

On Optimal Design for Regression Model (with Two Independent Variables) Considering Weighted Normal Distribution

Habib Jafari*, Shima Pirmohamadi, Sedighe Parviz

Department of Statistics, Razi University, Kermanshsh, Iran

Abstract During sampling, a researcher may sometimes encounter a population, in which selection chance of each one of the samples is not equal; in such a situation, it is better to utilize weighted distribution instead of main distribution for data analysis. Accordingly, considering a weighted distribution (for a state in which response variable has such a distribution in a linear regression model) and also a multiple regression (two regressors) model, this paper proposed an appropriate design (for estimating the parameters) using the known optimality design (D-optimality) to obtain optimality of the presented design.

Keywords Optimal Criterion, Optimal Design, D-Optimal Design, Weighted Distribution

1. Introduction

Assume that, while sampling from a population, chance of selecting the samples is not equal. In such a case, samples would follow a weighted distribution. Weighted distributions which are obtained from a main distribution could have various weights; selection of weights depends on sample types and sampling conditions[6]. The weight which is used here has been known as a skewed length and occurs when the selection possibility of a distance is in proportion to length of that distance[7]. In this study, a linear regression model with two independent variables was considered so that its response variable had normal weighted distribution with the mentioned weight. In this state, to estimate available parameters in the presented model, a suitable design was introduced and a desirable optimality criterion (Section 4) was used to locally optimize the introduced design (in this article, non-linear models were considered and local optimization method was used considering the application of the presented criterion and dependence of Fisher information matrix on the parameters). Accordingly, the article structure is as follows:

In Section 2, linear regression model and weighted distribution for its response variable are presented. Fisher information matrix in normal weighted distribution is presented in Section 3. Locally D-optimal design for some of the partitioned spaces of parameters is given in Section 4 and Section 5 concludes the paper.

2. Linear Regression Model with Skewed Responses

Assume $y_j = \mathbf{x}_j^T \boldsymbol{\beta} + \varepsilon_j$; $j = 1, \dots, n$, in which y_j is a response variable, $\mathbf{x}_j = (1, x_{1j}, x_{2j})^T$ is vector of regression function, $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2)^T$ is vector of unknown parameters, ε_j is random error and n is sample volume. In the above presented regression model, response variable has normal distribution as $Y_j \sim N(\mathbf{x}_j^T \boldsymbol{\beta}, \sigma^2)$. Now, without losing generality, assume $\sigma^2 = 1$; thus, density function of Y_j random variable will be as:

$$f_{Y_j}(y_j) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(y_j - \mu_j)^2}, \quad (1)$$

in which $\mu_j = \mathbf{x}_j^T \boldsymbol{\beta}$.

Density function of weighted distribution is generally defined as[5]:

$$f_Y^w(y) = \frac{w(y)f_Y(y)}{E^f(w(Y))} \quad (2)$$

where $w(y) \geq 0$ and is called weight function so that

$$0 < E^f(w(Y)) < \infty.$$

Accordingly, $f_Y^w(y)$ has the conditions of density function.

As mentioned in Section 1, the weight which was studied here was weight of skewed length; i.e. $w(Y) = Y$; however, since Y has normal distribution, positive condition of $w(Y)$ will not always hold. Therefore, the truncate distribution of Y in $Y \geq 0$ is used, in which case density function of normal weighted distribution is as follows:

$$f_{Y_j}^w(y_j | \mathbf{x}_j, \boldsymbol{\beta}) = \frac{y_j}{\mathbf{x}_j^T \boldsymbol{\beta}} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(y_j - \mathbf{x}_j^T \boldsymbol{\beta})^2}. \quad (3)$$

In this relation, mean is regarded greater than 4[4].

* Corresponding author:

jafari_habib@yahoo.com (Habib Jafari)

Published online at <http://journal.sapub.org/am>

Copyright © 2013 Scientific & Academic Publishing. All Rights Reserved

According to what was mentioned in the introduction, a suitable design was considered in the following way in order to estimate passive parameters:

$$\xi = \begin{pmatrix} (x_{11} & x_{12} & \dots & x_{1m}) \\ (x_{21} & x_{22} & \dots & x_{2m}) \\ w_1 & w_2 & \dots & w_m \end{pmatrix} \in \mathcal{E}. \quad (4)$$

Then, the present design was optimized using D-optimality criterion (Section1-4). In Relation (4), \mathcal{E} is space design which is defined as $\mathcal{E} = \{\xi | \xi \in \mathcal{E}, \sum_{j=1}^m w_j; 0 \leq w_j \leq 1\}$. Considering definition of the criterion used in this study, Fisher information matrix should be calculated for each and every member of Design (4). Then, the information matrix related to the mentioned design will be calculated as:

$$\mathbf{M}(\xi, \beta) = \sum_{j=1}^m w_j \mathbf{M}(x_j, \beta) \quad (5)$$

where $p \leq m \leq \frac{p(p+1)}{2}$ and p is the number of unknown parameters[8].

3. Fisher Information Matrix

Fisher information matrix for each and every observation (related to Design (4)) in a non-weighted distribution is as follows:

$$\mathbf{M}(x_j, \beta) = E^f \left(-\frac{\partial^2 \log f(y | x_j, \beta)}{\partial \beta \partial \beta^T} \right) \quad (6)$$

And in a weight distribution, it refers to:[5]

$$\mathbf{M}^w(x_j, \beta) = \mathbf{M}(x_j, \beta) + \frac{\partial^2 \log h(\mu_j)}{\partial \mu_j^2} x_j x_j^T \quad (7)$$

where $h(\mu_j) = E^f(w(Y_j))$. Now, calculating $\mathbf{M}(x_j, \beta)$, $h(\mu_j)$ and x_j and using Function (1) and Equation (7), the following can be given:

$$\mathbf{M}^w(x_j, \beta) = \begin{pmatrix} 1 - \frac{1}{\mu_j^2} & x_{1j} - \frac{x_{1j}}{\mu_j^2} & x_{2j} - \frac{x_{2j}}{\mu_j^2} \\ x_{1j} - \frac{x_{1j}}{\mu_j^2} & x_{1j}^2 - \frac{x_{1j}^2}{\mu_j^2} & x_{1j}x_{2j} - \frac{x_{1j}x_{2j}}{\mu_j^2} \\ x_{2j} - \frac{x_{2j}}{\mu_j^2} & x_{1j}x_{2j} - \frac{x_{1j}x_{2j}}{\mu_j^2} & x_{2j}^2 - \frac{x_{2j}^2}{\mu_j^2} \end{pmatrix}$$

This matrix depends on unknown parameters and will be used in the following section to obtain the optimal design.

4. Locally D-optimal Design

Information matrix is very important in optimal designs since various types of optimality criteria are defined based on the information matrix. Among optimality criteria, D-criterion is used and focused more frequently due to its stability and that is why this criterion was also used in this study.

In general terms, D-optimality criterion is defined as $-\log \det \mathbf{M}(\xi, \beta)$ [1]. Since information matrix of weighted distribution is considered here, the aforesaid criterion will be defined as $-\log \det \mathbf{M}^w(\xi, \beta)$.

To calculate an optimal design, first, a design with the determined number of points should be considered[3]. Three, four, five and six point designs could be assumed based on the number of model's unknown parameters according to

Carathéodory's theorem[8]. The designs considered here included three and four point designs.

As mentioned in Section 3, $\mathbf{M}^w(x_j, \beta)$ depends on unknown parameters; therefore, D-optimal design was locally calculated in this study.

In Table 1, locally D-optimal design is calculated in terms of various values of parameters when the number of points in this design is equal to the number of parameters. Table 2 shows locally D-optimal design for a state in which the number of points is four.

In Tables 1 and 2, x_{ij}^* ; $i = 1, 2$ and $j = 1, 2, 3, 4$ and w_j^* ; $j = 1, 2, 3, 4$ are respectively points and weights of D-optimal design. Additionally, $x_{1j}^* \in \mathcal{X}_1$, $x_{2j}^* \in \mathcal{X}_2$, $w_j^* \in [0, 1]$ and $\beta_l \in [-2, 2]$; $l = 0, 1, 2$ in which $\mathcal{X}_1 = \mathcal{X}_2 = [-9, 9]$.

It is noteworthy that values of weights in these tables are rounded up to three decimal places.

Table 1. Locally D-optimal design with three points

β_0	β_1	β_2	$\begin{pmatrix} (x_{11}^* & x_{12}^* & x_{13}^*) \\ (x_{21}^* & x_{22}^* & x_{23}^*) \\ w_1^* & w_2^* & w_3^* \end{pmatrix}$
1	-2	0.8	$\begin{pmatrix} (-1) & (-9) & (1) \\ (1.25) & (0) & (9) \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix}$
		0.9	$\begin{pmatrix} (-1) & (-9) & (1) \\ (9) & (0) & (5.55) \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix}$
	-1.9	1	$\begin{pmatrix} (-1) & (-9) & (1) \\ (9) & (0) & (4.90) \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix}$
		1.1	$\begin{pmatrix} (-1) & (-9) & (1) \\ (9) & (0) & (4.45) \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix}$
	-1.8	1.2	$\begin{pmatrix} (-1) & (-9) & (1) \\ (9) & (0) & (3.91) \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix}$
		1.3	$\begin{pmatrix} (-1) & (-9) & (1) \\ (9) & (0) & (3.61) \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix}$
1.1	-1.7	1.4	$\begin{pmatrix} (-1) & (-9) & (1) \\ (9) & (0) & (3.28) \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix}$
		1.5	$\begin{pmatrix} (-1) & (-9) & (1) \\ (9) & (0) & (3.06) \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix}$

Table 1 shows that, if the number of points in the design is equal to the number of unknown parameters, in this case, the uniform design would be optimal[8]; however, points of the design would change considering the values for the unknown parameters.

Table 2. Locally D-optimal design with four points

β_0	β_1	β_2	$\begin{pmatrix} (x_{11}^* & x_{12}^* & x_{13}^* & x_{14}^*) \\ (x_{21}^* & x_{22}^* & x_{23}^* & x_{24}^*) \\ w_1^* & w_2^* & w_3^* & w_4^* \end{pmatrix}$
-1	2	-0.8	$\begin{pmatrix} (-1) & (9) & (0) & (2.90) \\ (-9) & (0) & (-9) & (1) \\ 0.250 & 0.250 & 0.250 & 0.250 \end{pmatrix}$
		-0.9	$\begin{pmatrix} (-1) & (9) & (0) & (2.95) \\ (-9) & (0) & (-9) & (1) \\ 0.250 & 0.250 & 0.250 & 0.250 \end{pmatrix}$
	1.9	-1	$\begin{pmatrix} (-1) & (9) & (0) & (3.15) \\ (-9) & (0) & (-9) & (1) \\ 0.251 & 0.251 & 0.249 & 0.249 \end{pmatrix}$
		-1.1	$\begin{pmatrix} (-1) & (9) & (0) & (3.21) \\ (-6.27) & (0) & (-9) & (1) \\ 0.210 & 0.210 & 0.290 & 0.290 \end{pmatrix}$
-1.1	1.8	-1.2	$\begin{pmatrix} (-1) & (9) & (0) & (3.50) \\ (-5.75) & (0) & (-9) & (1) \\ 0.204 & 0.204 & 0.296 & 0.296 \end{pmatrix}$
		-1.3	$\begin{pmatrix} (-1) & (9) & (0) & (3.55) \\ (-5.30) & (0) & (-9) & (1) \\ 0.199 & 0.199 & 0.301 & 0.301 \end{pmatrix}$
	1.7	-1.4	$\begin{pmatrix} (-1) & (9) & (0) & (3.82) \\ (-4.85) & (0) & (-9) & (1) \\ 0.194 & 0.194 & 0.306 & 0.306 \end{pmatrix}$
		-1.5	$\begin{pmatrix} (-1) & (9) & (0) & (3.88) \\ (-4.53) & (0) & (-9) & (1) \\ 0.194 & 0.194 & 0.306 & 0.306 \end{pmatrix}$

5. Conclusions

If selection chance of sampling is not equal, a weighted distribution is considered for analyzing the data. In this study, this weighted distribution was studied for the response variable in a multiple linear regression model (with two regressors). The focused weighted distribution was a normal weighted distribution with the weight function of skewed length.

In order to obtain D-optimal design, first, Fisher information matrix was calculated based on the weighted distribution so that the information matrix related to the

weighted distribution was obtained in terms of information matrix of the main distribution. Then, D-optimal design was locally calculated considering the dependence of information matrix on the model's unknown parameters. Locally D-optimal designs are mentioned in Tables 1 and 2 as specific states of D-optimal designs for some values of parameters.

In addition to three and four point designs, five and six point designs could be considered and locally D-optimal designs can be also calculated in such states.

ACKNOWLEDGEMENTS

This work is partially supported by the Cooperative 4407 knowledge Base New Ideas.

REFERENCES

- [1] Atkinson, A.C., Donev, A.N., Tobias, R.D. (2006). *Optimum Experimental Designs, with SAS*. Oxford University Press.
- [2] Dette, H., Melas, V.B., Wong, W.K. (2005). Optimal Design for Goodness-of-Fit of the Michaelis-Menten Enzyme Kinetic Function. *J. Am. Stat. Assoc* 100:1370-1381.
- [3] Ford, I., Torsney, B., Wu, G.F.J. (1992). The Use of a Canonical Form in the construction of Locally Optimal Designs for non-Linear Problems. *J R Stat Soc Ser B* 54:569-583.
- [4] Navarro, J., Ruiz, J.M., del Aguila, Y. (2001). Parametric Estimation from Weighted Samples. *Biom J* 43:297-311.
- [5] Ortiz, I., Martinez, I., Rodriguez, C., del Aguila, Y. (2009). Optimal Designs for Generalized Linear Models with Biased Response. *Metrika* 70:225-237.
- [6] Patil, G.P., Rao, C.R., (1978). Weighted Distribution and Size-Biased with Applications to Wildlife Populations and Human Families. *Biometrics* 34:179-189.
- [7] Satten, G.A., Kong, F., Wright, Glynn, S., Schreiber, G., (2004). How Special Is a Special Interval: Modeling Departure from Length-Biased Sampling in Renewal Processes. *Biostatistics* 5:145-151.
- [8] Silvey, S.D. (1980). Optimal Design an Introduction to the Theory for parameter Estimation. *London: Chapman and Hall*.