ISSN 1995-0802, Lobachevskii Journal of Mathematics, 2020, Vol. 41, No. 3, pp. 430–434.

Parameter Estimation of the Negative Binomial—New Weighted Lindley Distribution by the Method of Maximum Likelihood

Ampai Thongteeraparp[#1*] **and A. Volodin**[2**]

(Submitted by A. M. Elizarov)

[1] *Department of Statistics, Kasetsart University, Bangkok, 10903 Thailand*

[2] *Department of Mathematics and Statistics, University of Regina, Regina, SK, S4S 0A2 Canada*

Received July 2, 2019; revised August 31, 2019; accepted October 3, 2019

Abstract—In this article we investigate the parameter estimation of the Negative Binomial—New Weighted Lindley distribution. We are interested in the maximum likelihood method because it provides estimators with many superior properties, such as minimum variance and asymptotically unbiased estimators. The simulation study is performed in order to investigate the accuracy of the maximum likelihood estimators of the parameters of the Negative Binomial—New Weighted Lindley distribution.

DOI: 10.1134/S1995080220030178

Keywords and phrases: *method of maximum likelihood, count data analysis, negative binomial—new weighted Lindley distribution.*

1. INTRODUCTION

The estimation of parameters is a statistical inferential procedure that is exceptionally important for analysing and modeling data. If we are dealing with a parametric model and trying to fit our data to this model, it is necessary to assign values to the parameters of the model. Choosing an appropriate estimator that provides reliable results is important; hence, there are several optimal criteria for the estimators that should be considered. In this article, we consider the maximum likelihood method of parameter estimation (MLE), which is most frequently used for a parameters estimation. The MLE is a popular technique for estimating a parameter of a given model; as its name suggests, the MLE is meant to maximize the likelihood function. It was introduced by Fisher (1922) [1] and has been broadly used since.

The Lindley distribution was generalized by many researchers in recent years. Ghitany et al. [2] investigated the Lindley distribution in the context of reliability analysis. Subsequently, a weighted Lindley (WL) distribution was proposed for modelling survival data. The WL distribution has the property that the hazard rate (mean residual life) function exhibits upside-down bathtub and monotone (increasing or decreasing) shapes. The main advantage of the WL distribution is in its wide applications to real survival data.

The New Weighted Lindley (NWL) distribution was recently introduced by Asgharzadeha et al. [3]. It is a two parameter continuous lifetime distribution defined as follows. A random variable X follows the NWL distribution with parameters $\alpha > 0$ and $\beta > 0$, if its probability density function is

$$f(x) = \frac{\beta^2(1+\alpha)^2}{\alpha\beta(1+\alpha)+\alpha(2+\alpha)}(1+x)\,(1-\exp(-\alpha\beta x))\exp(-\beta x), \quad x > 0,$$

where $\alpha > 0$ and $\beta > 0$.

*E-mail: fsciamu@ku.ac.th

**E-mail: andrei@uregina.ca

Corresponding author.

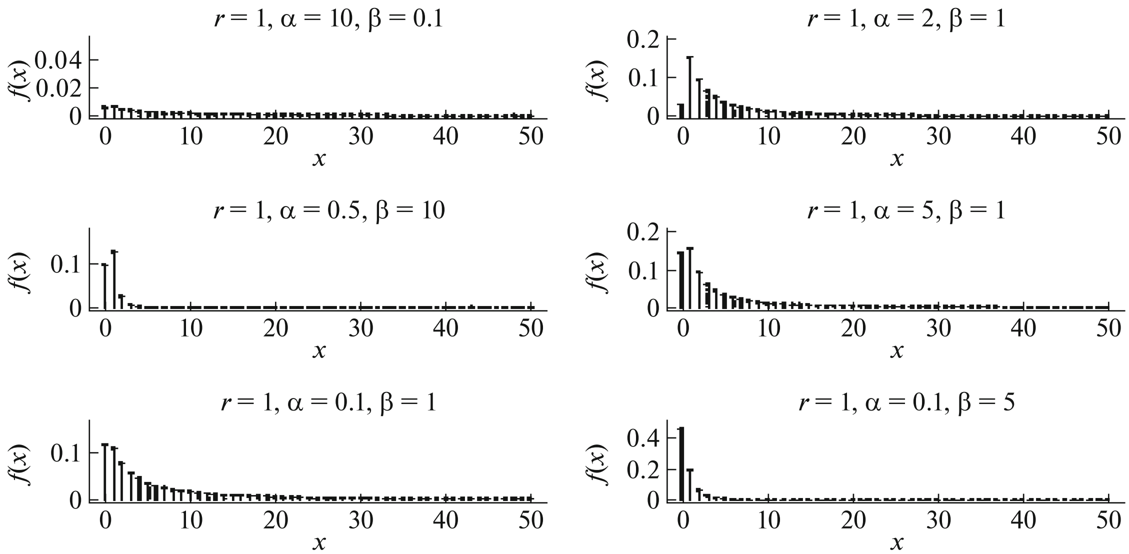

Fig. 1. The probability mass function of the NB-NWL distribution of some specified values of r, α, and β.

The Negative Binomial—New Weighted Lindley (NB-NWL) distribution is proposed by Samutwachirawong (2017) [5] as follows. Let $X|\lambda$ be a random variable following the Negative Binomial distribution with parameters r and $p = \exp(-\lambda)$; that is, $X|\lambda \sim NB(r, p = \exp(-\lambda))$. If λ has the NWL distribution with parameters α and β (denoted as $\lambda \sim NWL(\alpha, \beta)$), then X is called a *NB-NWL random variable*.

The probability mass function of X is given by

$$f(x; r, \alpha, \beta) = \binom{r+x-1}{x} \sum_{j=0}^{x} \binom{x}{j} (-1)^j B, \quad x = 0, 1, 2, \dots,$$

where

$$B = B(r, \alpha, \beta) = \frac{\beta^2(1+\alpha)^2}{\alpha\beta(1+\alpha) + \alpha(2+\alpha)} \left\{ \frac{\beta + r + j + 1}{(\beta + r + j)^2} - \frac{\beta(1+\alpha) + r + j + 1}{[\beta(1+\alpha) + r + j]^2} \right\},$$

and $\alpha > 0$, $\beta > 0$, and r is a positive integer.

Below we present some typical probability mass function plots of this distribution.

The usefulness of the NB-NWL distribution is illustrated in Samutwachirawong (2017) [5] by its application to modeling of the number of hospitalized patients with diabetes at Ratchaburi hospital, Thailand. The result of the study shows the good fit of NB-NWL distribution for this purpose (based on the log-likelihood value and p-value of the Kolmogorov—Smirnov test). Obviously, the NB-NWL distribution is a flexible distribution to model count data.

In this article, we investigate the performance of the maximum likelihood parameter estimation. Some specified values of parameters are fixed, and simulated data is considered. The simulations were performed to compare the true value of the parameters of the NB-NWL distribution and their maximum likelihood estimations. Simulations are performed in R language, Version 3.4.1 [4].

2. MAIN RESULTS

2.1. Parameter Estimation

Here the parameter estimation for the NB-NWL distribution via the MLE method is discussed. If $X \sim NB - NWL(r, \alpha, \beta)$, then the likelihood function of parameters r, α, and β can be written as

$$L(r, \alpha, \beta; X) = \prod_{i=1}^{n} \left[\binom{r + x_i - 1}{x_i} \sum_{j=0}^{x_i} \binom{x_i}{j} (-1)^j B \right].$$

Table 1. Scenarios of the simulation study for the parameter estimates of the NB-NWL distribution

Scenarios	r	α	β	$E(X)$	$Var(X)$
1	1	10	0.1	0.49	1.21
2	1	2	1.0	2.71	14.49
3	1	0.5	10	5.14	60.87
4	1	5	1.0	6.05	57.20
5	1	0.1	1.0	5.25	22.78
6	1	0.1	5.0	10.21	60.14

Consequently, the associated log-likelihood function is

$$\mathcal{L} = \log L(r, \alpha, \beta; X) = \sum_{i=1}^{n} [\log \Gamma(r, x_i) - \log \Gamma(r) - \log \Gamma(x_i + 1)] + \sum_{i=1}^{n} \log \left[\sum_{j=0}^{x_i} \binom{x_i}{j} (-1)^j B \right],$$

where $\Gamma(t) = \int_0^\infty x^{t-1} e^{-x} dx$ is the Gamma function, and we apply the well known identity $\Gamma(n) = (n-1)!$.

The MLEs for the parameters are obtained by differentiating the log-likelihood function with respect to r, α, and β, which provides the following score equations:

$$\frac{\partial \mathcal{L}}{\partial r} = \sum_{i=1}^{n} \Psi(r - x_i) - n\Psi(r) + \sum_{i=1}^{n} \left(\frac{\sum_{j=1}^{x_i} \binom{x_i}{j} (-1)^j \frac{\partial B}{\partial r}}{\sum_{j=0}^{x_i} \binom{x_i}{j} (-1)^j B} \right),$$

$$\frac{\partial \mathcal{L}}{\partial \alpha} = \sum_{i=1}^{n} \left(\frac{\sum_{j=1}^{x_i} \binom{x_i}{j} (-1)^j \frac{\partial B}{\partial \alpha}}{\sum_{j=0}^{x_i} \binom{x_i}{j} (-1)^j B} \right), \quad \frac{\partial \mathcal{L}}{\partial \beta} = \sum_{i=1}^{n} \left(\frac{\sum_{j=1}^{x_i} \binom{x_i}{j} (-1)^j \frac{\partial B}{\partial \beta}}{\sum_{j=0}^{x_i} \binom{x_i}{j} (-1)^j B} \right),$$

where $\Psi(t) = \frac{\Gamma'(t)}{\Gamma(t)}, t \geq 0$, is the digamma function.

The maximum likelihood estimators $\hat{r}, \hat{\alpha}$, and $\hat{\beta}$ for the parameters r, α, and β, respectively, are obtained by equating the score equations to zero. These estimates cannot be found in closed form, and a numerical method can be employed to obtain the solutions. The MLE estimators $\hat{r}, \hat{\alpha}$, and $\hat{\beta}$ can be obtained by solving the resulting equations simultaneously using the *optim* function in R language.

2.2. Random Variate Generation of NB-NWL Distribution

A random variate X from the $NB - NWL(r, \alpha, \beta)$ can be generated by the following algorithm.

1. Generate U from the uniform distribution $U(0, 1)$.
2. Set $\lambda = -\frac{1}{\beta} \log(1 - U^{1/\alpha})$. Then λ follows the $NWL(\alpha, \beta)$ distribution.
3. Generate X from $NB(r, p = \exp(\lambda))$ distribution.

2.3. Simulation Results

The following are six scenarios for our simulation study.

The results of the simulation study of the NB-NWL distribution for all six scenarios are presented in Tables 2–3. We found that, in all cases, the MSE of the parameter estimates decreases as sample size increases.

Table 2. Scenario 1 (true parameters $r = 1, \alpha = 10, \beta = 0.1$), Scenario 2 (true parameters $r = 1, \alpha = 2, \beta = 1$) and Scenario 3 (true parameters $r = 1, \alpha = 0.5, \beta = 10$)

n	Parameter	Scenario 1			Scenario 2			Scenario 3		
		estimate	Var	MSE	estimate	Var	MSE	estimate	Var	MSE
20	r	1.1653	2.3768	2.4041	8.2991	38.3054	49.1895	5.7178	19.5551	26.9413
	α	6.0760	5.7539	6.0857	7.2345	18.3679	23.3608	0.5444	9.3652	11.1727
	β	0.5265	2.3539	4.5251	2.3312	1.7218	1.8315	3.0028	1.3735	1.6207
50	r	2.6532	3.3961	6.1292	6.6146	24.0552	26.6623	4.2175	4.8497	6.3321
	α	5.2249	9.5775	9.6532	6.5797	13.3695	15.8650	0.59267	6.1867	7.0455
	β	2.6008	3.2735	3.6344	2.2957	0.7430	0.8304	3.3333	1.0130	1.0408
100	r	1.2797	1.7044	1.7826	1.0713	8.7470	8.7521	1.6500	0.9254	1.3478
	α	8.5214	6.5140	6.5145	3.4622	4.4752	4.6888	0.4563	2.1256	2.3921
	β	2.2122	2.7089	2.7540	1.0977	0.1762	0.1858	9.4132	0.8310	0.8385
200	r	1.3618	1.2535	1.3844	1.0800	1.4390	1.4454	1.0685	0.3785	0.3832
	α	10.3750	4.7081	5.4738	2.0939	0.9557	0.9645	0.5070	2.1256	2.1256
	β	0.18606	1.6630	1.6825	1.0977	0.1762	0.1858	10.0983	0.3180	0.3180

Table 3. Scenario 4 (true parameters $r = 1, \alpha = 5, \beta = 1$), Scenario 5 (true parameters $r = 1, \alpha = 0.1, \beta = 1$) and Scenario 6 (true parameters $r = 1, \alpha = 0.1, \beta = 5$)

n	Parameter	Scenario 1			Scenario 2			Scenario 3		
		estimate	Var	MSE	estimate	Var	MSE	estimate	Var	MSE
20	r	4.7156	5.1411	6.6188	2.6538	14.8658	21.9084	4.6725	5.6729	6.1252
	α	7.7702	14.7727	22.4466	2.6538	14.8658	21.9084	1.6135	21.4444	22.6843
	β	5.5979	2.7251	3.9304	5.7710	2.8516	3.4460	7.1033	5.7667	5.7774
50	r	3.8656	1.6124	1.7460	1.7749	8.3925	8.9930	1.1471	1.3645	1.3861
	α	5.7470	3.1760	3.7340	0.9314	15.9534	24.5466	1.7527	7.7427	7.8066
	β	4.8481	0.6350	0.7561	2.5078	1.7062	1.9640	6.0596	2.0796	2.0831
100	r	1.6902	0.6910	0.7272	1.0207	1.3167	1.3171	1.0952	0.6742	0.6833
	α	5.3164	0.7436	0.8437	0.8631	4.7826	8.2536	0.8905	4.1265	4.1347
	β	1.7450	0.4945	0.5545	1.2708	0.6347	0.7080	6.0370	1.2737	1.2750
200	r	1.5036	0.2486	0.2486	1.0101	0.8851	0.8852	1.0225	0.3470	0.3475
	α	5.0205	0.1996	0.2000	0.1163	2.4744	2.4747	0.1562	1.7924	1.7946
	β	1.5072	0.1156	0.1157	1.0713	0.5714	0.5765	1.0164	0.5214	0.5217

3. CONCLUSION

By generating random samples from the NB-NWL distribution, we performed the Monte-Carlo simulations for a comparison of the true value of the parameters of the NB-NWL distribution with their estimates by the method of maximum likelihood. Six scenarios for true parameter values were considered which are presented in Table 1 with the means and variances of the corresponding NB-NWL distribution. A simulation study for all scenarios was performed with data generated from this

NB-NWL distribution with sample sizes 20, 50, 100, and 200, respectively. We used the statistical software R to generate each sample of a fixed size and repeated this for 500 trials. The MSE of the parameter estimates decreased as sample size increased. For a small sample size ($n = 20$), the efficiency of parameter estimation seems to be poor.

REFERENCES

1. R. A. Fisher "On the mathematical foundations of theoretical statistics," Philos. Trans. R. Soc. London, Ser. A **222**, 309–368 (1922).
2. M. E. Ghitany, B. Atieh, and S. Nadarajah, "Lindley distribution and its application," Math. Comput. Simul. **78**, 493–506 (2008).
3. A. Asgharzadeha, S. B. Hassan, S. Nadarajah, and F. Sharafia, "A new weighted Lindley distribution with application," Brazil. J. Prob. Stat. **30** (1), 1–27 (2016).
4. R Core Team, *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, Vienna, Austria, 2015).
5. S. Samutwachirawong, "A negative binomial-new weighted Lindley distribution for count data and its application to hospitalized patients with diabetes at Ratchaburi hospital, Thailand," in *Proceedings of the 2nd International Conference of Multidisciplinary Approaches on UN Sustainable Development Goals (UNSDGs), Bangkok, Thailand, 28–29 December 2017* (2017), pp. 25–32.