Jurnal Riset Rumpun Matematika dan Ilmu Pengetahuan Alam (JURRIMIPA)
Vol.2, No.1 April 2023
e-ISSN: 2828-9390; p-ISSN: 2828-9382, Hal 76-86

Penaksiran Parameter Pada Distribusi Erlang Berdasarkan Metode Maksimum Likelihood Dengan Menggunakan Algoritma Newton Raphson Dan Fisher Scoring

Parameter Estimation on Erlang Distribution Based on Maximum Likelihood Method Using Newton Raphson Algorithm and Fisher Scoring

Meili Yanti
Universitas Sumatera Utara
Email : meiliynt5@gmail.com

Open Darnius
Dosen Universitas Sumatera Utara
Email : opendarnius@gmail.com

***Abstract**. The Erlang distribution is a special case of the Gamma distribution with the k shape parameter and the λ rate parameter. In this study, the parameter estimation of the Erlang distribution was carried out using the Maximum Likelihood method. In maximizing the function, an implicit and non-linear form is obtained, then it is solved using the Newton Raphson algorithm. Apart from Newton Raphon, the estimation of parameters was also carried out using the Fisher Scoring algorithm. The Fisher Scoring algorithm is similar to the Newton Raphson algorithm, the difference is that Fisher Scoring uses an matrix information. The result of parameter estimation in Erlang distribution using Newton Raphson algorithm which is applied to outgoing telephone call data that generated by Matlab R2010a software cannot be done simultaneously. Therefore, the parameter assessment is carried out on the k parameter first, then followed by the λ parameter estimation and the parameter* $\hat{k} = 7$ *and* $\hat{\lambda} = 0.6886812$ *are obtained. Meanwhile, the parameter estimation using the Fisher Scoring algorithm produces an equation that is not different from the Newton Raphson algorithm*

***Keywords**: Fisher Scoring Algorithm, Newton Raphson Algorithm, Erlang Distribution, Maximum Likelihood Method, Parameter Estimation*

Abstrak. Distribusi Erlang merupakan salah satu kasus khusus dari Distribusi Gamma dengan parameter bentuk k dan parameter tingkat λ. Pada penelitian ini dilakukan penaksiran parameter pada distribusi Erlang dengan metode Maksimum Likelihood. Dalam memaksimumkan fungsi diperoleh bentuk yang implisit dan tidak linier, maka diselesaikan dengan menggunakan algoritma Newton Raphson. Selain Newton Raphon, penaksiran parameter juga dilakukan dengan algoritma Fisher Scoring. Algoritma Fisher Scoring mirip dengan algoritma Newton Raphson, perbedaannya adalah Fisher Scoring menggunakan matriks informasi. Hasil penaksiran parameter pada distribusi Erlang dengan menggunakan algoritma Newton Raphson yang diaplikasikan pada data panggilan keluar telepon yang dibangkitkan melalui software Matlab R2010a tidak dapat dilakukan secara serentak. Oleh karena itu, penaksiran parameter dilakukan terhadap

Received Febuari 01, 2023; Revised Maret 01, 2023; April 01, 2023
* Meili Yanti, meiliynt5@gmail.com

parameter k terlebih dahulu kemudian dilanjutkan dengan penaksiran parameter λ dan diperoleh parameter $\hat{k} = 7$ dan $\hat{\lambda} = 0{,}6886812$. Sedangkan penaksiran parameter dengan algoritma Fisher Scoring menghasilkan bentuk persamaan yang tidak berbeda dengan algoritma Newton Raphson

Kata kunci: Algoritma Fisher Scoring, Algoritma Newton Raphson, Distribusi Erlang, Metode Maksimum Likelihood, Penaksiran Parameter

INTRODUCTION

Parameter estimation is part of inferential statistics which is a way to predict the characteristics of a population based on the sample taken. The assessment carried out must be justified which is stated with the level of confidence from the estimated results obtained. A good estimator must fulfill several characteristics of an estimator desired by an opportunity, namely unbiased, efficient, and consistent.

In its application, the estimation method is used to estimate the parameters of a probability distribution. The probability distribution consists of a discrete probability distribution and a continuous probability distribution. The Erlang distribution is a special case of the Gamma distribution which comes from a continuous probability distribution with parameters of the form $k \in N$ (N is a natural number) and rate parameter $\lambda > 0$ or can be written $X \sim ERL(k, \lambda)$ with X continuous random variable . The Erlang distribution can also be said to be an exponential family distribution with parameter λ (Oliver,2013). The choice of the Erlang distribution of the Gamma distribution system is explained by its use in queuing theory. Research on the characterization and measurement of weighted Erlang distribution information and concluded that the weighted Erlang distribution provides better results and estimates compared to the Exponential distribution on the data used (Sofi & Ahmad,2017)

There are several methods used in estimating the parameters including the Moment method, Maximum Likelihood method, and Bayes method. The maximum likelihood method is an efective and important approach for parameter estimation (Yang, *et all*, 2019). The Maximum Likelihood method is the method most often used in research. This is because the Maximum Likelihood method is related to numerical ability, especially in generating a solution point in an equation or function. If in maximizing a function an implicit and non-linear form is obtained, it can be solved using the Newton Raphson and Fisher Scoring algorithms. A study obtained the results of parameter

e-ISSN: 2828-9390; p-ISSN: 2828-9382, Hal 76-86

estimation from geographically weighted ordinal logistic regression using the Fisher information matrix with Maximum Likelihood (Widyaningsih et al, 2017)

Research on the Fisher Scoring method for incomplete data and stated that researchers tend to be interested in using the Fisher Scoring method compared to the Expectation Maximization algorithm in an empirical context (Takai, 2020). In addition, a study also uses Prior Information to estimate the parameters of the Erlang distribution (Sarma et al, 2019).

LITERATURE REVIEW

Parameter Estimation and Opportunity Distribution

Parameters are values that follow reference information or information that can explain the limits or certain parts of a system of equations. Estimation is the whole process that uses an estimator to produce an estimate of a parameter. An estimate is a statement about a known population parameter based on a population from a sample taken at random from the population in question. So with this estimation, the population parameter estimation can be known (Harinaldi, 2005). The probability distribution can be interpreted as the magnitude of the probability of each outcome arising from a random experiment.

Gamma Distribution

Let X is a continuous random variable with Gamma distribution with parameters α and β, the probability density function is given as follows:

$$f(x) = \begin{cases} \frac{1}{\beta^{\alpha}\Gamma(\alpha)} x^{\alpha-1} e^{-\frac{x}{\beta}}, & x > 0, \alpha > 0, \beta > 0 \\ 0 & \text{others} \end{cases}$$

$\alpha =$ shape parameter, dan $\beta =$ scale paremeter (Walpole,2012).

The Gamma function is defined by:

$$\Gamma(\alpha) = \int_0^{\infty} x^{\alpha-1} e^{-x} dx \quad \text{for } \alpha > 0 \text{ (Harinaldi, 2005)}$$

Erlang Distribution

Continuous random variable X is said to have Erlang distribution with shape parameter $(k \in N)$ and rate parameter $(\lambda > 0)$ is $X \sim ERL(k,\ \lambda)$. If the probability density function is written with the symbol $f(x)$ so:

$$f(x) = \begin{cases} \frac{\lambda^k}{(k-1)!} x^{k-1} e^{-\lambda x}; & x > 0, \lambda > 0 \text{ dan } k \in N \\ 0 & \text{others (Veerarajan, 2008)} \end{cases}$$

Maximum Likelihood Method

The Maximum Likelihood method is one way to estimate unknown parameters. Let $X_1, X_2, \dots, X_n$ is a random sample of size n from a distribution with pdf $f(x|\,\theta)$, dependent on $\theta \in \Omega$, Ω is called the parameter space (Bain & Engelhardt, 1992). Therefore $X_1, X_2, \dots, X_n$ is a random variable, shared pdf of $X_1, X_2, \dots, X_n$ with $X_1 = x_1$, $X_2 = x_2$, …, $X_n = x_n$ can be expressed as:

$$f(x_1, x_2, \dots, x_n \,|\, \theta) = f(x_1|\theta) f(x_2|\theta) \dots f(x_n|\theta)$$

Then the Likelihood function of the sample is:

$$L(x_1, x_2, \dots, x_n \,|\, \theta) = L(\theta)$$

Newton Raphson Algorithm and Fisher Scoring Algorithm

Newton Raphson's algorithm is derived from a quadratic approximation of the Loglikelihood objective function through the derivative of a linear estimate (Bain & Engelhardt, 1992). Newton Raphson's algorithm estimates $l'(\theta^{(r)})$ using the Taylor series expansion by $\theta^{(r)}$ as follows:

$$l'(\theta) \approx l'\left(\theta^{(r)}\right) + \mathrm{H}\left(\theta^{(r)}\right)\left(\theta - \theta^{(r)}\right)$$

$\mathrm{H}\left(\theta^{(r)}\right)$ is the Hessian matrix or the second derivative matrix of the sum of the squares of each parameter defined by:

$$H(\theta) = \begin{bmatrix} \frac{\partial^2 l}{\partial\theta_1\partial\theta_1} & \frac{\partial^2 l}{\partial\theta_1\partial\theta_2} & \cdots & \frac{\partial^2 l}{\partial\theta_1\partial\theta_n} \\ \frac{\partial^2 l}{\partial\theta_2\partial\theta_1} & \frac{\partial^2 l}{\partial\theta_2\partial\theta_2} & \cdots & \frac{\partial^2 l}{\partial\theta_2\partial\theta_n} \\ \vdots & \vdots & & \vdots \\ \frac{\partial^2 l}{\partial\theta_n\partial\theta_1} & \frac{\partial^2 l}{\partial\theta_n\partial\theta_2} & \ddots & \frac{\partial^2 l}{\partial^2\theta_n} \end{bmatrix}$$

Assuming the Hessian matrix is invertible or has an inverse, then the equation can be modified to:

$$\theta^{(r+1)} = (\theta^{(r)}) - H(\theta^{(r)})^{-1} l'(\theta^{(r)})$$

$l'(\theta^{(r)})$ is the first derivative of Loglikelihood.

The Fisher Scoring algorithm is similar to Newton Raphson's algorithm, the difference is that Fisher Scoring uses an information matrix.

Phone Outgoing Call Duration Data

In this study, parameter estimation will be carried out on the duration data of outgoing telephone calls generated by customers on telephone traffic. The data was obtained through simulation of data generation using Matlab R2010a software.

Table 1. Phone Outgoing Call Duration Data

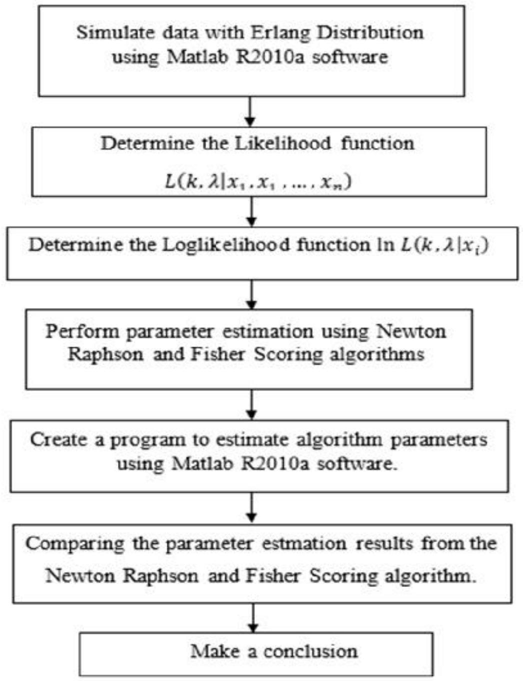
No	Call Duration (minutes)
1	4,1888
2	8,6825
3	6,4625
4	10,3163
5	6,5107
6	4,1151
7	6,2962
8	9,1402
9	5,4026
10	8,0142
11	4,8778
12	3,1993
13	7,2547
14	6,9805
15	7,2218
16	8,2237
17	7,0202
18	8,1617
⋮	⋮
120	14,3077

Jurnal Riset Rumpun Matematika dan Ilmu Pengetahuan Alam (JURRIMIPA)
Vol.2, No.1 April 2023
e-ISSN: 2828-9390; p-ISSN: 2828-9382, Hal 76-86

RESEARCH METHODS

This research was conducted in the following steps:

Figure 1. Research Methods



The data used in this study are secondary data, namely data taken from existing sources. The data in this study comes from the results of data generation simulations with the help of Erlang-distributed Matlab R2010a software.

RESULT AND DISCUSSION

Estimation of Erlang Distribution Parameters Based on Maximum Likelihood Method

$$f(x) = \frac{\lambda^k}{(k-1)!} x^{k-1} e^{-\lambda x} \; ; \; x > 0, \; \lambda > 0 \text{ dan } k \in N$$

Next, the value of the Likelihhod and Loglikelihood functions will be sought from the Erlang distribution.

The Likelihood function of the Erlang distribution is:

$$L(k,\lambda|x_1,x_1,\ldots,x_n) = \prod_{i=1} L_{(k,\lambda|x_i)}$$

$$= \prod_{i=1}^{n} f(x_i)$$

$$= \prod_{i=1}^{n} \frac{\lambda^k}{(k-1)!} x^{k-1} e^{-\lambda x}$$

$$= \left(\frac{\lambda^k}{(k-1)!}\right)^n \prod_{i=1}^{n} x_i^{k-1} e^{-\lambda \sum_{i=1}^{n} x_i}$$

The Loglikelihood function of the Erlang distribution is:

The L[illegible] $l(k, \lambda | x_1, x_1, \dots, x_n) = ln\, L(k,\lambda|x_i)$

$$= ln\left[\prod_{i=1}^{n} L(k,\lambda|x_i)\right]$$

$$= ln\left[\left(\frac{\lambda^k}{(k-1)!}\right)^n \prod_{i=1}^{n} x_i^{k-1} e^{-\lambda \sum_{i=1}^{n} x_i}\right]$$

$$= nk\, ln\, \lambda - n\, ln\, (k-1)! + (k-1)\sum_{i=1}^{n} ln x_i - \lambda \sum_{i=1}^{n} x_i$$

Newton Raphson's Algorithm

The second derivative of the Loglikelihood functon with respect to parameters k and λ or commonly called the Hessian matrix is obtained, namely:

$$H(\theta^{(r)}) = \begin{bmatrix} \frac{\partial}{\partial k^2} l(k,\lambda|x_1, x_1, \dots, x_n) & \frac{\partial}{\partial k \partial \lambda} l(k,\lambda|x_1, x_1, \dots, x_n) \\ \frac{\partial}{\partial k \partial \lambda} l(k,\lambda|x_1, x_1, \dots, x_n) & \frac{\partial}{\partial \lambda^2} l(k,\lambda|x_1, x_1, \dots, x_n) \end{bmatrix}$$

$$= \begin{bmatrix} -n\psi(1,k) & \frac{n}{\lambda} \\ \frac{n}{\lambda} & -\frac{nk}{\lambda^2} \end{bmatrix}$$

Then the equation of Newton Raphson's algorithm with Erlang distribution is:

$$\theta^{(r+1)} = (\theta^{(r)}) - \mathrm{H}(\theta^{(r)})^{-1} l'(\theta^{(r)})$$

$$\binom{k^{r+1}}{\lambda^{r+1}} = \binom{k^r}{\lambda^r} - \begin{bmatrix} -n\psi(1, k^r) & \frac{n}{\lambda^r} \\ \frac{n}{\lambda^r} & -\frac{nk^r}{(\lambda^r)^2} \end{bmatrix}^{-1} \begin{bmatrix} n ln \lambda^r - n\psi(k^r) + \sum_{i=1}^{n} ln\, x_i \\ \frac{nk^r}{\lambda^r} - \sum_{i=1}^{n} x_i \end{bmatrix}$$

Jurnal Riset Rumpun Matematika dan Ilmu Pengetahuan Alam (JURRIMIPA)
Vol.2, No.1 April 2023
e-ISSN: 2828-9390; p-ISSN: 2828-9382, Hal 76-86

After the Newton Raphson equation is known, we will look for the estimated value of the k and λ parameters of the Erlang distribution on the outgoing telephone call data. The parameter estimation process using Matlab R2010a software produces divergent iterations. This happens because there is a ψ function on the first derivative and the second derivative of the Loglikelihood function. The ψ function comes from the derivative of $\ln(k-1)!$ to the parameter k. The ψ function does not apply if k is negative which means that the factorial value cannot be determined. Therefore, the estimation of the parameter k is carried out based on the Maximum Likelihood method only, and the estimation of the parameter λ is carried out using the Newton Raphson algorithm. The following is the iteration result of parameter estimation $\hat{\lambda}$ with Newton Raphson algorithm.

Table 2. Iteration Results Parameter

Iterasi	[illegible]
0	15
1	2,3583446
2	0,8961033
3	115,4959012
4	4,0829355
5	1,3435294
6	0,1482646
7	0,3816560
8	0,7987290
9	3,3461581
10	1,1880751
11	-0,2518887
12	-0,3186000
13	0,3239115
14	0,6442588
15	0,6881162
16	0,6886837
17	0,6886812

Fisher Scoring Algorithm

The following is the negative expected value of the second derivative of the Loglikelihood function of the Erlang Distribution or called the information matrix $(I(\theta))$.

$$I(\lambda_r k_r) = -E\begin{bmatrix} -n\psi(1,k) & \frac{n}{\lambda} \\ \frac{n}{\lambda} & -\frac{nk}{\lambda^2} \end{bmatrix}$$

$$= -\begin{bmatrix} -n\psi(1,k) & \frac{n}{\lambda} \\ \frac{n}{\lambda} & -\frac{nk}{\lambda^2} \end{bmatrix}$$

$$= \begin{bmatrix} n\psi(1,k) & -\frac{n}{\lambda} \\ -\frac{n}{\lambda} & \frac{nk}{\lambda^2} \end{bmatrix}$$

So that the equation form of the Fisher Scoring algorithm with Erlang distribution is obtained:

$$\theta^{(r+1)} = (\theta^{(r)}) + I(\theta^{(r)})^{-1} l'(\theta^{(r)})$$

$$\binom{k^{r+1}}{\lambda^{r+1}} = \binom{k^r}{\lambda^r} + \begin{bmatrix} n\psi(1,k) & -\frac{n}{\lambda} \\ -\frac{n}{\lambda} & \frac{nk}{\lambda^2} \end{bmatrix}^{-1} \begin{bmatrix} nln\lambda^r - n\psi(k^r) + \sum_{i=1}^{n} lnx_i \\ \frac{nk^r}{\lambda^r} - \sum_{i=1}^{n} x_i \end{bmatrix}$$

From the general equation above, it is known that the general equation of the Fisher Scoring algorithm is the same as the general equation of the Newton Raphson algorithm. In calculating the value of the Information matrix which is the negative value of the expected matrix formed by the second derivative of the Loglikelihood function, there is no random variable X, thus causing the value of the information matrix to be constant.

CONSLUSION

The estimation of the parameters $\hat{k}$ and $\hat{\lambda}$ on the Erlang distribution with the Newton Raphson algorithm cannot be carried out simultaneously because of the ψ function which if $\hat{k}$ is negative then the iteration process cannot be carried out. Therefore, the estimation of the parameter $\hat{k}$ is carried out using the Maximum Likelihood method, then the parameter estimation $\hat{\lambda}$ is carried out using the Newton Raphson algorithm. There is no difference in the form of the Newton Raphson algorithm and the Fisher Scoring algorithm in the Erlang distribution. This is due to the absence of random variables in the calculation of the value of the information matrix, so there is no better algorithm found between the Newton Raphson algorithm and the Fisher Scoring algorithm.

REFFERENCES

Bain, L. and Engelhardt. 1992. *Introduction to Probability and Mathematical Statistics. 2nd ed*. California: Duxbury Press.

Harinaldi. 2005. *Prinsip-Prinsip Statisktik untuk Teknik dan Sains*. Jakarta: Erlangga.

Millar, R. B. 2011. *Maximum Likelihood Estimation and Inference: with example in R, SAS and ADMB*. New Zealand: John Wiley & Sons.

Oliver, C. I. 2013. *Markov Processes for Stochastic (Second Edition).* USA: Elsever.

Sarma, B. B. R., Kumar, V. V., Lalitha, and S. V. N. L. 2019. Estimation of Parameters in Erlang Distributrion using Prior Information. *International Journal of Recent Technology and Engineering (IJRTE)*, 8: 4969-4971. DOI: 10.35940/ijrte.B1072.078219.

Sofi, M. and Ahmad, S. P. 2017. Parameter Estimation of Weighted Erlang Distribution Using R Software. *International Journal of Modern Mathematical Science*. 7(6). ISSN: 2225-0522

Takai, K. 2020. Incomplete-data Fisher Scoring Method with Steplength Adjusment. *International Journal of Computational Physics*. DOI: 10.1007/s11222-020-09923-z.

Veerarajan, T. 2008. *Probability, Statistics, and Random Processes*. New Delhi: Tata McGraw-Hill.

Walpole, E. R. 2012. *Probability & Statistics for Engineers & Scientists*. Boston: Prentice Hall.

Widyaningsih, P., Saputro, D. R. S., and Putri, A. N. 2017. Fisher Scoring Method for Parameter Estimation of Geographically Weighted Ordinal LOGISTIC Regression (GWOLR) Model. *International Conference on Mathematics: Education, Theory and Application*. DOI: http://dx.doi.org/10.1088/1742-6596/855/1/012060

Yang, F., Ren, H., and Hu, Z. 2019. Maximum Likelihood Estimation for Three-Parameter Weibull Distribution Using Evolutionary Strategy. *Mathematical Problems in Engineering (Research Article)*. DOI: https://doi.org/10.1155/2019/6281781.