Journal of Statistical Planning and Inference 174 (2016) 1–10

Contents lists available at ScienceDirect

Journal of Statistical Planning and Inference

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

Aberration in qualitative multilevel designs

Roberto Fontana [a], Fabio Rapallo [b,*], Maria Piera Rogantin [c]

[a] Department of Mathematical Sciences, Politecnico di Torino, Italy

[b] Department of Sciences and Technological Innovation, Università del Piemonte Orientale, Italy

[c] Department of Mathematics, Università di Genova, Italy

ARTICLE INFO

Article history:
Received 16 October 2015
Received in revised form 28 January 2016
Accepted 15 February 2016
Available online 22 February 2016

MSC:
62K15
13P10

Keywords:
Algebraic statistics
Complex coding
Counts of factor levels
Fractional factorial designs
Generalized word-length pattern
Indicator function

ABSTRACT

Generalized Word Length Pattern (GWLP) is an important and widely-used tool for comparing fractional factorial designs. We consider qualitative factors, and we code their levels using the roots of the unity. We write the GWLP of a fraction $\mathcal{F}$ using the polynomial indicator function, whose coefficients encode many properties of the fraction. We show that the coefficient of a simple or interaction term can be written using the counts of its levels. This apparently simple remark leads to major consequence, including a convolution formula for the counts. We also show that the *mean aberration* of a term over the permutation of its levels provides a connection with the variance of the level counts. Moreover, using *mean aberrations* for symmetric s^m designs with s prime, we derive a new formula for computing the GWLP of $\mathcal{F}$. It is computationally easy, does not use complex numbers and also provides a clear way to interpret the GWLP. As case studies, we consider non-isomorphic orthogonal arrays that have the same GWLP. The different distributions of the *mean aberrations* suggest that they could be used as a further tool to discriminate between fractions.

1. Introduction

In design of experiments, Generalized Word-Length Pattern (GWLP) is an important tool for comparing fractional factorial designs. GWLP comes out from the Minimum Aberration criterion and from the notion of Word-Length Pattern (WLP). WLP was introduced in Fries and Hunter (1980) for regular designs with binary factors, and generalized to the non-regular case in Tang and Deng (1999). In Suen et al. (1997) the definition of WLP is extended to symmetrical multilevel regular fractions. For a regular fraction $\mathcal{F}$ of a full-factorial design $\mathcal{D}$ with m factors, the WLP of $\mathcal{F}$ is the sequence $A(\mathcal{F}) = (A_1(\mathcal{F}), A_2(\mathcal{F}), \dots, A_m(\mathcal{F}))$, where A_j is the number of defining words with length j. Such a measure of the degree of aliasing can be easily interpreted in the regular case. The WLP has been generalized for non-regular asymmetrical designs by Xu and Wu (2001) and named as GWLP, but it has a less evident meaning than in the regular case.

The aberration and the GWLP through the polynomial indicator function of the fraction $\mathcal{F}$ have been introduced in Li et al. (2003) and Cheng and Ye (2004) for two- and three-level cases respectively. In those papers the aberration of a simple or interaction term is defined as the square of the module of the corresponding coefficient of the indicator function, and the jth element $A_j(\mathcal{F})$ of the GWLP is the sum of the aberrations of the terms of order $j, j = 1, \dots, m$.

As demonstrated in Xu and Wu (2001), the GWLP does not depend on the choice of a particular orthonormal basis of the functions defined over $\mathcal{D}$, while the aberration does. Pistone and Rogantin (2008) use the complex coding of the factor levels to express the basis of the functions, and in particular of the indicator function. With this coding the coefficients of

* Corresponding author. Tel.: +39 0131 360 223; fax: +39 0131 360 250.
E-mail addresses: roberto.fontana@polito.it (R. Fontana), fabio.rapallo@uniupo.it (F. Rapallo), rogantin@dima.unige.it (M.P. Rogantin).

http://dx.doi.org/10.1016/j.jspi.2016.02.003
0378-3758/

the indicator function are related in a simple manner to many interesting properties of the fraction and allows us to define aberration and GWLP in a clear way. The complex coding is particularly useful in the case of qualitative factors, as assumed in this work. To simplify the computation, avoiding the use of complex numbers, Fontana and Pistone (2013) represent the coefficients using the counts of the levels appearing in each simple or interaction term. As general references for GWLP and its properties, the reader can refer to Mukerjee and Wu (2006) and Chen and Cheng (2012).

The practical use of the GWLP to discriminate among different designs is well known. Given two designs $\mathcal{F}_1$ and $\mathcal{F}_2$, the Generalized Minimum Aberration (GMA) criterion consists in the sequential minimization of the GWLP. $\mathcal{F}_1$ is better than $\mathcal{F}_2$ if there exists j such that $A_1(\mathcal{F}_1) = A_1(\mathcal{F}_2), \dots, A_j(\mathcal{F}_1) = A_j(\mathcal{F}_2)$ and $A_{j+1}(\mathcal{F}_1) < A_{j+1}(\mathcal{F}_2)$. Despite the fact that the GMA criterion is widely applied, the statistical meaning of the elements of the GWLP is somewhat unclear. In the original work of Xu and Wu (2001), the GWLP of a symmetrical design is written as the MacWilliams transform of the distance distribution. This result has been generalized in Qin and Ai (2007) to the case of multilevel designs. Under a different point of view, Grömping and Xu (2014) write the first non-zero element of the GWLP as the sum of the R^2 coefficients of suitably defined linear models. The connection between GWLP and Discrete Discrepancy has been investigated in Qin and Fang (2004). Katsaounis and Dean (2008) describe methods for screening the non-equivalence of designs, faster than those based on the indicator function. However, such methods mainly apply to two-level designs.

In this work we use the expression of the GWLP via the aberrations of the interaction terms of a given order. In turn, the aberrations are computed using only the level counts of the corresponding terms. We fully exploit such new expressions in two directions. First, we establish a convolution formula for the counts of the terms, in symmetrical s^m designs with s a prime number. Second, we introduce the *mean aberration* of a simple or interaction term, over the permutations of its levels. The mean aberration has a very simple expression, and it is easy to compute and to explain. Indeed, we show that the mean aberration is proportional to the variance of the level counts.

Moreover, we prove that for symmetrical s^m designs, s prime, the jth element $A_j(\mathcal{F})$ of GWLP is the sum of the mean aberrations of the terms of order $j, j = 1, \dots, m$ and therefore the mean aberrations produce an alternative decomposition of the GWLP. In our knowledge, the proposed formula is the simplest over all alternative expressions in literature. Nevertheless, in general, this property does not hold.

The paper is organized as follows. In Section 2 a short review of the algebraic theory of factorial designs is given. In Section 3 the convolution formula that expresses the relationships among the level counts is obtained. In Section 4 the mean aberration is defined, and its connection with the GWLP are studied. The main result of this paper states that the GWLP can be computed as the sum of mean aberrations (see Proposition 8). Section 5 is devoted to the comparison of fractions that have the same GWLP but different distributions of the mean aberrations. Finally in Section 6 we briefly describe some directions for future work.

2. Algebraic characterization of fractional designs

In this section, for ease in reference, we present some relevant results of the algebraic theory of fractional designs. The interested reader can find further information, including the proofs of the propositions, in Fontana et al. (2000) and Pistone and Rogantin (2008).

Let us consider an experiment which includes m factors.

Let us code the s_j levels of the jth factor by the s_jth roots of the unity $\omega_k^{(s_j)} = \exp\left(\sqrt{-1}\frac{2\pi}{s_j}\,k\right)$, $k = 0, \dots, s_j - 1$, $j = 1, \dots, m$. We denote such a factor by Ω_{s_j}, $\Omega_{s_j} = \left\{\omega_0, \dots, \omega_{s_j-1}\right\}$.

As $\alpha = \beta \bmod s$ implies $\omega_k^\alpha = \omega_k^\beta$, it is useful to introduce the residue class ring $\mathbb{Z}_s$ and the notation $[k]_s$ for the residue of $k \bmod s$. For integer α, we obtain $(\omega_k)^\alpha = \omega_{[\alpha k]_s}$. We also have $\omega_h\omega_k = \omega_{[h+k]_s}$. We drop the sub-$s$ notation when there is no ambiguity.

We denote by $\mathcal{D}$ the full factorial design with complex coding:

$$\mathcal{D} = \mathcal{D}_1 \times \cdots \mathcal{D}_j \cdots \times \mathcal{D}_m \quad \text{with } \mathcal{D}_j = \Omega_{s_j};$$

the cardinality of the full factorial design is $\#\mathcal{D} = \prod_{j=1}^m s_j$.

We denote by L the exponent set of the complex coded design $\{0, \dots, s_j - 1\}, j = 1, \dots, m$:

$$L = \mathbb{Z}_{s_1} \times \cdots \times \mathbb{Z}_{s_m}.$$

Notice that L is both the exponent set of the complex coded design and the integer coded design. The elements of L are denoted by $\alpha, \beta, \dots$:

$$L = \left\{\alpha = (\alpha_1, \dots, \alpha_m) : \alpha_j = 0, \dots, s_j - 1, j = 1, \dots, m\right\};$$

$[\alpha - \beta]$ is the m-tuple $\left([\alpha_1 - \beta_1]_{s_1}, \dots, \left[\alpha_j - \beta_j\right]_{s_j}, \dots, [\alpha_m - \beta_m]_{s_m}\right)$. The computation of the jth element is in the ring $\mathbb{Z}_{s_j}$.

In order to use polynomials to represent all the functions defined over $\mathcal{D}$, including counting functions, we define