North Asian International Research Journal of Sciences, Engineering & I.T.

ISSN: 2454-7514 Vol. 3, Issue-8 August-2017

IRJIF. I.F. : 3.821

OBTAINING PARAMETER ESTIMATE FROM THE WATSON PROBABILITY DISTRIBUTION

A. M. YAHAYA *[1], N. P. DIBAL [2] AND H. R. BAKARI [3]

[123]Department of Mathematics and Statistics, Faculty of Science University of Maiduguri, Nigeria
*Corresponding:abdullstatii@gmail.com

ABSTRACT:

This paper addresses parameter estimate from Watson probability distribution. The Watson distribution is one of the simplest distributions for analyzing axially symmetric data. This distribution has gained some attention in recent years due to its modeling capability. The aim of this paper was to employ Maximum Likelihood Estimation (MLE) jointly with a numerical Method (Newton Raphson method) to obtain parameter estimate from the Watson model. This study used simulation study to generate data sets for different sample sizes using R statistical software. Standard errors were also computed and a 5% wald-confidence interval was constructed for the distribution. The result of the study revealed the method of maximum likelihood estimation (MLE) jointly with a numerical method (Newton Raphson method) is efficient for obtaining parameter estimate from Watson model. The study recommends that the method of maximum likelihood estimation (MLE) jointly with a numerical method (Newton Raphson) should be employed to obtain parameter estimate from intractable models.

***Keywords:* Parameter estimation, Watson model, Maximum Likelihood Estimation, Newton Raphson Method.**

1. INTRODUCTION

In statistical inference, there are two broad categories for the interpretation of probability and these are the Bayesian inference and the frequentist inference. These views differ on the fundamental nature of probability. Frequentist inference loosely define probability as the limit of an event's relative frequency in a large number of trials, and only in the context of experiments that is random and well-defined. Bayesian inference, on the other hand, assigns probabilities to any statement even when a random process is not involved. In Bayesian inference, probability is a means of representing an individual's degree of belief in a statement, or given evidence.

In frequentist approach, a general method for estimating the specific parameters is called Maximum Likelihood (ML). The Maximum Likelihood is perhaps the most versatile method for fitting statistical models to data. In typical applications, the goal is to use a parametric statistical model to describe a set of data or a process that generate a set of data. The appeal of ML stems from the fact that it can be applied to a wide range of statistical models and kinds of data (e.g., continuous, discrete, categorical, censored, truncated, etc.), where other popular methods like least squares do not in general provide a satisfactory method of estimation. Indeed, when assuming an underlying normal distribution, the least squares estimates of regression coefficients are equivalent to ML estimates. The ML method is, however, much more general because it allows one to use other distributions as well as more general assumptions about the model and the form of the data.

The maximum likelihood estimator was developed by [11] based on the work done by Karl Pearson who worked on several estimation methods.While Fisher agreed with Pearson that the method of moments is better than least sq uares, Fisher had an idea for an even better method where it took many years to fully conceptualize this method that ended up with the name maximum likelihood estimation.

In 1912, when Fisher was a third year undergraduate student, Fisher published a paper called "Absolute criterion for fitting frequency curves." The concepts in his paper were based on the principle of inverse probability which was later discarded. If any method can be considered comparable to inverse probability it is Bayesian estimation. Fisher was convinced that he had an idea for the superior method of estimation; criticism of his idea only fueled his pursuit which led to precise definition of estimation. In the end, his debates with other statisticians resulted in the creation of many statistical terms being used today, including the word "estimation" itself and even "statistics". Finally, Fisher defined the difference between probability and likelihood and put his final touches on maximum likelihood estimation in (1922), Fisher introduced likelihood in the context of estimation via the method of maximum likelihood, but subsequently he did not think of it as just a device to produce parameter estimates. The likelihood is a tool for an objective reasoning with data especially when dealing with the uncertainty due to the limited amount of information contained in the data. The likelihood function captures all the information in the data about a certain parameter, not just its maximizer. The obvious role of the maximum likelihood estimate (MLE) is to provide a point estimate for the parameter of interest; the purpose of having a point estimate is determined by the application area. In cases where a model parameter has a physical meaning, it is reasonable to ask what the best estimate is given the data. The uncertainty is in a way a nuisance and not part of the scientific question. Another important role is for simplifying a multi-parameter likelihood through a profile likelihood nuisance parameters are replaced by the MLEs.

Estimation is the process of determining approximate values for parameters of different population or events. How well the parameter is approximated depend on the method and the type of data. The method of maximum likelihood corresponds to many well-known estimation methods in statistics (such as; maximum likelihood, moments, least squares, Bayesian estimation etc) and ending particular parametric values that make the observed results the most probable (given the model). But 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. 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. Newton's method can be used to find solutions when no closed form exists and it converges quickly. Here the importance of an efficient estimator is reinforced since the platykurtic nature of an inefficient estimator diminishes the ability of the algorithm to converge. However, with the rapid increase of computer speed, maximum likelihood estimation has become easier and has increased in popularity.

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 [11]. 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 examples of this [15]. 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, [5] used Composite Likelihood methods for approximating the likelihood function and also [15]; [2], 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 [2] 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 [17]; [13]; [7] and [10]. The use of ABC posterior samples for conducting model comparison was studied in [8] and [16]. 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 [9]. Although [14] 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 [6]; [12]. [4] suggest the maximum product of spacing (MPS) method. This method can be applied to any univariate distribution. [4] 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. [1] applied the grouped-data likelihood approach to the shifted power transformation model of [3].

3. MATERIALS AND METHODS

The study employed MLE jointly with Numerical method (Newton Raphson method) to obtain the estimates, standard errors and 5 % Wald interval of the estimates of Watson probability distribution in simulation studies using R Statistical software.

3.1 Model Specification

Watson Probability Distribution: (3.1)

3.1.1 The maximum likelihood estimation method

It is highly desirable to have a method that is generally applicable to the construction of statistical estimators that have "good" properties. In this section we present an important method for finding estimators of parameters proposed by geneticist/statistician Sir Ronald A. [11] called the method of maximum likelihood. Even though the method of moments is intuitive and easy to apply, it usually does not yield "good" estimators. The method of maximum likelihood is intuitively appealing, because we attempt to find the values of the true parameters that would have most likely produced the data that we in fact observed. For most cases of practical interest, the performance of maximum likelihood estimators is optimal for large enough data. This is one of the most versatile methods for fitting parametric statistical models to data. First, we define the concept of a likelihood function.

3.1.2 Definition Let , be the joint probability (or density) function of n random variables, …, with sample values, …,. The likelihood function of the sample is given by, …, = (, …,;), , in a briefer notation].

It emphasize that is a function of for fixed sample values.

3.1.3 The Newton-Raphson method

The Newton-Raphson method, or Newton Method, is a powerful technique for solving equations numerically. Like so much of the differential calculus, it is based on the simple idea of linear approximation. The Newton Method, properly used, usually homes in on a root with devastating efficiency.

3.1.4 Using Linear Approximations to Solve Equations

Let be a well-behaved function, and let be a root of the equation = 0. Starting with an estimate of. From, It produce an improved- hoping-estimate. From, it produces a new estimate. From, it produces a new estimate. It go on until it is 'close enough' to -or until it becomes clear that it is getting nowhere. The above general style of proceeding is called iterative. Of the many iterative root-finding procedures, the Newton-Raphson method, with its combination of simplicity and power, is the most widely used. The initial estimate is sometimes called, but most mathematicians prefer to start counting at 0.Sometimes the initial estimate is called a "guess." The Newton Method is usually very good if is close to, and can be horrid if it is not. The "guess" should be chosen with care.

3.1.5 The Newton-Raphson Iteration

Let be a good estimate of and let since the true root is and the number h measures how far the estimate is from the truth. Since h is 'small, 'the linear (tangent line) approximation can be used to conclude that And therefore, unless is close to 0, It follows that The new improved (?) estimate is obtained from in exactly the same way as was obtained from Continue in this way. If is the current estimate, then the next estimate is given.

4. IMPLEMENTING THE ITERATIVE METHOD

4.1 Watson Probability distribution

Consider a Watson distribution with p.d. : (4.1), (4.2), (4.3)

Equating the derivatives to zero and solving the equation is difficult making direct analytical solutions intractable.

+ (4.4)

4.2 The Numerical Method Using Simulation

The score function is given in (4.4) while the second derivative is (4.5)

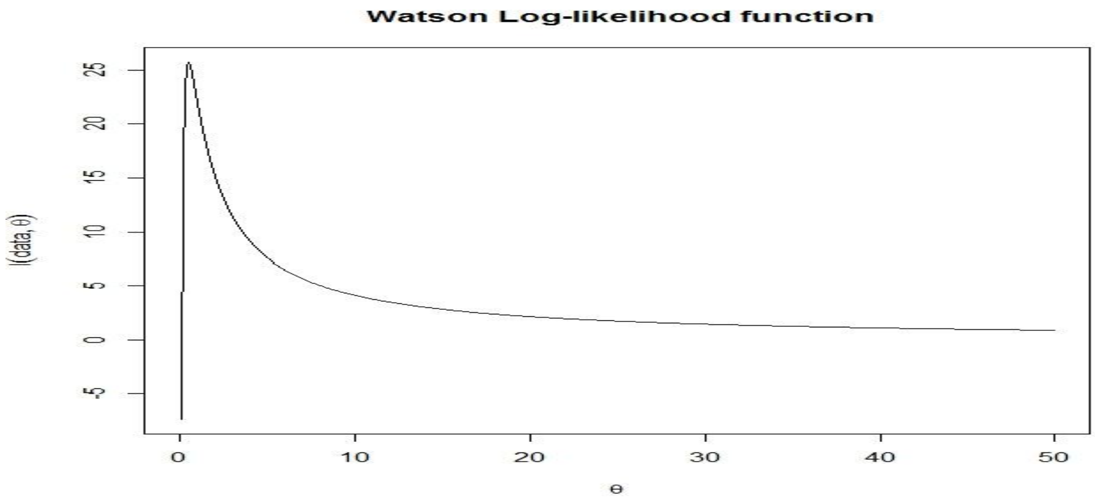

Figure 4.1: Loglikelihood graph of a Watson distribution.

Table 4.1: Watson distribution Result Table

N		S . E	Wald C . I
10	0.1536509	0.06046312	0.09318783,0.21411407
100	0.1536509	0.01912012	0.133453080,0.1727711
1000	0.156504a	0.006046287	0.1476041,0.1596967
10000	0.1536504	0.001912004	0.1517384,0.555624

5. DISCUSSION

From the simulation study the iterations obtained by applying Newton Raphson method in obtaining maximum likelihood estimate for Watson probability distribution shows a total of 11 iterations were carried out to obtain the maximum likelihood estimate. The solution converges at the 9th 2200.691 as the value of the log-likelihood and the value of the estimate which maximizes the function was 0.1546604 with gradient $3.183231 \times 10{-}6$. The variance which is the value of the second derivatives is 273541.

Table 4.1 shows the estimate which maximizes the likelihood function of Watson probability distribution for different sizes which ranges between 0.1536504 to 0.1536509 with standard errors reducing as the sample size increases (0.06046312 to 0.001912004).

6. CONCLUSION

Based on the result of the estimates obtained from the simulation study, it can be concluded that the method of maximum likelihood estimation method and numerical method (Newton Raphson method) can be employed to estimate parameter from intractable probability distribution model (Watson probability distribution).

7. RECOMMENDATION

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

- Parameter estimate from Watson distribution should be obtained using maximum likelihood estimation and 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] Box, G. E. P.,& Cox, D. R. (1964). An analysis of transformations. Journal of the Royal Statistical Society: Series B, 26: 211-252.

[4] 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.

[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] Jing, J., Koch, I. & Naito, K. (2012).Polynomial Histograms for Multivariate Density and Mode Estimation. Scandinavian Journal of Statistics 39:7596.

[12] Marin, J., Pudlo, P., Robert, C. P.,& Ryder, R. (2011).Approximate Bayesian Computational methodss. Statistics and Computing in press.

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

[14] 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.

[15] 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.

[16] Wilkinson, R. D. (2008).Approximate Bayesian computation (ABC) gives exact results under the assumption of error model. Arxiv preprint arXiv