Statistics and Applications {ISSN 2454-7395 (online)}
Volume 23, No. 2, 2025 (New Series), pp 71–88
http://www.ssca.org.in/journal

Three and Four Component Uniform Mixture Designs Based on Ellipsoidal Region - A Beginner's Training Manual

Poonam Singh and Himanshu Shukla
Department of Statistics, University of Delhi, Delhi-110007

Received: 20 January 2024; Revised: 30 August 2024; Accepted: 18 September 2024

Abstract

Talke and Borkowski (2012) discussed the generation of space-filling uniform designs in unit hypercube and gave measures of uniformity based on distance criteria. In this paper, a method is proposed for constructing uniform designs in ellipsoidal region and the designs are constructed for 2 and 3 dimensional ellipsoidal regions. The uniformity of the constructed designs is measured by distance based approaches. Further, these designs are used to obtain 3- and 4-component uniform mixture design using a transformation proposed for this purpose. The uniformity of the constructed mixture designs is measured by the DM_2 criterion.

Key words: Uniform design; Ellipsoidal Region; Mixture Experiments; Distance based criteria; DM_2 criterion.

AMS Subject Classifications: 62K20

1. Introduction

Mixture experiments involve blending of two or more ingredients to form end products. In these experiments, the response or characteristic of interest depends on the proportions of the components or ingredients present in the mixture and not on the total amount of the mixture. For example, the corrosion resistance or the strength of stainless steel depends on the proportions in which iron, copper, nickel, and chromium are mixed in the alloy. In a mixture experiment with s components, if x_i represents the proportion of the i^{th} ingredient in the mixture, then

$$0 \leq x_i \leq 1, i = 1, 2, ..., s \tag{1}$$

with $\sum_{i=1}^{s} x_i = x_1 + x_2 + \cdots + x_s = 1$

As a result, the factor space reduces to a $(s-1)$ dimensional simplex $S_{(s-1)}$, where

$$S_{(s-1)} = \{x = (x_1, x_2, ..., x_s) : \sum_{i=1}^{s} x_i = 1, x_i \geq 0\}. \tag{2}$$

Corresponding Author: Himanshu Shukla
Email: himanshushukla1903@gmail.com

The concept of mixture experiments was introduced by Quenouille (1953). Scheffé (1958, 1963) gave designs and models for mixture experiments. The simplex-lattice designs and simplex-centroid designs of Scheffé (1958, 1963) are boundary point designs except for the overall centroid or face centroids. For more details on mixture experiments see Cornell (2011).

Fang (1980) and Wang and Fang (1981) introduced uniform designs and discussed the use of quasi monte carlo method in their constructions. Uniform designs are space filling designs where design points are scattered uniformly over the experimental region. These designs can be easily obtained by using U-type designs defined on a unit hypercube. A U-type design can be obtained by using good lattice point (glp) method of Korobov (1959 a, b). Hua and Wang (1981) and Fang and Wang (1994) discussed the glp method and used it in the construction of uniform designs. Uniform design obtained using the glp method tends to have low discrepancy (Hua and Wang, 1981). Borkowski and Piepel (2009) introduced distance based criteria to measure the uniformity of a design. Talke and Borkowski (2016) gave a method to measure uniformity of a design in a spherical region. Ning *et al.* (2011a) gave DM_2 criterion to measure uniformity of a mixture design. For the comprehensive study on uniform design one can refer to Prescott (2008), Ning *et al.* (2011b), Fang *et al.* (2018), Adriana *et al.* (2022), Zhou and Xu(2014).

In many industrial and experimental situations, the experimentation at the vertices of the simplex region is not possible. Also, limited resources, time, and budget does not allow the exploration of entire simplex. In such cases, instead of exploring the entire simplex, the region is restricted to any subset of the simplex. In literature, spherical and ellipsoidal regions are considered. Thompson and Myers (1968) proposed a technique for obtaining mixture experimental design over some ellipsoidal region of interest with in the simplex. Mandal *et al.* (2015) obtained optimum mixture designs used for parameter estimation in Scheffé (1958) models on the ellipsoidal region within the simplex. The problem of combinatorial drug experiment on lung cancer was conducted by Al-Shyoukh *et al.* (2011) for systematically qualitative characterization of cellular responses induced by multiple signals. Singh and Shukla (2023) proposed a method for obtaining the design in s-dimensional spherical region using the design in 2-dimensional spherical region given by Talke and Borkowski (2016).

In this paper, we propose a method for the construction of uniform designs in ellipsoidal region and introduce a transformation to obtain uniform mixture designs from uniform designs in ellipsoidal region. The NT-net design based on glp generators is used to obtain the U-type design for this purpose. The design is then used to obtain design in $(s-1)$ dimensional ellipsoidal region from which uniform designs for s component mixture experiments are generated using the proposed transformation. The design for 3 and 4 component mixture experiments are obtained for illustrative purposes.

First, the designs on 2 and 3 dimensional ellipsoidal region are obtained, then from these designs uniform designs for 3 and 4 component mixture experiments are obtained. The uniformity of the designs in ellipsoidal region is measured using distance based approach and that of the mixture designs is obtained using the DM_2 criterion. DM_2 criterion has an explicit calculation formula therefore it is quite useful in measuring the uniformity of the design. Borkowski and Piepel (2009) discussed an approach using G function to obtain uniform experimental design for the s component mixtures from the design in $(s-1)$ dimensional unit hypercube, but due to nonlinear transformation of the G function the pattern of the uniformity is not preserved in the simplex. Here, in this paper, two transformations are used to obtain designs in the simplex and it is observed that the pattern of uniformity is preserved from the ellipsoidal region to simplex.

In Section 2, various measures of uniformity are discussed. In Section 3, a method for obtaining designs in ellipsoidal region is developed and the particular cases of 2 and 3 dimensions are discussed in detail. Section 4 includes a transformation that is used to construct uniform mixture designs in s dimensions from the uniform designs in $(s-1)$ dimensional ellipsoidal region and the designs for 3 and 4 components are obtained. The conclusion of the paper is given in Section 5. Section 6 gives guidance to the beginners in this field.

2. Measures of uniformity

Fang (1980) and Wang and Fang (1981) were the first to apply the idea of Number-Theoretic methods (NTM) to experimental designs. Number-theoretic method or Quasi Monte Carlo method is a combination of number-theory and numerical analysis and has a variety of applications in statistics. Fang and Wang (1994) discussed tha glp set and some measures of uniformity of a design. Borkowski and Piepel (2009) introduced three distance-based criteria, (i) root mean square distance (RMSD), (ii) average distance (AD) and (iii) maximum distance (MD).

Generally Monte Carlo sampling method is used to compute these values. In this method two sets a training set and a sampling set are taken and the design points of these sets are obtained using Number-Theoretic net (NT-net). From the given training set, a sampling set with smaller number of runs is obtained and the value of uniformity measure of the sampling set is evaluated based on the training set. However, in this article, we are generating different sampling sets based on different glp generator and the training set is separately generated by an appropriate glp generator. The criteria values of different sampling sets are evaluated using the training set.

Suppose $D = \{x_j = (x_{j1}, x_{j2}, ..., x_{js}); j = 1, 2, ..., n\}$ represents a sampling set *i.e.* a design with s components and n runs for which the distance based uniformity measures are to be computed and $T = \{t_k = (t_{k1}, t_{k2}, ..., t_{ks}), k = 1, 2, ..., N\}$ represent a training set consisting of $N(> n)$ runs and s components. Then D and T matrices of order $n \times s$ and $N \times s$ respectively, where

$$\mathrm{D} = \begin{bmatrix} x_{11} & x_{12} & \cdots & \cdots & x_{1s} \\ x_{21} & x_{22} & \cdots & \cdots & x_{2s} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{n1} & x_{n2} & \cdots & \cdots & x_{ns} \end{bmatrix} and \quad \mathrm{T} = \begin{bmatrix} t_{11} & t_{12} & \cdots & \cdots & t_{1s} \\ t_{21} & t_{22} & \cdots & \cdots & t_{2s} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ t_{N1} & t_{N2} & \cdots & \cdots & t_{Ns} \end{bmatrix}.$$

Let $d_s^2(x_j, t_k)$ denote the Euclidean distance between the points x_j and t_k scaled between 0 and 1, then

$$d_s^2(x_j, t_k) = \sum_{i=1}^{s} (x_{ji} - t_{ki})^2$$

and the three distance-based criteria are defined as follows:

(1) $RMSD = \sqrt{\frac{1}{N}\sum_{k=1}^{N} min_{1\leq j\leq n} d_s^2(x_j, t_k)}$

(2) $AD = \frac{1}{N}\sum_{k=1}^{N} min_{1\leq j\leq n} d_s(x_j, t_k)$

(3) $MD = max_{1\leq k\leq N}(min_{1\leq j\leq n} d_s(x_j, t_k))$.

The above mentioned distance based measures can also be applied for mixture experiments where $\sum_{i=1}^{s} x_{ji} = 1$ and $\sum_{i=1}^{s} t_{ki} = 1$ to measure the uniformity of the design. When the number of runs is small, the uniformity of a mixture design can be measured using the DM_2-discrepancy which is generalization of star discrepancy (Ning *et al.*, 2011a). Star discrepancy given by Weyl (1916) is a basic and simplest measure of uniformity. Ning *et al.* (2011a) gave the following analytical expression to compute the DM_2-discrepancy value of a design P_n with n runs defined over the s dimensional simplex.

$$\begin{aligned} &\mathrm{DM}_2\left(\mathrm{P_n}\right) = \\ &\left(\frac{\sqrt{s}}{(s-1)!}\right)^{1/2}\left\{C_{n,s} - \frac{2(s-1)!}{n}\sum_{i=1}^{n}\sum_{(\tau_2,\ldots,\tau_s)\in[0,1]^{s-1}} a_\tau \cdot (x_{i1})^{\tau_1} \cdot \prod_{j=2}^{s} x_{ij}^{\tau_j} + \right. \\ &\left.\frac{1}{n^2}\sum_{i=1,k=1}^{n}\left(\max\left(1-\sum_{j=2}^{s}\max\left(x_{ij}, x_{kj}\right), 0\right)\right)^{s-1}\right\}^{1/2} \end{aligned} \tag{3}$$

where

$$\begin{aligned} &[0,1]^{s-1} = \{(t_1, t_2, \ldots, t_{s-1}) : t_i = 0 \text{ or } 1\}, \\ &C_{n,s} = ((s-1)!)^3 2^{s-1}/(2(s-1)!)\prod_{k=0}^{s-2}(2s+k-1), \\ &a_\tau = (s-1)!/\left(2(s-1) - \sum \tau_i\right)!, \\ &\text{and } \tau_1 = 2(s-1) - \sum_{2}^{s}\tau_j. \end{aligned}$$

3. Uniform designs in ellipsoidal region

Consider an s dimensional ellipsoidal region defined as

$$\begin{aligned} E_n^s\left(\boldsymbol{y_0}\right) &= E_n^s\left(a_1, a_2, \ldots, a_s\right) \\ &= \left\{(y_1, y_2, \ldots, y_s) : \frac{\left(y_1 - y_{01}\right)^2}{a_1^2} + \frac{\left(y_2 - y_{02}\right)^2}{a_2^2} + \cdots + \frac{\left(y_s - y_{0s}\right)^2}{a_s^2} \leq 1\right\} \end{aligned} \tag{4}$$

where $y_0 = (y_{01}, y_{02}, ..., y_{0s})$ is the center, n is the number of points in the region and $a_1^2 + a_2^2 + \cdots + a_s^2 \leq 1$.

When $y_0 = (0, 0, ..., 0)$ and $a_1^2 + a_2^2 + \cdots + a_s^2 = 1$, it represents the standard ellipsoidal region denoted by $E_n^s(1)$,

$$E_n^s\left(1\right) = \left\{ (y_1, y_2, \ldots, y_s) : \frac{\left(y_1\right)^2}{a_1^2} + \cdots + \frac{\left(y_s\right)^2}{a_s^2} \leq 1 \right\}. \tag{5}$$

To obtain uniform designs in ellipsoidal region, we first obtain U-type design for s dimensional unit hypercube using the good lattice point (glp) method and then convert this into a design in ellipsoidal region using some transformation.

Construction 1:

The steps for generating the uniform design are as follows:

Step 1: For a given number of runs n, obtain the initial generating vector $H_n = (h_1, h_2, \ldots, h_k)$, with $1 \leq h_t < h_{t'} < n$ for $t < t'$, such that the greatest common divisor $\gcd(n, h_t) = 1$ for $t = 1, 2, ..., k$.
Step 2: For a given number of factors $s (\leq k)$, select a generator $(h_{t_1}, h_{t_2}, ..., h_{t_s})$; $h_{t_j} < h_{t_{j'}}$ for $t_j < t_{j'}$; $h_{t_1} = 1$ from H_n and generate an $n \times s$ NT-net design $Z = (z_{ij})$; $i = 1, 2, ..., n$ and $j = 1, 2, ..., s$ where $z_{ij} = \frac{2iu_{ij}-1}{2n}$ with $u_{ij} = h_{ij}$ (mod n).
Step 3. Obtain the design $y = (y_1, y_2, \cdots, y_s)$ in s dimensional ellipsoidal region $E_n^s(1)$ by using the transformation

$$\begin{aligned} y_{i1} &= a_{i1} \prod_{k=2}^{s} cos(2\pi z_{ik}), \\ y_{ij} &= a_{ij} sin(2\pi z_{ij}) \prod_{k=j}^{s} cos(2\pi z_{ik}); j = 2, 3, \ldots, s, \\ y_{is} &= a_{is} sin(2\pi z_{is}), \end{aligned} \tag{6}$$

where $a_{i1}, a_{i2}, \cdots, a_{is}$ are axis half lengths satisfying $\sum_{j=1}^{s} a_{ij}^2 = 1$.

Step 4. Compute the criteria values RMSD, AD, and MD of the designs obtained in step 3 for all the possible generators.
Step 5. Select the designs with minimum values of the RMSD, AD, and MD criteria value.

Note that there are many ways to obtain a_{ij}'s satisfying $\sum_{j=1}^{s} a_{ij}^2 = 1$. In this paper, Number theoretic–net (NT-net) approach is used. The NT-net design lies between 0 and 1. Some of the values are near 0 and some are near 1. The harmonic mean is one of the ways to choose a_{ij}'s as it gives less weightage to larger values and large weightage to smaller values to maintain balance and equal importance of each pair of a_{ij}'s in obtaining the a_{ij}'s. If there is availability of the normally distributed data then the axis of the ellipsoid can be obtained by using the concept of the confidence region given by Tzeng and Berns (2005).

The N points of the training set used in computation of criteria value in step 4 are generated using a generator chosen arbitrarily.

In step 5, if the generator giving minimum value of RMSD and AD is different from the generator giving minimum value of MD, then it may be considered as optimal generator whereas if the generator giving the smallest value of AD is also different from the generator giving minimum value of RMSD, then the generator for RMSD is considered as optimal generator as RMSD is least affected by variation of points near the boundary.

It may also be noted that for large n, the number of design generators are large therefore all non-equivalent design generators can be considered to generate designs and the design with minimum value of criteria is selected. For saving time, one can also use a good lattice point with a power generator in Step 4.

According to Talke and Borkowski (2016), equivalent glp generators are those generators whose design points can be found by column permutation of other glp generators. Let $H_n = (h_1, h_2, ..., h_k)$ be the initial generating vector with $1 \leq h_t < h_{t'} < n$ for $t < t'$ such that the greatest common divisor $gcd(n, h_t) = 1$ for $t = 1, 2, ..., k$. In a two dimensional unit hypercube $[0, 1]^2$, the two glp generators $(1, h_i)$ and $(1, h_j)$ are said to be equivalent if

$$(h_i \times h_j)(mod\ n) = 1.$$

Similarly, in a three dimensional unit hypercube $[0, 1]^3$, a glp generator $(1, h_i, h_j)$ is said to be equivalent to $(1, h_k, h_l)$ and $(1, h_m, h_s)$ with $h_j > h_i$ and $h_k > h_i$, if it satisfies the five conditions given below:

$$\begin{aligned}
&(i)\ (h_i \times h_k)mod(n) = 1\\
&(ii)\ (h_j \times h_m)mod(n) = 1\\
&(iii)\ (h_i \times h_l \times h_m)mod(n) = 1\\
&(iv)\ (h_l \times h_s)mod(n) = 1\\
&(v)\ (h_j \times h_k \times h_s)mod(n) = 1.
\end{aligned}$$

Remark: 1. Fang and Wang (1994) have defined a glp set as a set $P_n = \{x_k = (x_{k1}, x_{k2}, ..., x_{ks})\}$ of lattice points with smallest discrepancy among all possible generating vectors $(n; h_1, h_2, ..., h_s)$ for given n and $s(< n)$. They also showed that $h_1 = 1$ can always be assumed in a generator $(h_1, h_2, \ldots, h_s)$. However, in general for a given n there can be more than 's' h_i's say k in the initial generating vactor $H_n = (h_1, h_2, \ldots, h_k)$. This is described in Talke and Borkowski (2012). This leads to choosing a generating vector of size s out of the k elements in H_n which is possible in m= ${}^{k-1}C_{s-1}$ ways if $h_1 = 1$ is fixed. In the proposed method, we have first obtained the set $H_n = (h_1, h_2, \ldots, h_k)$ with all possible h_i's satisfying $1 \leq h_i < h_j < n$ for $i < j$ with $gcd(n, h_i) = 1; i = 1, 2, \ldots, k(< n)$ and then selected a generating vector of size s from it for a given s.

2. Fang and Wang (1994) have also discussed various methods to measure the uniformity of the design but distance-based approach of Talke and Borkowski (2016) is not discussed in this book. In the proposed method, we have used three distance-based criteria to measure the uniformity of the designs in ellipsoidal region. Obiri *et al.* (2020) has mentioned that with increase in the number of experimental factors or dimensions of the design space, it is difficult to compute discrepancy, therefore, the distance-based approach is more suitable. In the proposed method, instead of obtaining a glp set in unit hypercube, the discrepancy values

of the designs in ellipsoidal region are computed and the best design is selected. Further, as mentioned in Talke and Borkowski (2016), the design points obtained by the transformation given by Fang and Wang (1994, Section 1.5) does not preserve uniformity of the design in 2-dimensional ball B_2 despite the fact that set of points obtained from C^2 being uniformly scattered in C^2. Thus, a different transformation is provided by Talke and Borkowski (2016) to obtain design points in B_2. The transformation given in Fang and Wang (1994) depends on the value of (r_i, θ_i) to obtain designs. In the proposed method, the transformation used for obtaining the design in ellipsoidal region depends on the value of a_{ij}'s and z_{ij}'s for $i = 1, 2, \ldots, n$; $j = 1, 2, \ldots, s$ where $Z = (z_{ij})$ is the design in unit hypercube and a_{ij}'s are the axis half-length satisfying $\sum_{j=1}^{s} a_{ij}^2 = 1$. The value of r_i used in Fang and Wang (1994) depends only on the number of design points whereas the value of a_{ij} in the proposed transformation depends on the number of design points and the number of components. Here, a_{ij}'s are obtained by considering the harmonic mean which gives equal weightage to all the NT-net design points whereas in transformation of Fang and Wang (1994) chosen r_i's gives higher weightage to higher value of r_i and lower weightage to smaller value of r_i.

The specific transformation for s=2 and s=3 are discussed in Section 3.1 and 3.2 respectively.

3.1. Designs in two dimensional ellipsoidal region

Consider a two dimensional standard ellipsoidal region given by

$$E_n^2(1) = \left\{ (y_1, y_2) : \frac{(y_1)^2}{a_1^2} + \frac{(y_2)^2}{a_2^2} \leq 1 \right\}$$

where $a_1^2 + a_2^2 = 1$.

Let $z_n = \{z_n(i) = (z_{i1}, z_{i2}); i = 1, 2, \ldots, n\}$ be an n point design obtained after Step 1 of the Construction 1, then the uniform design in 2 dimensional ellipsoidal region $E_n^2(1)$ can be obtained using the transformation:

$$E_n^2(1) = \left\{ (y_{i1}, y_{i2}) = (a_{i1}\,\mathrm{Cos}\,(2\pi z_{i2})\,, a_{i2}\,\mathrm{Sin}\,(2\pi z_{i2}))\,; a_{i1}^2 + a_{i2}^2 = 1; i = 1, 2, \ldots, n \right\} \tag{7}$$

where $a_{i2} = \frac{2z_{i1}z_{i2}}{z_{i1}+z_{i2}}$ and $a_{i1} = \sqrt{1 - a_{i2}^2}; i = 1, 2, \ldots, n.$

Example 1: Consider a design with s=2 and n=21. For n=21, the design can be obtained using the generators from the initial candidate generating vector H_{21} with $H_{21} = \{1, 2, 4, 5, 8, 10, 11, 13, 16, 17, 19, 20\}$. Different generating vectors for constructing designs in unit hypercube $C^2 = [0, 1]^2$ are given in Table 1.

Table 1: List of glp generators for design in C^2

(1,2)	(1,4)	(1,5)	(1,8)	(1,10)	(1,11)	(1,13)	(1,16)	(1,17)	(1,19)	(1,20)

For each of these eleven glp generators listed in Table 1, a design is obtained in 2 dimensional ellipsoidal region. The design points of these designs are displayed in Figure 1.

From the Figure 1, it can be observed that the design points in designs obtained using glp generators (1,2), (1,4), (1,8), (1,19) are scattered more uniformly in comparison to the

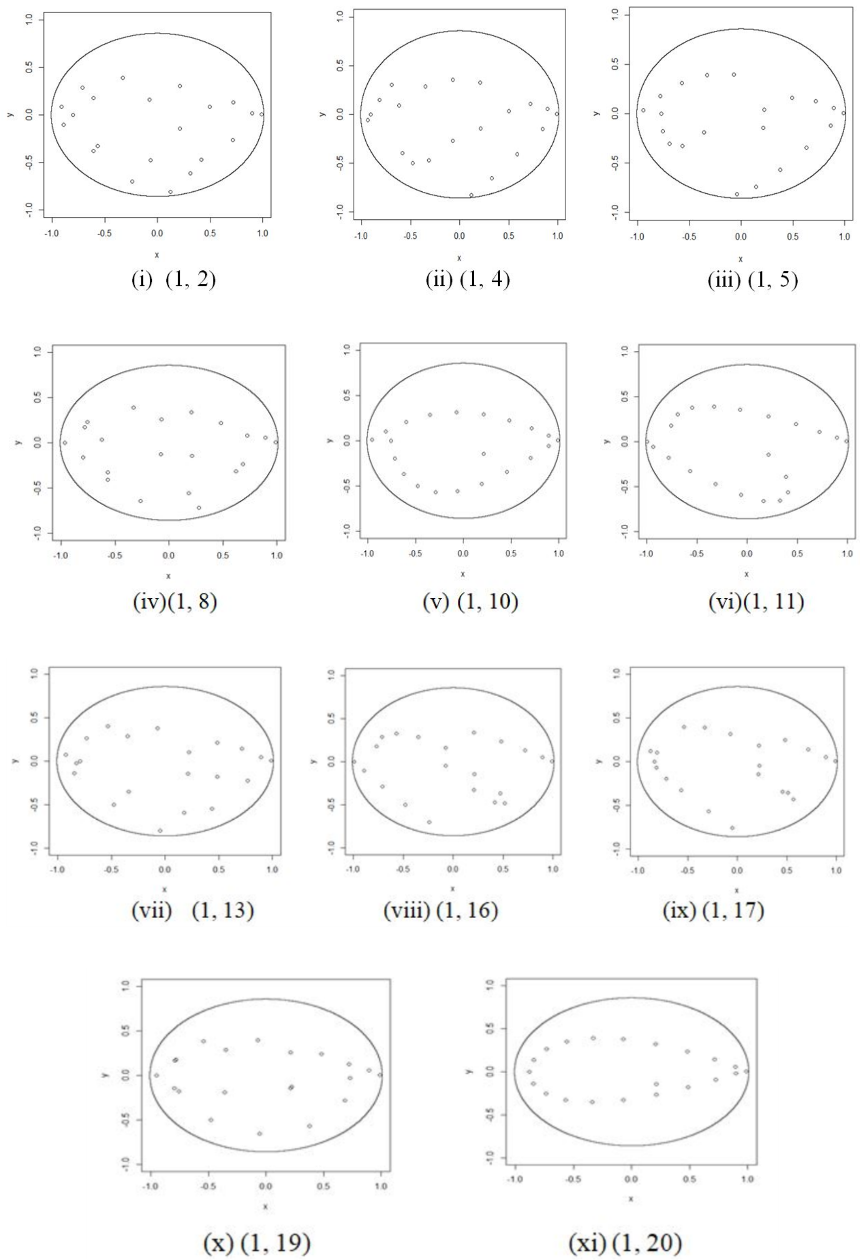

Figure 1: Plot of design points in designs obtained using generators (i) to (xi)

other designs. Scatteredness of these four designs are almost same so for choosing the most uniformly scattered design, the design with minimum value of criterion used can be selected.

Out of eleven generators only seven generators are non-equivalent. The generators (1,2), (1,4), (1,5), (1,8) are equivalent to (1,11), (1,16), (1,17), (1,19) respectively. The values of the three distance based criteria of uniformity for all the eleven designs are given in Table 2. For computing the criteria values a design, the points of the training set are obtained by using the generator (1,13) and the values are computed on the basis of 3000 evaluation points.

From the Table 2, it can be observed that the generator (1,2) gives the smallest

Table 2: Criteria values for all possible 21 run designs in 2 dimensional ellipsoidal region

Design Generator	RMSD	AD	MD
(1,2)	0.120193	0.103221	0.341422
(1,4)	0.127202	0.106910	0.328640
(1,5)	0.130708	0.105731	0.372629
(1,8)	0.122616	0.103479	0.317779
(1,10)	0.126546	0.104291	0.408531
(1,11)	0.129088	0.105089	0.394096
(1,13)	0.125127	0.105774	0.340943
(1,16)	0.143172	0.115458	0.425651
(1,17)	0.141105	0.115841	0.416019
(1,19)	0.123441	0.103987	0.350676
(1,20)	0.179616	0.132331	0.552201

value of 'RMSD' and 'AD' and the 'MD' value of the design obtained by generator (1,8) is minimum. Therefore, the generator (1,2) can be taken as optimal generator. It can also be observed from Figure 1 that the design points obtained by generator (1,2) are clustered less in comparison to that obtained by the generator (1,8).

3.2. Designs in three dimensional ellipsoidal region

Consider a three dimensional standard ellipsoidal region given by

$$E_n^3\,(1) = \left\{(y_1, y_2, y_3) : \frac{(y_1)^2}{a_1^2} + \frac{(y_2)^2}{a_2^2} + \frac{(y_3)^2}{a_3^2} \leq 1\right\}$$

where $a_1^2 + a_2^2 + a_3^2 = 1$.

Let $z_n = \{z_n(i) = (z_{i1}, z_{i2}, z_{i3}); i = 1, 2, ..., n\}$ be the design in three dimensional unit hypercube obtained after step 1 of the construction 1, then the design in three-dimensional ellipsoidal region with n runs can be obtained using transformation

$$\begin{aligned} E_n^3(1) = \{&(y_{i1}, y_{i2}, y_{i3}) - (a_{i1}\operatorname{Cos}(2\pi z_{i2})\operatorname{Cos}(2\pi z_{i3}), a_{i2}\operatorname{Sin}(2\pi z_{i2})\operatorname{Cos}(2\pi z_{i3}), a_{i3}\operatorname{Sin}(2\pi z_{i3});\\ &a_{i1}^2 + a_{i2}^2 + a_{i3}^2 = 1; i = 1, 2, ..., n \end{aligned} \tag{8}$$

with $a_{i2} = \frac{2z_{i1}z_{i3}}{z_{i1}+z_{i3}}, a_{i3} = \frac{2z_{i1}z_{i2}}{z_{i1}+z_{i2}}$ and $a_{i1} = \sqrt{1-(a_{i2}^2+a_{i3}^2)}; i = 1, 2, \ldots, n.$

Example 2: Consider a design with s=3 and n=34. For n=34, the design can be obtained using the generators from the initial candidate generating vector H_{34} with $H_{34} = \{1, 3, 5, 7, 9, 11, 13, 15, 19, 21, 25, 27, 29, 31, 33\}$. For n=34, we have 35 different non-equivalent glp generators. The design obtained using all non-equivalent glp generators can be transformed into designs in 3 dimensional ellipsoidal region using the transformation (8). The criteria values of all these designs are computed and are given in Table 3. For computing the criteria values a design, the points of the training set are obtained by using the generator

(1,13,27) and the values are obtained on the basis of 8000 evaluation points.

Table 3: Criteria values for all non-equivalent designs in 3 dimensional ellipsoidal region

Design Generator	RMSD	AD	MD
(1,3,5)	0.19366	0.17805	0.48729
(1,3,7)	0.18659	0.16590	0.59109
(1,3,9)	0.19181	0.17475	0.58436
(1,3,11)	0.19181	0.17475	0.58436
(1,3,13)	0.20657	0.18613	0.55194
(1,3,15)	0.19796	0.17902	0.49603
(1,3,19)	0.20066	0.18114	0.50394
(1,3,21)	0.19043	0.17420	0.55360
(1,3,25)	0.21661	0.19547	0.52553
(1,3,27)	0.20181	0.18323	0.49835
(1,3,29)	0.20210	0.18420	0.50646
(1,3,31)	0.34922	0.28035	0.99897
(1,3,33)	0.23548	0.20737	0.70056
(1,5,7)	0.20323	0.18267	0.60436
(1,5,9)	0.19300	0.17361	0.51461
(1,5,11)	0.19723	0.17434	0.54075
(1,5,19)	0.21192	0.18817	0.32439
(1,5,21)	0.21578	0.19718	0.66106
(1,5,27)	0.20802	0.18742	0.59109
(1,5,29)	0.35220	0.28477	1.01637
(1,5,31)	0.19961	0.18250	0.51467
(1,5,33)	0.22281	0.19422	0.66455
(1,9,13)	0.20751	0.18886	0.57064
(1,9,15)	0.19687	0.18306	0.46457
(1,9,21)	0.20429	0.18215	0.57527
(1,9,25)	0.34638	0.28098	0.44235
(1,9,31)	0.20321	0.18469	0.64441
(1,9,33)	0.21291	0.18723	0.61685
(1,11,13)	0.20094	0.17888	0.40342
(1,11,25)	0.20902	0.18785	0.54557
(1,11,27)	0.20441	0.18483	0.51833
(1,11,29)	0.20375	0.18224	0.53322
(1,13,21)	0.34145	0.27485	0.99529
(1,13,25)	0.20427	0.18413	0.52374
(1,13,27)	0.22557	0.19462	0.64485

From Table 3, it can be observed that the design with generators (1,3,7) has least 'RMSD' and 'AD' value whereas 'MD' value of design with generator (1,5,19) is less. Therefore, we consider all the generators equivalent to (1,3,7) and (1,5,19) and out of all the equivalent design, the design with minimum criteria value should be chosen. The distance based criteria values for designs with generator equivalent to (1, 3, 7) and (1,5,19) are given in Table 4.

Table 4: Criteria value of designs equivalent to designs with generators (1,3,7) and (1,5,19) in E^3_{34} (1)

Generator	Equivalent Generators	RMSD	AD	MD
(1,3,7)	**(1,3,7)**	**0.18659**	**0.1659**	0.59109
	(1,23,25)	0.19469	0.17739	0.56514
	(1,5,15)	0.21131	0.18756	0.54959
(1,5,19)	**(1,5,19)**	0.21192	0.18817	**0.32439**
	(1,7,31)	0.19808	0.17776	0.55356
	(1,9,11)	0.20558	0.18323	0.50081
	(1,9,31)	0.20321	0.18469	0.64441

From Table 4, it can be observed that the generator (1,3,7) gives the smallest value of 'RMSD' and 'AD', whereas the 'MD' value is minimum for the design obtained by generator (1,5,19). Therefore, the generator (1,3,7) can be taken as optimal generator. The scatter plots of the designs obtained by the generator (1,3,7) and (1,5,19) are displayed in Figure 2.

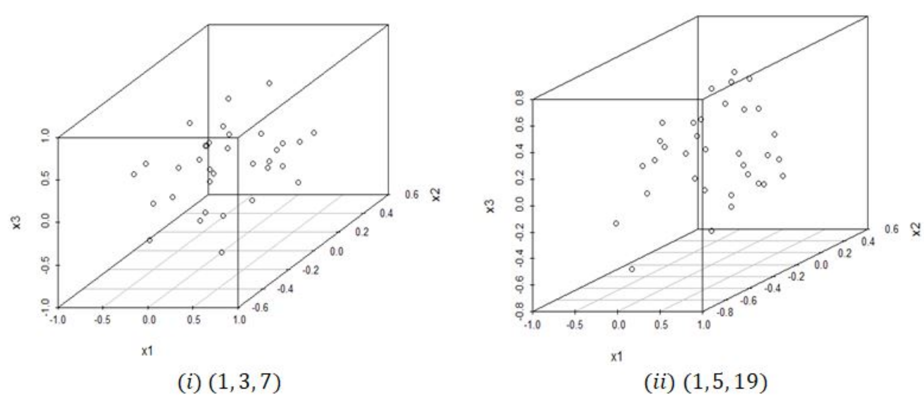

Figure 2: Scatter plot of the designs obtained by generator (i) (1,3,7) and (ii) (1,5,19)

From Figure 2, it can be observed that the design based on generator (1, 3, 7) is more uniform.

4. Uniform mixture designs for simplex

The designs in 2 and 3 dimensional ellipsoidal region can be obtain using the transformation (7) and (8) by using the methods discussed in Section 3. These designs are used to obtain the designs for mixture experiments as described in construction 2.

Construction 2:

Let $y = (y_1, y_2, ..., y_{s-1})$ represents the design in (s-1) dimensional ellipsoidal region, then a mixture design can be obtained from this using the following transformation:

For $i = 1, 2, ..., s-1$, define $\xi_i = \frac{e^{y_i}}{\sum_{i=1}^{s-1} e^{y_i}}$ and make the following transformation:

$$
\begin{aligned}
x_1 &= 1 - \xi_1, \\
x_i &= \prod_{j=1}^{i-1} \xi_j (1 - \xi_i); i = 2, 3, ...(s-1), \\
x_s &= \prod_{i=1}^{s-1} \xi_i
\end{aligned}
\tag{9}
$$

then $x = (x_1, x_2, ..., x_s)$ is the required s component mixture design.

4.1. Designs for three component mixtures

Consider the 2 dimensional ellipsoidal region and the two designs obtained by using the glp generators (1,2) and (1,8). Then,, the two 3 component mixture designs corresponding to these designs are obtained by using the transformation (9). The design in 2 dimensional ellipsoidal region obtained using generator (1,2) along with corresponding mixture design is given in Table 5 and that obtained using the generator (1,8) are given in Table 6.

Table 5: Design in two dimensional ellipsoidal region and three component mixture design using generator (1,2)

Design Point	Design in Ellipsoidal Region		Mixture Design		
	y_1	y_2	x_1	x_2	x_3
1	0.90039	0.01550	0.29216	0.50103	0.20680
2	0.49749	0.08660	0.39870	0.36156	0.23974
3	-0.07372	0.16323	0.55896	0.19451	0.24652
4	-0.60717	0.17769	0.68673	0.09814	0.21513
5	-0.91427	0.08572	0.73106	0.07233	0.19661
6	-0.89357	-0.10444	0.68764	0.09757	0.21479
7	-0.56645	-0.32669	0.55965	0.19391	0.24644
8	-0.06550	-0.48002	0.39783	0.36261	0.23956
9	0.41926	-0.47187	0.29088	0.50286	0.20627
10	0.71505	-0.26396	0.27309	0.52840	0.19851
11	0.98781	0.00677	0.27269	0.52898	0.19833
12	0.71890	0.13303	0.35758	0.41270	0.22972
13	0.21118	0.30723	0.52399	0.22658	0.24942
14	-0.33192	0.38897	0.67280	0.10706	0.22014
15	-0.71054	0.28749	0.73067	0.07254	0.19679
16	-0.80287	0.00000	0.69059	0.09574	0.21368
17	-0.60784	-0.38156	0.55633	0.19684	0.24683
18	-0.23948	-0.70300	0.38615	0.37681	0.23704
19	0.12391	-0.80979	0.28218	0.51527	0.20255
20	0.31319	-0.61497	0.28330	0.51366	0.20304
21	0.21449	-0.14549	0.41096	0.34696	0.24207

For the two, 3 component mixture designs given in Table 5 and Table 6, the DM_2 values are calculated which are given in Table 7.

Table 6: Design in two dimensional ellipsoidal region and three component mixture design using generator (1,8)

Design Point	Design in Ellipsoidal Region		Mixture Design		
	y_1	y_2	x_1	x_2	x_3
1	-0.62287	0.03490	0.65876	0.11645	0.22480
2	-0.07409	-0.12989	0.48606	0.26414	0.24981
3	0.72784	0.08097	0.34370	0.43073	0.22557
4	-0.96825	0.00000	0.72477	0.07575	0.19948
5	0.68812	-0.23447	0.28443	0.51204	0.20353
6	-0.07212	0.26117	0.58256	0.17426	0.24318
7	-0.56645	-0.32669	0.55965	0.19391	0.24644
8	0.98784	0.00665	0.27266	0.52903	0.19831
9	-0.75553	0.22801	0.72781	0.07409	0.19810
10	0.18219	-0.55978	0.32257	0.45891	0.21852
11	0.48412	0.21651	0.43349	0.32093	0.24558
12	-0.79955	-0.16141	0.65433	0.11949	0.22618
13	0.62012	-0.31476	0.28194	0.51562	0.20245
14	-0.33192	0.38897	0.67280	0.10706	0.22014
15	-0.26427	-0.64275	0.40649	0.35225	0.24126
16	0.89329	0.05651	0.30221	0.48691	0.21088
17	-0.78236	0.16924	0.72144	0.07760	0.20097
18	0.27639	-0.72169	0.26932	0.53389	0.19679
19	0.20888	0.33608	0.53176	0.21925	0.24899
20	-0.56868	-0.40866	0.53992	0.21167	0.24841
21	0.21449	-0.14549	0.41096	0.34696	0.24207

Table 7: DM_2 value of the 3 component mixture designs

Design Generator	DM_2 value
(1,2)	0.20811
(1,8)	0.21414

From the Table 7, it can be observed that the criteria value of design based on generator $(1, 2)$ has less DM_2 value and therefore it provides design with better uniformity.

4.2. Designs for four component mixtures

Consider the 3 dimensional ellipsoidal region and the designs obtained using the glp generators (1,3,7) and (1,5,19). Then, the two four component mixture designs corresponding to these two designs are obtained by using the transformation (9). The designs in 3 dimensional ellipsoidal region obtained using the generator (1,3,7) along with corresponding mixture design is given in Table 8 and that obtained using the generator (1,5,19) are given in Table 9.

For the two, 4 component mixture designs given in Table 8 and Table 9, the DM_2 values are calculated and are given in Table 10.

Table 8: Design in three dimensional ellipsoidal region and four component mixture design using generator (1,3,7)

Design Point	Design in Ellipsoidal Region			Mixture Design			
	y_1	y_2	y_3	x_1	x_2	x_3	x_4
1	0.32315	0.00440	0.02285	0.59476	0.28585	0.08356	0.03583
2	-0.41776	-0.05388	0.04178	0.75138	0.15968	0.05393	0.03501
3	0.00000	-0.10460	-0.06848	0.64720	0.24069	0.07517	0.03693
4	-0.18455	0.05609	-0.14718	0.69790	0.18601	0.07969	0.03641
5	-0.87265	0.01175	0.01864	0.82935	0.10013	0.04118	0.02934
6	-0.17402	-0.00316	0.24197	0.72989	0.18356	0.05111	0.03544
7	0.65706	0.14242	0.12940	0.54290	0.33221	0.09121	0.03368
8	0.21510	0.20547	-0.24715	0.61836	0.23739	0.10958	0.03468
9	0.08141	-0.19930	-0.32187	0.58735	0.28405	0.09315	0.03546
10	0.58516	-0.05417	0.11568	0.53553	0.35064	0.08077	0.03306
11	0.00000	0.00000	0.46683	0.72183	0.20079	0.04305	0.03433
12	-0.84944	-0.10222	0.02136	0.81818	0.11203	0.03948	0.03031
13	-0.30481	-0.18389	-0.16548	0.69494	0.20004	0.06820	0.03683
14	0.09769	0.36078	-0.20959	0.67066	0.18823	0.10693	0.03418
15	-0.29917	0.10470	0.15968	0.75489	0.15513	0.05508	0.03489
16	0.12267	-0.03837	0.41722	0.68689	0.22965	0.04837	0.03508
17	0.72112	-0.04459	0.04478	0.49328	0.38732	0.08863	0.03076
18	0.19336	0.09501	-0.50589	0.58390	0.25918	0.12447	0.03245
19	-0.18046	-0.45959	-0.35989	0.61422	0.27320	0.07628	0.03630
20	0.00000	-0.13930	0.39171	0.70145	0.22101	0.04329	0.03425
21	-0.11109	0.12545	0.65402	0.77355	0.16148	0.03335	0.03162
22	-0.29956	0.25188	-0.06935	0.74966	0.14156	0.07450	0.03428
23	-0.13234	-0.01170	-0.02828	0.69116	0.20123	0.07073	0.03688
24	0.41720	0.42709	-0.07987	0.61807	0.23461	0.11310	0.03423
25	0.22552	0.14049	0.22332	0.65710	0.23490	0.07105	0.03695
26	0.07597	-0.24125	0.34616	0.67088	0.25025	0.04486	0.03401
27	0.37766	-0.45551	-0.13673	0.50800	0.38678	0.07427	0.03095
28	0.00000	0.00000	-0.58311	0.60909	0.23810	0.11947	0.03334
29	-0.44680	-0.23511	-0.18059	0.71758	0.18385	0.06224	0.03633
30	-0.22162	-0.10609	0.62197	0.77514	0.16811	0.02709	0.02967
31	0.03699	0.34538	0.59061	0.75614	0.16291	0.04661	0.03434
32	-0.16099	0.59138	-0.38497	0.74499	0.11702	0.10987	0.02813
33	0.11301	-0.09508	-0.90978	0.53954	0.28827	0.14368	0.02851
34	0.96210	-0.09052	-0.09091	0.41104	0.46790	0.09619	0.02488

From the Table 10, it can observed that the criteria value for design using generator (1,3,7) has less DM_2 value and hence it provides more uniformly scattered design points.

5. Conclusion

In this paper, a method is proposed to construct uniform designs in ellipsoidal region and the designs in 2 and 3 dimensional ellipsoidal regions are obtained. It is observed

Table 9: Design in three dimensional ellipsoidal region and four component mixture design using generator (1,5,19)

Design Point	Design in Ellipsoidal Region			Mixture Design			
	y_1	y_2	y_3	x_1	x_2	x_3	x_4
1	-0.64748	-0.02036	-0.00724	0.79032	0.12737	0.04957	0.03274
2	-0.14593	0.04845	0.04592	0.70812	0.18841	0.06688	0.03659
3	0.46334	-0.03106	-0.10664	0.54033	0.33079	0.09538	0.03350
4	-0.16031	-0.01150	0.17158	0.71863	0.18950	0.05585	0.03602
5	-0.03201	-0.04087	-0.21982	0.64538	0.22997	0.08801	0.03663
6	-0.33227	0.08514	0.23184	0.76613	0.15084	0.04889	0.03413
7	0.75380	0.02321	-0.01646	0.48573	0.38689	0.09705	0.03032
8	-0.47411	-0.24313	0.05108	0.74687	0.17240	0.04617	0.03455
9	-0.34555	0.02579	0.02550	0.74352	0.16112	0.05992	0.03545
10	0.74003	-0.09205	-0.15444	0.45769	0.41433	0.09960	0.02837
11	-0.48051	-0.07523	0.30184	0.78661	0.14510	0.03643	0.03186
12	0.00000	0.16408	-0.43473	0.64611	0.20632	0.11376	0.03381
13	0.00000	0.00000	0.52155	0.72860	0.19774	0.03998	0.03368
14	0.29277	0.05266	-0.07405	0.59669	0.27537	0.09218	0.03575
15	-0.21416	-0.24806	0.19510	0.71201	0.20782	0.04541	0.03476
16	-0.32505	0.46508	-0.17310	0.77105	0.11344	0.08473	0.03079
17	0.72112	-0.04459	0.04478	0.49328	0.38732	0.08863	0.03076
18	-0.53092	-0.05777	0.15530	0.78219	0.14166	0.04320	0.03294
19	-0.07507	0.44871	-0.38620	0.70769	0.14804	0.11337	0.03089
20	0.31058	-0.05921	0.60236	0.66993	0.25480	0.04201	0.03326
21	-0.12218	-0.05377	-0.12884	0.67365	0.21231	0.07707	0.03697
22	-0.02910	-0.07000	0.32151	0.70413	0.21184	0.04873	0.03530
23	-0.16991	0.28774	-0.40189	0.70356	0.15756	0.10623	0.03265
24	0.50097	0.03714	0.35557	0.59897	0.29989	0.06607	0.03507
25	-0.19690	-0.67197	-0.18881	0.61977	0.29033	0.05544	0.03446
26	-0.05284	0.56681	-0.07181	0.73955	0.13439	0.09384	0.03221
27	0.42584	-0.03292	0.38274	0.61388	0.29188	0.05938	0.03485
28	-0.36576	-0.28817	-0.13497	0.70062	0.20252	0.06033	0.03652
29	0.00000	0.11247	0.35913	0.71840	0.19286	0.05295	0.03579
30	0.00000	0.00000	-0.54480	0.61240	0.23737	0.11646	0.03377
31	0.19961	0.04542	0.63163	0.70566	0.22008	0.04059	0.03367
32	-0.15955	-0.56293	-0.58508	0.56924	0.30680	0.08907	0.03489
33	-0.17986	0.44889	0.39813	0.78530	0.12826	0.05336	0.03308
34	0.96210	-0.09052	-0.09091	0.41104	0.46790	0.09619	0.02488

Table 10: DM_2 value of the 4 component mixture designs

Design Generator	DM_2 value
(1,3,7)	0.2942798
(1,5,19)	0.2977485

that the designs chosen by 'RMSD' and 'AD' criteria have better uniformity as compared to designs selected by 'MD' criterion. This may be because 'MD' is greatly influenced by points near the boundary and it fails to provide the representative points for the overall space. A transformation to construct mixture designs from the design in ellipsoidal region is also given in this paper. Designs for 3 and 4 component mixtures are constructed and the DM_2 criterion is used to measure the uniformity of these designs. The mixture design obtained by using the design selected in ellipsoidal region by 'RMSD' and 'AD' criteria has less discrepancy value as compared to that obtained by using the design selected by 'MD' criterion. The design which was found best in ellipsoidal region also possesses greater uniformity in simplex region. Hence, the transformation proposed in this paper preserves the uniformity when a design is transformed from the ellipsoidal region to simplex region.

6. Guidance to the beginners

The researchers who are new in this area can start with the understanding of the basic terminology and concepts provided in Section 1 and the references cited therein. They can begin with clearly identifying the number of components of the mixture and the region of interest. Here, an ellipsoidal region is considered. A suitable method needs to be developed for obtaining the design in the region of interest. Here, in Section 3, a method for obtaining the design in s-dimensional ellipsoidal region is given in steps 1-3 of Construction 1. The next step is to select a suitable measure of uniformity to obtain the criteria values for the generated designs and select the designs with lowest discrepancy. Section 2 can be referred for the distance-based approach and DM_2-discrepancy. The constructed designs for (s-1)-dimensional ellipsoidal region can then be utilized to generate designs for s component mixtures by using the construction 2 given in Section 4 and the best design is selected on the basis of DM_2-discrepancy described in Section 2. The method discussed in the paper is provided in the form of following steps:

1. Obtain initial generating vector $H_n = (h_1, h_2, ..., h_k)$ for a given n.
2. For the given number of factor s, obtain all possible generators of size s from H_n with $h_1 = 1$. This leads to $m = {}^{k-1}C_{s-1}$ distinct generators.
3. For each of the m generators, obtain the design $Z = (z_{ij})$, $i = 1, 2, \ldots, n$; $j = 1, 2, ..., s$ in unit hypercube.
4. For each of the m generators, obtain the design Y in the ellipsoidal region from Z.
5. Compute the RMSD, AD, and MD value of all the Y designs and select the designs with lowest criteria values and note their corresponding generators.
6. For each of m generators, obtain the mixture design X for s component mixture from all design Y in (s-1) dimensional ellipsoidal region.
7. Compute the DM_2 discrepancy value for all the mixture designs and select the design with lowest value of DM_2-discrepancy.

Acknowledgements

The authors are thankful to the editors and reviewers for their valuable comments and suggestions that helped in improvement of the papers. This research work supported by Faculty Research Programme (FRP) grant vide Ref. No./IoE/2023-24/12/FRP from Institute of Eminence (IoE), University of Delhi, Delhi.

Conflict of interest

The authors do not have any financial or non-financial conflict of interest to declare for the research work included in this article.

References

Adriana, M. M., Saúl, Z. M., Luis, M. G., and Carlos, A. C. C. (2022). Uniform mixture design via evolutionary multi-objective optimization, *Swarm and Evolutionary Computation*, **68**, 100979.

Al-Shyoukh, I., Yu, F., Feng, J., Yan, K., Dubinett, S., Ho, C. M., and Sun, R. (2011). Systematic quantitative characterization of cellular responses induced by multiple signals. *BMC Systems Biology*, **5**, 1–17.

Borkowski, J. J. and Piepel, G. F. (2009). Uniform designs for highly constrained mixture experiments. *Journal of Quality Technology*, **41**, 35–47.

Cornell, J. A. (2011). *Experiments with Mixtures: Designs, Models, and the Analysis of Mixture Data*. John Wiley and Sons.

Fang, K. T. (1980). The uniform design: Application of number-theoretic methods in experimental design. *Acta Mathematicae Applicatae Sinica*, **3**, 363–372.

Fang, K. T. and Wang, Y. (1994). *Number-theoretic Methods in Statistics*. CRC Press.

Fang, K.T., Liu, M.Q., Qin, H., and Zhou, Y.D. (2018). *Theory and Application of Uniform Experimental Designs*, Science Press and Springer, Beijing.

Hua, L. K. and Wang, Y. (1981). *Applications of Number Theory to Numerical Analysis*. Springer and Science Press, Berlin and Beijing.

Korobov, N. M. (1959a). The approximate computation of multiple integrals. *Doklady Akademii Nauk SSSR*, **124**, 1207–1210.

Korobov, N. M. (1959b). Computation of multiple integrals by the method of optimal coefficients. *Vestnik Moskovskogo Universiteta Sectia de Matematica, Astronomie, Fizica, Chimie*, **4**, 19–25.

Mandal, N. K., Pal, M., Sinha, B. K., and Das, P. (2015). Optimum mixture designs in a restricted region. *Statistical Papers*, **56**, 105–119.

Ning, J. H., Zhou, Y. D., and Fang, K. T. (2011a). Discrepancy for uniform design of experiments with mixtures. *Journal of Statistical Planning and Inference*, **141**, 1487–1496.

Ning, J.H., Fang, K.T., and Zhou, Y.D. (2011b). Uniform design for experiments with mixtures, *Communications in Statistics-Theory and Methods*, **40**, 1734–1742.

Obiri, M. Y. (2020). *Space-Filing Designs for Mixture/Process Variable Experiments* (Doctoral dissertation, Montana State University).

Prescott, P. (2008). Nearly uniform designs for mixture experiments. *Communications in Statistics—Theory and Methods*, **37**, 2095–2115.

Quenouille, M. H. (1953). *The Design and Analysis of Experiments*. Charles Griffin and Company, London, England.

Scheffé, H. (1958). Experiments with mixtures. *Journal of the Royal Statistical Society: Series B (Methodological)*, **20**, 344–360.

Scheffé, H. (1963). The simplex-centroid design for experiments with mixtures. *Journal of the Royal Statistical Society: Series B (Methodological)*, **25**, 235–251.

Singh, P. and Shukla, H. (2023). Uniform mixture designs using designs in 2-dimensional spherical region. *International Journal of System Assurance Engineering and Management*, **14**, 1888–1897.

Thompson, W. O. and Myers, R. H. (1968). Response surface designs for experiments with mixtures. *Technometrics*, **10**, 739–756.

Talke, I. S. and Borkowski, J. J. (2012). Generation of space-filling uniform designs in unit hypercubes. *Journal of Statistical Planning and Inference*, **142**, 3189–3200.

Talke, I. S. and Borkowski, J. J. (2016). Measures of uniformity for space-filling uniform designs in a spherical region. *International Journal of Experimental Design and Process Optimisation*, **5**, 23.

Tzeng, D. Y. and Berns, R. S. (2005). A review of principal component analysis and its applications to color technology. *Color Research & Application: Endorsed by Inter-Society Color Council, The Colour Group (Great Britain), Canadian Society for Color, Color Science Association of Japan, Dutch Society for the Study of Color, The Swedish Colour Centre Foundation, Colour Society of Australia, Centre Français de la Couleur*, **30**, 84–98.

Wang, Y. and Fang, K. T. (1981). A note on uniform distribution and experimental design, *KeXue TongBao*, **26**, 485–489.

Weyl, H. (1916). über die gleichverteilung der zahlem mod eins. *Annals of Mathematics*, **77**, 313–352.

Zhou, Y. D. and Xu, H. (2014). Space-filling fractional factorial designs. *Journal of the American Statistical Association*, **109**, 1134–1144.