BMC JOURNAL OF SCIENTIFIC RESEARCH
A Multidisciplinary Reviewed Research Journal
ISSN: 2594-3421

Prediction of Birth Weight of Newborn Nepalese' Babies by Maximum Likelihood Estimation

Govinda Prasad Dhungana*[1]
Email: govinda.dhungana@bimc.tu.edu.np

Abstract

The first weight of newborn babies is a vital indicator of their growth and progress in the performance of their life span. Normal birth weight children are healthier than low birth weight and overweight. Total 1111 data were taken from the 2016 NDHS data set and the R environment used to estimate the average birth of newborn babies.

The maximum likelihood estimate of normal distribution is used to determine the average (SD) weight of newborn babies. The estimated value is validation by KS test and pdf plot with normal curve. Prevalence of low birth weight and overweight from the standard normal distribution has also been predicted.

The mean weight of newborn infants of the children is (2.94 kg) 2940 grams and the standard deviation is (0.573 kg) 573 grams. The percentage of low birth weight and high birth weight is 22.04% and 3.14% respectively. The average weight of Nepalese newborn babies is ordinary, but the prevalence of low birth weight is still high in Nepal.

Keyword: Low birth weight, Maximum likelihood Estimation, Newborn baby, Nepal, Parameters

1. Introduction

Birth weight is the first weight of baby, taken just after he or she is born. It is the crucial measure for new born babies defined as normal weight (2.5 kg to 4.0 kg), low birth weight (< 2.5 kg) and high birth weight (> 4.0 kg) [1]. Normal birth weight is healthier in their future life, while low and overweight babies are more suspected and infected with various diseases [2]. Also, infants with overweight may be at higher risk of birth injury. The average birth weight of babies of European heritage and Canada is 3.5 kilograms, while the standard range is between 2.5 kilograms and 4.5 kilograms [3], in the United States (US), it is between 3.0 kg and 3.5 kg and the weight of Indian newborn birth is between 2.5 kg and 3.5 kg [4]. The birth weight of newborn babies varies with different maternal factors such as, maternal age, mother's weight during pregnancy, blood sugar, blood pressure, pregnancy complications, uterine condition, substance abuse and nutritional intake during pregnancy. Nepal is one of the least developed countries with poverty and inadequate health services that lead to high child morbidity and mortality rates [5].

The median weight of the Nepalese male infants was 3.1 kg, while the female infants were 2.9 kg [5]. The percentage of low birth weight in Nepal was 11.9% [6]. The

* Mr. Dhungana is Lecturer in Statistics at Birendra Multiple Campus (Tribhuvan University)
[1] Currently he is Ph.D Scholar, Deen Dayal Upadhaya Gorakhpur University, Gorakhpur, India

Global Nutrition Target by 2025 (World Health Assembly [WHA]) is a reduction of 30% in low birth weight and no increase in childhood overweight [6].

The prevalence of abnormal (low birth weight as well as overweight) is still high in Nepal. In the current situation, the prevalence is estimated by the direct technique (population base prevalence), but this paper focuses on estimating the prevalence by theoretical distribution. The theoretical approach is Maximum Likelyhood Estimation (MLE) method is used to predict the prevalence of abnormal weight. In statistics, MLE is a method of estimating the parameters of the statistical model used in the observations, by finding the parameter values that maximize the likelihood of making observations based on the parameters. The aim of this study is therefore to estimate the prevalence of the average birth weight of newborn babies in Nepal by using the MLE technique of normal distribution.

2. Methods and Materials

Since the birth weight is measured by Kg, which is continuous in nature. Continuous data measured by different theoretical distributions, of which the most popular theoretical distribution is a normal distribution having a probability density function (pdf),

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, \qquad -\infty < x < \infty \tag{1}$$

Where, mean (μ) and standard deviation (σ) are both unknown parameters of birth weight of newborn babies. Therefore, these two unknown parameters are estimated by the MLE technique of normal distribution. The MLEs of the unknown's parameters of the distribution based on $\underset{\sim}{x} = \left(x_1, \ldots, x_n\right)$ observed sample value with of set of parameters $\ell(\mu, \sigma \mid \underset{\sim}{x})$ is monotonic increase function. Hence, the log likelihood function of the parameter $\ell(\mu, \sigma)$ is given by;

$$\ell\text{n}(\ell) = -n\ \ell\text{n}\left(\sigma\sqrt{2\pi}\right) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu)^2 \tag{2}$$

Maximum likelihood estimators of the parameters have obtained by partial differentiating w.r.t.to parameters and equating to zero, we have

$$\frac{\partial \ln \ell}{\partial \mu} = \frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i - \mu) = 0 \tag{3}$$

$$\frac{\partial \ln \ell}{\partial \sigma} = -\frac{n}{\sigma} + \frac{1}{\sigma^3}\sum_{i=1}^{n}(x_i - \mu)^2 = 0 \tag{4}$$

Therefore, we estimate the unknown parameters μ (mean birth weight) and σ (standard deviation of birth weight) by maximizing the log-likelihood function of equation (2) directly using *optim* () function of R software used as Newton-Raphson iterative technique [7, 8].

For interval estimation of model parameters, vector $\delta = (\mu, \sigma)$ and the corresponding MLE of δ is $\hat{\delta} = \left(\hat{\mu}, \hat{\sigma}\right)$, are asymptotic normally distributed i.e. $((\hat{\delta} - \delta) \to N_2(0, (I(\delta))^{-1})$.

Where, I (δ) is the fisher's information matrix. In practice, it is useless that the MLE has asymptotic variance $I(\delta))^{-1}$ because we do not know δ . Hence, we approximate the asymptotic variance by "plugging in "the estimated value of the parameters. The common procedure is to use observed information matrix $O(\hat{\delta})$ (as an estimated of the information matrix I (δ)) given by

$$O(\hat{\delta}) = -\begin{pmatrix} \frac{\partial^2 \ell}{\partial \mu^2} & \frac{\partial^2 \ell}{\partial \mu \partial \sigma} \\ \frac{\partial^2 \ell}{\partial \mu \partial \sigma} & \frac{\partial^2 \ell}{\partial \sigma^2} \end{pmatrix} = -H(\delta)_{|\delta=\hat{\delta}} \tag{5}$$

Where, elements of the observed information matrix are:

$$\frac{\partial^2 \ln \ell}{\partial \mu^2} = -\frac{n}{\sigma^2} \tag{6}$$

$$\frac{\partial^2 \ln \ell}{\partial \sigma^2} = -\frac{2n}{\sigma^2} \tag{7}$$

$$\frac{\partial^2 \ln \ell}{\partial \mu \, \partial \sigma} = -\frac{2}{\sigma^3} \sum_{i=1}^{n} (x_i - \mu) = 0 \tag{8}$$

The inverse of Hessian matrix is called variance –covariance matrix which is obtained by Newton Raphson algorithm to maximize the likelihood produces from observed information matrix. Therefore, it is given by

$$\left(-H(\delta)_{|\delta=\hat{\delta}}\right)^{-1} = \begin{pmatrix} Var(\hat{\mu}) & cov(\hat{\mu},\hat{\sigma}) \\ cov(\hat{\mu},\hat{\sigma}) & \mathrm{var}(\hat{\sigma}) \end{pmatrix} \tag{9}$$

Finally, from the asymptotic normality of MLE's, approximate 100 (1- α) % confidence interval for μ and σ can be constructed as

$$\hat{\mu} \pm z_{\alpha/2}\sqrt{\mathrm{var}(\hat{\mu})} \text{ , and } \quad \hat{\sigma} \pm z_{\alpha/2}\sqrt{\mathrm{var}(\hat{\sigma})} \tag{10}$$

Where, $z_{\alpha/2}$ is the upper percentile of standard normal variate.

3. Data Analysis

Nepal Demographic and Health Survey (NDHS) 2016 is a national representative quantitative population-based survey. This study was completed from 19 June 2016 to 31 January 2017 [1]. The data set is available from *https://www dhsprogram com/ data/dataset_admin* . A total of 1146 birth weight of newborn babies was recorded in available dataset. After download the dataset, only newborn birth weight was extracted from the dataset and checked for completeness and accuracy. To detect the outliers using the following expression.

Outliers < { Q_1 -1.5× ($Q_3 - Q_1$)} and { Q_3 +1.5× ($Q_3 - Q_1$)} < Outliers (11)

There were 35 (3.05%) outliers were identified from the dataset. These outliers were eliminated and final sample was 1111 taken for study purpose.

4. Results

4.1 Exploratory Data Analysis (EDA)

Exploratory data analysis (EDA) is a method to analyzing data sets and it summarizes their main characteristics. We have shown that the basic descriptive statistic whereas mean and median was approximately same. The nature of data is more or less normal from the descriptive statistics (Table 1).

Table 1: Descriptive statistics of birth weight of newborn babies (n=1111)

Min	Q_1	Mean	Median	Q_3	Max	Skewness	Kurtosis
1.10	2.50	2.94	3.00	3.50	4.30	-0.400	2.88

4.2 Estimated value of Parameters

We estimated the value of parameters with standard error (SE) by maximizing the log-likelihood function given in equation (2) directly using *optim* () function in R. Therefore, the estimated value of μ and σ with 100 (1- α) % interval of birth weight of newborn babies was in Table 2. The maximized value of log likelihood was $\ell(\hat{\mu},\hat{\sigma})$ = - 957.8475.

Table 2: Estimated value of parameters of birth weight of newborn babies

Parameters	MLE	SE	95% CI	P-value
$\hat{\mu}$	2.94339	0.01719	2.91 to 2.98	<0.001
$\hat{\sigma}$	0.57305	0.01216	0.55 to 0.59	<0.001

In figure 1, we plot the profile of negative log-likelihood function of $\hat{\mu}$ and $\hat{\sigma}$ for the given data set. It is clear that the likelihood equation have a unique solution.

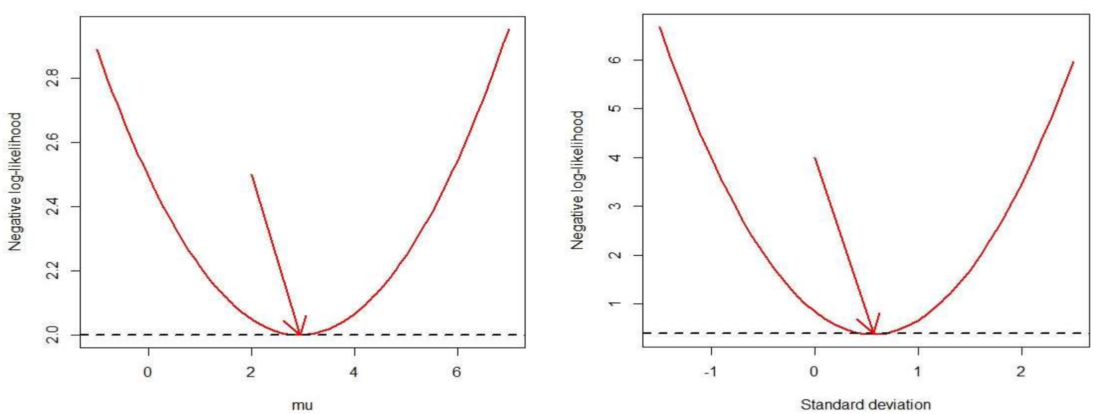

Figure .1 Profile log-likelihood function of $\hat{\mu}$ (left panel) and $\hat{\sigma}$ (right panel)

4.3 Model Validation

To verify the validity of the model, we measure the Kolmogorov-Smirnov (KS) test. It measures the difference between the empirical distribution function and the fitted distribution function when the parameters are obtained by MLE. The KS result was 0.1253 ($p < 0.001$). It was found that there was no perfect match for theoretical and empirical distribution functions. However, the data were fitted on histogram with the predicted value of MLE the normal distribution. This indicates that our dataset was approximately normal (Figure 2).

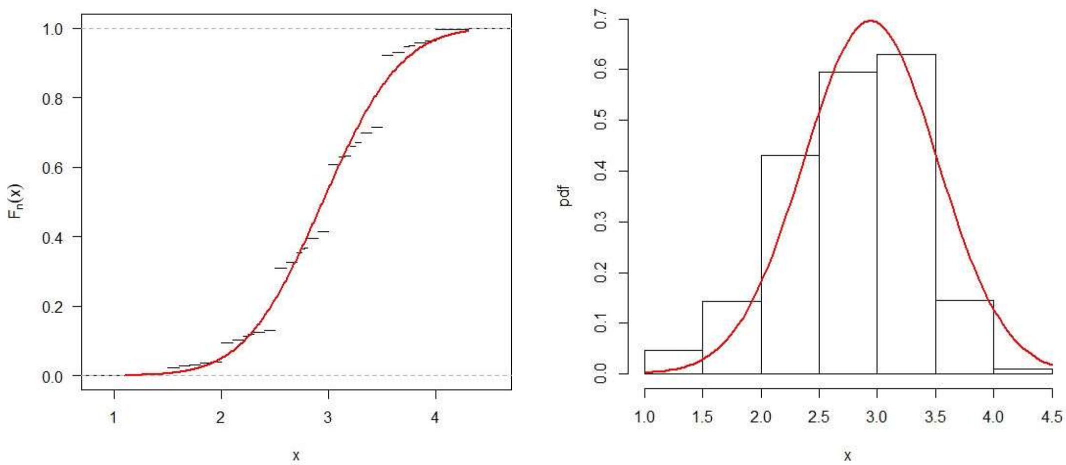

Figure 2: The empirical and fitted distribution function (left panel), histogram with predicted normal distribution (right panel)

4.3 Prediction of Low birth weight and overweight

Now, we predict the low birth weight of newborn babies of Nepal .Let X denote the low birth weight of newborn children and WHO provide a cutoff point of low birth weight is < 2.5 kg of newborn babies. i.e. $p(X< 2.5)$ =?

We have estimated the parameters by MLE and the value of standard normal distribution is $Z = \frac{X-\hat{\mu}}{\hat{\sigma}}$ = -0.77. Therefore, we have, P (z < - 0.77) = $p(-\infty < z < 0) - p(-\infty < z < -0.77)$ = 0.2206=22.06% (Figure 3, left panel). Finding revels that approximately 22.06% of the newborn child of babies will be low birth weight. Similarly, we have to predict the overweight newborn babies because WHO provide a cutoff point of overweight is > 4 Kg of newborn babies. i.e. $P(X > 4)$ =?

Likewise, standard normal distribution, $Z = \frac{X-\hat{\mu}}{\hat{\sigma}}$ =1.86, then, P (z >1.86) = $p(0 < z < \infty) - p(0 < z < 1.86)$ = 0.0314 =3.14% (Figure 3, right panel). Finding revels that approximately 3.14% of newborn child will be overweight.

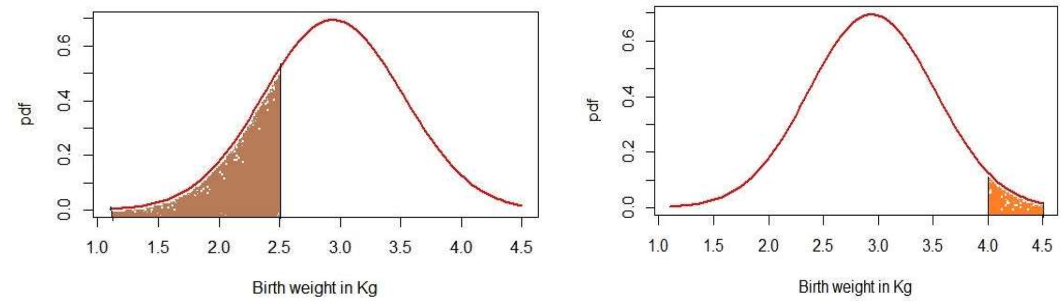

Figure 3: Prediction of low birth weight (left panel), overweight (right panel)

5. Discussion:

The present study focuses on estimating the parameters, mean and standard deviations of the birth weight of newborn babies by MLE of normal distribution. The average weight of newborn babies is 2.9 kg or 2900 grams and the standard deviation is 0.573 kg or 573 grams. The median weight of newborn babies is 3.0 kg or 3000 grams and the first quartile is 2500 grams and the third quartile is 3500 grams. The median birth weight of females was 2900 grams and males was 3010 grams in Patan Hospital [5]. Similarly, another study conducted by Sreeramareddy, et al. [9], found that the average birth weight was 3029 ± 438 grams. Likewise, a study conducted by Khanal et al. [10], the average birth weight was 3024 (SD = 654.5) grams. The results of all findings are identical with the present study because there is no disparity in the average weight of newborn babies in different periods of time. Likewise, other characteristics such as maternal age, age at pregnancy, nutritional status are more or less similar, all over the country.

In this study, the predicted prevalence of low birth weight is 22.04 % and overweight is 3.14 %. A study conducted by Sharma et al. [11], the estimate of low birth weight prevalence was (12-32) %. Likewise, another study conducted in Chitwan [12], out of 220 newborn babies, the prevalence of LBW was 23.6 % (95 % CI: 21.88 % to 25.32%). The findings of these studies are similar to the results of the present study. There is no difference in the predicated prevalence of low birth weight of the present study and the prevalence of LBW which is calculated from the population base study. Furthermore, these studies are adopted in limited sample size and a small area, but result is an approximate equal with the present study, which implies that the prediction methodology is more robust. Similarly, LBW estimated prevalence in globally that 15–20% of all births [13, 14], which is approximate same with the present study.

Some other study like as, Khanal et al. [10], showed that the prevalence of LBW was 12.1% (95% CI: 10.6%-13.7%) and NDHS 2016 reports the prevalence of low birth weights was 12.0%. These findings are contradictory with the finding of the present study. There may be variations due to prediction methods. The present study predicts the prevalence of theoretical method which is used for forecasting future prevalence, but the research of Khanal et al., and NDHS reported population base prevalence. The next cause may be Khanal et al. was used pooled data from the 2006 and the 2011 Nepal Demographic and Health Surveys (NDHS) whereas the present study focused on 2016 NDHS data set. There may also be varied, since the p-value of the KS test is < 0.001, which shows that our proposed model does not satisfy the goodness of fit, reflect that the prediction value has difference than the observed prevalence value.

6. Conclusion

The maximum likelihood estimate of normal distribution is used to determine the average (SD) weight of newborn babies. The value of log likelihood is $\ell(\hat{\mu},\hat{\sigma}) = -957.8475$. It is clear that the likelihood function has a unique solution. To verify the validity of the model, Kolmogorov-Smirnov (KS) test is 0.1253 ($p < 0.001$). It is found that there is no perfect fit of theoretical and empirical distribution functions, reflect that some deviation happens to predict the prevalence of low birth weight and overweight. The mean weight of newborn babies of children is (2.94 kg) 2940 grams and the standard deviation is (0.573 kg) 573 grams. The percentage of low birth weight and high birth weight is 22.04 % and 3.14 %, respectively. The average weight of newborn babies in Nepal is normal, but the prevalence of low birth weight is still high in Nepal. In order to prevent low birth weight, the Government of Nepal focused on health education in maternal health and nutrition in adolescents as well as in newly married women in school education. In the same way, other awareness programs such as reproductive age of mothers, weight of mothers during pregnancy, workload during pregnancy, pregnancy complications, substance abuse and consumption of food and nutrition during pregnancy etc. are provided to adolescent girls and newly married women by different programs.

Reference:

[1] Ministry of Health, Nepal; New ERA; and ICF (2017). *Nepal Demographic and Health Survey 2016 Kathmandu,* Nepal: Ministry of Health, Nepal

[2] Health Topics, available on https://medlineplus.gov

[3] Birth weight Wikipedia, available on https://en.wikipedia.org, wiki, birth weight, 2019

[4] Birth statistics, available on https://www.verywellfamily.com/baby-birth-weight-statistics-2633630

[5] Ulstein, M., Rana, G., Yangzom, K., Gurung, R., Karki, A., & Pradhan, U. (1988). Some fetal and pregnancy parameters in Nepal. *Acta obstetricia et gynecologica Scandinavica,* 67(1), 47-52.

[6] Department of Health services (DoHS), Annual report 2017/18

[7] Braun, W. J., & Murdoch, D. J. (2016). *A first course in statistical programming with R.* Cambridge University Press.

[8] Team, R. C. (2019). A language and environment for statistical computing. Vienna,

Austria: R Foundation for Statistical Computing; 2012. *URL https://www. R-project. org.*

[9] Sreeramareddy, C. T., Chuni, N., Patil, R., Singh, D., & Shakya, B. (2008). Anthropometric surrogates to identify low birth weight Nepalese newborns: a hospital-based study. *BMC pediatrics, 8(1), 16.*

[10] Khanal, V., Zhao, Y., & Sauer, K. (2014). Role of antenatal care and iron supplementation during pregnancy in preventing low birth weight in Nepal: comparison of national surveys 2006 and 2011. *Archives of Public Health, 72(1), 4.*

[11] Sharma, S. R., Giri, S., Timalsina, U., Bhandari, S. S., Basyal, B., Wagle, K., & Shrestha, L. (2015). Low birth weight at term and its determinants in a tertiary hospital of Nepal: a case-control study. *PloS one, 10(4), e0123962.*

[12] Bansal, P., Garg, S., & Upadhyay, H. P. (2019). Prevalence of low birth weight babies and its association with socio-cultural and maternal risk factors among the institutional deliveries in Bharatpur, Nepal. *Asian Journal of Medical Sciences, 10(1), 77-85.*

[13] Cutland, C. L., Lackritz, E. M., Mallett-Moore, T., Bardají, A., Chandrasekaran, R., Lahariya, C., & Muñoz, F. M. (2017). Low birth weight: Case definition & guidelines for data collection, analysis, and presentation of maternal immunization safety data. *Vaccine,* 35(48Part A), 6492.Vaccine. 2017 Dec 4; 35(48Part A): 6492–6500. doi: 10.1016/j.vaccine.2017.01.049

[14] Rejali, M., Mansourian, M., Babaei, Z., & Eshrati, B. (2017). Prediction of low birth weight delivery by maternal status and its validation: Decision curve analysis. *International journal of preventive medicine, 8.*