Journal of the Egyptian Mathematical Society 25 (2017) 212–215

Contents lists available at ScienceDirect

Journal of the Egyptian Mathematical Society

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

Original article

Optimal designs for probability-based optimality, parameter estimation and model discrimination

N.M. Kilany

Faculty of Science, Menoufia University, Shebin El Koom, Egypt

ARTICLE INFO

Article history:
Received 15 July 2016
Revised 3 December 2016
Accepted 3 January 2017
Available online 26 January 2017

MSC:
62-XX
62Kxx
62K05

Keywords:
D-optimality
P-optimality
KL-optimality
Kullback–Leibler distance
Compound criteria

ABSTRACT

In this paper, a new compound optimality criterion will be introduced. This criterion called PDKL-optimality. The proposed criterion aimed to introduce designs satisfy maximum probability of success, efficient parameter estimation and true model. An equivalence theorem is stated and proved for PDKL-optimality criterion.

1. Introduction

The D-optimality criterion is the common criterion for achieving efficient parameter estimation. For more details about D-optimality, see [1–3]. D-optimal designs are conventional optimizations based on a chosen optimality criterion and the model that will be suitable.

In the literature, there are several optimality criteria for discriminate between models (Ds -, T- and KL-criteria). Each of these criteria becomes applicable under certain condition and situation. In the case of the experimenter want to discriminate between nested models, the Ds-criterion can be applied. Thus, for two nested regression models which differ by $s > 1$ parameters. T-optimality criterion introduced in [4,5] is a different method for discriminating between models. This criterion is useful for two or more regression models and applied on linear or nonlinear models. However, T-criterion must be used to discrimination homoscedastic models with Gaussian errors. Uciński and Bogacka [6] introduced an extension of T-criterion for non-homoscedastic errors. For discriminating between more generalized models with random errors following any distribution [7, 8] introduced the KL-criterion, which depends on the Kullback–Liebler distance. Moreover, the T-criterion is a special case of the KL-criterion in the homoscedastic case and the generalization provided by Uciński and Bogacka [6] (in the heteroscedastic case), when the error distribution is normal. Finally, the KL-criterion can be used when the rival models are nested or not, homoscedastic or heteroscedastic, and in the case of the distribution has normal error.

Sometimes, experimenters wish to maximize the probability of an outcome. To this aim, McGree and Eccleston [9] have proposed a P-optimality criterion, which provide a maximum probability of observing outcome. Moreover, there are situation when an experimenter may be interested to achieve multiple objectives. For this aim, a PDKL-optimality criterion will be derived in this paper. This criterion proposed a method of compound criteria to achieve designs to hold an efficient parameter estimation, true model and a high probability of favorite outcome.

The paper is organized as follows: Section 2 introduced a simple review for D-, KL-, P- optimum designs. Compound design criteria DKL- and DP-optimum designs are presented Section 3. Finally, a new criterion called PDKL-optimality will be derived and an equivalence theorem is proved in Section 4.

E-mail address: neveenkilany@hotmail.com

http://dx.doi.org/10.1016/j.joems.2017.01.002

2. D-, KL-, P-optimum designs

2.1. KL-optimum designs

López-Fidalgo et al. [8] introduced a criterion for discrimination between two models, which consider a generalization of T-criterion for the case of non-normal models. This criterion called KL-criterion and it is definition depend on the Kullback–Leibler distance between two statistical models.

Let y be a random variable and let $f_1(y, \mathrm{x}, \theta_1)$ and $f_2(y, \mathrm{x}, \theta_2)$ be two rival probability density functions of y, $\mathrm{x} \in \chi$ and on a vector of unknown parameters, $\theta_i \in \Theta_i, i = 1, 2$. Assuming that $f_1(y, \mathrm{x}, \theta_1)$ is the "true" model, then the KL distance between the true model $f_1(y, \mathrm{x}, \theta_1)$ and other model $f_2(y, \mathrm{x}, \theta_2)$ is

$$\mathcal{I}(f_1, f_2, \mathrm{x}, \ \theta_2) = \int_\chi f_1(y, \mathrm{x}; \theta_1) log\frac{f_1(y, \mathrm{x}; \theta_1)}{f_2(y, \mathrm{x}; \theta_2)} dy, \qquad \mathrm{x} \in \chi$$

where, an experimental condition x generated by the experimenter from a design ξ is a random variable (or a random vector) belongs to an experimental domain $\chi \subset \mathbb{R}^m, m \geq 1$.

The KL-optimality criterion is given by

$$I_{21}(\xi) = \min_{\theta_2 \in \Theta_2} \int_\chi \mathcal{I}(f_1, f_2, \mathrm{x}, \ \theta_2)\xi(d\mathrm{x}) \tag{1}$$

The KL-optimum design is the design maximizes $\mathrm{I}_{21}(\xi)$ and denoted by $\xi^*{}_{\mathrm{KL}}$.

For regular design $\xi^*{}_{21}$, López-Fidalgo et al. [8] prove that $\xi^*{}_{21}$is a KL-optimum design if f $\psi_{21}(x, \xi^*{}_{21}) \leq 0$, $x \in \chi$, where,

$$\psi_{21}(x, \xi) = \mathcal{I}\left(f_1, f_2, \mathrm{x}, \ \hat{\theta}_2\right) - \int_\chi \mathcal{I}\left(f_1, f_2, \mathrm{x}, \ \hat{\theta}_2\right)\xi(dx)$$

is the directional derivative of $I_{21}(\xi)$. The KL-efficiency of a design ξ relative to the optimum design $\xi^*{}_{21}$is

$$Eff_{21}(\xi) = \frac{\mathrm{I}_{21}(\xi)}{\mathrm{I}_{21}\left(\xi^*_{21}\right)} \tag{2}$$

2.2. D-optimum designs

D-optimality is the vital design criterion, introduced by [10], which interested of the quality of the parameter estimates. The idea of D-optimality depends on maximization of logarithm the determinant of the information matrix $M(\xi, \ \theta)$,$log|M(\xi, \ \theta)|$, or equivalently, minimizes logarithm determinant of the inverse of information matrix, $log|M^1(\xi, \ \theta)|$. In the general context for D-optimality [11] redefined the D-optimality criterion as follows:

$$\Phi_{D_i}[M_i(\xi, \theta_i)] = \begin{cases} log|M_i(\xi, \theta_i)| & \text{if } |M_i(\xi, \theta_i)| \text{ is nonsingular,} \\ -\infty & \text{if } |M_i(\xi, \theta_i)| \text{ is singular} \end{cases} \tag{3}$$

where $|M_i((\xi, \theta_i), \xi)| = \sum_{\mathrm{x} \in \chi} J_i(\mathrm{x}, \theta_i)\xi(\mathrm{x})$ is the information matrix corresponding to the probability density function $f_i(y, \mathrm{x}; \theta_i)$, $i = 1, 2$ and $J_i(\mathrm{x}, \theta_i)$ is the Fisher's information matrix for a single observation on y at x.

A design $\xi^*_{D_i}$ is a D-optimum design iff $\psi_{D_i}(\mathrm{x}, \xi^*_{D_i}) \leq 0$, $x \in \chi$, where

$$\psi_{D_i}(\mathrm{x}, \xi) = tr\left[M_i^{-1}(\xi, \theta_i)J_i(\mathrm{x}, \theta_i)\right] - q_i, \ \ i = 1, 2$$

is the directional derivative of the D-criterion function. The D-efficiency of any design ξ is given by

$$Eff_{D_i}(\xi) = \left(\frac{|M(\xi, \theta_i)|}{\left|M\left(\xi^*_{D_i}, \theta_i\right)\right|}\right)^{1/q_i} \quad i = 1, 2. \tag{4}$$

where q_i is the number of parameters for each model.

2.3. P-optimum designs

Often, experimenters request to obtain a maximum probability of an outcome. To this aim, McGree and Eccleston [9] have proposed a P-optimality criterion. P-optimality criterion is a criterion aimed to maximize a function of the probability of observing a particular outcome.

One of the forms of P-optimality which defined as a maximization of a weighted sum of the probabilities of success, which is defined as follows:

$$\Phi_P(\xi) = \sum\nolimits_{j=1}^n \pi_j(\boldsymbol{\theta}, \xi_j)w_j, \quad \text{for} \quad j = 1, 2, \ldots\ldots, n$$

where, $\pi_j(\boldsymbol{\theta}, \xi_j)$ is the j-th probability of success given by ξ_j and w_j is the experimental effort relating to the j-th support point. In this criterion, design weights have been included and will play a role in maximizing the probabilities.

For two rival models $f_1(y, x, \theta_1)$ and $f_2(y, x, \theta_2)$, we can defined the P-optimality criterion by the following function

$$\Phi_{P_i}(\xi) = \sum_{j=1}^n \pi_{ij}(\theta_i, \xi_j)w_j, \ \ i = 1, 2 \tag{5}$$

where $\pi_{ij}(\theta_i, \xi_j)$ is the j-th probability of success in the model $f_i(y, x; \theta_i)$ and θ_i are the parameters for the two possible models. A design $\xi^*_{P_i}$ is a P-optimum design for high probability of success for the model $f_i(y, x; \theta_i)$ iff $\psi_{P_i}(x, \xi^*_{P_i}) \leq 0$, $x \in \chi$, where

$$\psi_{P_i}\left(x, \xi^*_{P_i}\right) = \frac{\Phi_{P_i}(x) - \Phi_{P_i}\left(\xi^*_{P_i}\right)}{\Phi_{P_i}\left(\xi^*_{P_i}\right)}$$

is the directional derivative of $\Phi_{P_i}(\xi)$. The *P*- efficiency of a design ξ relative to the optimum design ξ^*_P is

$$Eff_{P_i}(\xi) = \frac{\sum_{j=1}^n \pi_{ij}(\theta_i, \xi_j)w_j}{\sum_{j=1}^n \pi_{ij}\left(\theta_i, \xi^*_{P_i}\right)w_j}, \ \ i = 1, 2 \tag{6}$$

3. Compound design criteria

There are situations when a practitioner may be interested in a multiple objectives. To achieve the possible objectives, compound criteria can be used. A compound criterion optimizes a combination of multiple objective functions molded by maximizing a weighted product of efficiencies. In this Section, the DKL- and DP-compound criteria will be presented. The aim of DKL-optimality is to obtain an efficient parameter estimation and true model and DP-optimality aimed to obtain an efficient parameter estimation with probability based optimality.

3.1. DKL-optimum designs

Tommasi [12] introduced the DKL-optimality criterion for dual objective; discrimination between two rival models and efficient estimation for their parameters. For discrimination between $f_1(y, \mathrm{x}; \theta_1)$ and $f_2(y, \mathrm{x}; \theta_2)$ models, two possible KL-criteria have been considered, namely $I_{21}(\xi)$ and $I_{12}(\xi)$, excepting the case of nested models, where the largest model must be considered as the true model.

The DKL-optimality defined as follows

$$\Phi_{DKL}(\xi) = \left(\frac{\mathrm{I}_{21}(\xi)}{\mathrm{I}_{21}\left(\xi^*_{21}\right)}\right)^{\alpha_1}\left(\frac{\mathrm{I}_{12}(\xi)}{\mathrm{I}_{12}\left(\xi^*_{12}\right)}\right)^{\alpha_2}\left(\frac{\left|M_1(\boldsymbol{\theta}, \xi)\right|}{\left|M_1\left(\boldsymbol{\theta}, \xi^*_{D_1}\right)\right|}\right)^{\alpha_3/q_1} \times \left(\frac{\left|M_2(\boldsymbol{\theta}, \xi)\right|}{\left|M_2\left(\boldsymbol{\theta}, \xi^*_{D_2}\right)\right|}\right)^{1-\alpha_1-\alpha_2-\alpha_3/q_2} \tag{7}$$