AP-Optimum Designs for Minimizing the Average Variance and Probability-Based Optimality

Authors: N. M. Kilany
– Faculty of Science, Menoufia University
Menoufia, Egypt. (neveenkilany@hotmail.com)

W. A. Hassanein
– Faculty of Science, Tanta University
Tanta, Egypt. (wfanwar@science.tanta.edu.eg)

Abstract:

- The purpose of this paper is to introduce a new class of compound criteria and optimum designs that provide a specified balance between minimizing the average variance and high probability of a desired outcome. The proposed criterion called AP-optimality that combines A-optimality and P-optimality and address this issue for generalized linear models. An equivalence theorem for this criterion is provided and two numerical examples are presented for different GLMs to illustrate the achieved dual properties.

Key-Words:

- *Optimum design, A-optimality, P-optimality, Compound criteria*

1. INTRODUCTION

The type of any design is always an option regardless of the type of model we wish to fit (for example, first order, first order plus some interactions, full quadratic, cubic, etc.) or the objective specified for the experiment. The design of experiments for generalized linear models (GLMs) has received considerable attention in recent years, for example the research by *Woods et al.* [9]. To some extent, this has been in response to design issues raised by researchers in experimental sciences, such as new technologies (for example genomics and areas of modern biology), where the inherent characteristics of data in these fields lead to the consideration of GLMs for analysis and consequently design. GLMs are non-linear models and, as such, pose substantial challenges in terms of design, in particular in the need to have information on the model parameters prior to designing an experiment to estimate these parameters. Much of the research into design for GLMs has concentrated on quite small models: one or two variables and 'simple' optimality criteria, such as D-optimality, which is concerned solely with parameter estimation. However, the paper by *Woods et al.* [9] investigated complex models for binary data with several variables over a number of models in the form of a compound criterion called product design optimality. Historically, most optimal design criteria have been concerned with parameter estimation, and more recently some have combined the notions of parameter estimation and model discrimination (for example, DT-optimality, *Atkinson* [1]). Examples of other compound criteria can be found in *Waterhouse* [8] where criteria are described that also yield designs that offer efficient parameter estimation and model discrimination.

A-optimality criterion corresponds to minimize the variance of the asymptotic distribution of the maximum likelihood estimate of that parameter, employed that criterion of optimality is the one that involves the use of Fisher's information matrix. For linear models with one discrete factor and additive general regression term the problem of characterizing A-optimal design measures for inference on treatment effects, the regression parameters and all parameters will be considered. While, P-optimal design maximizes the average probability of success of a given design.

The aim of this paper is to derive method for designing experiments from which minimizing average variance of the parameter estimates can be obtained, while at the same time maximizing the probability of a particular event that is of importance to experimenter. This paper is organized as follows; Section 2 is devoted to represent the optimum design background. In Section 3, a simple review for A – and P – optimum designs is introduced. In Section 4, the AP-optimum design is proposed to achieve the dual goals of minimizing the average variance and maximizing the average of the probability of observing an outcome. Moreover, the equivalence theorem is derived. Two numerical examples are given in Section 5 to illustrate the method and the value of the proposed criterion in meeting the dual aims.

2. OPTIMUM DESIGN PRELIMINARIES

Consider the generalized linear models GLMs

$$
\begin{aligned}
&E(Y) = \mu = g^{-1}(X\beta) \\
&\eta = g(X\beta)
\end{aligned}
\tag{2.1}
$$

which is defined by the distribution of the response Y, a matrix of independent variables (predictors) X , a vector of unknown parameters β and a linear predictor η and two functions:

1. A link function $g(.)$ that describes how the mean, $E(Y_i) = \mu_i$ depends on the linear predictor $g(\mu_i) = Y_i$.
2. A variance function that describes how the variance, $Var(Y_i)$ depends on the mean

$$Var(Y_i) = \phi(V(\mu)) \tag{2.2}$$

where the dispersion parameter ϕ is a constant.

In GLMs, the errors or noise ϵ_i have relaxed assumptions where it may or may not have normal distribution. GLMs are commonly used to model binary or count data. Some common link functions are used such that the identity, logit, log and probit link to induce the traditional linear regression, logistic regression, Poisson regression models.

An approximate (continuous) design is represented by the probability measure ξ over the design space δ. If the design has trials at n distinct points in δ , it can be written as

$$\xi = \left\{ \begin{matrix} x_1 & x_2 \ldots\ldots & x_n \\ w_1 & w_2 \ldots\ldots & w_n \end{matrix} \right\} \tag{2.3}$$

A design ξ defines, for $i = 1, \ldots, n$, the vector of support-point $x_i \in \chi$ related to y_i, where χ is a compact experimental domain and the experimental weights w_i corresponding to each x_i, where $\sum_{i=1}^{n} w_i = 1$. The design space can be then expressed as $\delta = \{\xi_i \in X^n \times [0,1]^n : \sum_{i=1}^{n} w_i = 1\}$. Such designs are called approximate or continuous designs.

3. A- and P-OPTIMUM DESIGNS

3.1. A-optimum design

A-optimality criterion introduced by *Chernoff* [2]; who showed that the employed criterion of optimality is the one that involves the use of Fisher's information matrix. For the case where it is desired to estimate one of the p parameters in the information matrix, this criterion corresponds to minimize the variance of the asymptotic distribution of the maximum likelihood estimate of that parameter.

A-optimality minimizes the average variance of the parameter estimates. Alternatively, it can be expressed as the following form;

$$\Phi_A(\xi) = \min_{x_i, i=1,\dots n} tr(X^T X) \tag{3.1}$$

For a discussion on an A-optimal designs for binary models, see *Sitter and Wu* [6], *Zhu and Wong* [11]. *Yang* [10] introduced A-optimal designs for generalized linear models with two parameters which are logistic, probit and double exponential models.

The equivalence theorem states that, the derivative function

$$f^T(x)\, M^{-2}(\boldsymbol{\theta}, \xi) f(x) \le tr[M^{-1}(\boldsymbol{\theta}, \xi)], \quad x \in \chi \tag{3.2}$$

where M is the information matrix and the equality holds only if $\xi = \xi_A^*$, $x \in \xi_A^*$

A-efficiency of a design ξ is defined as:

$$Eff_A(\xi) = \frac{tr\left[M^{-1}(\boldsymbol{\theta}, \xi_A^*)\right]}{tr\left[M^{-1}(\boldsymbol{\theta}, \xi)\right]} \tag{3.3}$$

where ξ_A^* is A-optimal.

3.2. P-optimum designs

McGree and *Eccleston* [5] have offered a P-optimality criterion, which is defined as a criterion that maximizes a function of the probability of observing a particular outcome. One of the forms of P-optimality which defined is concerned with the maximization of a weighted sum of the probabilities of success. The form of this criterion is

$$\Phi_P(\xi) = \sum_{i=1}^{n} \pi_i\left(\boldsymbol{\theta}, \xi_i\right) w_i \tag{3.4}$$

where, $\pi_i\left(\boldsymbol{\theta}, \xi_i\right)$ is the i-th probability of success given by ξ_i and w_i is the experimental effort relating to the i-th support point. In this criterion, design weights have been included and will play a role in maximizing the probabilities.

Let ξ_P^* be the design maximizing (3.4). Under some regularity conditions, *McGree* an *Eccleston* [5] proved an equivalence theorem for P-optimum designs, in which the derivative function $\psi_P\left(x, \xi_P^*\right) \leq 0$, $x \in \chi$, where

$$\psi_{P_A}\left(x, \xi_P^*\right) = \frac{\Phi_P\left(x\right) - \Phi_P\left(\xi_P^*\right)}{\Phi_P\left(\xi_P^*\right)} \tag{3.5}$$

is the directional derivative of $\Phi_p\left(\xi\right)$. The P- efficiency of a design ξ relative to the optimum design ξ_P^* is

$$Eff_{P_A}\left(\xi\right) = \frac{\sum_{i=1}^{n} \pi_i\left(\boldsymbol{\theta}, \xi_i\right) w_i}{\sum_{i=1}^{n} \pi_i\left(\boldsymbol{\theta}, \xi_{P_A}^*\right) w_i} \tag{3.6}$$

4. AP-OPTIMUM DESIGN

There is a situation when an experimenter may be interested to achieve multiple objectives. For this aim, we will construct a design that combine A-optimality with P-optimality. The new criterion will be called AP-optimality. This criterion offers a method of achieving minimizing the average variance and a high probability of a desired outcome.

The AP-optimality criterion is given by the following weighted geometric mean of efficiencies:

(4.1)

$$\{Eff_A(\xi)\}^{\alpha}\{Eff_P(\xi)\}^{1-\alpha} = \left(\frac{tr[M^{-1}(\boldsymbol{\theta}, \xi_A^*)]}{tr[M^{-1}(\boldsymbol{\theta}, \xi)]}\right)^{\alpha} \left(\frac{\sum_{i=1}^{n} \pi_i\left(\boldsymbol{\theta}, \xi_i\right) w_i}{\sum_{i=1}^{n} \pi_i\left(\boldsymbol{\theta}, \xi_P^*\right) w_i}\right)^{1-\alpha}$$

where the coefficients $0 \leq \alpha \leq 1$ .When $\alpha = 0$, we obtain P-optimality and when $\alpha = 1$, we obtain A -optimality. To clarify the structure of the design criterion, take log in (4.1) yields,

$$\begin{aligned} &\alpha \log(tr\left[M^{-1}\left(\boldsymbol{\theta}, \xi_A^*\right)\right]) - \alpha \log(tr\left[M^{-1}\left(\boldsymbol{\theta}, \xi\right)\right]) + \\ &(1-\alpha) log \sum_{i=1}^{n} \pi_i\left(\boldsymbol{\theta}, \xi_i\right) w_i - (1-\alpha) log \sum_{i=1}^{n} \pi_i\left(\boldsymbol{\theta}, \xi_P^*\right) w_i \end{aligned} \tag{4.2}$$

The terms involving ξ_A^* and ξ_P^* are constants when a maximum is found over ξ. Many bibliographical references presented the concept of this maximization method such as Dette [4]; *Atkinson* [1]; *Tommasi* [7]; and *McGree* and *Eccleston* [5]. So that the criterion to be maximized is

$$\Phi_{AP}(\xi) = -\alpha \log(tr\left[M^{-1}(\boldsymbol{\theta}, \xi)\right]) + (1-\alpha) log \sum_{i=1}^{n} \pi_i(\boldsymbol{\theta}, \xi_i) w_i \tag{4.3}$$

The negative sign for the first term on the right hand side of (4.3) arises because the average variance is minimized. Designs maximizing (4.3) are called AP-optimum and denoted ξ_{AP}^*.

The equivalence theorem is stated as follows;

Theorem 4.1. *For AP-optimal design, ξ_{AP}^*, the following three statements are equivalent.*

1. *A necessary and sufficient condition for a design ξ_{AP}^* to be AP-optimum is fulfillment of the inequality $\psi_{AP}(x, \xi_{AP}^*) \leq 1,\ x \in \chi$, where the derivative function of* (4.3) *is given by*
(4.4)
$$\psi_{AP}(x, \xi_{AP}^*) = \alpha \left(\frac{f^T(x)\, M^{-2}(\boldsymbol{\theta}, \xi_{AP}^*) f(x)}{\Phi_A(\xi_{AP}^*)} \right) + (1-\alpha) \left(\frac{\Phi_P(x) - \Phi_P(\xi_{AP}^*)}{\Phi_P(\xi_{AP}^*)} \right)$$
2. *The upper bound of $\psi_{AP}(x, \xi_{AP}^*)$ is achieved at the points of the optimum design.*
3. *For any non optimum design ξ, that is a design for which $\Phi_{AP}(\xi) < \Phi_{AP}(\xi_{AP}^*)$, $sup_{x\in\chi}\, \psi_{AP}(x, \xi_{AP}^*) > 1$.*

Proof: Since $0 \leq \alpha \leq 1$, ψ_{AP} is a convex combination of logarithm of two design criteria. Therefore, the AP-criterion satisfies the conditions of convex optimum design theory and an equivalence theorem applies. Because of the way the terms in (4.4) have been scaled, the upper bound of ψ_{AP} over $x \in \chi$ is one, achieved at the points of the optimum design. Furthermore, ψ_{AP} is the linear combination of the directional derivatives given by A-optimality and P-optimality. Thus, the theorem has been proved. □

5. APPLICATIONS TO GENERALIZED LINEAR MODELS

In this Section, the AP-optimality criterion is applied to two types of generalized linear models, Logit and probit models, for binary data. The data were based on the work given by in *Corana* et.al. [3]. The A-, P-, and the proposed compound AP- efficiencies are calculated and the optimal designs are obtained to illustrate the main objective of the compound criterion that allow both minimizing the average variance of the parameter estimates plus increasing the probability of the desired outcome.

Example 5.1. Logit Model

The considering logit model has two main factor effects besides the interaction with initial parameter estimates $\theta = [1, -2, 1, -1]^T$ with $x_j \in [-1, 1]$ as follows;

$$Log\left(\frac{\pi}{1-\pi}\right) = 1 - 2x_1 + x_2 - x_1x_2 \tag{5.1}$$

AP-optimal designs and their A- and P-efficiencies for $\alpha = 0,\ 0.25,\ 0.5,\ 0.75\ ,\ 1$ are obtained and presented in Table 1.

α	x_1	x_2	w_i	π_i	A_{eff}	P_{eff}
0	-1.000	1.000	1.000	0.9933	n/a	1
0.25	1.0000	-1.000	0.0835	0.2689	0.822183	0.8060
	0.8020	1.000	0.0999	0.3999		
	-1.000	-1.000	0.1983	0.7311		
	-0.3980	1.000	0.6182	0.9596		
0.5	1.000	-1.000	0.1570	0.2689	1	0.6644
	1.000	1.000	0.1600	0.2889		
	-1.000	-1.000	0.2802	0.7311		
	-0.1059	1.000	0.4028	0.9103		
0.75	1.000	1.000	0.2121	0.2689	0.741011	0.5826
	1.000	-1.000	0.2121	0.2689		
	-1.000	-1.000	0.2740	0.7311		
	0.0148	1.000	0.3017	0.8761		
1	1.000	-1.000	0.2500	0.2689	0.864115	0.5352
	1.000	1.000	0.2500	0.2689		
	-1.000	-1.000	0.2500	0.7311		
	0.0680	1.000	0.2500	0.8577		

Table 1: AP-optimum design and their A-and P- efficiencies for the Logit model at different values of α

Table (5.1) shows the designs that maximize the AP-criterion. It can be noticed that there is little changes in the design points with high variation in design weights. That is, the P_{eff}'s are increased through the given designs as well as the probability of success is increased. Figure (1) illustrates the A- and P-efficiencies for $\alpha = 0,\ 0.25,\ 0.5,\ 0.75\ and\ 1$. The dot-dashed line represents the A-efficiency of the designs, and the solid line shows their P-efficiencies. The following A-optimal design has a P-efficiency of 0.6644.

$$\xi_A^* = \left\{ \begin{array}{ccc} 1.0000 & -1.000 & 0.1570 \\ 1.0000 & 1.000 & 0.1600 \\ -1.000 & -1.000 & 0.2802 \\ -0.1059 & 1.000 & 0.4028 \end{array} \right\}$$

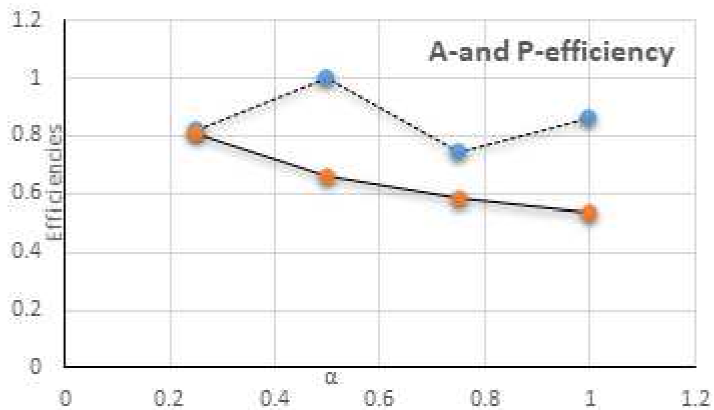

Figure 1: A- and P-efficiencies of AP-optimal designs for different values of α.

By using the AP-criterion and choosing $\alpha = 0.25$, we are able to increase the P-efficiency to 0.806, while achieving a A-efficiency of 0.822183. The AP-optimal design is

$$\xi_{AP}^* = \left\{ \begin{array}{ccc} 1.000 & -1.000 & 0.0835 \\ 0.802 & 1.000 & 0.0999 \\ -1.000 & -1.000 & 0.1983 \\ -0.398 & 1.000 & 0.6182 \end{array} \right\}$$

Example 5.2. Probit Model

In the following Example, the AP-optimlaity criterion is applied to the probit model. The response variable is modelled via three main factor effects with initial parameters $\beta = [1, -0.5, 1, -1]$, with $x_j \in [-1, 1]$,

$$\Phi^{-1}(\pi) = 1 - 0.5\ x_1 + x_2 - x_3 \tag{5.2}$$

Table (5.2) include the main results of the designs and their A- and P- efficiencies for $\alpha = 0, 0.2,\ 0.35,\ 0.5,\ 0.75,\ 1$. Figure (2) illustrates the A- and P-efficiencies for $\alpha = 0, 0.2,\ 0.35,\ 0.5,\ 0.75,\ 1$. Using the compound criteria AP-criterion, at $\alpha = 0.5$, we can see that the A-efficiency and P-efficiencies have very close high efficiencies, 0.982127 , 0.983084, respectively.

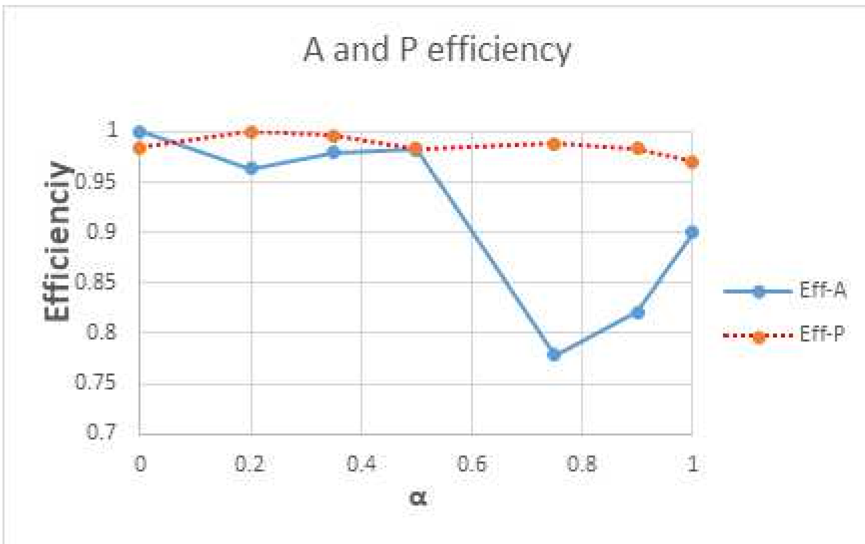

Figure 2: A- and P-efficiencies of AP-optimal designs for different values of α.

Hence, the AP-optimal design which satisfy the dual problem is obtained as:

$$\xi^*_{AP} = \left\{ \begin{array}{ccc} -1.0000 & -1.0000 & 1.0000 \\ -05140 & 1.0000 & -1.0000 \\ -0.4709 & 1.0000 & 1.0000 \\ -0.4395 & -1.0000 & -1.0000 \\ -0.0373 & -1.0000 & 1.0000 \\ 0.0372 & 1.0000 & 1.0000 \\ 0.4709 & -1.0000 & 1.0000 \\ 1.0000 & -1.0000 & 1.0000 \\ 1.0000 & 1.0000 & 1.0000 \end{array} \right\}$$

α	$\boldsymbol{x_1}$	$\boldsymbol{x_2}$	$\boldsymbol{x_3}$	$\boldsymbol{w_i}$	$\boldsymbol{\pi_i}$	$\boldsymbol{A_{eff}}$	$\boldsymbol{P_{eff}}$
0	-1.0000	-1.0000	1.0000	0.0708	0.3085	1	0.98427
	-0.6296	1.0000	-1.0000	0.1105	0.9988		
	-0.5423	1.0000	1.0000	0.1600	0.8980		
	-0.5423	-1.0000	-1.0000	0.1600	0.8980		
	-0.0186	-1.0000	1.0000	0.0658	0.1611		
	0.0186	1.0000	1.0000	0.0658	0.8389		
	0.0186	-1.0000	-1.0000	0.0658	0.8389		
	0.5423	-1.0000	1.0000	0.1600	0.1020		
	1.0000	-1.0000	-1.0000	0.0708	0.6915		
	1.0000	1.0000	1.0000	0.0708	0.6915		
0.2	-1.0000	-1.0000	1.0000	0.0608	0.3085	0.964075	1
	-0.5368	-1.0000	-1.0000	0.2315	0.8980		
	-0.5368	1.0000	1.0000	0.2315	0.8980		
	-0.5244	1.0000	-1.0000	0.1232	0.9987		
	0.5368	-1.0000	1.0000	0.2315	0.1020		
	1.0000	-1.0000	-1.0000	0.0608	0.6915		
	1.0000	1.0000	1.0000	0.0608	0.6915		
0.35	-1.0000	-1.0000	1.0000	0.0630	0.3085	0.979698	0.996276
	0.5027	1.0000	-1.0000	0.1213	0.9987		
	-0.4894	-1.0000	-1.0000	0.2299	0.8925		
	-0.4894	1.0000	1.0000	0.2299	0.8925		
	0.4894	-1.0000	1.0000	0.2299	0.1075		
	1.0000	-1.0000	-1.0000	0.0630	0.6915		
	1.0000	1.0000	1.0000	0.0630	0.6915		
0.5	-1.0000	-1.0000	1.0000	0.0618	0.3085	0.982127	0.983084
	-0.5140	1.0000	-1.0000	0.1085	0.9987		
	-0.4709	1.0000	1.0000	0.2144	0.8925		
	-0.4395	-1.0000	-1.0000	0.2241	0.8888		
	-0.0373	-1.0000	1.0000	0.0245	0.1635		
	0.0372	1.0000	1.0000	0.0245	0.8365		
	0.4709	-1.0000	1.0000	0.2144	0.1075		
	1.0000	-1.0000	-1.0000	0.0662	0.6915		
	1.0000	1.0000	1.0000	0.0618	0.6915		

0.75	-1.0000	-1.0000	1.0000	0.0504	0.3085	0.778262	0.988816
	-1.0000	-1.0000	-1.0000	0.0143	0.9332		
	-1.0000	1.0000	1.0000	0.0143	0.9332		
	-1.0000	1.0000	-1.0000	0.1306	0.0089		
	-0.0101	-1.0000	-1.0000	0.2251	0.8413		
	-0.0101	1.0000	1.0000	0.2251	0.8413		
	0.0101	-1.0000	1.0000	0.2251	0.1587		
	1.0000	-1.0000	1.0000	0.0143	0.0668		
	1.0000	-1.0000	-1.0000	0.0504	0.6915		
	1.0000	1.0000	1.0000	0.0504	0.6915		
0.9	-1.0000	-1.0000	1.0000	0.0533	0.3085	0.820936	0.983858
	-0.9942	1.0000	-1.0000	0.1197	0.9987		
	-0.8203	-1.0000	-1.0000	0.0212	0.9207		
	-0.8203	1.0000	1.0000	0.0212	0.9207		
	-0.0348	-1.0000	-1.0000	0.2189	0.8461		
	-0.0348	1.0000	1.0000	0.2189	0.8461		
	0.0348	-1.0000	1.0000	0.2189	0.1539		
	0.8203	-1.0000	1.0000	0.0212	0.0793		
	1.0000	1.0000	1.0000	0.0533	0.6915		
	1.0000	-1.0000	-1.0000	0.0533	0.6915		
1	-1.0000	-1.0000	1.0000	0.0550	0.3085	0.899469	0.970774
	-0.9344	1.0000	-1.0000	0.0924	0.9989		
	-0.6464	-1.0000	-1.0000	0.0515	0.9066		
	-0.6464	1.0000	1.0000	0.0515	0.9066		
	-0.0612	-1.0000	-1.0000	0.1960	0.8485		
	-0.0612	1.0000	1.0000	0.1960	0.8485		
	0.0612	-1.0000	1.0000	0.1960	0.1515		
	0.6464	-1.0000	1.0000	0.0515	0.0934		
	1.0000	-1.0000	-1.0000	0.0550	0.6915		
	1.0000	1.0000	1.0000	0.0550	0.6915		

Table 2: AP-optimum design and their A-and P- efficiencies for the Probit model at different values of α

6. CONCLUSION

The criterion AP-optimum design introduced here provides a new compound criterion that yield minimum of the average variance of the parameter estimates plus a high probability of observing a particular outcome. The equivalence theorem is stated and proved for AP-optimum design. Two illustrated examples are presented for logit and probit models. The results indicate the potentiality of using the proposed AP-optimality criterion.

REFERENCES

[1] Atkinson, A. C. (2008). DT–Optimum Designs for Model Discrimination and Parameter Estimation, *Journal of Statistical Planning and Inference*, **138**, 56–64.

[2] Chernoff, H. (1953). Locally Optimal Designs for Estimating Parameters, *Annals of Mathematical Statistics*, **24**, 586–602.

[3] Corana, A.; Marchesi, M.; Martini, C. and Ridella, S. (1987). Minimizing Multimodal Functions of Continuous Variables with the 'Simulated Annealing' Algorithm, *ACM Trans. Math. Software*, **13**, 262–280.

[4] Dette, H. (1993). On a Mixture of the D- and D_1-Optimality Criterion in Polynomial Regression, *Journal of Statistical Planning and Inference*, **35**, 233–249.

[5] Mcgree, J. M. and Eccleston, J. A. (2008). Probability-Based Optimal Design, *Australian and New Zealand Journal of Statistics*, **50**, 1, 13–28.

[6] Sitter, R. and Wu, C. (1993). Optimal Designs for Binary Response Experiments: Fieller, D, and A criteria, *Scandinavian Journal of Statistics*, **20**, 329-342.

[7] Tommasi (2009). Optimal Designs for Both Model Discrimination and Parameter Estimation, *Journal of Statistical Planning and Inference*, **139**, 4123-4132.

[8] Waterhouse, T. H. (2005). *Optimal Experimental Design for Nonlinear and Generalized Linear Models*, (PhD Thesis). University of Queensland, Australia.

[9] Woods, D. C.; Lewis, S. M.; Eccleston, J. A. and Russell, K. G. (2006). Designs for Generalized Linear Models with Several Variables and Model Uncertainty, *Technometrics*, **48**, 284–292.

[10] Yang, M. (2008). A-optimal Designs for Generalized Linear Models with Two Parameters, *Journal of Statistical Planning and Inference*, **138**, 624 – 641.

[11] Zhu, W.and Wong, W. (2000). Multiple-Objective Designs in a Dose-Response Experiment, *Journal of Biopharmaceutical Statistics*, **10**, 1, 1-14.