Statistics and Probability Letters 104 (2015) 123–132

Contents lists available at ScienceDirect

Statistics and Probability Letters

journal homepage: www.elsevier.com/locate/stapro

Mixture discrepancy on symmetric balanced designs

A.M. Elsawah[a,b,*], Hong Qin[b]

[a] Department of Mathematics, Faculty of Science, Zagazig University, Zagazig 44519, Egypt

[b] Faculty of Mathematics and Statistics, Central China Normal University, Wuhan 430079, China

ARTICLE INFO

Article history:
Received 18 February 2015
Received in revised form 16 April 2015
Accepted 7 May 2015
Available online 15 May 2015

MSC:
62K05
62K15

Keywords:
Mixture discrepancy
Uniform designs
Symmetric balanced designs
Lower bounds

ABSTRACT

Mixture discrepancy is more reasonable than other discrepancies for measuring the uniformity from different aspects such as the intuitive view, the uniformity of sub-dimension projection, the curse of dimensionality and the geometric property of the kernel function. In this paper, we discuss in depth the mixture discrepancy as the uniformity measure for symmetric two-, three- and four-level balanced designs.

1. Introduction

Uniform design has been widely applied in many fields, such as manufacturing, system engineering, pharmaceutics and natural sciences. Uniform design is spread experimental points uniformly on the experimental domain. Much effort has been spent on the measurement of the design uniformity and the construction of uniform designs. While there exist several discrepancy measures, which can be used in evaluating the uniformity of designs, the problem of finding designs with minimum discrepancy on a super-rectangle is still very difficult due to the computational intractability of finding a globally optimal solution, and also due to the lack of good benchmarks. It is found that, if restricted on balanced lattice designs, or called U-type designs, the problem becomes relatively easier. It is for this reason that, we only discuss in this paper symmetrical U-type designs. A symmetrical balanced (U-type) design $D(n; s^m)$ corresponds to an $n \times m$ matrix $X = (x_1, \ldots, x_m)$ such that each column x_i takes values from a set of s integers, say $0, 1, \ldots, s-1$, equally often. By mapping $f : x \to (2x+1)/(2s)$, $x = 0, 1, \ldots, s-1$, the n runs are transformed into n points in $\Omega^m = [0, 1]^m$. Throughout this paper the class of n runs and m factors with s-level balanced (U-type) designs, denoted as $U(n; s^m)$.

The measure of uniformity plays a key role in the construction of uniform designs. There are several different discrepancies defined. Recently, Zhou et al. (2013) pointed out some unreasonable phenomena associated with the commonly used discrepancies in the literature such as the centered L_2-discrepancy (CD) and the wrap-around L_2-discrepancy (WD). They proposed a new measurement known as the mixture discrepancy (MD). They proposed that MD satisfies 7 criteria for assessing measures of uniformity introduced in Fang et al. (2006). Furthermore Zhou et al. (2013) pointed out some limitations of CD and WD through examples. That is, CD covers the points near the center insufficiently and this measurement will cause

* Corresponding author at: Department of Mathematics, Faculty of Science, Zagazig University, Zagazig 44519, Egypt.
E-mail addresses: a.elsawah@zu.edu.eg, a_elsawah85@yahoo.com (A.M. Elsawah).

http://dx.doi.org/10.1016/j.spl.2015.05.007
0167-7152/

some problems when data is high dimensional, *WD* is not sensitive for each level shift in a certain sense. These limitations may lead to some unreasonable results. *MD* performs well under these situations, compared to *CD* and *WD*. Meanwhile, this new discrepancy has a clear geometric meaning and a simple computational formula. This new discrepancy has many good properties and can avoid weaknesses of *CD* and *WD*. The analytical formula of the mixture discrepancy is given as follows. For a design $\mathfrak{D} \in U(n; s^m)$, its mixture discrepancy value, denoted as $MD(\mathfrak{D})$, can be expressed in the following form:

$$[MD(\mathfrak{D})]^2 = \left(\frac{19}{12}\right)^m - \frac{2}{n}\sum_{i=1}^{n}\prod_{k=1}^{m}\Delta_{ik} + \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\prod_{k=1}^{m}\nabla_{ijk}, \tag{1.1}$$

where $\nabla_{ijk} = \frac{15}{8} - \frac{1}{4}|x_{ik} - \frac{1}{2}| - \frac{1}{4}|x_{jk} - \frac{1}{2}| - \frac{3}{4}|x_{ik} - x_{jk}| + \frac{1}{2}|x_{ik} - x_{jk}|^2$, $\Delta_{ik} = \frac{5}{3} - \frac{1}{4}|x_{ik} - \frac{1}{2}| - \frac{1}{4}|x_{ik} - \frac{1}{2}|^2$, $x_{ik} = (2d_{ik} + 1)/2s$, $d_{ik} \in \{0, 1, \ldots, s-1\}$, $i = 1, \ldots, n$, $k = 1, \ldots, m$.

In this paper, we discuss in depth the mixture discrepancy for symmetric two-, three- and four-level balanced designs and give new analytical expressions. Based on these new formulations, we present new lower bounds of this discrepancy for symmetric two-, three- and four-level balanced designs, which can be used as benchmarks for searching uniform (optimal) designs. We describe necessary conditions for the existence of a uniform design meeting these lower bounds. On the other hand, our results give new lower bounds, which are more useful and sharper than the lower bounds of Zhou et al. (2013) and Ke et al. (2015) for two-level balanced designs and Ke et al. (2015) for three-level balanced designs.

2. Mixture discrepancy on two-level balanced designs

In this case, the variable Δ_{ik} can only take one possible value, i.e., $\left(\frac{305}{192}\right)$ and the variable Δ_{ijk} can also take two possible values, i.e., $\left(\frac{7}{4}\right)$ when $x_{ik} = x_{jk}$ or $\left(\frac{3}{2}\right)$ when $x_{ik} \neq x_{jk}$. Thus, for two-level balanced designs the formula (1.1) can be simplified to

$$[MD(\mathfrak{D})]^2 = \left(\frac{19}{12}\right)^m - 2\left(\frac{305}{192}\right)^m + \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\left(\frac{7}{4}\right)^{\vartheta_{ij}}\left(\frac{3}{2}\right)^{m-\vartheta_{ij}}$$

$$[MD(\mathfrak{D})]^2 = \left(\frac{19}{12}\right)^m - 2\left(\frac{305}{192}\right)^m + \frac{1}{n}\left(\frac{7}{4}\right)^m + \frac{1}{n^2}\left(\frac{3}{2}\right)^m\sum_{i=1}^{n}\sum_{j\neq i}^{n}\left(\frac{7}{6}\right)^{\vartheta_{ij}}, \tag{2.1}$$

where $\vartheta_{ij} = \sharp\left\{(i,j) : x_{ik} = x_{jk},\ k = 1, 2, \ldots, m\right\}$, $\sum_{i=1}^{n}\sum_{j\neq i}^{n}\vartheta_{ij} = nm(\ell - 1)$, $2\ell = n$, $\vartheta_{ii} = m$ and $\sharp\{A\}$ is the cardinality of the set A.

From (2.1) we will give the lower bound of $[MD(\mathfrak{D})]^2$ depends on the fact in the following lemma.

Lemma 1 (*Elsawah and Qin, 2015c*)**.** *Suppose $\sum_{i=1}^{n} z_i = c$ and z_i's are nonnegative, then for any positive γ, we have*

$$\sum_{i=1}^{n}\gamma^{z_i} \geq \gamma^{\sigma}(p + q\gamma),$$

where p and q are integers such that $p + q = n$, $\sigma = \lfloor \frac{c}{n} \rfloor$, $p\sigma + q(\sigma + 1) = c$ and $\lfloor \varepsilon \rfloor$ means the largest integer contained in ε.

Now, from (2.1) and Lemma 1 after simple arrangements as in Elsawah and Qin (2015a), we have the following lower bound of $[MD(\mathfrak{D})]^2$.

Theorem 1. *For any design $\mathfrak{D} \in U(n; 2^m)$, let $2\ell = n$, $\alpha + \beta = n(n-1)$, $\delta = \left\lfloor \frac{m(\ell-1)}{n-1} \right\rfloor$ and $\alpha\delta + \beta(\delta + 1) = nm(\ell - 1)$. Then, we have*

$$[MD(\mathfrak{D})]^2 \geq \mathcal{LB}_{2:1},$$

where

$$\mathcal{LB}_{2:1} = \left(\frac{19}{12}\right)^m - 2\left(\frac{305}{192}\right)^m + \frac{1}{n}\left(\frac{7}{4}\right)^m + \frac{1}{n^2}\left(\frac{3}{2}\right)^m\left(\frac{7}{6}\right)^{\delta}\left[\alpha + \beta\left(\frac{7}{6}\right)\right].$$

Proof. Since $\sum_{i=1}^{n}\sum_{j\neq i}^{n}\vartheta_{ij} = nm(\ell - 1)$ is a constant for any balanced design $\mathfrak{D} \in U(n; 2^m)$, from Lemma 1, it is straightforward to show $\sum_{i=1}^{n}\sum_{j\neq i}^{n}\left(\frac{7}{6}\right)^{\vartheta_{ij}} \geq \left(\frac{7}{6}\right)^{\delta}\left[\alpha + \beta\left(\frac{7}{6}\right)\right]$, which completes the proof.

Recently, Zhou et al. (2013) and Ke et al. (2015) obtained the following lower bounds of $[MD(\mathfrak{D})]^2$ for two-level balanced designs respectively.

Proposition 1 (*Zhou et al., 2013*)**.** *For any design $\mathfrak{D} \in U(n; 2^m)$, let $\omega_{n,i} = n \pmod{2^i}$. Then, we have*

$$[MD(\mathfrak{D})]^2 \geq \mathcal{LB}_{2:2},$$