Volume 2, Issue 9 September 2016 ISSN NO: 2454 - 7514

North Asian International Research Journal Consortium

North Asian International Research Journal Of Science, Engineering and Information Technology

IRJIF IMPACT FACTOR: 3.821

NAIRJC JOURNAL PUBLICATION

North Asian International Research Journal Consortium

Welcome to NAIRJC

ISSN NO: 2454 -7514

North Asian International Research Journal of Science, Engineering & Information Technology is a research journal, published monthly in English, Hindi. All research papers submitted to the journal will be double-blind peer reviewed referred by members of the editorial board. Readers will include investigator in Universities, Research Institutes Government and Industry with research interest in the general subjects

Editorial Board

Address: -North Asian International Research Journal Consortium (NAIRJC) 221 Gangoo, Pulwama, Jammu and Kashmir, India - 192301, Cell: 09086405302, 09906662570, Ph. No: 01933-212815,
Email: nairjc5@gmail.com, nairjc@nairjc.com, info@nairjc.com Website: www.nairjc.com

IRJIF IMPACT FACTOR: 3.821

OBTAINING PARAMETER ESTIMATE FROM THE TRUNCATED POISSON PROBABILITY DISTRIBUTION

A. M. YAHAYA [1], N. P. DIBAL [2], H. R. BAKARI [3] & ADEGOKE T.M [4]

[1234] Department of Mathematics and Statistics, Faculty of Science, University of Maiduguri, Nigeria

ABSTRACT:

In this paper, an estimate for the Maximum likelihood Estimator in truncated Poisson probability distribution was obtained by employing the Maximum Likelihood Estimation method jointly with the Newton Raphson Method. The aim of this study was to use Maximum Likelihood Estimation (MLE) and a numerical Method (Newton Raphson method) to obtain parameter estimate from the truncated Poisson Probability distribution since the Maximum Likelihood Estimator of the truncated Poisson Probability distribution does not have a close form solution. The methods were demonstrated on real life data and simulation study using data sets generated by the R statistical software for different sample sizes. The standard errors were also computed and a 5% wald-confidence interval was constructed for the distribution. The result of the study shows that the method of maximum likelihood estimation jointly with the numerical method (Newton Raphson method) is capable of providing efficient estimate from the Truncated Poisson probability distribution in both simulation and real life data. The study recommends that parameter estimates from the truncated Poisson probability distribution should be obtained using maximum likelihood estimation jointly with Newton Raphson method.

Keywords: ***Parameter estimation, Truncated Poisson probability distribution, Maximum Likelihood Estimation, Newton Raphson Method.***

1. INTRODUCTION

In probability theory, the zero-truncated Poisson (ZTP) distribution is a certain discrete probability distribution whose support is the set of positive integers. This distribution is also known as the conditional Poisson distribution or the positive Poisson distribution. It is the conditional probability distribution of a Poisson-distributed random variable, given that the value of the random variable is not zero. Thus it is impossible for a ZTP random variable to be zero.

This study is concerned with maximum likelihood. In statistics, maximum-likelihood estimation (MLE) is a method of estimating the parameters of a statistical model. When applied to a data set and given a statistical model, maximum-likelihood estimation provides estimates for the model's parameters. But with more complicated models, maximum likelihood alone may not result in a closed form solution. Analytic expressions for maximum likelihood estimators in complex models are usually not easily available, and numerical methods are needed. In the MLE-based methods, as the basic estimating equation is not in closed form, it can be solved only numerically. There are several typical MLE-based methods for solving this equation, such as the secant method, the bisection method and the Newton-Raphson method. However, in both the secant method and the bisection method, the convergence rates are very low; in the Newton-Raphson method, the convergence rates are very fast, it has to compute both the basic estimating function and its derivative at each iterative step. Furthermore, these above MLE-based methods require either initial values or trial computation of the estimated parameters. The present study used MLE and Newton Raphson Method. The choice of this method is based on the inability of the maximum likelihood method alone to obtain Parameter estimates from two-parameter Weibull distribution. Thus, since one of the research objectives seek to obtain the estimate of the parameter(s) of the two-parameter Weibull probability distribution which Maximum Likelihood Estimate alone cannot obtain, the researcher therefore used of MLE and Newton Raphson Method to obtain their estimates.

2. REVIEW OF RELATED LITERATURE

The likelihood function tells us how likely the observed sample is function of the possible parameter values. Thus, maximizing the likelihood function for the data gives the parameter values for which the observed sample is most likely to have been generated, that is, the parameter values that ''agree most closely'' with the observed data (R.A. Fisher,1920).

Modern applied statistics deals with many settings in which the point wise evaluation of the likelihood function is impossible or computationally difficult. Areas such as financial modelling, genetics, geostatistics, neurophysiology and stochastic dynamical systems provide numerous example of this (Pritchard *et al*., 1999). It is consequently difficult to perform any inference (classical or Bayesian) about the parameters of the model. Various approaches to overcome this difficulty have been proposed, Cox and Reid (2004) used Composite Likelihood methods for approximating the likelihood function and also Pritchard *et al.*, 1999*;* Beaumont *et al*.*, 2002,* applied Approximate Bayesian Computational methods for approximating the posterior distribution for obtaining estimates of parameter. It is well-known that ABC produces a sample approximation of the posterior distribution (Beaumont *et al.*, 2002) in which there exist a deterministic approximation error in addition to Monte Carlo variability. The quality of the approximation to the posterior and theoretical properties of the estimators obtained with ABC have been studied in Wilkinson (2008); Marin *et al*. (2011); Dean *et al*. (2011) and Fearnhead and Prangle (2012). The use of ABC posterior samples for conducting model comparison was studied in Didelot *et al*. (2011) and Robert *et al*. (2011). Using this sample approximation to characterize the mode of the posterior would in principle allow (approximate) maximum a posteriori (MAP) estimation. Furthermore, using a uniform prior distribution, under the parameters of interest, over any set which contains the MLE will lead to a MAP estimate which coincides with the MLE. In low-dimensional problems if we have a sample from the posterior distribution of the parameters, we can estimate its mode by using either nonparametric estimators of the density or another mode seeking technique such as the mean-shift algorithm (Fukunaga and Hostetler, 1975). Although Marjoram *et al*. (2003) noted that (ABC) can also be used in frequentist applications, in particular for maximum-likelihood estimation this idea does not seem to have been developed. Alternative nonparametric density estimators which could also be considered within the AMLE context have been proposed recently in Cule *et al*. (2010); Jing *et al*. (2012). Cheng and Amin (1983) suggest the maximum product of spacing (MPS) method. This method can be applied to any univariate distribution. Cheng and Traylor (1995) point out the drawbacks of the MPS method owing to the occurrence of the tied observations and numerical effects involved in ordering the cdf when there are explanatory variables in the model. Atkinson, Pericchi *et al.* (1991) apply the grouped-data likelihood approach to the shifted power transformation model of Box and Cox (1964).

3. MATERIALS AND METHODS

The study employed MLE jointly with Numerical method (Newton Raphson method) to obtain the estimates, standard errors and Wald interval of the truncated Poisson distribution in both simulation study and real life data using R software.

3.1 The maximum likelihood estimation method

The maximum likelihood estimation method had been used in special cases by Gauss in 1812 but a full description of properties and a presentation of its application were performed 100 years later by Ronald Fisher [1]. Nowadays, the maximum likelihood method is the most popular estimation technique, mainly for its good theoretical properties. See other literature [5, 12, 18, 20] for the existence and the uniqueness of the maximum likelihood estimates for discussed distributions. The idea of the maximum likelihood method is based on the assumption that observed data are the most likely outcome of a random experiment in respect to the considered probability distribution. In the discussed method the key role plays the likelihood function specified as the probability of observed data depending on the values of distribution parameters.

3.1.1 Truncated Poisson Probability Distribution

The truncated Poisson distribution is a continuous probability distribution named after a French Mathematician Simeon- Denis Poisson (1781-1840) who discovered it in 1838. It is used to model count data for which the value zero cannot occur.

The truncated Poisson distribution likelihood function L is

Table 3.1

Distribution	$f(x, \psi)$
Truncated Poisson	$f(x; \psi) = \frac{\psi^x e^{-\psi}}{x!\,(1 - e^{-\psi})}: \quad x = 1,2,3,4,5, \ldots.$

Table 3.2

Distribution	$L(x, \psi)$
Truncated Poisson	$\prod_{i=1}^{n} \frac{\psi^x e^{-\psi}}{x!\,(1 - e^{-\psi})}$

The maximum likelihood estimators of the distribution parameter is found by maximizing the likelihood functions L (actually it logarithms) with respect to parameter value. Maximum likelihood estimate of the truncated Poisson probability distribution is therefore the solution of equation which is obtained by equating partial derivative of In (L) to zero presented in Table3.3 making direct analytical solution intractable. The discussed problem has no explicit algebraic solution, therefore numerical calculation is required.

4. THE NUMERICAL METHOD (SIMULATION STUDY)

4.1 Truncated Poisson distribution

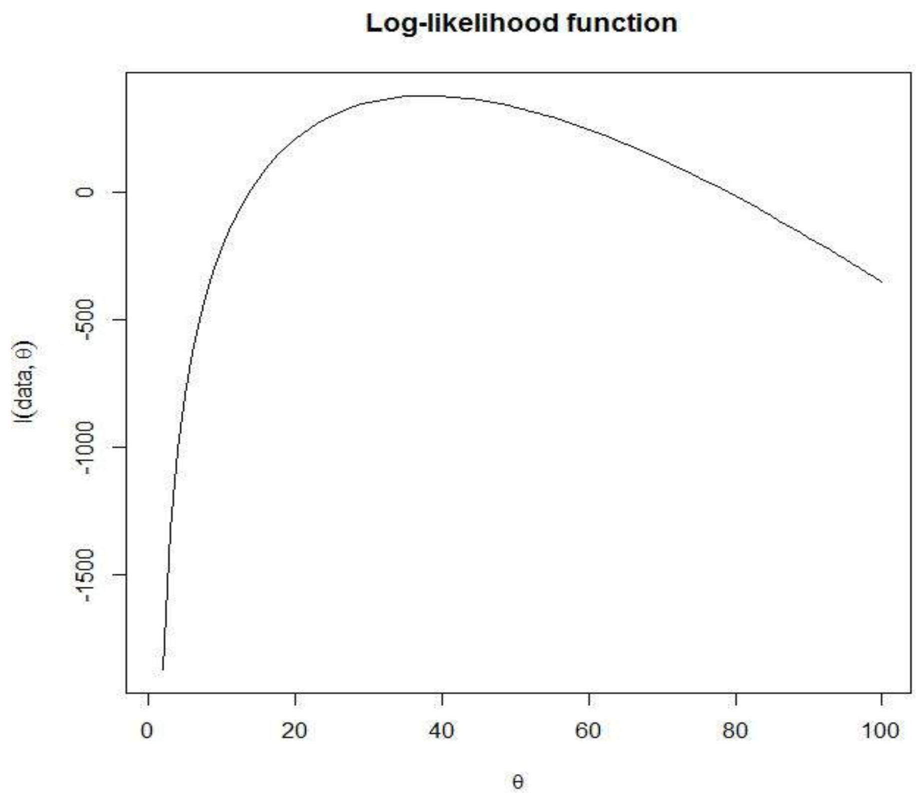

Figure 4.1: Log likelihood graph of a truncated poison distribution.

Table 4.1: Truncated Poison Distribution Result Table

N	$\widehat{\psi}\, mle$	S.E($\widehat{\psi}$)	Wald C.I for $\widehat{\psi}$
10	1.318373	0.04625436	1.131081, 2.056168
100	1.318373	0.2237425	1.132031, 2.055743
1000	1.318373	0.07187667	1.13343, 2.063654
10000	1.318373	0.02275701	1.1354, 2.096547

4.2 The Numerical Method (Real Life Data)

4.2.1Truncated Poisson Probability Distribution

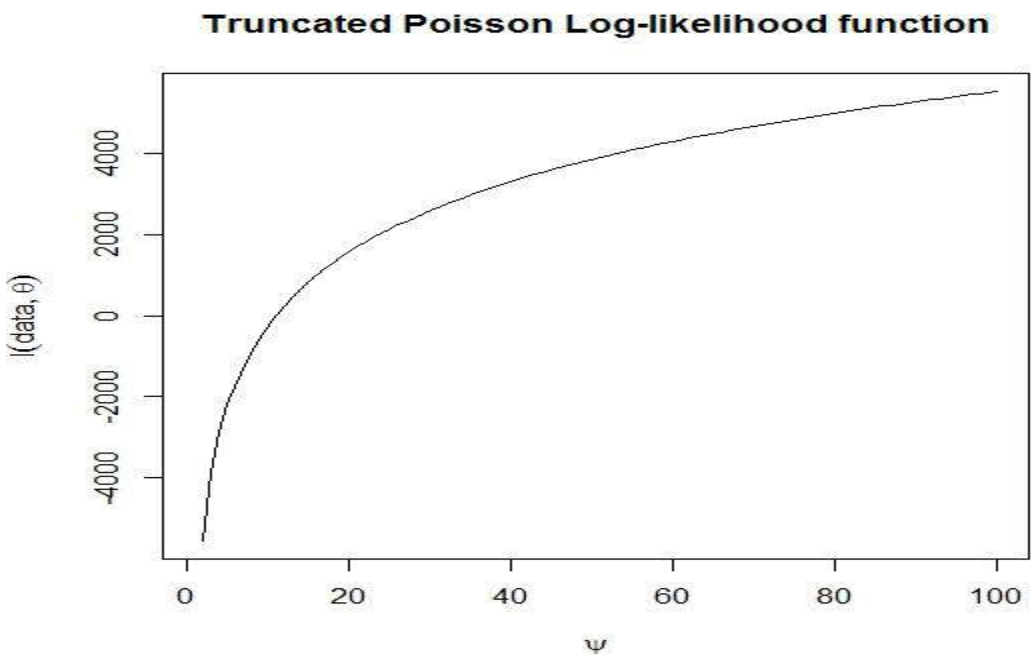

Figure 4.2: Log likelihood graph of a truncated poison distribution.

Table 4.2: Truncated Poison Distribution Result Table

$\widehat{\psi}\, mle$	S.E($\widehat{\psi}$)	Wald C.I
403.8331	8.204813	395.6283, 412.0379

4.3 Discussion

By applying the Newton Raphson method using simulation study, a total of 5 iterations were performed to obtain the maximum likelihood estimates of truncated Poisson probability distribution. Convergence was achieved at the 9^{th} returning -5.669706 as the value of the log-likelihood and the value of the estimate which maximizes the function is 3.048175 with gradient -1.48604×10^{-5}. The variance which was the value of the second derivative is 2.920648

Table 4.1 shows the estimate which maximizes the likelihood function of truncated Poisson probability distribution for different sizes which ranges between 1.318373 and 1.318373 with standard error reducing as the sample size increases (0.02275701 to 0.4625436).

By applying the Newton Raphson method using real life data, a total of 5 iterations were performed to obtain the maximum likelihood estimates of truncated Poisson probability distribution. Convergence was achieved at the 9th returning -12117.43 as the value of the log-likelihood and the value of the estimate which maximizes the function was 403.8331 with gradient -3.378232×10^{-7}. The variance which is the value of the second derivative is 0.01485466.

Table 4.10 shows the estimate which maximizes the likelihood function of truncated Poisson probability distribution is (403.8331) with standard error (8.204813).

5. CONCLUSION

Parameter estimates from truncated Poisson probability distribution can easily be obtained by the method of maximum likelihood estimation jointly with numerical approach (Newton Raphson method) with the help of computer. The result of the study shows that the Maximum Likelihood Estimation (MLE) jointly with the Newton Raphson method is capable of estimating the parameter truncated Poisson probability distribution.

6. RECOMMENDATION

Based on the results drawn from this study, the following recommendation was made:

- Parameter estimate from truncated Poisson probability distribution should be obtained using Maximum Likelihood jointly with Newton Raphson method.

REFERENCES

[1] Atkinson, A. C.,Pericchi, L. R.,& Smith, R. L. (1991). Grouped likelihood for the shifted power transformation. *Journal of the Royal Statistical Society*: Series B, 53: 473- 482.

[2] Beaumont, M. A., Zhang, W., & Balding, D. J.(2002). Approximate Bayesian computation in population genetics.*Genetics*, 2 (162), 2025-2035.

[3] Cheng, R. C. H. & Amin, N. A. K. (1983). Estimating parameters in continuous univariate distributions with a shifted origin. *Journal of the Royal Statistical Society: Series B,* 45: 394-403.

[4] Cheng, R. C. H. & Traylor, L. (1995). Non-regular maximum likelihood problems. *Journal of the Royal Statistical Society*: Series B, 57: 3-44.

[5] Cox, D. R. & Reid, N. (2004). A note on pseudolikelihood constructed from marginal densities., Biometrika, 2(91), 729-737.

[6] Cule, M. L.,Samworth, R. J. & Stewart, M. I. (2010). Maximum likelihood estimation of a multi-dimensional log-concave density. *Journal Royal Statistical Society* B 72: 545600.

[7] Didelot, X., Everitt, R. G., Johansen, A. M. and Lawson, D. J. (2011). Likelihood-free estimation of model evidence.*Bayesian Analysis*, vol. 6,49-76.

[8] Fukunaga, K. & Hostetler, L. D. (1975). The Estimation of the Gradient of a Density Function, with Applications in Pattern Recognition. IEEE Transactions on Information Theory 21: 3240.

[9] Fearnhead, P. &Prangle, D. (2012).”Constructing Summary Statistics for Approximate Bayesian Computation: Semi-automatic ABC (with discussion). Journal of the Royal Statistical Society Series B (Methodology) in press.

[10] Fisher, R. A. (1922). On the mathematical foundations of theoretical statistics. *Philosophical Transactions of the Royal Society of London*: Series A, 222: 309-368.

[11] Marjoram, P., Molitor, J., Plagnol, V.,&Tavare, S. (2003). Markov chain Monte Carlo without likelihoods. Proceedings of the National Academy of Sciences USA: 1532415328.

[12] Pritchard, J. K., Seielstad, M. T., Perez-Lezaun, A., & Feldman, M. T. (1999).Population Growth of Human Y Chromosomes: A Study of Y Chromosome Microsatellites. Molecular Biology and Evolution 16: 1791 1798.

[13] Robert, C. P., Cornuet, J., Marin, J. & Pillai, N. S. (2011). Lack of confidence in ABC model choice. Proceedings of the National Academy of Sciences of the United States of America 108: 1511215117.

Publish Research Article

Dear Sir/Mam,

We invite unpublished Research Paper,Summary of Research Project,Theses,Books and Book Review for publication.

Address:- North Asian International Research Journal Consortium (NAIRJC)
221, Gangoo Pulwama - 192301
Jammu & Kashmir, India
Cell: 09086405302, 09906662570,
Ph No: 01933212815
Email:- nairjc5@gmail.com, nairjc@nairjc.com , info@nairjc.com
Website: www.nairjc.com