Statistics and Probability Letters 114 (2016) 111–118

Contents lists available at ScienceDirect

Statistics and Probability Letters

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

Construction of Sudoku-based uniform designs with mixed levels

Hongyi Li [a,b], Kashinath Chatterjee [c], Bo Li [a], Hong Qin [a,*]

[a] College of Mathematics and Statistics, Central China Normal University, Wuhan 430079, China

[b] Normal College, Jishou University, Jishou 416000, China

[c] Department of Statistics, Visva-Bharati University, Santiniketan, India

A R T I C L E I N F O

Article history:
Received 17 March 2016
Received in revised form 20 March 2016
Accepted 20 March 2016
Available online 30 March 2016

MSC:
62K15
62K10
62K99

Keywords:
Sudoku design
Generalized discrete discrepancy
Mirror mapping
Foldover technique
Uniformity

A B S T R A C T

This paper considers the derivation of a new lower bound for the generalized discrete discrepancy of designs involving mixed level factors. Using mirror mapping and foldover techniques, a simple and efficient construction method of Sudoku-based mixed level uniform designs is proposed. Moreover, the properties of such Sudoku-based mixed level uniform designs are investigated.

1. Introduction

The research on the construction of uniform designs is attractive. In this regard, Li et al. (2014) considered the construction and analysis of Sudoku-based uniform designs with equal level factors. It is well known that Sudoku designs originated from a popular game named Sudoku puzzle. It is clear from the definition that Sudoku designs are U-type designs. These designs are so constructed that they have control of blocks as compared with the general U-type designs. This enables extensive application of these designs in various fields. Because of the special structure of Sudoku designs, many researchers paid attention to the study of these designs in the recent years. Xu et al. (2011) considered the construction of space-filling designs based on Sudoku designs and these designs achieve maximum uniformity in univariate and bivariate margins. Fontana (2011) studied the fractional factorial design generation based on Sudoku designs. Subramani (2012) extended the Sudoku designs to orthogonal (Graeco) Sudoku square designs similar to orthogonal (Graeco) Latin square designs and presented a simple method of constructing Graeco Sudoku square designs of odd order. Using graph theoretic techniques, Fontana (2014) developed a simple algorithm for uniform random sampling of Latin squares and Sudoku designs. Li et al. (2014) considered the construction and analysis of Sudoku-based uniform designs with equal level factors.

* Corresponding author.
E-mail address: qinhong@mail.ccnu.edu.cn (H. Qin).

http://dx.doi.org/10.1016/j.spl.2016.03.016
0167-7152/

A literature review reveals that the uniformity criterion received a great deal of attention in the field of design of experiments. Based on this criterion, Fang and Wang (1994) pioneered the study of uniform designs. Uniform designs are model independent and spread experiment points evenly over the experimental domain. The uniformity criterion favors designs with the smallest uniformity and can be applied to both regular and nonregular factorials. Fang et al. (2005) provided uniform designs which are useful space-filling designs. There are many measures to assess the uniformity of various designs. Among them, the discrete discrepancy measure due to Hickernell and Liu (2002) and Qin and Fang (2004) possesses nice properties. Chatterjee and Qin (2011) extended this measure to the generalized discrete discrepancy.

The present paper aims at obtaining a new lower bound for the generalized discrete discrepancy measure for designs with mixed levels. This lower bound can be used as a benchmark for searching uniform designs. Moreover, this paper aims at providing a simple but efficient method for the construction of Sudoku-based uniform designs with mixed level factors and also aims at studying the uniformity aspect of those designs measured by the generalized discrete discrepancy.

The paper is organized as follows. Section 2 provides the notations and preliminaries. In Section 3, a new lower bound for the generalized discrete discrepancy is provided. A construction algorithm of Sudoku-based designs is proposed in Section 4. Some illustrative examples are also provided in Section 4. Based on the squared discrete discrepancy measure, Section 5 deals with the derivation of the relationship between the constructed uniform designs and their subdesigns. We conclude with some comments in Section 6.

2. Notations and preliminaries

Here we use the symbol $\mathcal{D}(n; {m_1}^{s_1}{m_2}^{s_2})$ to denote a class of designs involving n runs and s $(=s_1 + s_2)$ factors where, without loss of generality, the first s_1 factors are each at m_1 levels and the last s_2 factors are each at m_2 levels. A design $d \in \mathcal{D}(n; {m_1}^{s_1}{m_2}^{s_2})$ is an $n \times s$ matrix where the entries of the first s_1 columns are taken from the set $\{1, 2, \dots, m_1\}$ and the entries of the last s_2 columns are taken from the set $\{1, 2, \dots, m_2\}$. It is to be noted that the rows of the matrix represent runs and the columns represent the factors. If each entry of every column of d appears equally often, the design d is then called a U-type design with mixed levels. Let $\mathcal{U}(n; {m_1}^{s_1}{m_2}^{s_2})$ be the class of such designs. Moreover, we use the symbol $\mathcal{U}(n; m^s)$ to denote the class of U-type designs where we have $m_1 = m_2 = m$. Designs belonging to the class $\mathcal{U}(n; m^s)$ are called symmetric U-type designs.

In this paper, the generalized discrete discrepancy is used as the measure of uniformity. For any design $d \in \mathcal{U}(n; {m_1}^{s_1} {m_2}^{s_2})$, the generalized discrete discrepancy value is denoted by $DD(d; a_1, b_1, a_2, b_2)$. Following Chatterjee and Qin (2011), we have

$$[DD(d; a_1, b_1, a_2, b_2)]^2 = -\left[\frac{a_1 + (m_1 - 1)b_1}{m_1}\right]^{s_1}\left[\frac{a_2 + (m_2 - 1)b_2}{m_2}\right]^{s_2} + \frac{{b_1}^{s_1}{b_2}^{s_2}}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\left(\frac{a_1}{b_1}\right)^{\psi_{ij}^{(1)}}\left(\frac{a_2}{b_2}\right)^{\psi_{ij}^{(2)}}, \quad (1)$$

where $a_1 > b_1 > 0$, $a_2 > b_2 > 0$, a_1, b_1, a_2, b_2 are all constants and $\psi_{ij}^{(t)}$ is the coincidence number between the ith and the jth rows of d corresponding to the s_t factors each at m_t levels, $t = 1, 2$.

When $a_1 = a_2 = a$, $b_1 = b_2 = b$, the generalized discrete discrepancy is exactly the same as the discrete discrepancy proposed by Fang et al. (2003) and Qin and Fang (2004). We use $DD(d; a, b)$ instead of $DD(d; a_1, b_1, a_2, b_2)$. By (1), $DD(d; a, b)$ can be rewritten as

$$[DD(d; a, b)]^2 = -\left[\frac{a + (m_1 - 1)b}{m_1}\right]^{s_1}\left[\frac{a + (m_2 - 1)b}{m_2}\right]^{s_2} + \frac{b^{s_1+s_2}}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\left(\frac{a}{b}\right)^{\psi_{ij}(d)}, \quad (2)$$

where $\psi_{ij}(d)$ is the coincidence number between the ith and the jth rows of d.

Moreover, when $m_1 = m_2 = m$, $s_1 + s_2 = s$ in (2), $[DD(d; a, b)]^2$ can be expressed as

$$[DD(d; a, b)]^2 = -\left[\frac{a + (m - 1)b}{m}\right]^{s} + \frac{b^s}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\left(\frac{a}{b}\right)^{\psi_{ij}(d)}. \quad (3)$$

The following definition will be helpful in developing the rest of the paper.

Definition 1. Suppose $p = m \times l$, where p, m and l are all integers larger than one. If the design X with p blocks is composed of a p order matrix which includes p distinct characters and the p characters appear exactly once in each row, each column and each block with order $l \times m$, then the design X is called as a Sudoku design with order p.

Note that the Sudoku design with order p can be decomposed l column blocks with order $p \times m$ or m row blocks with order $l \times p$. For example, Sudoku designs X_6 with order 6, X_8 with order 8 and X_9 with order 9 are respectively given in Tables 1–3. X_6 includes 3 column blocks with order 6×2 or 2 row blocks with order 3×6, X_8 includes 2 column blocks with order 8×4 or 4 row blocks with order 2×8, and X_9 includes 3 column blocks with order 9×3 or 3 row blocks with order 3×9.

3. A new lower bound for generalized discrete discrepancy

This section considers the derivation of a new lower bound for the generalized discrete discrepancy.