Scientific
Research

The Construction of Locally D-Optimal Designs by Canonical Forms to an Extension for the Logistic Model

Irene García Camacha Gutiérrez, Raúl Martín Martín

Mathematics Department, Castilla-La Mancha University, Toledo, Spain
Email: Irene.GarciaCamacha@uclm.es, Raul.MMartin@uclm.es

Abstract

Logistic regression models for binary response problems are present in a wide variety of industrial, biological, social and medical experiments; therefore, optimum designs are a valuable tool for experimenters, leading to estimators of parameters with minimum variance. Our interest in this contribution is to provide explicit formulae for the D-optimal designs as a function of the unknown parameters for the logistic model $\log\left(\eta(x,\theta)\big/\left(1-\eta(x,\theta)\right)\right)=\alpha+\beta x+\delta q$ where q is an indicator variable. We have considered an experiment based on the dose-response to a fly insecticide in which males and females respond in different ways, proposed in Atkinson *et al*. (1995) [1]. To find the D-optimal designs, this problem has been reduced to a canonical form.

Keywords

Binomial Response; D-Optimal Design; Canonical Form

1. Introduction

There are many natural phenomena or external factors to which males and females respond differently; this feature happens in most live species and has received several book length treatments. The interest of the present work focuses on a particular insect specie: flies. The experiment consists of supplying a dose of insecticide on and analysing its effectiveness. The characterization of this process is the impossibility of identifying the gender of the flies before and during the treatment application. The behaviour is studied on the total population and, due to the experimental differences on the response, it is considered that sex is distributed according to binomial

with success probability $\dfrac{1}{2}$. The theory of optimal design is used to calculate the optimal dose levels with a determinate probability of death.

To model the experiment, the logistic model for binary data was chosen. Denoting η as the probability of death, x as the logarithm of dose (in micrograms per millilitre) and y as the number of deaths, it must be:

$$y \mid x \sim Bi(1, \eta) \text{ with } \eta(x, \theta) = \frac{e^{\alpha + \beta x + \delta q}}{1 + e^{\alpha + \beta x + \delta q}}, a^* \leq x \leq b^*,$$

where q is a factor with values 0 for males and 1 for females and being $\theta = (\alpha, \beta, \delta)$ the unknown parameters vector. This model can be linearized by the logarithm of the probability ratio: $\log(\eta/(1-\eta)) = \alpha + \beta x + \delta q = \mu$ Atkinson *et al.* (1995) [1]. A wide range of dichotomous response-mechanisms can be expressed in terms of the previous model.

2. Optimum Design

The main objective of optimal experimental design is to select where and how many trials are necessary to be collected in order to achieve an optimal estimation of the model parameters. The observations will be taken into experimental region, χ, which will be in most cases a closed interval of the real line, $[a^*, b^*]$. On χ it is defined an approximated design as

$$\xi = \begin{Bmatrix} x_1 & x_2 & \cdots & x_n \\ p_1 & p_2 & \cdots & p_n \end{Bmatrix},$$

where p_i denotes a probability measure with support on $x_1, \cdots, x_n$. The number of support points is guaranteed to be finite, or at least it is possible to find one finite equivalent, as consequence of Caratheodory's theorem. Due to the continuity of probability measure chosen, it may occur not to obtain an exact number of trials for each dose level. However, the theory assures us that when the number of observations is sizeable, the integer approximation of the values is enough good.

To proceed with this type of problems, a certain functional from the called Fisher information matrix is required to achieve its maximum value. According to the considered model, it is built the information matrix to a single observation x on an insect of known sex as

$$M_{i,j}(x, \theta) = \omega(x, \theta) \frac{\partial \mu}{\partial \theta_i} \frac{\partial \mu}{\partial \theta_j},$$

being $\omega(x, \theta) = \lambda(x, \theta) = \eta(1 - \eta)$ the variance used in fitting a generalized model by weighted least squares (McCullagh and Nelder, 1989 [2]). Nevertheless, the matrix just defined does not take into account the differences on the response previously mentioned. In spite of the experimental limitations about the lack of sex knowledge, it is possible to modify the above matrix to consider this uncertainly as it is shown:

$$M(x, \theta) = 0.5\omega_M(x, \theta) \begin{pmatrix} 1 & x & x^2 \\ x & x^2 & x \\ 1 & x & 1 \end{pmatrix} + 0.5\omega_F(x, \theta) \begin{pmatrix} 1 & x & 0 \\ x & x^2 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

In this work, the function of the information matrix chosen is D-optimality criteria. This criterion minimizes the volume of the confidence ellipsoid of the parameters and it is given by $\phi_D[M(\xi, \theta)] = |M(\xi, \theta)|$ or equivalently $\log|M(\xi, \theta)|$ which has the advantage of being a concave function. This optimality criterion has received much attention in the literature because of its direct interpretation (Silvey, 1980 [3]).

3. Canonical Forms

The use of canonical forms in the present paper greatly simplifies the process to obtain the optimal one. If the

problem is transformed through a suitable choice from χ to Z, the dependence of the optimal design on the true value of θ for given design space χ will be replaced, in the transformed problem for arbitrary Z, which will vary with θ. Then, if we are able to solve the transformed problem, we have implicitly solved the original design problem (Ford *et al.* (1992) [4]). The invariance of the design criteria chosen by the transformation is an indispensable requirement for performing this canonical version of the problem.

First, we consider the chosen model as a distribution function which is denoted by $\eta(\theta, x) = F(\alpha + \beta x + q\delta)$.

Let's suppose the insect sex is known. The following argument is valid for both males and females, so we will denote with the factor q an individual regardless of gender. Applying the change of variable $z = \alpha + \beta x$, the original problem can be reformulated as

$$\eta(z, \theta') = \frac{e^{z+q\delta}}{1+e^{z+q\delta}},$$

with $\alpha + \beta a^* \le z \le \alpha + \beta b^*$ or $\alpha + \beta a^* \ge z \ge \alpha + \beta b^*$ depending on β's sign. The information matrix can be built using the chain's rule for n dose levels as below:

$$M_x(\xi) = \sum_{i=1}^{n} \frac{p_i H(x_i) H(x_i)^t}{F(z_i + q\delta)(1 - F(z_i + q\delta))}$$

$$= \sum_{i=1}^{n} p_i \omega(z_i + q\delta) \begin{pmatrix} 1 \\ x_i \\ q \end{pmatrix} \begin{pmatrix} 1 & x_i & q \end{pmatrix} = \sum_{i=1}^{n} p_i v(x_i) v(x_i)^t,$$

being

$$H(x_i) = \left\{ \frac{\partial F(z_i + q\delta)}{\partial(z_i + q\delta)} \frac{\partial(z_i + q\delta)}{\partial \alpha}, \frac{\partial F(z_i + q\delta)}{\partial(z_i + q\delta)} \frac{\partial(z_i + q\delta)}{\partial \beta}, \frac{\partial F(z_i + q\delta)}{\partial(z_i + q\delta)} \frac{\partial(z_i + q\delta)}{\partial \delta} \right\}^t$$

the partial-derivative vector from the chain's rule, f the probability density function and $v(x_i) = \sqrt{\omega(z_i + q\delta)} \begin{pmatrix} 1 & x_i & q \end{pmatrix}^t$. Note we have defined

$$\omega(z_i + q\delta) = \frac{f(z_i + q\delta)^2}{F(z_i + q\delta)(1 - F(z_i + q\delta))}$$

abusing notation. The previous transformation can be expressed in matrix form as

$$\begin{pmatrix} 1 \\ z \\ q \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ \alpha & \beta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ x \\ q \end{pmatrix} = B \begin{pmatrix} 1 \\ x \\ q \end{pmatrix},$$

and defining,

$$g(z) = Bv(x)$$

$$= \frac{f(z_i + q\delta)}{\sqrt{F(z_i + q\delta)(1 - F(z_i + q\delta))}} \begin{pmatrix} 1 & 0 & 0 \\ \alpha & \beta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ x \\ q \end{pmatrix}$$

$$= \frac{f(z_i + q\delta)}{\sqrt{F(z_i + q\delta)(1 - F(z_i + q\delta))}} \begin{pmatrix} 1 \\ z \\ q \end{pmatrix},$$

the partial-derivative vector $v(x)$ can be expressed in terms of z through $v(x) = B^{-1}g(z)$ since B is invertible for being $\beta \ne 0$. Let us write the information matrix using the formula for change of basis as $M_x(\xi) = B^{-1}M_z(\xi)B$ (Fedorov, 1972 [5]). Since the D-optimal criterion does not vary by non-singular linear

transformations of the design space, the maximization problem of $M_x(\xi)$ determinant reduces to maximize $M_z(\xi)$. Hereafter, it will only work with the information matrix depending on z and then it will carry the inverse change out for solving the original problem without loss of concept. Note the parameter dependence has been considerably reduced:

$$M_z(\xi) = \sum_{i=i}^{n} p_i \omega(z_i + q\delta) \begin{pmatrix} 1 \\ z_i \\ q \end{pmatrix} (1 \quad z_i \quad q) = \sum_{i=1}^{n} p_i g(z_i) g(z_i)^t.$$

By adding uncertainty about sex, the information matrix to the design $\xi = \begin{Bmatrix} z_1 & z_2 & \cdots & z_n \\ p_1 & p_2 & \cdots & p_n \end{Bmatrix}$ with

$\sum_{i=1}^{n} p_i = 1$ for n dose levels results

$$M_z(\xi) = 0.5 \sum_{i=1}^{n} p_i \left[g_F(z_i) g(z_i)^t + g_M(z_i) h(z_i)^t \right]$$

being,

$$g_F(z_i) = \begin{pmatrix} \sqrt{\omega(z_i + \delta)} \\ z_i \sqrt{\omega(z_i + \delta)} \\ \sqrt{\omega(z_i + \delta)} \end{pmatrix} \text{ and } g_M(z_i) = \begin{pmatrix} \sqrt{\omega(z_i)} \\ z_i \sqrt{\omega(z_i)} \\ 0 \end{pmatrix}.$$

The above expression can be reformulated into a more general formula

$$M_z(\xi) = 0.5 \sum_{i=1}^{2n} p_i g(z_i) g(z_i)^t, \tag{1}$$

where the z_i's are repeated to consider the information by each possibility about sex since their partial derivatives are different. The novelty of this paper is described in the previous lines. The simplified formulation of the information matrix will allow us to achieve analytical expressions for the optimal weights for several cases in the next section.

4. Results

The aim of our study is to select where and how many observations must be collected in order to reduce the volume of confidence ellipsoid of the parameters as possible. For this purpose, it is necessary to obtain an expression of the determinant of the information matrix which will be maximized later. To calculate the determinant to the information matrix written as (1), it is possible to apply the following formula to its explicit expression given by Ardanuy *et al*. (1999) [6]:

$$|M_z(\xi)| = \left| \sum_{i=1}^{2n} p_i g(z_i) g(z_i) \right| = \sum_{i_1 < \cdots < i_m} p_{i_1} \cdots p_{i_m} \cdot \left| \left\{ g_k(z_{i_j}) \right\} \right|^2, \tag{2}$$

being m the number of parameters and i_j the elements of S_{2n}, the symmetrical group of $2n$-order permutations. We focus the study to calculate optimal designs to a small size of observations, two and three design points. The reason for this choice is that the upper and lower bound for the number of point suffers a slight modification due to each trial has two reading. Then the system will be non-singular if there are at least $\frac{m+1}{2}$ point since m is odd and the Caratheodory upper's bound will be halved with $\frac{m(m+1)}{4}$. The only possibilities are the considered in the present work.

4.1. Two Points Design

Using the formula (2), it is possible to obtain the following expression:

$$|M_z(\xi)| = 0.5\left[p^2(1-p)\begin{vmatrix} \sqrt{\omega(z_1+\delta)} & \sqrt{\omega(z_1)} & \sqrt{\omega(z_2+\delta)} \\ z_1\sqrt{\omega(z_1+\delta)} & z_1\sqrt{\omega(z_1)} & z_2\sqrt{\omega(z_2+\delta)} \\ \sqrt{\omega(z_1+\delta)} & 0 & \sqrt{\omega(z_2+\delta)} \end{vmatrix}^2\right.$$

$$+ p^2(1-p)\begin{vmatrix} \sqrt{\omega(z_1+\delta)} & \sqrt{\omega(z_1)} & \sqrt{\omega(z_2)} \\ z_1\sqrt{\omega(z_1+\delta)} & z_1\sqrt{\omega(z_1)} & z_2\sqrt{\omega(z_2)} \\ \sqrt{\omega(z_1+\delta)} & 0 & 0 \end{vmatrix}^2$$

$$+ p(1-p)^2\begin{vmatrix} \sqrt{\omega(z_1+\delta)} & \sqrt{\omega(z_2+\delta)} & \sqrt{\omega(z_2)} \\ z_1\sqrt{\omega(z_1+\delta)} & z_2\sqrt{\omega(z_2+\delta)} & z_2\sqrt{\omega(z_2)} \\ \sqrt{\omega(z_1+\delta)} & \sqrt{\omega(z_2+\delta)} & 0 \end{vmatrix}^2 \tag{3}$$

$$\left.+ p(1-p)^2\begin{vmatrix} \sqrt{\omega(z_1)} & \sqrt{\omega(z_2+\delta)} & \sqrt{\omega(z_2)} \\ z_1\sqrt{\omega(z_1)} & z_2\sqrt{\omega(z_2+\delta)} & z_2\sqrt{\omega(z_2)} \\ 0 & \sqrt{\omega(z_2+\delta)} & 0 \end{vmatrix}^2\right]$$

$$= 0.5\left[p^2(1-p)(A+B) + p(1-p)^2(C+D)\right]$$

$$= 0.5\left[p^2(1-p)E + p(1-p)^2 F\right],$$

where E and F are the squares of the determinants which result from combining the column matrices $g_F(z_i), g_F(z_i)$ with $i=1,2$ and operating conveniently. An analytical expression to the optimal weights can be obtained calculating the critical points for the last expression:

$$\frac{\partial|M_z(\xi)|}{\partial p} = \left[3(F-E)p^2 + 2(E-2F)p + F\right] = 0 \Rightarrow p^* = \frac{2F-E\pm\sqrt{F^2-EF+E^2}}{3(F-E)}.$$

To compute the optimal points z_1, z_2 will be enough to replace p^* in the determinant expression and to use a maximization routine. Note the parameter dependence still persists due to non-linearity of the model. To avoid such dependence, Chernoff (1953) [7] suggests providing a prior guess for the parameters $\theta_0 = (\alpha_0, \beta_0, \delta_0)$. In that sense, the achieved designs in the present work will be called locally D-optimal. The parameter estimates were obtained in Atkinson *et al.* (1995) [1]. The last step will be to calculate the value of the original variables through inverse change $x_i^* = \dfrac{z_i^* - \alpha_0}{\beta_0}$. The obtained result is shown in **Table 1**.

4.2. Three Points Design

It is known from other previous works that it can consider a symmetrical design for the present case. Let's us assume such design $\xi = \left\{\begin{matrix} z_1 & z_2 & z_3 \\ p & 1-2p & p \end{matrix}\right\}$. Applying the determinant formula and proceeding analogously as (3), it is possible to achieve a tractable expression for the determinant and an analytical formula for the optimal weights in a way:

$$\frac{\partial|M_z(\xi)|}{\partial p} = \left[3Ap^2 + 2Bp + C\right] = 0 \Rightarrow p^* = \frac{-B\pm\sqrt{B^2-3AC}}{3A},$$

being *A*, *B* and *C* the squares of the determinants which result from applying the formula to its fast calculation grouped conveniently. Taking into account the previous considerations, the results are shown in **Table 2**.

An advantage of working with approximate designs is that the optimality of a design can be easily checked. Kiefer and Wolffowitz (1960) [8] provided a main result for optimal experimental design theory: the general equivalence theorem. The central point of this theorem establishes a necessary condition to check whether a proposed design is *D*-optimal or not. It consists of verifying if the standardized variance of the prediction, which is known as sensitivity function and is defined as

$$d(x,\xi) = f^t(x)M^{-1}(\xi)f(x),$$

is equal to number of parameters in the support points of the design. However, for the present case it is not possible to apply directly the above formula due to lack of knowledge about gender. It requires to use an extension of the equivalence theorem motivated by Chaloner and Larntz (1989) [9]:

$$d(x,\xi^*) = 0.5w_M(x)f_M^t(x)M^{-1}(\xi)f_M(x) + 0.5w_F(x)f_F^t(x)M^{-1}(\xi)f_F(x) \le m.$$

The results shown in **Figure 1** allows us to validate the optimal designs proposed in this work. As we can observe from the figures, the points where the sensitivity function intersects the number-parameter line represent the optimal points obtained with the procedure described in this paper.

5. Conclusions

In this paper, it is proposed the use of canonical forms to solve a problem non-standard of optimal experimental designs laid out by Atkinson *et al.* (1995) [1] upon calculating the optimum dose of a fly insecticide. The main difficulty arises by adding uncertainty about gender since they differ in the response and the experiment only senses applied on the whole population. The witty transformation of the problem to a canonical version reduces the parameter dependence leading to analytical expression of the optimal weights. From these, we are able to compute D-optimal designs for several cases. In particular, it is constructed optimal designs for two and three dose levels.

Regarding future work, we will try to take advantage of the transformation geometry, $G(Z)$, for identifying

Table 1. Locally D-optimal designs for two points design.

$\alpha_0 = 1.804$, $\beta_0 = 1.1757$, $\gamma_0 = -1$		
x_i^*	−1.757	−0.018
p_i^*	0.5	0.5

Table 2. Locally D-optimal designs for three points design.

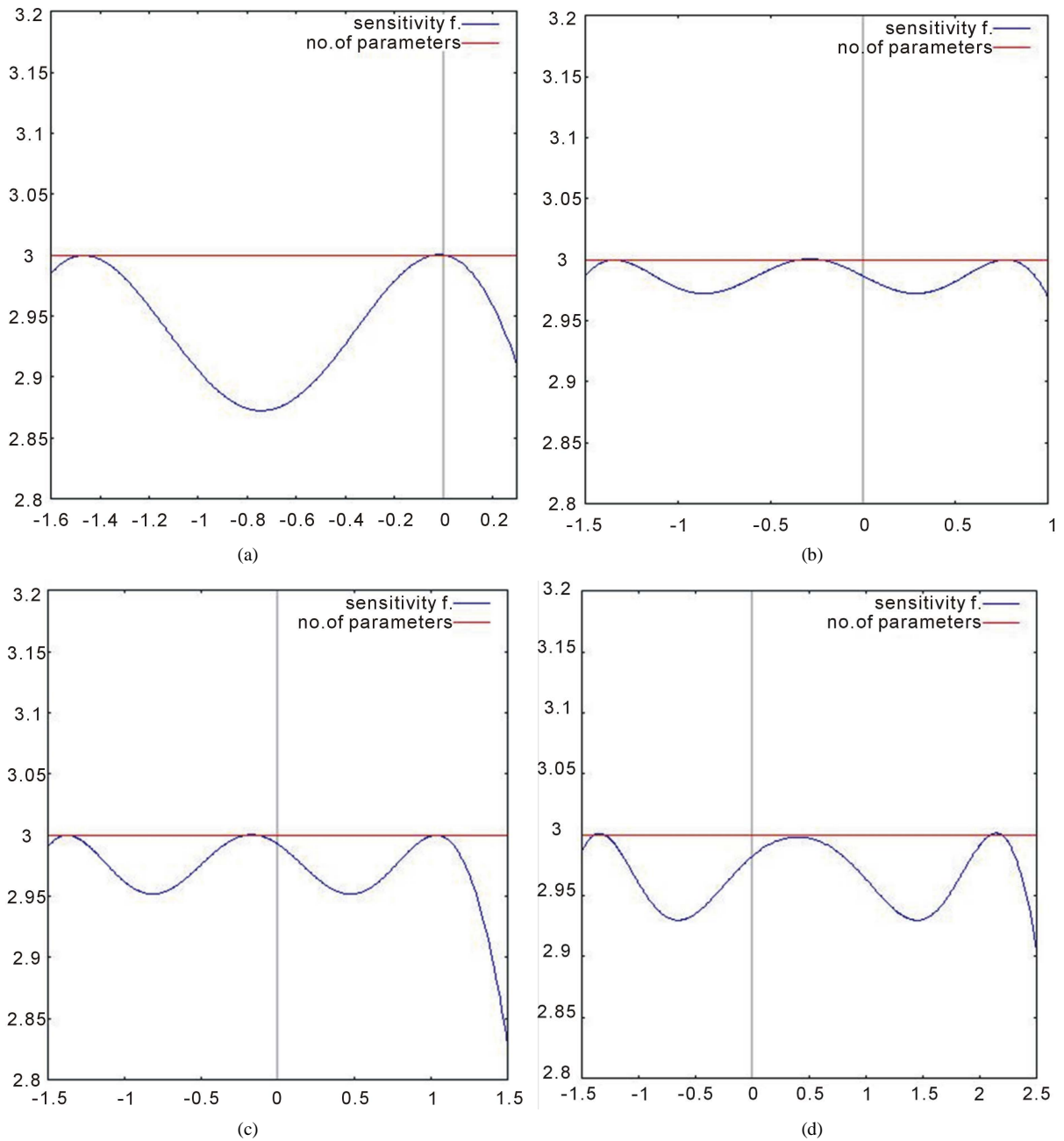| $\alpha_0 = 1.804$, $\beta_0 = 1.1757$				
$\gamma_0 = -2.599$	x_i^*	−1.347	−0.287	0.772
	p_i^*	0.375	0.250	0.375
$\gamma_0 = -3$	x_i^*	−1.378	−0.173	1.032
	p_i^*	0.339	0.322	0.339
$\gamma_0 = -5$	x_i^*	−1.349	0.396	2.141
	p_i^*	0.316	0.368	0.316

Figure 1. Testing locally D-optimal designs ($\alpha_0 = 1.804$, $\beta_0 = 1.1757$). (a) 2 points: $\gamma_0 = -1$; (b) 3 points: $\gamma_0 = -2.599$; (c) 3 points: $\gamma_0 = -3$; (d) 3 points: $\gamma_0 = -5$.

the support point. It is known that these are the points of contact between G and the smallest ellipsoid centred on the origin containing G (Sibson, 1972 [10], Silvey and Tittetington, 1973 [11], Silvey, 1980 [3], and Torsney and Musrati, 1993 [12]). However, this procedure must be adapted non-trivially to add two readings by observation with their corresponding probabilities.

References

[1]   Atkinson, A.C., Demetrio, C.G.B. and Zochhi, S.S. (1995) Optimun Dose Levels When Males and Females Differ en Response. *Applied Statistics*, **44**, 213-226.

[2]   McCullagh, P. and Nelder, J.A. (1992) Generalized Linear Models. Chapman and Hall, London.

[3]   Silvey, S.D. (1980) Optimal Design. Chapman and Hall, London. http://dx.doi.org/10.1007/978-94-009-5912-5

[4]   Ford, I., Tosney, B. and Wu, C.F.J. (1992) The Use of a Canonical form in the Construction of Locally Optimal Designs for Non-Linear Problems. *Journal of the Royal Statistical Society*: *Series B*, **54**, 569-583.

[5]   Fedorov, V.V. (1972) Theory of Optimal Experiments. Academic Press, New York.

[6]   Ardanuy, R., Lopez-Fidalgo, J., Laycock, P.J. and Wong, W.K. (1999) When Is an Equally-Weihted Design D-Optimal? *Annals of the Institute of Statistical Mathematics*, **51**, 531-540. http://dx.doi.org/10.1023/A:1003954207112

[7]   Chernoff, H. (1953) Locally Optimal Designs for Estimating Parameters. *The Annals of Mathematical Statistics*, **24**, 586-602. http://dx.doi.org/10.1214/aoms/1177728915

[8]   Kiefer, J. and Wolfowitz, J. (1960) The Equivalence of Two Extremun Problems. *Canadian Journal of Mathematics*, **12**, 363-366.

[9]   Chaloner, K. and Larntz, K. (1989) Optimal Bayesian Design Applied to Logistic Regression Experiments. *Journal of Statistical Planning and Inference*, **21**, 191-208. http://dx.doi.org/10.1016/0378-3758(89)90004-9

[10]   Sibson, R. (1972) Discussion on Results in the Theory and Construction of D-Optimun Experimental Designs (by H. P. Wynn). *Journal of the Royal Statistical Society*, **34**, 174-175.

[11]   Silvey, S.D. and Titterington, D.M. (1973) A Geometrical Approach to Optimal Design Theory. *Biometrika*, **60**, 21-32. http://dx.doi.org/10.1093/biomet/60.1.21

[12]   Torsney, B. and Musrati, A.K. (1993) On the Construction of Optimal Designs with Applications to Binary Response and to Weighted Regression Models. *Model Oriented Data Analysis*, Physica-Verlag, Heidelberg.