ARTICLE IN PRESS

Journal of Statistical Planning and Inference ▮ (▮▮▮▮) ▮▮▮–▮▮▮

Contents lists available at ScienceDirect

Journal of Statistical Planning and Inference

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

Maximin power designs in testing lack of fit

Douglas P. Wiens

Department of Mathematical and Statistical Sciences; University of Alberta, Edmonton, Alberta, Canada T6G 2G1

ARTICLE INFO

Article history:
Received 26 February 2018
Received in revised form 25 June 2018
Accepted 21 July 2018
Available online xxxx

MSC:
primary 62K05
secondary 62F35
62J05

Keywords:
F-test
Finite design space
Maximin
Minimax
Non-centrality
Power

ABSTRACT

In a previous article (Wiens, 1991) we established a maximin property, with respect to the power of the test for Lack of Fit, of the absolutely continuous uniform 'design' on a design space which is a subset of $\mathbb{R}^q$ with positive Lebesgue measure. Here we discuss some issues and controversies surrounding this result. We find designs which maximize the minimum power, over a broad class of alternatives, in discrete design spaces of cardinality N. We show that these designs are supported on the entire design space. They are in general not uniform for fixed N, but are asymptotically uniform as $N \to \infty$. Several examples with N fixed are discussed; in these we find that the approach to uniformity is very quick.

© 2018 Elsevier B.V. All rights reserved.

1. Introduction

In Wiens (1991), henceforth referred to as [W], we studied the uniform 'design', as applied to design spaces S_c that are subsets of $\mathbb{R}^q$ – intervals, hypercubes, etc. – with positive Lebesgue measure. We call such design spaces *continuous*, to distinguish them from the finite, discrete design spaces considered in this article.

The uniform design on S_c is the absolutely continuous measure, with constant density $1/\int_{S_c} d\boldsymbol{x}$. Of course such a design must be approximated in order to implement it in an actual experiment. A contribution of [W] was that, in a sense made precise there and detailed in Section 2, the uniform design possesses an optimality property in the class of all designs on S_c – it maximizes the minimum power of the standard F-test for Lack of Fit (LOF) of a fitted linear regression model, with the minimum taken over a broad class of alternatives.

The theory in [W] has been adapted to justify the use of discrete uniform designs in numerous applications in the sciences. For its application to drug combination studies see the series of papers Tan et al. (2003), Fang et al. (2008), Tan et al. (2009), Fang et al. (2009) and Fang et al. (2015). The ideas in [W] have gained traction in the theory of artificial neural networks – see Zhang et al. (1998) – and reduced support vector machines – see Lee and Huang (2007). The theory has been extended to nonparametric regression models – Xie and Fang (2000) – and, also allowing for heteroscedasticity, by Biedermann and Dette (2001) and Bischoff and Miller (2006).

The continuous nature of S_c in this context has been controversial. Indeed Bischoff (2010) argues that it allows for classes of alternative regression models – as used both in [W] and in Biedermann and Dette (2001) – that are too broad for the optimality property to be asymptotically meaningful (when the continuous uniform design is viewed as the limit of discrete uniform designs); he proposes a restricted interpretation.

E-mail address: doug.wiens@ualberta.ca.

https://doi.org/10.1016/j.jspi.2018.07.007
0378-3758/© 2018 Elsevier B.V. All rights reserved.

That the richness of classes of alternatives as in [W] makes discrete designs inadmissible was noted in Wiens (1992, p. 355), where we state ‘ *Our attitude is that an approximation to a design which is robust against more realistic alternatives is preferable to an exact solution in a neighbourhood which is unrealistically sparse.*’ This remains our view. Nonetheless, in this article we suggest an alternate approach that we feel is less controversial. We take a *finite* design space $S = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N\}$ – here N can be arbitrarily large, allowing for at least a close approximation of the space of interest in an anticipated application. We obtain exact designs for small values of N – these are non-uniform – and show that the maximin designs are asymptotically uniform, as $N \to \infty$. Theory and examples show that this limit is approached very quickly.

In the next section we outline the mathematical framework, provide a reduction of the maximin problem to a simpler minimax problem, and prove the asymptotic optimality of the uniform design. Some solutions with N fixed are given in Section 3. Proofs are in the Appendix. The computing code is available from the author's personal web site.

2. Preliminaries

As far as possible we use notation as in [W], to which we refer the reader for background material relating the standard F-test for lack of fit to properties of the design. We denote by λ the uniform probability measure on S, viz. $\lambda(\boldsymbol{x}_i) = 1/N$, $i = 1, \ldots, N$. To facilitate comparisons with [W] we now write $\lambda_c(\boldsymbol{x})$ for the continuous uniform design on a continuous design space S_c.

For a design ξ on S we write $\xi_i = \xi(\boldsymbol{x}_i)$. An implementable design with n observations requires that $n\xi_i$ be an integer; we shall loosen this restriction and allow ξ to be any probability distribution on S. In particular, we include λ as a possible design.

For p-dimensional regressors $\boldsymbol{z}(\boldsymbol{x})$ we entertain a class of departures

$$E[Y(\boldsymbol{x})] = \boldsymbol{z}'(\boldsymbol{x})\boldsymbol{\theta} + f(\boldsymbol{x}) \quad \left(f \in \mathcal{F}_\eta^+\right), \tag{1}$$

(the ‘full’ models, in LOF terminology) from the fitted (‘reduced’) regression model

$$E[Y(\boldsymbol{x})] = \boldsymbol{z}'(\boldsymbol{x})\boldsymbol{\theta}. \tag{2}$$

In [W] we defined this class of functions on S_c by (i) $\int_{S_c} f^2(\boldsymbol{x})\, d\lambda_c(\boldsymbol{x}) \geq \eta^2$ and (ii) $\int_{S_c} \boldsymbol{z}(\boldsymbol{x}) f(\boldsymbol{x})\, d\lambda_c(\boldsymbol{x}) = \boldsymbol{0}_{p\times 1}$. We now adopt an analogous definition of $\mathcal{F}_\eta^+$ on the discrete design space S, viz.,

$$\int_S f^2(\boldsymbol{x})\, d\lambda(\boldsymbol{x}) = \frac{1}{N}\sum_{i=1}^N f^2(\boldsymbol{x}_i) \geq \eta^2, \tag{3a}$$

$$\int_S \boldsymbol{z}(\boldsymbol{x}) f(\boldsymbol{x})\, d\lambda(\boldsymbol{x}) = \frac{1}{N}\sum_{i=1}^N \boldsymbol{z}(\boldsymbol{x}_i) f(\boldsymbol{x}_i) = \boldsymbol{0}_{p\times 1}. \tag{3b}$$

Condition (3a) enforces a separation between the fitted and alternate models, so that the test has positive power, and (3b) ensures the identifiability of the regression parameters under (1), via

$$\boldsymbol{\theta} \overset{def}{=} \arg\min_{\boldsymbol{t}} \sum_{i=1}^N \left(E[Y(\boldsymbol{x}_i)] - \boldsymbol{z}'(\boldsymbol{x}_i)\boldsymbol{t}\right)^2.$$

This defines $\boldsymbol{\theta}$ uniquely, in the presence of (3b) and the requirement, made here, that the matrix $\mathbf{Z}_{N\times p} = [\boldsymbol{z}(\boldsymbol{x}_1) \vdots \cdots \vdots \boldsymbol{z}(\boldsymbol{x}_N)]'$ be of full column rank. We write $f_i = f(\boldsymbol{x}_i)$ and define $\boldsymbol{f} = (f_1, \ldots, f_N)'$ and $\boldsymbol{D}_\xi = diag(\xi_1, \ldots, \xi_N)$. Define as well

$$\boldsymbol{b}_{f,\xi} = \int_S \boldsymbol{z}(\boldsymbol{x}) f(\boldsymbol{x})\, d\xi(\boldsymbol{x}) = \sum_{i=1}^N \boldsymbol{z}(\boldsymbol{x}_i) f(\boldsymbol{x}_i)\xi_i = \mathbf{Z}'\boldsymbol{D}_\xi \boldsymbol{f},$$

$$\boldsymbol{B}_\xi = \int_S \boldsymbol{z}(\boldsymbol{x})\boldsymbol{z}'(\boldsymbol{x})\, d\xi(\boldsymbol{x}) = \sum_{i=1}^N \boldsymbol{z}(\boldsymbol{x}_i)\boldsymbol{z}'(\boldsymbol{x}_i)\xi_i = \mathbf{Z}'\boldsymbol{D}_\xi\mathbf{Z},$$

and assume that $\boldsymbol{B}_\xi$ is non-singular. Then as at (2.2) of [W] the non-centrality parameter (NCP) of the F-statistic for testing the LOF of the fitted model (2), with alternatives of the form (1), and using a design ξ, is proportional to

$$\mathcal{B}(f,\xi) = \boldsymbol{f}'\boldsymbol{D}_\xi\boldsymbol{f} - \boldsymbol{b}'_{f,\xi}\boldsymbol{B}_\xi^{-1}\boldsymbol{b}_{f,\xi}.$$

The power of the test is an increasing function of the NCP, as long as the F-statistic is stochastically increasing in this parameter. This monotonicity is well known to hold in finite samples under a Gaussian error distribution, and is at least asymptotically valid otherwise, under mild conditions.

In its alternate form

$$\mathcal{B}(f,\xi) = \sum_{i=1}^N \left(f(\boldsymbol{x}_i) - \boldsymbol{z}'(\boldsymbol{x}_i)\boldsymbol{B}_\xi^{-1}\boldsymbol{b}_{f,\xi}\right)^2\xi_i,$$

Download English Version:

https://daneshyari.com/en/article/11020325

Download Persian Version:

https://daneshyari.com/article/11020325

Daneshyari.com