INDIAN JOURNAL OF SCIENCE AND TECHNOLOGY

RESEARCH ARTICLE

OPEN ACCESS
Received: 07-08-2023
Accepted: 13-09-2023
Published: 25-10-2023

Citation: Swetha N, David J (2023) A – Optimal Design for Poisson Regression Model Using Square Root Link. Indian Journal of Science and Technology 16(39): 3407-3413. https://doi.org/10.17485/IJST/v16i39.1999

* **Corresponding author**.
swethanagalingam29@gmail.com

Funding: None

Competing Interests: None

Published By Indian Society for Education and Environment (iSee)

ISSN
Print: 0974-6846
Electronic: 0974-5645

A – Optimal Design for Poisson Regression Model Using Square Root Link

N Swetha[1]*, Joshua David[2]

1 Research Scholar, Department of Statistics, Madras Christian College, Chennai, 600059, Tamil Nadu, India
2 Assistant Professor, Department of Statistics, Madras Christian College, Chennai, 600059, Tamil Nadu, India

Abstract

Objective: The main objective of this study is to identify the optimal operating conditions by determining the point of least variation on response variables, crucial for various aspects of product quality such as yield or strength. **Method:** Newton-Raphson method is used and the solution to these formulas are obtained numerically using R-software. The parameters in the context of A-optimal designs is a necessary step i.e., θ= (α, β, μ). **Findings:** Our findings show that A-optimal designs outperform conventional design criteria in terms of lowering parameter estimate variance. Moreover, we illustrate how the A-optimal designs can lead to a more efficient allocation of resources statistical models, optimizing the use of time and resources while ensuring robust statistical conclusions. **Novelty and Applications :** In the context of generalized linear models, most of the recent studies were on logistic regression models and many of them focused on optimal experimental designs with concentration on D optimality. In this research, Poisson regression models were considered for A-optimization with square root link. This approach proves to be particularly valuable as it accounts for the model's dependence on unknown parameters, making it a useful technique for practical applications. **Suggestions**: This methodology is quite general and may be applied to find A-optimal designs for other models, like the Compound Poisson, Gamma and inverse Gaussian model, with log link or other types of optimal designs.

Keywords: AOptimality; Poisson Regression Model; Square Root Link; Experimental Design; Parameter Estimation

1 Introduction

Optimal experimental designs are increasingly being used to save money while obtaining the most accurate statistical results. The focus of this research is on model-based optimum designs, which entail the use of a statistical model with unknown parameters in conjunction with a design criterion. These designs are based on probabilistic constructs and use a criterion stated as a convex function of Fisher's information matrix[1]. This method seeks to find a design that minimizes the criterion within the constraints of the design space. For implicit models into optimal approximate design principles, one can refer to design resource by [2]. Additionally, the practical applications and optimality criteria of optimal design concepts across various fields have been highlighted and recent contributions in this area include works by [3–6].

In recent years, the field of biomedical [7] and clinical trials has paid more attention to optimal experimental designs for Poisson regression models. The dependence of design support points on unknown parameters of the Fisher information matrix is a significant difficulty in the building and evolution of design for generic nonlinear models [8–10] for non-linear optimal design. The best designs for generalized linear models can't be found without knowing the parameters [11,12]. Locally optimal designs for multivariate generalized linear model was given by [13]. [14,15] for count data with Poisson models and the Rasch Poisson model and with gamma block effects.

In order to reduce sampling uncertainty, the A-Optimality Criterion has been successfully applied in the best experimental design. Pool-based active learning of Poisson regression classifiers has been added to its list of uses. The A-Optimality Criterion has appealing theoretical properties, and practical evaluations have shown how well it supports active learning for the square root link regression model. The fundamental idea behind determining optimal operating conditions involves identifying the point of minimal variance in response variables, which correspond to various quality attributes of a product, such as yield or strength.

It is highly beneficial that the optimal design is based on the unknown true model parameters for the Poisson regression model. One possible solution to this problem is to use A-Optimal designs for scaled data, which are aimed particularly at calculating the percentile of the response curve. Non-linear predictors are frequently used in pharmacokinetics research. Despite significant advances in optimum experimental design, a significant research gap exists in the domain of A-optimality design. While other optimality criteria, such as D-, V E-, and G-optimality [3,16–20] have received a lot of attention, A-optimality is still relatively unknown. [21,22] provide detailed discussions on optimal designs for linear mixed effect models. [23,24] suggested an alternate approach to optimal design for non-linear regression models. Several researchers have investigated several optimal designs to improve parameter estimation in generalized linear models, including [25]. However, non-linear models provide more theoretical and computational challenges than linear models do. As a result, designing optimally for generalized linear models has become increasingly challenging. Generalized linear models have applications in experimental domains like constructions, clinical trials psychology, and engineering [4,7,26,27]. As a result, more research is required to determine the best designs for these models.

1.1 Comparative study

In this section, we will look at the essential components, advantages, limitations, and practical challenges related with A-optimal designs with a comparative study. A-optimal designs provide a useful framework for designing experiments and optimizing collecting data efforts in a variety of domains. Their capacity to maximize the precision of parameter estimations makes them a vital tool for increasing the quality of statistical analyses and facilitating informed decision-making.

[16] In this paper they have demonstrated how its output can be used to find new analytic A-optimal design. While the optimization of experimental designs for Poisson regression models has gained a lot of attention, there is still a significant research gap when it comes to A-optimal designs customized to the square root link function. When it comes to optimality criteria, the existing literature has primarily concentrated on traditional gamma and inverse Gaussian models with inverse link function. Filling this research gap is critical for realizing the full potential of the square root link function and enabling analysts to better understand the complexities of count data while determining statistically robust inferences. As they have mentioned that this methodology is quite general and may be applied to find analytical A-optimal designs for other models, like the Poisson model with the log link, square root link function, or other types of optimal designs, we have done the comparative study with this paper.

[17]In this article, they have compared A-optimal screening designs to D-optimal screening designs. They have provided the evidence that the A-optimality criterion produces more desirable screening designs than the D-optimality criterion, and that A-optimal designs generally perform better in terms of other optimality criteria than D-optimal designs. D optimal designs aims to minimize the determinant of the variance-covariance matrix directly. However, in some cases, this can lead to extreme values and instability. A-Optimal designs takes a slight different approach by minimizing the trace of the inverse of the variance-covariance matrix. This approach balances efficiency and stability. Therefore in our study we have focused on A-optimality in order to estimate the parameters with high precision.

In [19], used the quasi-information as an approximation to the Fisher information. V-criterion, which aims to minimize the mean variance of prediction of the mean response, is used to compute optimal experimental designs. And they have achieved locally optimal designs using this criterion in two unique scenarios: a Poisson linear regression model with random intercepts or random slopes. However, the literature on A-optimal designs for Poisson regression models using square root link remains conspicuously sparse. This gap signifies the absence of tailored methodologies for generating A-optimal designs that leverage the unique statistical properties of the square root link function. Addressing this gap is crucial to expand the applicability of A-optimal designs, enabling researchers in fields such as ecology and epidemiology to make more accurate and efficient inference from their count data.

2 Methodology

A-Optimal design is a concept in experimental design and statistics. It is the method of selecting experimental settings or data points in such a way that the variance of parameter estimations in a statistical model is minimized. In other words, A-Optimal designs are intended to provide the most efficient and precise estimation of model parameters based on the available data. It should be noted that A-Optimal designs are only one form of optimality criterion in experimental design. Other criteria include D-Optimal, E-Optimal, and G-Optimal designs, each having its own focus and mathematical properties. The optimality criterion chosen is determined by the study objectives and the properties of the statistical model being utilized.

Even if the volume of a confidence ellipsoid is small (as strived for using D-optimality criterion) all variances of the parameter estimates are not necessarily small. As an alternative, A-optimality strives for minimizing the sum of the variances of the parameter estimators. The diagonal elements of inverse of the standardized information matrix are proportional to the asymptotic variance of the ML estimator of θ. The design that minimizes

Ψ {M (ξ, θ)} = tr [$M^{-1}(\xi, \theta)$] is called A- Optimal.

Let $Y_1, Y_2 \ldots Y_n$ be independent Poisson-distributed response variables for n experimental units, with each Yi's density written as

$$\frac{\lambda_i{}^x}{x!} e^{-\lambda_i}, \qquad \lambda_i > 0$$

Then, using the given ”square root link” function g, the expected value of response Yi is related to the predictors Xi:

$$E[Y_i/x_i] = \mu_i = g(x_i, \beta) = g(\eta)$$

Where,

$$\mu = \eta^2$$

β is a set of parameters that must be estimated. The variance of the response is calculated as follows

$$Var[Y_i/x_i] = V(\mu_i)$$

Beyond the mean and variance relations stated above, the expected information is independent of the distribution's form. The jk^{th} element of the expected information matrix for GLMs are as follows:

$$I_{jk} = \sum_i^N \frac{x_{ij} x_{jk}}{V(\mu_i)} \left(\frac{\partial \mu_i}{\partial \eta_i} \right)^2$$

In matrix notation this would be

$$I = X'WX$$

Where W is a diagonal matrix having values

$$w_i = \frac{1}{V(\mu_i)} \left(\frac{\partial \mu_i}{\partial \eta_i} \right)^2$$

As a result, the asymptotic variance covariance matrix matrix.

The simplest form of Poisson model using the square root link function is given as follows.

$$\mu = \eta^2$$

2.1 A- Optimal design with Square root link

The linear predictor is defined by

$$\eta(x) = \alpha + \beta(x-\mu)^2 \tag{1}$$

The introduction of parameters in the context of A-optimal designs is a necessary step since A-optimal designs seek to minimize the variance-covariance matrix of parameter estimations. Parameters are θ= (α, β, μ).

Where α controls the height of the response, β denotes width of the curve, μ denotes optimum response.

The square root link

$$\eta(x) = \sqrt{\mu_i} \tag{2}$$

The information matrix under the Poisson model, is a given observation at x

$$I(\theta,x) = v(x)\left(\frac{\partial\eta(x)}{\partial\theta}\right)\left(\frac{\partial\eta(x)}{\partial\theta}\right)'$$

Where,

$$\left(\frac{\partial\eta(x)}{\partial\theta}\right) = \begin{pmatrix} 1 \\ (x_i-\mu)^2 \\ -2\beta(x_i-\mu) \end{pmatrix}$$

And

$$v(x) = \frac{1}{V(Y)}\left(\frac{\partial\mu_i}{\partial\eta_i}\right)^2 = \mu \tag{3}$$

Finding the point of maximum or minimum response, i.e. estimating the parameters, is a key application for this model. The optimal design is a set of points and weights that best optimize the selected criterion function; the criterion function is often connected to the precision of parameter estimations, such as the size of a confidence interval of parameter estimators' sum of the variance. The standardized information matrix is employed in the criterion function. The standardized information matrix for a certain design is the weighted total of the contributions from each of the n design points.

$$M(\theta,\xi) = \sum_i^n w_i v(x_i)\left(\frac{\partial\eta(x)}{\partial\theta}\right)\left(\frac{\partial\eta(x)}{\partial\theta}\right)'$$

$$\sum_i^n w_i\mu^2 \begin{pmatrix} 1 & (x_{ij}-\mu)^2 & -2\beta_j(x_{ij}-\mu) \\ (x_{ij}-\mu)^2 & (x_{ij}-\mu)^4 & -2\beta(x_{ij}-\mu)^3 \\ -2\beta_j(x_{ij}-\mu) & -2\beta(x_{ij}-\mu)^3 & 4\beta^2{}_j(x_{ij}-\mu)^2 \end{pmatrix}$$

There are several methods to the practice of determining the optimal design. These include algorithms, analytical, numerical and graphical methods.

3 Results and Discussion

From Eq (1), we have the linear predictor to be

$$P(y=|\,\chi,\beta,\mu) = \eta(x) = \alpha + \beta(x-\mu)^2$$

In the square root link predictive model, the response variable y represents class labels. The parameter μ, referred to as the "location parameter," is assigned the value of x such that P(y=|χ, β, μ). The parameter x, known as the "scale parameter," indicates the change in probability relative to x. In our experiments involving square root link predictive modeling, the Poisson regression employs more than two parameters. The model takes the following parametric form,

$$P(y=|\,\chi,\beta) = \pi^*(\mathrm{x}) = \left(\alpha - x^2\right)^2 \tag{4}$$

The term $x'\beta$ denotes the dot product of two vectors: x, the predictors, and β,the model parameters. In this paper, we choose to approach the model from a regularization standpoint rather than adopting the Bayesian viewpoint. Our Gaussian priors on the parameters do not incorporate any domain-specific knowledge but can be regarded as a plausible interpretation of the parametric form of the regularization.

3.1 Parameter Estimation

For the Square Root Link Poisson Regression Predictive Model, an objective function is derived. The true parameter values are denoted as β, while the maximum likelihood (Ml) estimates of the model parameters are represented as β. assuming that the parameterization of the model has no effect on the response distributions, the true parameters can be determined to be those that approach as the training set size increases. The regularized Inverse regression model's observed Fisher information matrix is designated by the notation I (X, θ) where X denotes the training set predictor matrix, often called the design matrix. The observations x_n are vectors formed from the rows of the training set matrix X. The objective function to minimize the variance.

$$\sum_{n\in \text{ Pool}} \text{Var}\left(\sigma\left(x'n\widetilde{\beta}\right)\right) = \sum_{n\in \text{ Pool}} \text{E} \mid \sigma\left(x'n\widetilde{\beta}\right)\right) - \sigma\left(x'n\widetilde{\beta}\right)^2 \tag{5}$$

In other words, we want the model predictions over the entire pool to be as close as possible to the predictions of the "true" model, in the squared loss sense, which minimizes the prediction variance over a single observation.

Using the Taylor series approximation we have:

$$\text{Var}\left(\sigma\left(x'_{nn}\widetilde{\beta}\right)\right) \simeq \text{Var}\left(c'_n\widetilde{\beta} - \beta\right)\right) \simeq c'_n I(X,\widetilde{\beta}) - 1_{cn} \tag{6}$$

Defining $A_n = c_n c'_n$ and A= ∑n A_n we drive a formula for minimizing the variance over the pool:

$$\begin{aligned} \sum_{n\in \text{ Pool}} c'_n I(X,\widetilde{\beta}) - 1_{cn} &= \sum_{n\in \text{ Pool}} \text{tr}\left\{A_n I\left(X,\widetilde{\beta}^{-1}\right\} \\ &= \text{tr}\left\{AI\left(X,\widetilde{\beta}^{-1}\right\} = \phi(\text{X},\text{y}) \end{aligned} \tag{7}$$

The objective functions for A-optimally in square root link predictive regression involve a matrix that provides the A matrix. In the literature, this matrix is often denoted as A (θ) to explicitly indicate its dependency on the model parameters (or the labeling of the training set), which we have not explicitly included in our notation. Instead, we utilize the notation ϕ(X, y) to demonstrate the criterion's dependency on the response value of the link function. Considering the parameters values to be α=1 and β= -0.6 and the corresponding x value is, x=1.06017, we have the following solution,

Ψ {M (ξ, θ)} = [tr M^{-1} (ξ, θ)] = 0.714

In the table we have considered the parameters θ= (α, β, μ). Here for fixed x=1.06017 and α=0 and for varying β and μ values, we have got the above table values. Newton-Raphson method is used and the solution to these formulas are obtained numerically using R-software.

Figure 1 For fixed α=0, and for varying β and μ values were considered, Our findings show that when dealing with optimal designs using a square root link function, there is a linear relationship between the function of mean values and predicted values; that is, as the function of mean values increases, the predictors decrease, resulting in the emergence of an asymptotic information matrix.

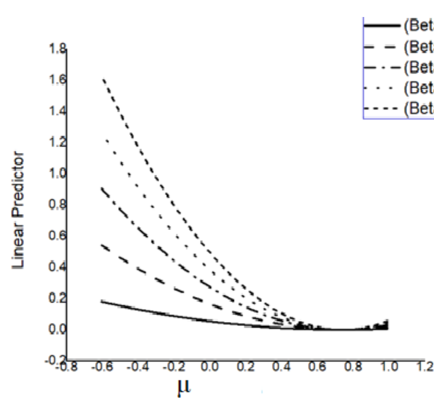

Fig 1. A-Optimal design (linear predictor vs mean values)

Table 1. Support points of the locally A-Optimal design for the Poisson model for various values of μ_i

μ_i	β =0.1	β =0.3	β =0.5	β =0.7	β =0.9
0.989999	0.000492	0.001477	0.002461	0.003446	0.004431
0.980292	0.000638	0.001914	0.003190	0.004466	0.005742
0.970774	0.000799	0.002397	0.003995	0.005594	0.007192
0.961438	0.000974	0.002924	0.004874	0.006823	0.008773
0.943296	0.001365	0.004097	0.006829	0.009561	0.012293
0.934479	0.001579	0.004739	0.007899	0.011058	0.014218
0.925826	0.001804	0.005414	0.009024	0.012633	0.016243
0.917331	0.002040	0.006120	0.010201	0.014282	0.018362
0.908991	0.002285	0.006856	0.011427	0.015998	0.020569
0.908991	0.002285	0.006856	0.011427	0.015998	0.020569
0.900801	0.002539	0.007619	0.012699	0.017778	0.022858
0.892757	0.002802	0.008408	0.014013	0.019618	0.025224
0.884856	0.003073	0.009220	0.015367	0.021514	0.027661
0.877093	0.003351	0.010055	0.016758	0.023462	0.030165
0.869465	0.003636	0.010910	0.018184	0.025457	0.032731
0.791615	0.007212	0.021636	0.036060	0.050484	0.064909
0.783593	0.007649	0.022948	0.038247	0.053546	0.068845
0.775570	0.008099	0.024298	0.040498	0.056697	0.072896
0.767548	0.008562	0.025688	0.042813	0.059939	0.077064
0.751503	0.009527	0.028582	0.047637	0.066692	0.085747
0.743480	0.010029	0.030087	0.050146	0.070204	0.090262
0.735458	0.010543	0.031631	0.052718	0.073806	0.094893
0.727435	0.011071	0.033213	0.055356	0.077498	0.099640

4 Conclusion

The study identifies the optimal operating conditions by determining the point of least variation on response variables. We accomplished this by employing a standardized information matrix concerning the asymptotic variance of maximum likelihood (ML) estimators. The results reveal that when dealing with optimal designs utilizing a square root link function, there exists linear relationship between the function of mean values and predicted values. To elaborate, as the function of mean values experiences an increase, the predictors decrease, consequently leading to the emergence of an asymptotic information matrix. This investigation successfully estimated the minimum variance of the surface point. This approach proves to be particularly valuable as it accounts for the model's dependence on unknown parameters, making it a useful technique for practical applications and as we have mentioned that this methodology is quite general and may be applied to find A-optimal designs for other models, like the compound Poisson, Gamma and inverse Gaussian model, with log link or other types of optimal designs.

Our future study is to investigate how A-optimal designs can be adapted to accommodate more complex statistical models, such as non-linear models or models with multiple response variables. This could involve developing new algorithms or optimization techniques. So that it can advance the field of A-optimal design and contribute to its broader applicability across various scientific and engineering disciplines.

References

1) Jung Y, Lee I. Optimal design of experiments for optimization-based model calibration using Fisher information matrix. *Reliability Engineering & System Safety*. 2021;216:107968–107968. Available from: https://doi.org/10.1016/j.ress.2021.107968.
2) Belmiro PM, Duarte, Atkinson AC, José FO, Granjo, Nuno MC, et al. Optimal Design of Experiments for Implicit Models. *Journal of the American Statistical Association*. 2021;117(539):1424–1437. Available from: https://doi.org/10.1080/01621459.2020.1862670.
3) Liu X, Yue R, Zhang Z, Wong WK. G-optimal designs for hierarchical linear models: an equivalence theorem and a nature-inspired meta-heuristic algorithm. *Soft Computing*. 2021;25(21):13549–13565. Available from: https://doi.org/10.1007/s00500-021-06061-0.
4) Liu L, Li J, Deng Y, Yang Z, Huang K, Zhao S. Optimal design of multi-layer structure composite containing inorganic hydrated salt phase change materials and cement: Lab-scale tests for buildings. *Construction and Building Materials*. 2021;275:122125–122125. Available from: https://doi.org/10.1016/j.conbuildmat.2020.122125.

5) Borrotti M, Sambo F, Mylona K. Multi-objective optimisation of split-plot designs. *Econometrics and Statistics*. 2023;28:163–172. Available from: https://doi.org/10.1016/j.ecosta.2022.04.001.
6) García-Ródenas R, García-García JC, López-Fidalgo J, Ángel Martín-Baos J, Wong WK. A comparison of general-purpose optimization algorithms for finding optimal approximate experimental designs. *Computational Statistics & Data Analysis*. 2020;144:106844–106844. Available from: https://doi.org/10.1016/j.csda.2019.106844.
7) Wong WK, Zhou J. Using CVX to Construct Optimal Designs for Biomedical Studies with Multiple Objectives. *Journal of Computational and Graphical Statistics*. 2022;32(2):744–753. Available from: https://doi.org/10.1080/10618600.2022.2104858.
8) Goos P, Syafitri U, Sartono B, Vazquez AR. A nonlinear multidimensional knapsack problem in the optimal design of mixture experiments. *European Journal of Operational Research*. 2020;281(1):201–221. Available from: https://doi.org/10.1016/j.ejor.2019.08.020.
9) Rahman MS, Shaikh AA, Ali I, Bhunia AK, Fügenschuh A. A Theoretical Framework for Optimality Conditions of Nonlinear Type-2 Interval-Valued Unconstrained and Constrained Optimization Problems Using Type-2 Interval Order Relations. *Mathematics*. 2021;9(8):908. Available from: https://doi.org/10.3390/math9080908.
10) Qiu K, Wang R, Xie Z, Zhu J, Zhang W. Comment on "Optimal design of chiral metamaterials with prescribed nonlinear properties" . 2021. Available from: https://doi.org/10.1007/s00158-020-02833-8.
11) Pijyan A, Zheng Q, Hong HG, Li Y. Consistent Estimation of Generalized Linear Models with High Dimensional Predictors via Stepwise Regression. *Entropy*. 2020;22(9):965. Available from: https://doi.org/10.3390/e22090965.
12) Idais O. A note on locally optimal designs for generalized linear models with restricted support. *Statistics & Probability Letters*. 2020;159:108692. Available from: https://doi.org/10.1016/j.spl.2019.108692.
13) Idais O. Locally optimal designs for multivariate generalized linear models. *Journal of Multivariate Analysis*. 2020;180:104663. Available from: https://doi.org/10.1016/j.jmva.2020.104663.
14) Grahoff U, Holling H, Schwabe R. D-optimal design for the Rasch counts model with multiple binary predictors. *British Journal of Mathematical and Statistical Psychology*. 2020;73(3):541–555. Available from: https://doi.org/10.1111/bmsp.12204.
15) Schmidt M, Schwabe R. Optimal designs for Poisson count data with Gamma block effects. *Journal of Statistical Planning and Inference*. 2020;204:128–140. Available from: https://doi.org/10.1016/j.jspi.2019.05.002.
16) Chen PY, Chen CC, Chen YS, Kee W, Wong. Numerical Methods for Finding A-optimal Designs Analytically. *Econometrics and Statistics*. 2022. Available from: https://doi.org/10.1016/j.ecosta.2022.09.005.
17) Jones B, Allen-Moyer K, Goos P. A-optimal versus D-optimal design of screening experiments. *Journal of Quality Technology*. 2021;53(4):369–382. Available from: https://doi.org/10.1080/00224065.2020.1757391.
18) Özdemir A, Turkoz M. Development of a D-optimal design-based 0–1 mixed-integer nonlinear robust parameter design optimization model for finding optimum design factor level settings. *Computers & Industrial Engineering*. 2020;149:106742. Available from: https://doi.org/10.1016/j.cie.2020.106742.
19) Niaparast M, Mehrmansour S, Schwabe R. V-optimality of designs in random effects Poisson regression models. *Metrika*. 2023;86(8):879–897. Available from: https://doi.org/10.1007/s00184-023-00896-3.
20) Olamide EI, Fasoranbaku OA, Adebola FB. E-optimal Experimental Designs for Poisson Regression Models in Two and Three Variables. *Tanzania Journal of Science*. 2021;47(3):999–1006. Available from: https://doi.org/10.4314/tjs.v47i3.11.
21) Jiang H, Yue R. D-optimal population designs in linear mixed effects models for multiple longitudinal data. *Statistical Theory and Related Fields*. 2021;5(2):88–94. Available from: https://doi.org/10.1080/24754269.2021.1884444.
22) Goudarzi M, Khazaei S, Jafari H. D-optimal designs for linear mixed model with random effects of Dirichlet process. *Communications in Statistics - Simulation and Computation*. 2021;p. 1–10. Available from: https://doi.org/10.1080/03610918.2021.1987467.
23) Masoudi E, Holling H, Duarte BPM, Wong WK. A Metaheuristic Adaptive Cubature Based Algorithm to Find Bayesian Optimal Designs for Nonlinear Models. *Journal of Computational and Graphical Statistics*. 2019;28(4):861–876. Available from: https://doi.org/10.1080/10618600.2019.1601097.
24) Feng Y, Liu Q, Yao Q, Zhao G. Model Averaging for Nonlinear Regression Models. *Journal of Business & Economic Statistics*. 2022;40(2):785–798. Available from: https://doi.org/10.1080/07350015.2020.1870477.
25) Li K, Yang J. Score-matching representative approach for big data analysis with generalized linear models. *Electronic Journal of Statistics*. 2022;16(1):16. Available from: https://doi.org/10.1214/21-ejs1965.
26) Pardo S. Generalized Linear Models. 2020. Available from: https://doi.org/10.1007/978-3-030-43328-4_9.
27) Brown VA. An Introduction to Linear Mixed-Effects Modeling in R. *Advances in Methods and Practices in Psychological Science*. 2021;4(1). Available from: https://doi.org/10.1177/2515245920960351.