Journal of Research in Mathematics Trends and Technology (JoRMTT) Vol. 3, No. 1, 2021 | 20-24

talenta PUBLISHER

Journal of Research in Mathematics Trends and Technology

Application of Fisher-scoring Algorithm on Parameter Estimation of Normal Distributed Data

S. A. Purba[1*], G. H. D. Sinaga[2], and J. E. Simarmata[3]

[1]Mathematics Study Program, Universitas HKBP Nommensen Pematangsiantar, Pematangsiantar, 21136, Indonesia
[2]Mechanical Engineering Study Program, Universitas HKBP Nommensen Pematangsiantar, Pematangsiantar, 21136, Indonesia
[3]Mathematics Education Study Program, Universitas Timor, Kefamenanu,85613, Indonesia

Abstract. In statistics, parameter estimation is the estimation of a population using sample data. A population data certainly has a certain distribution. Fisher-scoring is a form of Newton's method which is commonly used in solving the maximum likelihood equation. The focus of this research is to estimate distributed data using the Fisher-scoring algorithm.

Keyword: Fisher-scoring, Maximum Likehood, Estimation Parameter

Abstrak. *Dalam statistik estimasi parameter adalah pendugaan suatu populasi dengan menggunakan data sampel. Suatu data populasi tentunya memiliki distribusi tertentu. Fisher-scoring merupakan bentuk dari metode Newton yang biasa digunakan dalam menyelesaikan persamaan maksimum likelihood Fokus pada penelitian ini adalah melakukan estimasi data berdistribusi dengan menggunakan algoritma Fisher-scoring.*

Kata Kunci: *Estimasi Parameter, Maximum Likelihood, Fisher-scoring*

Received 15 February 2021 | Revised 25 February 2021 | Accepted 28 February 2021

1. Introduction

Parameter estimation is the estimation of population parameter values (eg mean, standard deviation, proportion, etc.) in the data or samples taken from the population. Population parameters can be known by estimating parameters through sample data. Maximum Likelihood Estimation (MLE) is an estimation method using a distribution approach and maximizing the likelihood function. Bain and Engelhard [1] said that the maximum likelihood method uses a parameter value in space as an estimator of an unknown value parameter.

Maximum likelihood estimation is a method that maximizes the probability function to obtain the estimator parameter with the maximum probability. Purba [2] said that the maximum probability estimate is an estimate of the parameters that follows a certain distribution.

*Corresponding author at: Mathematics Study Program, Universitas HKBP Nommensen Pematangsiantar, Pematangsiantar, 21136, Indonesia

E-mail address: switamyangnithapurba@gmail.com

 e-ISSN: 2656-1514, DOI: 10.32734/jormtt.v3i1.8345
Journal Homepage: https://talenta.usu.ac.id/jormtt

Ehlers [3] said that the Fisher-scoring algorithm has an assessment with the Newton-Raphson algorithm, but the difference is Fisher-scoring using the expected value of the second derivative matrix. Smyth [4] said the Fisher-scoring algorithm is a form of development of the Newton-Raphson method by replacing the Hessian matrix $H(\beta)$ with the Information matrix $I(\beta)$. Smyth [5] said that Fisher-scoring is linearly convergent, at a rate that depends on the relative difference between observed and expected information.

There are several studies on parameter estimation using the Maximum Likelihood method by researchers. In previous studies, parameter estimates have been carried out for several distributions with various models including Apriliani et al [6] estimating the Gamma distribution parameters on type II censored data using the Fisher-scoring algorithm, Kurniasih [7] estimating parameters on nonlinear statistical models with maximum likelihood. The focus of this research is to estimate the parameter values in normally distributed data based on the Fisher-scoring algorithm.

2. Methodology

In this research, parameter estimation of data is normally distributed based on Fisher-scoring algorithm. With the following stages:

1. Taking a sample of normally distributed data.
2. Find the likelihood and loglikelihood functions of normally distributed data.
3. Estimating parameters with maximum likelihood estimation based on Fisher-scoring algorithm.
4. Build Newton-Raphson Algorithm in Matlab program and parameter estimation using SPSS.
5. Comparing the estimation results of data parameters with Normal distribution based on Newton-Raphson with Matlab and SPSS.

3. Result and Discussion

In this study, parameter estimation will be carried out on normally distributed data (see Table 1). A data is said to be normally distributed if it has a significance value greater than 0.05. Checking the significance value of the data using the SPSS program. With the help of SPSS obtained a significance value of 0.124 (0.124 > 0.05) so the data is normally distributed.

Table 1. Normal Distribution Data

No	Data
1	683
2	1307
3	1704
4	4729
5	4175
6	2665
7	2041
8	2017
9	1929
10	1867
11	1828
12	1432
13	1176
14	756
15	534
16	312
17	173

3.1. Maximum Likelihood on Normal Distributed Data Estimation

The following is a joint probability density function (pdf) of normal distributed data with parameters μ and σ^2:

$$f(y_i) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{1}$$

The likelihood function of the normal distribution is

$$\begin{aligned} L(\theta \mid y_1, y_2, \dots, y_N) &= \prod_{i=1}^{N} L(\theta \mid y_i) = \prod_{i=1}^{N} f(y_i) = \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \\ &= (2\pi\sigma^2)^{-\frac{N}{2}} \exp\left(-\sum_{i=1}^{N} \frac{(y_i-\mu)^2}{2\sigma^2}\right) \end{aligned} \tag{2}$$

The loglikelihood function of the normal distribution is

$$\begin{aligned} l(\theta \mid y_1, y_2, \dots, y_N) &= \log L(\theta \mid y_1, y_2, \dots, y_N) = \log\left[\prod_{i=1}^{N} L(\theta \mid y_i)\right] \\ &= \ln\left[(2\pi\sigma^2)^{-\frac{N}{2}} \exp\left(-\sum_{i=1}^{N} \frac{(y_i-\mu)^2}{2\sigma^2}\right)\right] \\ &= -\frac{N}{2}\log(2\pi\sigma^2) - \sum_{i=1}^{N} \frac{(y_i-\mu)^2}{2\sigma^2} \end{aligned} \tag{3}$$

In this research, parameter estimation of μ and σ^2 that will be solve with Fisher-scoring algorithm in Matlab.

3.2. Fisher-scoring to Estimate the Parameter Estimation of Normal Distributed Data

This is the Fisher-scoring algorithm of normal distributed data.

$$\theta^{(k+1)} = \theta^{(k)} + I\left(\theta^{(k)}\right)^{-1} l'\left(\theta^{(k)}\right) \tag{4}$$

$$\theta^{(k+1)} = \begin{pmatrix} \mu \\ \sigma^2 \end{pmatrix}^k - \begin{bmatrix} \frac{N}{\sigma^2} & 0 \\ 0 & \frac{N}{2\sigma^4} \end{bmatrix}^{-1} \begin{bmatrix} \frac{1}{\sigma^2}\sum_{i=1}^{N}(y_i - \mu) \\ -\frac{N}{2\sigma^2} + \sum_{i=1}^{N}\frac{(y_i-\mu)^2}{2\sigma^4} \end{bmatrix} \tag{5}$$

3.3. Build Fisher-scoring Algorithm on Matlab Program

The following is a flowchart of the problem.

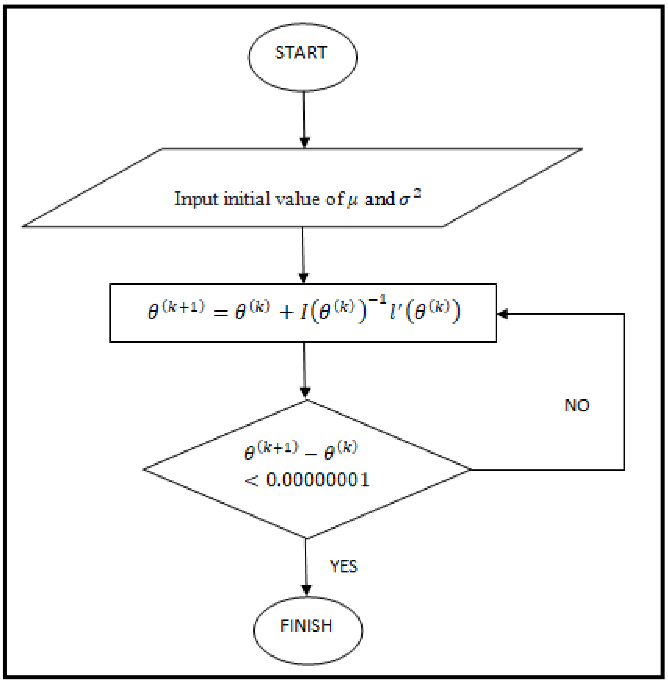

Figure 1. Flow Chart of Fisher-scoring Algorithm

From the algorithm that has been built, the Fisher-scoring iteration results are as follows.

Table 2. Fisher Scoring Algorithm Iteration Results with Matlab

Iteration	μ	σ^2
1	1725,1764706	3669402,4705882
2	1725,1764706	1445795,2041522

Table 3. Fisher Scoring Algorithm Results with SPSS

μ	σ^2
1725,18	1536157,404

4. Conclusion and Future Research

Based on the research, it is sufficient that the parameters using the Fisher-scoring algorithm on the second parameter are not significant, where the parameter estimation on one of the parameters there is a big enough difference. Therefore, it is necessary to do further research on data with different distributions and algorithms.

REFERENCES

[1] L. J. Bain and M. Engelhardt, *Introduction to Probability and Mathematical Statistics*, 2nd ed. Cengage Learning, 1992.

[2] S. A. Purba, "Estimasi Parameter Data Berdistribusi Normal Menggunakan Maksimum Likelihood Berdasarkan Newton Raphson," *J. Sains Dasar*, vol. 9, no. 1, pp. 16–18, 2021.

[3] R. Ehlers, "Maximum likelihood estimation procedures for categorical data," University of Pretoria, 2002.

[4] G. K. Smythe, "Optimization," *Encyclopedia of Environmetrics*, vol. 3. John Wiley & Sons, Chichester, pp. 1481–1487, 2002.

[5] G. K. Smyth, "Partitioned Algorithms for Maximum Likelihood and Other Non-linear Estimation," *Stat. Comput.*, vol. 6, pp. 201–216, 1996.

[6] Y. R. Apriliani, N. K. Dwidayati, and A. Agoestanto, "Estimasi Parameter Distribusi Gamma pada Data Tersensor Tipe II Menggunakan Algoritma Fisher-scoring," *J. Math.*, vol. 10, pp. 31–34, 2021.

[7] E. Kurniasih, "Estimasi Parameter pada Model Statistik Non-linier secara Maximum Likelihood," Universitas Islam Negeri Maulana Malik Ibrahim, 2014.