Ghana J. Sci. **63** (2), 2022, 12 - 20 https://dx.doi.org/10.4314/gjs.v63i.2

A-AND D-OPTIMAL DESIGNS IN REAL LIFE DATA USING IMPERIALIST COMPARATIVE ALGORITHM

O. K. Bodunwa* & A. I. Adewole
(O. K. B.: Department of Statistics, Federal University of Technology, Akure, Nigeria; A. I. A.: Department of Mathematical Sciences, Tai Solarin University of Education, Ijebu- Ode, Ogun-State Nigeria.)
**Corresponding author's email: okbodunwa@futa.edu.ng*

ABSTRACT

In this paper, we used Imperialist Competitive Algorithm (ICA) to get A-and D-optimal designs for two models; one with nonlinear and the other linear model were found for two life data. It was noted that the number of iterations for D-optimal design for both models was smaller than that of A-optimal design. This implies that getting a D-optimal design is preferable and saves time and energy compared to A-optimal since they are addressing the same issues related to the variance- covariance function. The Efficiency Lower Bound and the Sensitivity function embedded in ICA which support the general equivalent theorem in getting optimal design were also found in this study. Another advantage of using ICA algorithm is that it does not require the design space or the region of uncertainty to be discretized which means that the search for the support points of the optimal design is not restricted to the grid point. This was also confirmed here that the optimal design for A and D- designs were not restricted to the grid points.

Keywords: Imperialist Competitive Algorithm, Efficiency Lower Bound, Sensitivity function, A-optimal design, D-optimal design.

Introduction

In an experiment with set of trials, where the result of each trial is an observation, a design is choosing the points in which the trials will be conducted. Experimentation is the process of planning a study to meet specified objectives which constitutes a foundation of the empirical sciences (Zhu, 2012). One major advantage of an experiment is its ability to control the experimental conditions, as well as to determine the variables to include in a study (Fackle, 2008). In recent years, optimal experimental designs have replaced standard designs from catalogs in many applications. Optimal designs of experiments are a set of designs that produce maximum efficiency with respect to some statistical criteria. Here, efficiency refers to the degree of the worth of experimental design, that is, the precision of variances of estimators when dealing with the estimation of model parameters. Classification of optimal designs is based on the concept that provides a methodical approach to obtaining the possible best or highly efficient design using available data with respect to the situation being considered.

Optimal experimental design is an essential area of scientific research. The fundamental idea underlying optimality theory of a design is that Statistical inference about quantities of interest can be improved by optimally selecting levels of the control variables at low costs. Multiple optimality criteria exist for selecting the optimal

experimental designs, these includes: D-optimal design, which minimizes the volume of the joint confidence region on the vector of regression coefficients. A-optimal criterion deals with only the variances of the regression coefficients. It minimizes the sum of the main diagonal elements of $(X'X)^{-1}$. This is called the trace of $(X'X)^{-1}$. Thus, an A-optimal design minimizes the sum of the variances of the regression coefficients. In a nutshell, the D-and A-optimality criteria focus on precise estimation of the model parameters and are, therefore, estimation oriented. A design is said to be G-optimal if it minimizes the maximum scaled prediction variance over the design region; V-optimal criterion considers the prediction variance at a set of points of interest in the design region, it minimizes the average prediction variance over the design region. I-optimal design minimizes the average prediction variance of the model over a region of movement parameters, therefore, the I- and G-optimality criteria focus on precise predictions of the response and are, therefore, prediction oriented.

In this article, we focus on the performance of A- and D-optimality criteria with application to two real-life data on the rate of transport of sulfite ions from blood cells suspended in a salt solution which were obtained by W. H. Dennis and P. Wood at the University of Wisconsin, and analyzed by (Sredni, 1970). The chloride concentration (tabe) was determined from a continuous curve generated from electrical potentials and the Kinematic Viscosity data obtained which were obtained and discussed in Linssen (1975). In the model, f is predicted ln(viscosity), x_1 is temperature, and x_2 is pressure. D-and A-optimality criteria are, therefore, more suitable to comparing screening experiments than the prediction oriented I- and G-optimality criteria. D-optimum design is independent of model parameters in an experiment involving effective dose levels in generating optimum design for one-variable first-order Poisson model. Investigations on the dependence of D-optimal designs concerning model parameters of the research interest with a quadratic term in one variable, as well as with additive two-variable with interaction term was considered and the performances of certain appealing standard designs examined (Russell *et al* 2009).

In literatures, optimal design is found from the theory when the model is nonlinear with one or two factors. This approach encounters mathematical difficulties when the model has more factors. Under such situations, our experience is that the classical optimization numerical techniques fail to find the locally optimal design or they become very inefficient. This is because as the number of factors in the model increases, the number of parameters in the model also increases. Consequently, the number of design points for the optimal design increases, resulting in having substantially many more variables to optimize. Thus, the design problem becomes quickly high-dimensional and also non-separable when factors interact with one another. Premature convergence can become a severe issue since solutions can easily get trapped in local optima.

In view of this, different algorithms on how to get optimal design are now increasing applied for complicated optimization problems.

The first algorithms for creating designed experiments using a computer employed the D-optimality criterion and produced highly D-efficient designs (see, e.g., (Atkinson, 1992); (Cook & Nachtrheim, 1980); (Mitchell, 2000)).

Algorithms and software for generating A-optimal designs are scarce compared with algorithms and software for generating

D-optimal designs. Likewise, there is far more literature on D-optimal designs than on A-optimal designs. One reason for this is that computer searches for A-optimal designs require the computation of the inverse of the information matrix, while searches for D-optimal designs only require the determinant of the information matrix to be computed. The latter is computationally less expensive than the former, as a result of which computing D-optimal designs requires less time than computing A-optimal designs. Another reason why A-optimal design are less commonly used than D-optimal designs is that, unlike D-optimal designs, A-optimal designs are not invariant to changes in the scale or the coding of the factors. A third reason for the relative lack of popularity of A-optimal designs is that D-optimal designs are often said to perform well with respect to other optimality criteria (Atkison & Donev, 1992) (Goos, 2002).

Particle Swarm Optimization (PSO) which was proposed by (Elberhart & Kennedy, 1995). This algorithm, like (Kennedy, 2011) is one such algorithm which has been recently used to solve various optimal design. (Qiu *et al.,* 2014) Used the algorithm to solve the optimization problems in animal instincts to design efficient biomedical studies via PSO. (Chen *et al.,* 2015)

However, the D-optimal design problems in these papers have only 3 or fewer factors in the statistical model and so premature convergence may not be an issue. Other algorithms were developed to address the problem of factors that are more than 3.

Differential Evolution (DE) is an algorithm from a family of gradient-free algorithms-evolutionary algorithms. Mutation, crossover and selection are three fundamental operations in DE (Das & Suganthan, 2011). One advantage that DE has over other evolutionary algorithms is that it has fewer control parameters and works well in handling numerical optimization problems (Zhang *et al.*, 2017). Compared with PSO, DE can alleviate the premature convergence issue moderately since most of the mutation strategies of DE do not exert the selective pressure onto the current best solution. However, based on the studies of DE variants for solving high-dimensional problems, there is no specially designed mechanism to explore various but novelty regions in the search space and to preserve the diversity of the population. (Masoudi *et al.,* 2016) Used ICA to find the minimax and standardized maximin optimal designs.

Jones *et al.*, (2020) investigated the performance of A and D-optimal designs for screening experiment and persuaded the experimenters to choose A-optimal designs rather than D-optimal. Their reasons is that A-optimal criterion is more consistent with the screening objective than the D-optimality criterion. (Bodunwa *et al.,* 2019) Found the D-optimal design in regression models with the problem of heteroscedasticity in the application of Kinematic Viscosity data. This was done for different models and three correction methods were used for the heteroscedasticity.

Experimental

Materials and method

In this section, we define the two optimality criteria we use in this article. The definitions of the A and D-optimality criteria goes thus:
Considering the model

$$Y = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \varepsilon \quad (1)$$

In vector notation, we can write the model as

$$Y = f^T(x)\beta + \varepsilon \quad (2)$$

Assuming that all n random errors are independent normally distributed random variables with zero mean and variance σ^2, the best unbiased estimator for the parameter

vector β is the ordinary least-squares estimator

$$\hat{\beta} = (x^T x)^{-1} x^T Y \tag{3}$$

The variance-covariance matrix of this estimator is

$$\sigma^2 (x^T x)^{-1} \tag{4}$$

The inverse of the matrix

$$\sigma^{-2} (x^T x) \tag{5}$$

This is the information matrix and summarizes the information content of the experimental design concerning the parameters in the model. A-optimal criterion searches for the minimization of the trace of the inverse of information matrix, i.e.,

$$\min_{x_i, i=1,\cdots,n} trace\ (X'X^{-1}). \tag{6}$$

The A-optimal criterion is equivalent to the minimization of the average variance of the p parameter estimates in $\hat{\beta}$ in regression model. As a result, an A-optimal design minimizes the sum of the diagonal elements of the ordinary least-squares estimator's variance– covariance matrix. Because that sum is called the trace of the variance–covariance matrix, an A-optimal design minimizes $(X^T X)^{-1}$.

The relative A-efficiency of a design with model matrix X1 with respect to a design with model matrix X2 is defined as

$$\frac{trace(X_2^T X_2)^{-1}}{trace(X_1^T X_1)^{-1}} \tag{7}$$

A relative A-efficiency larger than one indicates that the former design is better than the latter in terms of A-optimality.

D-optimal design maximizes the determinant of the Fisher information matrix, $|X^T X|$. Because this determinant is inversely related to the volume of the p dimensional confidence ellipsoid about the parameters in β, a D-optimal design minimizes that volume for a given number of runs n. We use the relative D-efficiency to compare the quality of two designs with model matrices X1 and X2 in terms of the D-optimality criterion. The relative D-efficiency of a design with model matrix X1 with respect to a design with model matrix X2 is defined as

$$D_{eff} = \left(\frac{|X_1{}^T X_1|}{|X_2{}^T X_2|} \right)^{1/p} \tag{8}$$

A relative D-efficiency larger than one indicates that the former design is better than the latter in terms of the D-optimality criterion.

Imperialist competitive algorithm was used to get A-optimal design and D-optimal design conditioned for some set of data in real life. The first data was on the rate of transport of sulfite ions from blood cells suspended in a salt solution which were obtained by W. H. Dennis and P. Wood at the University of Wisconsin, and analyzed by Sredni (1970). The chloride concentration (%) was determined from a continuous curve generated from electrical potentials. A model was derived from the theory of ion transport and the same model was used here as

$$f(x.\theta) = \theta_1 \left(1 - \theta_2 e^{-\theta_3 x}\right) \tag{9}$$

Where f is the predicted chloride concentration and x is the time. This data was coded between -1 to +1 as often used in literatures.

The second data was on the kinematic viscosity of a lubricant, in stokes, as a function of temperature (o_c), and pressure in atmospheres (atm) as discussed in Linssen (1975) where y is predicted ln(viscosity), x_1 is temperature, and x_2 is pressure. The model used was

$$y_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + e_i \quad (10)$$

The imperialist competitive algorithm (ICA) is a computational technique employed in solving different kinds of optimization problems. It is a mathematical model and computer assimilation of human social evolution. The ICA can be defined as a form of meta-heuristic algorithm designed for solving optimization problems. In design of experiments, the imperialist competitive algorithm is basically used to solve optimal design problems pertaining to non-linear models. The package" ICAOD" authored by (Masoudi E. , Holling, Wong, & Kim, 2020) in R version 1.0.1 was used for findings locally D-optimal designs for linear and nonlinear models. It should be used when a vector of initial estimates is available for the unknown model parameters. Locally optimal designs may not be efficient when the initial estimates are far away from the true values of the parameters. Since the criteria of optimality for nonlinear models depends on the unknown parameters, the locally optimal design function deals with the parameter-dependency based on the information available for the unknown parameters.

Results and discussion

This section describes the results of the optimal designs using two real life data sets. The first data was on the rate of transport of sulfite ions from blood cells suspended in a salt solution which were obtained by W. H. Dennis and P. Wood at the University of Wisconsin, and analyzed by Sredni (1970). The chloride concentration (%) was determined from a continuous curve generated from electrical potentials represented in equation (9) above. This data was coded between -1 to +1 as often used in literatures. Using the model in equation (9) with initial parameters of 1, 3, 5 respectively were used as constants for θ_1, θ_2 and θ_3 .

TABLE 1
The initialized values used.

ICA control parameters	Values
Inipars	(1, 3, 5)
L_x and U_x	(-1, 1)
K	3
Stop rule	equivalence
Check freq	50

Table 1 above shows the initialized values selected for getting locally D- and A-optimal designs in Imperialist Comparative Algorithm. These were selected based on examples from literatures. The *Inipars* represent the initial parameter on the model, L_x and U_x are lower and upper bound of predictors respectively, *K* is number of design points. Must be at least equal to the number of model parameters to avoid singularity of the FIM. *Stop rule* Denote the type of stopping rule which is the equivalence and the algorithm verifies the general equivalence theorem in every *Check freq* iterations.

Following the above description table, the result of locally optimum designs were presented generally shows the no of Iterations, the Design Points with associated Weight, the Efficiency Lower Bound (ELB) and the Criterion Values which is the log determinant of the inverse of Fisher Information Matrix (FIM).

TABLE 2

D-optimal design for predicted chloride concentration on time.

Iteration	**(Design Points) Weight**	**ELB**	**Sen. Fun**	**Crit. Val**
1	(-0.9426) 0.3584; (-0.7572) 0.3862; (0.8939) 0.2555	0.2571376	8.666907	-12.47912
10	(-0.9998) 0.2910; (-0.8089) 0.2896; (0.7950) 0.4193	0.866633	0.461673	-13.62527
50	(-1) 0.3333; (-0.8001) 0.3333; (0.9999) 0.3333	0.9998237	0.00052906	-13.68086

As shown in Table 2, as at the first 10 up to 49th iteration, the design points were yet to be optimal because the ELB are far from 1. Often, the worth of a design ξ is measured by its efficiency relative to the optimal design $\xi*$. For D-optimality, the D-efficiency of a design ξ is

If its D-efficiency is near 1, ξ is close to $\xi*$. If the theoretical optimal design $\xi*$ is unknown, the proximity of a design ξ to $\xi*$ can be determined from convex analysis theory. Specifically, its D-efficiency is at least $\exp(-\frac{\theta}{k})$ where θ is the maximum positive value of the sensitivity function across the entire design space and k is the dimension of the parameter (Pazman, 1986) and from the condition in which the optimal design points are reach was when the ELB is very close to 1. This was established by (Xu, Wong, Tan, & Xu, 2019) that says if the D-efficiency lower bound is close to 1, the design ξ is close to the D-optimal design $\xi*$. It was noted here that as the number of iteration was increased to 100, the algorithms automatically finished or stopped on 50th iteration. A general observation is that the value of sensitivity function converges as the number of iterations increases and the criterion value. Therefore, the D-optimal design for the equation (9) above is

$$\xi^* = \begin{Bmatrix} -1 & -0.8001 & 0.9999 \\ 0.333 & 0.333 & 0.333 \end{Bmatrix}$$

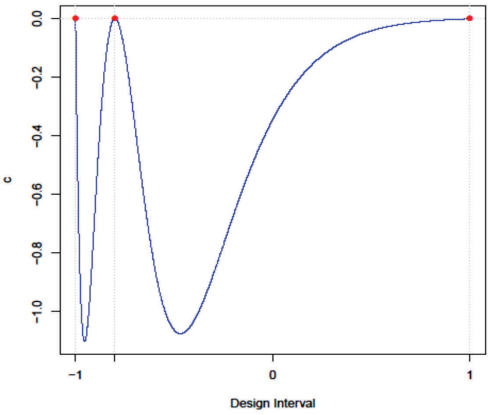

Fig. 1: D-optimal design - shows the derivative plot that confirms the optimality of the design point of the ICA generated designs.

TABLE 3

A-Optimal design for chloride.

Iteration	(Design Points) Weight	ELB	Sen. Fun	Crit. Val
1	(-0.9563) 0.0485; (-0.3136) 0.2545; (0.6576) 0.6970	0.08224637	33.47578	12.63548
10	(-0.9206) 0.0127; (-0.6398) 0.0541 (0.8263) 0.9332	0.1217664	21.63733	10.44512
50	(-0.9003) 0.00567; (-0.5423) 0.0432 (0.9965) 0.95113	0.8760041	3.58734232	10.2346
100	(-1) 0.00265; (-0.5192) 0.0494; (0.9999) 0.9479	0.9980996	0.005712154	10.20708
500	(-1) 0.0026; (-0.5192) 0.0494; (0.9999) 0.9479	0.9999931	2.07496e-05	10.20708
1000	(-1) 0.0026; (-0.5192) 0.0494; (1) 0.9479	0.9999931	2.07639e-05	10.20708

Table 3 is the equivalent of Table 2 in A-optimal criterion. The table shows that the number of Iteration for A-optimal is higher that of D-optimal. The design points as at 100 times remain the same with the associated weight but still displaying different ELB and sensitivity function with same criterion value. That same criterion values helps in generating the evolution plot in figure 3 that shows the best imperialist and the mean of it. The A-optimal design for the model in equation (9) is

$$\xi^* = \begin{Bmatrix} -1 & -0.5192 & 1 \\ 0.0026 & 0.0494 & 0.9479 \end{Bmatrix}$$

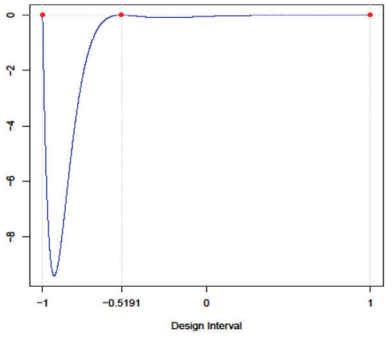

Fig. 2: Derivative plot of A-optimal design – shows the derivative plot that confirms the optimality of the design point of the ICA generated designs.

In the evolution plot in figure 3 below, the solid red line is the cost value of the best imperialist and the dotted blue line is the mean of all imperialists' costs. The evolution plots show how ICA's best imperialist quickly moves toward the optimal design after only a few iterations.

The second data set was on the kinematic viscosity of a lubricant, in stokes, as a function of temperature (o_c), and pressure in atmospheres (atm) as discussed in Linssen (1975) with the model (10) above.

This data was coded between -1 to +1 as often used in literatures. Both A- and D-optimal designs were also found here.

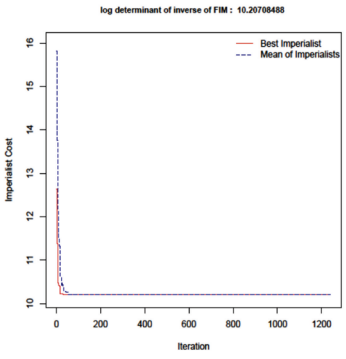

Fig. 3: Evolution plot

TABLE 4

D-Optimal design for real life on kinematic viscosity.

Iteration	(Design Points) Weight	ELB	Sen. Fun	Crit. Val
1	(0.9869, 0.6876) 0.36522; (-0.6243, -.96195) 0.26624; (-0.75962, 0.9244)0.1483; (0.3517, 0.1105)0.2202	0.3913577	4.665622	1.257633
10	(0.9337, 0.9958) 0.5156; (-0.9248, -0.9673) 0.2088; (-0.9988, 0.9930)0.1569; (-0.8836, 0.4275)0.1187	0.4683218	3.405851	0.6786509
50	(0.9990, 0.9999) 0.2746; (-0.9997, -0.9997) 0.2399; (-0.9998, 0.9989)0.2343; (-0.9949, 0.9962)0.2510	0.9554928	0.1397411	0.0095276

Here, since there are two predictor values, the design points were chosen from two points with the associated weight and the optimal design point were reached at 50th iteration time. The convergence of the sensitivity function and the criterion values were high and the D-efficiency lower bound was 95.5%. The D-optimal design for model (10) is

$$\xi^* = \begin{Bmatrix} (0.999, 0.999) & (-0.999, -0.9997) & (-0.9998, 0.9989) & (-0.9949, 0.9962) \\ 0.2746 & 0.2399 & 0.2343 & 0.2510 \end{Bmatrix}$$

TABLE 5

A-Optimal design for real life data on kinematic viscosity.

Iteration	(Design Points) Weight	ELB	Sen. Fun	Crit. Val
1	(0.9489, 0.6872) 0.36522; (-0.6237, -.96195) 0.26624; (-0.7596, 0.9238)0.1483; (0.3517, 0.1105)0.2202	0.1670044	14.9636	5.131856
10	(0.9328, 0.7627) 0.3059; (-0.8266, -0.9849) 0.3261; (-0.7686, 0.9821)0.1589; (-0.7014, 0.4947) 0.2090	0.4411568	3.800303	4.051697
100	(0.9999, 0.9999) 0.2502; (-0.9999, -0.9999) 0.2502; (-0.9999, 0.9999)0.2491; (-0.9999, 0.9999)0.2504	0.9955207	0.01349829	3.00005
500	(1, 1) 0.25; (-1, -1) 0.25; (-1, 1) 0.25; (1, -1) 0.25;	1	4.7835e-08	3

Table 5 shows the perfect design points as at 500 iteration for A-optimal design when the ELB is exactly 1 for the combination of minimum and maximum values of X_1 and X_2. A-optimal design for the model (10) is

$$\xi^* = \begin{Bmatrix} (1, 1) & (-1, -1) & (-1, 1) & (1, -1) \\ 0.25 & 0.25 & 0.25 & 0.25 \end{Bmatrix}$$

be restricted to the grid point. This was also confirmed here that the optimal design for A and D- designs were not restricted to the grid points.

Conclusion

In this paper, we used ICA to get A-and D-optimal designs for two models; one with nonlinear and the other linear model were found for two life data. It was noted that the number of iterations for D-optimal design for both models were smaller than that of A-optimal design. One of the advantages of using ICA algorithm is that it does not require the design space or the region of uncertainty to be discretized which mean that the search for the support points of the optimal design is not

References

Atkinson, A. C. (1992) A segmented algorithm for simulated annealing . Stat.Comput, 221- 230.

Atkison, A. C., & Donev, A. N. (1992) Optimum Experimental Designs. Oxford Statistical Science Series-8. Oxford University Press, Oxford.

Bodunwa, O. K., Fasoranbaku, O. A., & Ajiboye, A. S. (2019) Heteroscedatic D-optimal design in regression models with application in Kinematic Viscosity . FUTA Journal of reserach in sciences, 286 - 291.

Chen, R., Chang, S. P., Wang, W., Tung, H. C., & Wong, W. K. (2015) Minimax optimal designs via Particle Swarm Optimization methods. Statist. Comput, 975 - 988.

Cook, R. D., & Nachtrheim, C. J. (1980) A comparison of algorithms for constructing exact d-optimal designs. Technometrics, 315 - 324.

Das, S., & Suganthan, P. N. (2011) Differential evolution: A survey of the state- of -the- art. IEEETrans. Evol. Comput., 4 - 31.

Ehsan, M. (2016) Application of imperialist competitive algorithm to find. Computational Statistics and Data Analysis, 5-6.

Elberhart, R. C., & Kennedy, J. (1995) A new optimizer using particle swarm theory. In: Proceedings of the Sixth International Symposium on Micro Machine. Proceedings of the Sixth International Symposium on Micro Machine and Human Science, 39 -43.

Fackle, F. E. (2008) Optimal Design of Experiments for the Quadratic Logistic model. A Thesis submitted to the Department of Statistics, Stockholm University, Stockholm, in partial fulfillment of Doctor of Philosophy in Statistics.

Goos, P. (2002) The Optimal Design of Blocked and Split-plot Experiments. New York: Springer.

Jones, B., Allen-Moyer, K., & Goos, P. (2020) A-optimal Versus D-optimal design of screening experiments. Journal of Quality Technology.

Kennedy, J. (2011)``Particle swarm optimization," in Encyclopedia of Machine learning. Springer, 760 - 766.

Masoudi, E., Holling, H., & Wong, W. K. (2016) Application of Imperrialist competitive algorithm to find minimax and standardized maximin optimal design. Computational Statistics and Data Analysis.

Masoudi, E., Holling, H., Wong, W. K. & Kim, S. (2020) R. Package "ICAOD" Optimal Designs for Nonlinear Statistical Models by Imperialist Competitive Algorithm.

Mitchell, T. J. (2000) An algorithm for the construction of “D-Optimal” experimental designs. Technometrics, 48–54.

Pazman, A. (1986) Foundations of Optimum Experimental Design. Cham, Switzerland: Springer.

Qiu, J., Chen, R. B., Wang, W. & Wong, W. K. (2014). Using animal instincts to design efficient biomedical studies via particle swarm optimization. Swarm Evol. Comput., 1-10.

Russell, K. G., Woods, D. C. & Eccleston, J. A. (2009) D-optimal designs for Poisson regression models. Statistica Sinica, 721-730.

Sredni, J. (1970) Problems of design, estimation and lack of fit in model building. PhD Thesis, Department of Statistics, University of Wisconsin, Madison, Wisconsin.

Xu, W., Wong, W. K., Tan, K. C. & Xu, A. J. (2019) Findimg high dimensional D-optimal designs for Logistic Models via Differenrial Evolution. IEEEAccess.

Zhang, C., Lim, P., Qin, A. K. & Tan, K. C. (2017) Multiobjective deep belief network ensemble for remaining useful life estimation in prognostics. IEEE Tran. Neual Netw. Learn. Syst, 2306 - 2318.

Zhu, C. (2012) Construction of Optimal Design in polynomial regression models. . A thesis submitted to the Faculty of Graduates studies of The University of Manitoba in partial fulfilment of requirements for the Degree of Master of Science, Department of Statistics, Canada.

Received 14 Aug 22; revised 20 Jan 23.