resources at the same time while protecting environment. Sustainable agriculture is aimed at optimum management of natural resources for agriculture in order to quench the growing human needs at the same time maintaining the environmental quality by conservation of natural resources. Agricultural research is the backbone for ensuring sustainability of agriculture. Precise and planning of agricultural research is a prerequisite for finding valid inferences which would lead to formulation of efficient research policies for ensuring sustainability. Here lies the importance of statistics or specifically design of experiments for correct planning and designing of agricultural experiments according to situations. Situation specific precise designing and formulation of agricultural experiments will lead to accurate and precise inferences. The field of Mixture designs has huge applicability in agricultural research as several of the situations in agricultural experiments are that of mixture experiments. A mixture experiment is one in which the response is depends only upon the proportions of the mixture components and is independent on the amount of the mixture. A mixture is the final product obtained by mixing two or more components, the mixture components are called mixture ingredients. In a mixture experiment the sum total of the mixture component proportions are fixed as

*Corresponding author email id: arpan.stat@gmail.com

the ingredient proportions must add up to one. In case of agricultural experiments, the ingredients maybe several sources of nutrients used in making a fertilizer, it may be several proportions of the chemicals used in the formulation of a pesticide or an herbicide or the ingredients may be inputs at different crop growth stages in splits so that the total quantity of input applied to the crop is fixed. Similarly, in animal husbandry and veterinary feeding trials to study the response of feed to milk yield or body weight the total feed intake supplied to the animal is same and in combinations of different components (Murty and Das, 1968).

{Experimental Situation, Cornell (1981), pp.47} Four chemical pesticides *viz.* Vendex, Omite, Kelthane and Dibrom were sprayed on strawberry plants in an attempt to control mite population and a fixed dose say '*a*' of these four chemicals are applied in fifteen different combinations. Each chemical was applied individually and in combination with each of the others such that total dose of pesticide is same in all the fifteen treatment combinations. The objective of the experiment was to obtain the optimum combination of all the four pesticides to control mite population in strawberry.

In all the above examples the treatments are ingredients of the mixture and their proportions and the response is dependent only on the relative proportions of the ingredients present in the mixture. Thus we can infer that Mixture experiment methodology has wide application in agricultural research. Designing of a mixture experiment generally refers to selection of the proportions of the mixture ingredients in such a fashion such that a mathematical model can be fitted adequately to represent the mathematical relationship between response and ingredients of the mixture and at the same time ensure appropriate estimation of the parameters of the model. The design will also be able to find out combinations of the ingredients that will lead to optimum response. The components of the mixture geometrically represent a simplex. Mixture designs are essentially descriptions or layouts of the design points or mixture components. There are several mixture design available in literature few of which are Simplex Lattice Designs (Scheffé, 1958), Simplex Centroid Designs (Scheffé, 1963), Axial Designs (Chan *et al.,* 1999), Extreme Vertex Designs (McLean and Anderson, 1966) etc.

In agricultural sciences we often come across situations for mixture experiments in which the response exhibited is nonlinear. An experimental situation of the same has been given in (Cornell, 2002). For, situations exhibiting a nonlinear response, nonlinear models would be able to explain the situation in a much more efficient manner as compared to linear models. Besides, nonlinear models also have the advantage of having fewer parameters as compared to linear models. Most of the designs available in literature for Mixture experiments are linear in nature. Thus, there is a need to devise methodologies for obtaining designs for mixture experiments with nonlinear models. Algorithmic approach provides a great alternative for design construction for mixture experiments. Algorithms provide versatility in terms of number of runs of the design, model to be selected as well as the objective function to be chosen. Few Algorithms for Design generation already available in literature include the Detmax algorithm (Mitchell, 1974), KL Exchange algorithm (Atkinson and Donev, 1989), XVERT algorithm (Snee and Marquadt, 1974) and XVERT 1.0 algorithm (Nigam, 1983). There is also a lacunae for simple computer based solutions for design construction in mixture experiments adapting to different choices of models and continuous input variables. Toolkits like GOSSET, OPTEX procedure in SAS, statistical toolbox in MATLAB and AlgDesign package in R are well suited for design generation particularly for linear models with discrete factors however, they are not suitable for Nonlinear models and particularly with continuous factors there's a huge increase in the computational cost. In this study we have modified the Fedorov algorithm (1972) to obtain saturated locally D-optimal designs for mixture experiments with and without restrictions for a wide variety of models. The optimality criteria were satisfied using General Equivalence theorem extended to Nonlinear models by White (1973) and Whittle (1973). Lall *et al.* (2018a and c) have used algorithms to construct D-optimal designs for Exponential and Poisson regression model in Response Surface Experiments. Lall *et al.* (2018b) have also used algorithms to construct D-optimal designs for logistic model in Response Surface Experiments.

MATERIALS AND METHODS

For a mixture design with q components if the proportion of the i^{th} component is represented by x_i, then, the design space is represented by a q-1 dimensional simplex denoted by S_{q-1} and the following conditions hold true:

$$x_i \geq 0, \sum_{i-1}^{q} x_i = 1$$

These restrictions are fundamental for the formulation of any Mixture experiment design. The algorithmic method of design generation can be viewed as an optimization problem. The objective function can be chosen as per the criteria for design generation. In our example the objective function is the determinant of the Fishers Information Matrix fulfilling the D-optimality for obtaining the designs. Initially the *expand. grid* function of the R software is chosen to obtain a uniform candidate set of design points with minimal discrepancy fulfilling the model chosen and the constraints imposed. The objective of the algorithm is to look for the D-optimal design in the vicinity of the candidate set assuming that the optimal design exists in the vicinity. The Fedorov algorithm has been used to search the D-optimal design in the Fedorov algorithm each row of the design is replaced with new design points from the vicinity if there's an improvement in the objective function. This iteration is repeated until there is no improvement in the objective function and then the design obtained is the locally D-optimal design for that particular model. The algorithm also provides a flexibility in terms of the number of runs to be selected for the design for our case we have considered saturated designs. A saturated design has number of runs equal to the number of parameters to be estimated in the model. Saturated designs are advantageous in the sense that they are smaller designs and hence can lead to a significant reduction in computational cost. The step wise algorithmic procedure is mentioned below:

Steps of the Algorithm

1. The model to be fitted is chosen first.
2. In the next step the variables of the model as well as their range is specified.
3. In the third step a candidate set (CS) or a design space is generated consisting of a continuous grid of equidistant discrete points.
4. Then, the initial parameter guesses for the model and the number of runs required for the design is determined.
5. An initial design with a positive Fishers Information Matrix is chosen at random from the candidate set.
6. In the next step exchange of row coordinates of the design is performed iteratively with the coordinates from the vicinity. The coordinates are exchanged in such a fashion such that there is an improvement in the objective function.
7. All the above steps are repeated continuously until there is no such improvement observed in the objective function. At this step the design obtained is locally D-optimal design corresponding to the model selected and the constraints imposed.

The D-optimality criteria of the designs are verified using General Equivalence theorem.

RESULTS AND DISCUSSION

The above methodology has been used to obtain designs for linear as well as Nonlinear models. In all these designs the specific models are chosen, the initial parameter guesses of the parameters of the models are chosen, the number of runs for the design are determined and the algorithm is run. The designs obtained are specific to the model and the initial parameter guesses selected.

Two components mixture and three parameters: The design entails the situation with two mixing component and three parameters in the model considering the interaction term. Thus a saturated design would have 3 runs. The initial parameter guesses are given and the designs are obtained at different increments. The increments considered are 0.01, 0.02, 0.03, 0.04, 0.05, 0.06 and 0.10 respectively. The value of $|\mathbf{M}|$ shown in the table is the value of the Fishers Information Matrix obtained for the D-optimal design corresponding to the model and the particular increment. The designs obtained is mentioned in Table 1.

Here, we can see that the design obtained remains same for the increment size 0.01, 0.02, 0.05, 0.06 and 0.10 however it changes for increment size of 0.03 and 0.05. It is evident from the table that the size of the candidate set is an inverse function of the increment size as the increment is increased the candidate set becomes smaller and smaller.

Mixture designs for a logistic model: There often exist situations in agricultural problems where the response obtained is qualitative in nature, for eg. In baking a cake, the response may be the texture of the cake say firm or soft or in preparing a fruit punch the response maybe the taste of the punch. All the above cases are examples of mixture experiments with a qualitative response variable. In case of qualitative response variables, the relationship is usually Nonlinear in nature and the distribution is often Nonnormal. Here, we have taken such a situation where the response is qualitative. The response has been approximated using a three parameter logistic model for

Table 1: D-optimal Saturated Mixture Designs for Two Components Mixture

Model	$\eta = \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2;\ 0 \le (x_1, x_2) \le 1;\ x_1 + x_2 = 1$			
Increment	Size of the candidate set	D Optimal design obtained		$\lvert M \rvert$
		x_1	x_2	
0.01	101	0.5	0.5	0.002314
		0.0	1.0	
		1.0	0.0	
0.02	51	0.5	0.5	0.002314
		1.0	0.0	
		0.0	1.0	
0.03	34	0.49	0.51	0.002310
		1.0	0.0	
		0.0	1.0	
0.04	26	0.52	0.48	0.002307
		1.0	0.0	
		0.0	1.0	
0.05	21	0.5	0.5	0.002314
		1.0	0.0	
		0.0	1.0	
0.06	17	0.5	0.5	0.002314
		1.0	0.0	
		0.0	1.0	
0.10	11	0.5	0.5	0.002314
		1.0	0.0	
		0.0	1.0	

Table 2: Mixture designs for a logistic model

Model	$\eta = e\text{-}\phi^{i} / (1 + e\text{-}\phi_i)^2,\ \phi_i = \theta_1 x_{1i} + \theta_2 x_{2i} + \theta_3 x_{1i} x_{2i};\ x_1 + x_2 = 1$			
Increment	Size of the candidate set	D Optimal design obtained		$\lvert M \rvert$
		x_1	x_2	
0.01	101	0.0	1.0	0.00001097
		0.52	0.48	
		1.0	0.0	
0.02	51	0.52	0.48	0.00001097
		1.0	0.0	
		0.0	1.0	
0.03	34	0.52	0.48	0.00001096
		1.0	0.0	
		0.0	1.0	
0.04	26	0.52	0.48	0.00001096
		1.0	0.0	
		0.0	1.0	
0.05	21	0.5	0.5	0.00001093
		1.0	0.0	
		0.0	1.0	

a mixture model with two mixing component taking into consideration the interaction term. The algorithm has been run for five increments namely, 0.01, 0.02, 0.03, 0.04 and 0.05 with initial parameter guesses and the designs have been obtained in three runs. The D-optimal designs obtained for each of the increments has been mentioned in Table 2. The value of represents the value of the Fishers Information Matrix for the optimal design corresponding to each such increment.

CONCLUSION

The importance of agricultural research in the changing times of environmental conditions and with the importance on sustainability is immense. One of the key factors to ensure achievement of sustainability in agriculture is the correct planning and operation of agricultural research. Thus, the role of statistics and in particular design of experiments is crucial in correct designing of the agricultural experiments so that the inferences obtained from the experiments are valid and accurate. Since, a lot of situations in agricultural research from agronomical experiments to post harvest and food processing trials entail a mixture experiment situation. Thus, the correct application of mixture designs to agricultural situations is crucial to such problems. In this paper we have proposed an algorithmic methodology to obtain designs for mixture experiments under different situations. The algorithmic approach is versatile in the choice of the model, the number of runs of the design, constraints imposed on the model and also in optimality criteria of the designs. Thus the algorithmic methodology can act as an alternative for generation of designs for mixture experiments suiting to different conditions and situations. The application of the proposed algorithm can also help in minimization of the computational cost involved in design generation. All computations presented here were performed by developing suitable R-codes available with the authors.

ACKNOWLEDGEMENTS

The facilities provided by Director, ICAR-Indian Agricultural Statistics Research Institute is duly acknowledged.

REFERENCES

Atkinson, A.C. and A.N. Donev. 1989. The construction of exact D-optimum experimental designs with application to blocking response surface designs. *Biometrika*, 76(3), 515–526.

Becker, N.G. 1968. Models for the response of a mixture. *Journal of Royal Statistical Society, Series B*, 30: 349-358.

Chan, L.Y. and M.K. Sandhu. 1999. Optimal orthogonal block designs for a quadratic mixture model for three components. *Journal of Applied Statistics*, 26: 19–34.

Cornell, J.A. 1981. *Experiments with mixtures*. First Edition. Wiley, New York.

Cornell, J.A. 2002. *Experiments with mixtures*. Third Edition. Wiley, New York.

Fedorov, V.V. 1972. *Theory of optimal experiments*. New York: Academic Press.

Harender, K. and K. Malik. 2019. *Sustainable Agriculture, Research trends in Agriculture Sciences,* AkiNik Publications, New Delhi.

Lall, S.; S. Jaggi; E. Varghese; A. Bhowmik and C. Varghese. 2018c. Designs for fitting Poisson regression model. *Journal of Crop and Weed*, 14(1): 117-121.

Lall, S.; S. Jaggi; E. Varghese; C. Varghese and A. Bhowmik. 2018a. D-optimal designs for exponential and poisson regression models. *Journal of the Indian Society of Agricultural Statistics*, 72(1): 27-32.

Lall, S.; S. Jaggi; E. Varghese; C. Varghese and A. Bhowmik. 2018b. An algorithmic approach to construct D-optimal saturated designs for logistic model. *Journal of Statistical Computation and Simulation,* 88(6): 1191-1199.

McLean, R.A. and V.L. Anderson. 1966. Extreme vertices design of mixture experiments. *Technometrics*, 8(3): 447-456.

Mitchell, T.J. 1974. An algorithm for the construction of D-optimal experimental designs. *Technometrics,* 16(2): 203–210.

Murty, J.S. and M.N. Das. 1968. Design and Analysis of Experiments with Mixtures. *The Annals of Mathematical Statistics*, 39(5): 1517-1539.

Nigam, A.K. 1983. A new algorithm for extreme vertices designs for linear mixture models. *Technometrics*, 25(4), 367-371.

Scheffé, H. 1958. Experiments with mixtures. *Journal of Royal Statistical Society, Series B*, 20: 344-360.

Scheffé, H. 1963. Simplex-centroid designs for experiments with mixtures. *Journal of Royal Statistical Society, Series B*, 20: 344-360.

Snee, R.D. and D.W. Marquadt. 1974. Extreme vertices designs for linear mixture models. *Technometrics*, 16(3): 399-408.

White, L.V. 1973. An extension of general equivalence theorem to nonlinear models. *Biometrika*, 60(2): 345-348.

Whittle, P. 1973. Some general points in the theory of optimal experimental designs. *Journal of Royal Statistical Society, Series B*, 35: 123-130.

Received on December 2021; Revised on January 2022