



An Efficient Bayesian Optimal Design for Logistic Model

H. Talebi* and D. Poursina

University of Isfahan

Abstract. Consider a Bayesian optimal design with many support points which poses the problem of collecting data with a few number of observations at each design point. Under such a scenario the asymptotic property of using Fisher information matrix for approximating the covariance matrix of posterior ML estimators might be doubtful. We suggest to use Bhattcharyya matrix in deriving the information matrix, led to modified Bayesian D-optimal criterion. This criterion is used to obtain optimal design for logistic model. It is shown that the resulting optimal design is more efficient than design given by Chaloner and Larentz (1989) using ordinary Bayesian D-optimal criterion.

Keywords. Bhattcharyya matrix; logistic model; design support points; Bayesian D-optimal criterion; Bayesian optimal design.

MSC 2010: 62K05.

1 Introduction

Optimal design deals with the problem of choosing some appropriate support points at which one collects data by which he/she makes a desirable statistical inference about the parameters of postulated model. The desirability of a design is characterized by an optimality criterion based on a quantity of the parameter estimates such as the variance. Criteria in non-linear models depend on the parameters. To overcome this dependency problem, the authors

* Corresponding author

suggest some solutions which lead to locally, minimax and Bayesian optimal designs. In locally optimal design, the experimenter starts with an initial guess for each parameter. The resulting locally optimal designs depend on the guessed values, and hence, they are not robust against the preselected values. In Bayesian approach a prior distribution is imposed for each parameter, instead of an initial guess, which makes the optimal designs to be more robust than the locally optimal designs. We will focus on Bayesian designs in this paper. The Bayesian optimal designs has been considered by several authors, including Tsutakawa (1972, 1980), Chaloner (1993), Sun et al. (1996) and Dette et al. (2007). A unification theory of Bayesian optimal designs for non-linear models is discussed in Chaloner and Larntz (1989), who developed equivalence theorem from Whittle (1973).

In almost all of the research works in optimal design, including locally, minimax and Bayesian approach, the researchers commonly use the asymptotic approximation in deriving the optimal design; however, there are some exceptions. Russell et al. (2009) investigated the accuracy of approximation to the exact bias and covariance matrix for a two parameters simple logistic model with two design points and found out that this approximation is inefficient for small sample size. They suggested the mean square errors (MSE) criterion be employed for selecting an optimal design, where ML estimator of model parameter is replaced by maximum penalized likelihood (MPL) of Firth (1993). Poursina and Talebi (2012) also considered the approximation problem in small sample optimal design and used a more accurate approximation than the Cramer-Rao lower bound for variance of ML estimator. They used this approximation to calculate the Fisher information matrix of the parameters in a two parameters simple logistic model and proposed a new modified D-optimality criterion. In Bayesian approach, Tsutakawa (1972) pointed out the works of some authors who have dealt with reducing the MSE or asymptotic variance estimates. He suggested a modification in optimality criterion by adding a non-negative penalized-type constant to Fisher information in approximating the posterior variance of the model parameter in a single variable logistic model with one parameter. Sun et al. (1996) evaluated the accuracy of Tsutakawa (1972) approximate results for small to moderate sample sizes, numerically. They developed an exact solution to Bayesian optimal design and showed that the proposed asymptotic approximation of Tsutakawa (1972) is reasonably accurate. Sun and Tsutakawa (1997) improved this result by using a penalized expected posterior variance as optimality criterion.

We raise a new problem by considering the case, in which the total sample size is not necessarily small, yet, the asymptotic approximation might be invalid. As an example, for a fixed total sample size, consider the case in which there is a few observations at each design point, due to have a numerous number of support points. In Bayesian approach, due to assuming a prior distribution for the uncertainty assumption on parameters, the number of support points is not usually determined *a priori*. For more illustration, in such an approach, the number of support points increases as uncertainty in the prior distribution increases (Chaloner and Larntz, 1989). Braess and Dette (2007) provided an analytical solution to show such an occurrence, where by a more uncertainty they mean a prior distribution with a larger variance. Meanwhile, in the case of uniform prior, by a more uncertainty one means a distribution with a large support interval $[a,b]$ (see Chaloner and Larentz, 1989). It is noted that for a locally optimal design, as a special case with a degenerate prior, i.e. no uncertainty, it is common to come up with a minimal number of support points, that is, equal to the number of parameters in the underlying model. Li and Majumdar (2009) showed that this is always true under some conditions. Meanwhile, in Bayesian case, Chaloner (1993) had given sufficient conditions for obtaining a one point Bayesian D-optimal design, in a closed form, for a logistic model with one parameter. These conditions are true when the support of prior distribution is small. However; there are cases that the given sufficient conditions are not satisfied, that is, cases with a high uncertainty in prior distribution, which is not uncommon in practical situations. For such the cases, Braess and Dette (2007) established sufficient conditions under which the number of design points in Bayesian D-optimal designs increase arbitrarily. It is noted that asymptotic approximation of variance is valid for large enough number of support points; however, for the case with a fixed number of support points the asymptotic property is true for a large number of observations at each of these points (See McCullagh and Nelder, 1989, P. 119).

In this paper, we focus on the role of the number of support points in designing an experiment and consider the problem of Bayesian D-optimal design for a single explanatory variable in a two parameters logistic model. That is, we consider the problem of obtaining a design with a fixed, neither necessarily few nor large enough, number of support points and a relaxed total sample size. Under such a scenario, one might not be able to derive optimal design in a closed form nor to use the asymptotic properties of the estimators to approximate the variance by the inverse of the information

matrix, due to the absence of a large number of observations at each support point. To overcome this problem, we suggest to use a more accurate estimate than that of Cramer-Rao lower bound (CRB) in approximating the variance matrix.

In Section 2, by recalling some preliminary results on Fisher information matrix of the parameters for simple logistic model, we argue about the deficiency of this matrix in approximating the variance for small sample experiments and also for designs with numerous support points. The modified Fisher information is suggested based on Battachryya matrix and is used to approximate the variance of parameter estimates. In Section 3, we employ this new information matrix to propose a Bayesian D-optimality criterion, based on the criterion given by Chaloner and Larentz (1989). It is shown that this new criterion is a generalization of the origin. For different values of total sample sizes, some new optimal designs are presented based on prior distributions given by Chaloner and Larentz (1989). In Section 4, the performance of given new design is evaluated and compare with design obtained by Chaloner and Larentz (1989). The comparison is done through a simulation study. It will be shown that the new design is more efficient than the former, in particular, for small sample sizes; however, two designs converge in being the same as the sample size increases.

2 Preliminary Arguments

In the generalized linear models (GLM), let consider a function of response mean, μ , as a linear function of predictors such that

$$g(\mu) = \eta(X, \boldsymbol{\theta}) = X\boldsymbol{\theta},$$

where g is the link function, X is design matrix and $\boldsymbol{\theta}$ is a parameter set of interest. The variance of the maximum likelihood estimator (MLE), $\hat{\boldsymbol{\theta}}$, of $\boldsymbol{\theta}$, is approximated by $(X^T W X)^{(-1)}$, in which the diagonal matrix W is information matrix of $g(\mu)$.

As an example, for the binary responses with $Y_i \sim Bin(n_i, p_i)$ consider the logistic model with m distinct design points by

$$g(\mu_i) = \ln \left(\frac{p_i}{1 - p_i} \right) = \beta(x_i - \alpha). \quad (1)$$

Under model (1), for a larg m the variance of ML estimators of $\boldsymbol{\theta}$ is $\text{Cov}(\hat{\boldsymbol{\theta}}) =$

$(X^T W X)^{-1} \{1 + O(m^{-1})\}$, where the Fisher information matrix of parameters denoted by matrix $M(X, \theta)$, is:

$$M(X, \theta) = X^T W X = \sum_{i=1}^m \{n_i p_i (1 - p_i)\} \begin{pmatrix} \beta^2 & -\beta(x_i - \alpha) \\ -\beta(x_i - \alpha) & (x_i - \alpha)^2 \end{pmatrix}, \quad (2)$$

for $\theta = (\alpha, \beta)$.

This approximation is also true for fixed m when $n_i \rightarrow \infty$, in which the error term is replaced by $O(n_i^{-1})$ (see McCullagh and Nelder, 1989, P. 119).

So far, in this paper we have argued about the validity of using information matrix, as an approximation for the variance. That is, we addressed the case of having a moderate to large total sample size, yet the conditions for using the Fisher information matrix as an appropriate approximation for the inverse of variance matrix do not meet, due to have a design with a numerous number of design points. This often happens in dealing with the Bayesian optimal designs with a fixed total sample size, when prior distributions of the model parameters carry a large uncertainty or large location. Under such conditions, we suggest to use the Bhattacharyya matrix to obtain a more appropriate approximate for the inverse variance matrix. This requires a more understanding of this matrix. Readers are referred to Mohtashami et al. (2010) for more details.

Based on Bhattacharyya matrix a more exact asymptotic variance matrix, for logistic model, is given in Poursina and Talebi (2012). They showed that for a Bhattacharyya matrix of order $k = 2$, the modified information matrix of $g(\mu)$, say W^* , is a diagonal matrix with diagonal elements:

$$w_{ii}^* = \left\{ \frac{1}{n_i p_i (1 - p_i)} + \frac{(1 - 2p_i)^2}{4n_i (n_i - 1) p_i^2 (1 - p_i)^2} \right\}^{-1},$$

$$= (w_{ii}^{-1} + v_{ii})^{-1} \quad (3)$$

where w_{ii} is the i th diagonal element of W and $v_{ii} = \frac{(1-2p_i)^2}{4n_i(n_i-1)p_i^2(1-p_i)^2}$. By substituting W^* in information matrix of model parameters, we have:

$$M^*(X, \theta) = X^T W^* X = \sum_{i=1}^m w_{ii}^* \begin{pmatrix} \beta^2 & -\beta(x_i - \alpha) \\ -\beta(x_i - \alpha) & (x_i - \alpha)^2 \end{pmatrix}. \quad (4)$$

This modified information matrix will be used in proposing a new general criterion in the next section.

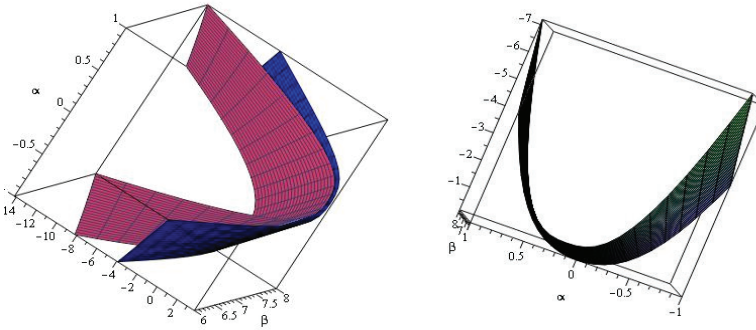


Figure 1. (a) Shows the $\log[\{M^*(\xi, \theta)\}]$, $\log[\{M(\xi, \theta)\}]$, (b) Shows $\log[\{M^*(\xi, \theta)\}] - \log[\{M(\xi, \theta)\}]$.

2.1 Design Construction

For a nonsingular design, ξ , Chaloner and Larentz (1989) gave the Bayesian D-optimal criterion by $\phi_1(M(\xi)) = E_{\theta}\{\log[\det\{M(\xi, \theta)\}]\}$. We generalize this criterion by replacing $M^*(\xi, \theta)$, given in (4), for $M(\xi, \theta)$ and denoted it by ϕ_1^* . Followings are more to understand ϕ_1^* in details.

- i) It is a general form of ϕ_1 in a sense if $n_i \rightarrow \infty$ or $p_i = \frac{1}{2}$, v_{ii} tends to zero and hence $w_{ii}^* = w_{ii}$, i.e., $\phi_1^* = \phi_1$. For more clarification, the integrands in criteria ϕ_1 and ϕ_1^* are plotted in Figure 1(a) for $N = 30$, based on a design with three points, given by Chaloner and Larentz (1989). The difference between these two curves is calculated and plotted in Figure 1(b). The difference is zero at $\alpha = 0$, that is, $p_i = \frac{1}{2}$. It is also noted that the difference values increase as p_i goes far from $\frac{1}{2}$. Similarly, one can show the equality of these two criteria graphically when $n_i \rightarrow \infty$.
- ii) By v_{ii} we, in fact, insert a penalty for approximating the variance of posterior mode estimator, proportioned to the number of observations, say n_i , at a design point. Clearly, the penalty term for a design point with small n_i is larger, which cause a deviance from Cramer-Rao lower bound in favor of having a more exact approximation for variance. If $n_i \rightarrow \infty$, then this terms goes to zero and hence the modified information tends to the ordinary information employed in the criterion of Chaloner and Larentz (1989), at this point.
- iii) The only difference between the ϕ_1^* and ϕ_1 is due to the M^* and M

which come from W^* and W , respectively. That is, the difference is related to the positive constant v_{ii} . Poursina and Talebi (2012), showed that M^* carries all regular conditions of M . This provides conditions in the general theory, developed by Chaloner and Larentz (1989), which allows the implication of equivalence theorem in Bayesian approach, so as here, by considering criterion ϕ_1^* . We advance this discussion, following Chaloner and Larentz (1989), used the general equivalence theorem of Whittle (1973), for non-linear models in Bayesian approach. Under the present conditions the equivalence Theorem is followed.

Theorem 1. *If ϕ^* is concave, then a ϕ^* -optimal design ξ_0 can be equivalently characterized by any of the three conditions*

- i) ξ_0 maximizes $\phi^*(\xi)$
- ii) ξ_0 maximizes $\sup_{x \in \chi} d(\xi, x)$
- iii) $\sup_{x \in \chi} d(\xi_0, x) = 0$

where,

$$d(\xi^*, x) = \int \text{tr}[\{M^{*-1}(\xi_i, \boldsymbol{\theta})\}M^*(x, \boldsymbol{\theta})p(\boldsymbol{\theta})d\boldsymbol{\theta}] - p,$$

in which p stands for the number of parameters in the model.

In this paper, we proceed to use this theorem for optimality checking, yet we replace modified information matrix, M^* , in $d(\xi, x)$. For more space saving we refer the reader to Poursina and Talebi (2012), who showed that the general equivalence theorem is still valid under such the modification.

Now, consider the modified criterion $\phi_1^*(M^*(\xi)) = E_{\boldsymbol{\theta}}\{\log[\det\{M^*(\xi, \boldsymbol{\theta})\}]\}$ for obtaining the optimal design in the two parameter logistic model with total sample size N as an optimality criterion. We investigate the design ξ^* with m support points x_i 's with corresponding weights w_i which maximizes $\phi_1^*(M^*(\xi))$ for independent uniform prior distributions of the parameters α and β . The desired optimal design is obtained by pursuing the following algorithm. In this algorithm neither the number of design points nor their weights are predetermined, although, the algorithm initiates with two support points.

- i) Set $i = 0$ and consider the non singular initial design ξ_i by taking two initial values in design region χ . It is noted that this design is minimally supported.

- ii) We take χ a compact design region such that all of the design points be interior points of this region. Let $x \in \chi$ be any support point of ξ_i . For a given small positive ϵ check the inequality:

$$\text{Max } d(\xi^*, x) \leq \epsilon,$$

If the inequality holds, then stop and conclude that ξ_i is optimal design. Otherwise, go to (iii).

- iii) Following Fedorov (1972) algorithm, add a new design point x_i to the design points such that

$$\begin{aligned} \arg \max_{x \in \chi} \int \text{tr}[\{M^{*-1}(\xi, \boldsymbol{\theta})\}M^*(x, \boldsymbol{\theta})p(\boldsymbol{\theta})d\boldsymbol{\theta}] \\ = \int \text{tr}[\{M^{*-1}(\xi_i, \boldsymbol{\theta})\}M^*(x_i, \boldsymbol{\theta})p(\boldsymbol{\theta})d\boldsymbol{\theta}]. \end{aligned}$$

The new design is equal to $\xi_{i+1} = \lambda_i \xi_i + (1 - \lambda_i)x_i$, where λ_i is a weighting sequence satisfies $\sum_{i \geq 1} \lambda_i = \infty$ and $\lim_{i \rightarrow \infty} \lambda_i = 0$.

- iv) Set i to $i + 1$, and go to part (ii).

It is noticed that by an iteration, in each stage a design point with a certain weight, which is depends on the total sample size, is merged to its nearest design point due to not be potentially a support point. In this algorithm neither the support points, nor their weights are held fixed, *a priori*. The codes, appeared in the Appendix, are written in Maple 13.

The resulting designs for $\alpha \in [-0.3, 0.3]$, and $\alpha \in [-1, 1]$, both for given $\beta \in [6, 8]$ are presented in Table 1. Figure 2 shows the plot of $d(\xi^*, x)$ in all design space. The Bayesian equivalence theorem, stated above, confirms the optimality of resulting designs, as it is also shown in Figure 2.

Table 1. The optimal design points with their weights (in parenthesis)

Total sample size	Design points	Design points
	$\alpha \in [-0.3, 0.3]$	$\alpha \in [-1, 1]$
20	-0.302030(0.331710)	—
	0.001892 (0.338745)	—
	0.303772(0.329545)	—
30	-0.303423(0.339688)	0.571691(0.187143)
	0.001918(0.323125)	-0.200061(0.206798)
	0.305023(0.337187)	0.199814(0.206798)
		-0.572535(0.187143)
		1.000052(0.104915)
	-1.002146(0.104915)	
50	-0.304707(0.348071)	0.571866(0.184162)
	0.000359(0.304451)	-0.199901(0.210511)
	0.305045(0.347478)	0.199980(0.210511)
		-0.571696(0.184162)
		-0.999922(0.105327)
	1.000204(0.105327)	
100	-0.304125(0.359719)	-0.572059(0.181662)
	0.002364(0.283901)	0.571680(0.181948)
	0.305883(0.356379)	-0.200004(0.212796)
		0.199756(0.212511)
		0.999190(0.105541)
	-0.999755(0.105541)	
500	-0.304470(0.369974)	0.977897(0.109254)
	0.000701(0.261762)	-0.977763 (0.109254)
	0.30480(0.368263)	0.564395(0.176807)
		-0.199571(0.213939)
		0.199967(0.213939)
	-0.564107 (0.176807)	

3 Performance Criterion

For different values of total sample size, some optimal designs are obtained and given in Table 1 based on new criterion. It is claimed that these designs are more efficient than their corresponding designs obtained in Chaloner and Larentz (1989). In order to compare these two sets of designs and show their

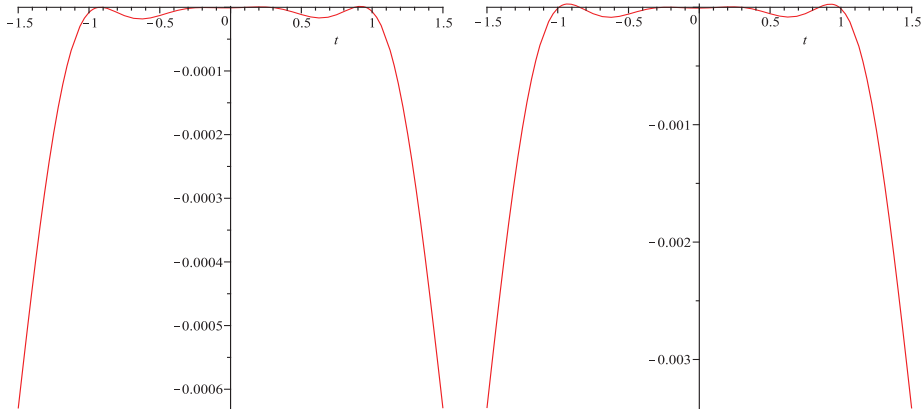


Figure 2. $d(\xi^*, x)$ in all design region for $N = 500$ and $N = 35$.

performances we suggest two quantities. They are, the ratio corresponding to the variance and coverage of an estimator, related to two designs. By coverage we mean the number of times that the parameter space includes the calculated estimates. The details are as the followings. Let

$$C_i = \frac{C_{i\xi^*}}{C_{i\xi}}, \quad V_i = \frac{V_{i\xi}}{V_{i\xi^*}},$$

where, C_i and V_i , are coverage and variance of parameter estimates for corresponding design, $i \in \{\alpha, \beta\}$, for designs ξ and ξ^* . We define $C = \frac{C_{\alpha\xi^*} + C_{\beta\xi^*}}{C_{\alpha\xi} + C_{\beta\xi}}$ and $V = V_{\alpha}V_{\beta}$ to consider the coverage and variance performances with respect to both of α and β , simultaneously. The design ξ^* is said to have a better performance than ξ in coverage and variance if $C > 1$ and $V > 1$, respectively.

4 Discussion: A Simulation Study

Using the above quantities and proposed performance criterion, we investigate the performance of the new optimal design ξ^* , in comparison with the well-known symmetric design, constructed by Chaloner and Larntez (1989), say ξ . This comparison is taken place by running a simulation study in Matlab R2008. Through this study, we simulate the posterior mode estimates

of the parameters, α and β , 10000 times and calculate their corresponding variance and coverage.

To accomplish simulation, first, the parameter values of α and β are randomly chosen from intervals which are the same as the support of their corresponding uniform priors. Then, the observations are generated, using the above designs, ξ^* and ξ , in order to calculate the parameter estimates. The calculated variance and coverage are given in Table 2 for different estimates. These results are shown that the new designs are more efficient than the former with respect to the variance; however, do not perform better in terms of coverage rate.

Table 2. Variance and coverage of parameters for $\alpha \in [-1, 1]$ and $\beta \in [6, 8]$ from uniform prior

Total Sample Size	$V_{\xi}(\alpha)$	$V_{\xi^*}(\alpha)$	$V_{\xi}(\beta)$	$V_{\xi^*}(\beta)$	$C_{\xi}(\alpha)$	$C_{\xi^*}(\alpha)$	$C_{\xi}(\beta)$	$C_{\xi^*}(\beta)$
30	0.4451	0.3715	1461.9	1390.2	9695	9590	3310	3051
50	0.328	0.3068	773.0172	701.67	9747	9612	3333	3169
60	0.3232	0.3017	583.563	556.36	9795	9742	3593	3555
80	0.3202	0.2989	363.273	343.54	9780	9722	3737	3715
100	0.3203	0.297	279.1251	256.45	9784	9740	4232	4138

The values of the efficiency criterion are presented in Table 3. Based on these values, design ξ^* performs better than ξ in terms of variance for all different values of total sample sizes, given in this Table. That is, ξ^* is superior in comparing with the optimum design obtained by Challoner and Larentz(1989), based on ordinary Fisher information matrix in terms of V .

Table 3. The ratio of variance and coverage for $\alpha \in [-1, 1]$ and $\beta \in [6, 8]$ uniform prior

Total Sample Size	V_{α}	V_{β}	C_{α}	C_{β}	V	C
30	1.198115747	1.051575313	0.989169675	0.921752266	1.259908942	0.972011
50	1.069100391	1.101681987	0.986149584	0.95079508	1.177808643	0.977141
60	1.071262844	1.048894601	0.994589076	0.98942388	1.123641813	0.993203
80	1.071261291	1.057440182	0.99406953	0.99411322	1.132794735	0.994082
100	1.078451178	1.088419185	0.995502862	0.97778828	1.173806953	0.990154

New optimal design leads to parameter estimate with smaller variance. The smaller variance compensate the loss of coverage which may occurs using this design. This means that V beats C in favor of the new design. That is,

the new design with a smaller variance is superior to the former one, which may have a better coverage. Two designs perform the same for large N , as it should be from the convergence of the Battacharyya bound (BB) to CRB. That is, the new modified Bayesian D-optimality proposed and presented in this paper, based on the BB obtained from Battacharyya matrix, is a more general form of the criterion given by Chaloner and Larentz (1989).

References

- Braess, D. and Dette, H. (2007). On the Number of Support Points of Maximin and Bayesian Optimal Designs. *The Annals of Statistics*, **35**, 772-792.
- Chaloner, K. and Larntz, K. (1989). Optimal Bayesian Design Applied to Logistic Regression Experiments. *J. Statist. Planning and Inference*, **21**, 191-208.
- Chaloner, K. (1993). A Note on Optimal Bayesian Design for Non-linear Problems. *J. Statist. Planning and Inference*, **37**, 229-235.
- Dette, H., Haines, L.M. and Imhof, L.A. (2007). Maximin and Bayesian Optimal Designs for Regression Models. *statistica sinica*, **17**, 463-480.
- Fedorov, V.V. (1972). Theory of Optimal Experiments (Review). *Biometrika*, **59**, 697-698.
- Firth, D. (1993). Bias Reduction of Maximum Likelihood Estimates. *Biometrika*, **80**, 27-38.
- Li, J. and Majumdar, D. (2009). Some Results on D-optimal Designs for Nonlinear Models with Applications. *Biometrika*, **96**, 487-493.
- McCullagh, P. and Nelder, J.A. (1989). *Generalized Linear Models, 2nd edition*. Chapman and Hall, London.
- Mohtashami Borzadaran, G.R., Rezaei Roknabadi, A. and Khorashadizadeh, M. (2010). A View on Bhattacharyya Bounds for Inverse Gaussian Distributions. *Metrika*, **72**, 151-161.
- Poursina, D. and Talebi, H. (2012). Modified D-Optimal Design for Logistic Model. *Journal of Statistical Computation and Simulation*. DOI:10.1080/00949655.2012.715311.
- Russell, K.G., Eccleston, J.A., Lewis, S.M. and Woods, D.C. (2009). Design Considerations for Small Experiments and Simple Logistic Regression. *Journal of Statistical Computation and Simulation*, **79**, 81-91.
- Sun, D., Tsutakawa, R.K. and Lu, W. (1996). Bayesian Design of Experiment for Quantal Responses: What is Promised Versus What is Delivered. *Journal of Statistical Planning and Inference*, **52**, 289-306.

Sun, D. and Tsutakawa, R.K. (1997). Bayesian Design for Dose-Response Curves with Penalized Risk. *Biometrics*, **53**, 1262-1273.

Tsutakawa, R.K. (1972). Design of Experiment for Bioassay. *Journal of the American Statistical Association*, **67**, 584-590.

Tsutakawa, R.K. (1980). Selection of Dose Levels for Estimating a Percentage Point of a Logistic Quantal Response Curve. *Applied Statistics*, **29**, 25-33.

Whittle, P. (1973). Some General Points in the Theory of Optimal Experimental Design. *J. Roy. Statist. Soc. Ser B*, **35**, 123-130.

Appendix

Maple codes which are written for design construction:

restart:

```
rβ :=? : rμ :=? : βL :=? : βU :=? : μL := -? : μU :=? : u1 :=
(μU - μL)/rμ : u2 := (βU - βL)/rβ : xL := -2 : xU := 2 : o :=? :
k := 2 : q := 500 : x[1] := -2 : w[1] := .5 : x[2] := 2 : w[2] := .5; N :=?
```

```
m:=proc(k)
```

```
local i,M:
```

```
M:=matrix(2,2):
```

```
M[1,1] := β2*sum((1/(o*w[i]*1/(1+e(-β*(x[i]-μ))))*(1-1/(1+e(-β*(x[i]-μ)))))
+(1-2*1/(1+e(-β*(x[i]-μ))))2/(4*o*w[i]*(o*w[i]-1)*1/(1+e(-β*(x[i]-μ))))2
*(1-1/(1+e(-β*(x[i]-μ))))2)(-1), i = 1..k);
```

```
M[1,2] := -β * sum((x[i] - μ) * (1/(o * w[i] * 1/(1 + e(-β*(x[i]-μ))))
*(1 - 1/(1 + e(-β*(x[i]-μ))))) + (1 - 2 * 1/(1 + e(-β*(x[i]-μ))))2/(4 * o * w[i] *
(o * w[i] - 1) * 1/(1 + e(-β*(x[i]-μ))))2
*(1 - 1/(1 + e(-β*(x[i]-μ))))2)(-1), i = 1..k) :
```

```
M[2,1] := M[1,2] :
```

```
M[2,2] := sum((x[i] - μ)2 * (1/(o * w[i] * 1/(1 + e(-β*(x[i]-μ))))
*(1 - 1/(1 + e(-β*(x[i]-μ))))) + (1 - 2 * 1/(1 + e(-β*(x[i]-μ))))2/(4 * o * w[i] *
```

$(o * w[i] - 1) * 1 / (1 + e^{(-\beta * (x[i] - \mu))})^2 * (1 - 1 / (1 + e^{(-\beta * (x[i] - \mu))}))^2)^{(-1)}, i = 1..k) :$

Matrix(M);end proc:

mm:=proc(t)

global x;

local M:M:=matrix(2,2):

$M[1, 1] := \beta^2 * (((1 / ((1 / (1 + e^{(-\beta * (t - \mu))}) * (1 - 1 / (1 + e^{(-\beta * (t - \mu))})))))) + (((1 - 2 * 1 / (1 + e^{(-\beta * (t - \mu))}))^2) / (4 * (o * 1 - 1) * 1 / (1 + e^{(-\beta * (t - \mu))})^2 * (1 - 1 / (1 + e^{(-\beta * (t - \mu))}))^2)))^{(-1)});$

$M[1, 2] := -\beta * ((t - \mu) * ((1 / ((1 / (1 + e^{(-\beta * (t - \mu))}) * (1 - 1 / (1 + e^{(-\beta * (t - \mu))})))))) + (((1 - 2 * 1 / (1 + e^{(-\beta * (t - \mu))}))^2) / (4 * (o * 1 - 1) * 1 / (1 + e^{(-\beta * (t - \mu))})^2 * (1 - 1 / (1 + e^{(-\beta * (t - \mu))}))^2)))^{(-1)});$

$M[2, 1] := M[1, 2] :$

$M[2, 2] := (((t - \mu)^2 * ((1 / ((1 / (1 + e^{(-\beta * (t - \mu))}) * (1 - 1 / (1 + e^{(-\beta * (t - \mu))})))))) + (((1 - 2 * 1 / (1 + e^{(-\beta * (t - \mu))}))^2) / (4 * (o * 1 - 1) * 1 / (1 + e^{(-\beta * (t - \mu))})^2 * (1 - 1 / (1 + e^{(-\beta * (t - \mu))}))^2)))^{(-1)});$

Matrix(M);

end proc:

for n from 1 to N do

sum1:=0:

for i from 1 to (rβ) do

for j from 1 to (rμ) do

$\mu := \mu L + u1/2 + u1 * (j - 1) :$

$\beta := \beta L + u2/2 + u2 * (i - 1) :$

$M1[r\beta * (i - 1) + j] := mm(t);$

$v[r\beta * (i - 1) + j] := MatrixInverse(m(k)) :$

$e[r\beta * (i - 1) + j] := multiply((M1[r\beta * (i - 1) + j], v[r\beta * (i - 1) + j]));$

$sum1 := sum1 + e[r\beta * (i - 1) + j] :$

end do:

end do:

$f := trace(sum1) * u1 * u2 * (1/4) :$

$chek := NLPsolve(f, t = xL..xU, method = branchandbound, maximize) :$

if (chek[1] + 500 * chek[1]^2 < 0.00001) then

```

print(finish) for i from 1 to k do printf(" nx[d]=f w[d]=f n=d",i,x[i],i,w[i],n);
od:n:=N::
else
k:=k+1:
W:=w:
X:=x:
x[k]:=rhs(chek[2][1]);
w[k]:=1/(n+1);
for i from 1 to k-1 do
w[i]:=1-1/(n+1)*W[i];
x[i]:=X[i];
end do;
for i from 1 to k-1 do
if(abs(x[i] - x[k]) < 0.1)then
x[i]:=(w[k]*x[k]+x[i]*w[i])/(w[k]+w[i]);
w[i]:=w[k]+w[i];
k:=k-1:i:=k:
end if:
od:
for i from 1 to k do
if(w[i] < 0.07)then
p:=1:delta1:=10:
for j from 1 to k do
if(j <> i)then
delta:=abs(x[i]-x[j]):
if(delta < delta1)then
p:=j:delta1:=delta:
end if:
end if:
od:
x[i]:=(x[i]*w[i]+x[p]*w[p])/(w[i]+w[p]):
w[i]:=w[i]+w[p]:
X:=x:W:=w:k:=k-1:
for l from p to k do
x[l]:=X[l+1]:w[l]:=W[l+1]:
od:w[l]:=1:
end if:
od:

```

```
end if:
for i from 1 to k do od:
printf(" nK is d n n is d",k,n);print(chek);
printf(" n***** n")
printf(" nx[d]=f w[d]=f",i,x[i],i,w[i]);;
end do:
```

H. Talebi

Department of Statistics,
University of Isfahan,
Isfahan, Iran.
email: *h-talebi@sci.ui.ac.ir*

D. Poursina

Department of Statistics,
University of Isfahan,
Isfahan, Iran.
email: *davood.poursina@gmail.com*