International Journal of Health, 3 (2) (2015) 34-37

International Journal of Health

Journal home page: www.sciencepubco.com/index.php/IJH
doi: 10.14419/ijh.v3i2.5033
Research Paper

A new probability model for estimation of child mortality for fixed parity

Sonam Maheshwari[1]*, Brijesh P. Singh [2], Puneet Kumar Gupta [3]

[1]Department of Community Medicine, SRMS-IMS, Bareilly-243202
[2]Faculty of Commerce & DST-CIMS, Banaras Hindu University, Varanasi
[3]Economics and Statistics Office, Vikas Bhawan, Rampur-244901
**Corresponding author E-mail:maheshwarisonam2@gmail.com*

Abstract

In demography, child mortality is useful as a sensitive index of a nation's health conditions and as guided for the structuring of public health schemes. In the present study, we proposed a probability model for the number of child loss among females for a fixed parity. The application of the model proposed in the paper is illustrated through its application to the data from Madhya Pradesh from National Family Health Survey-III (NFHS-III). Finally, we show that proposed model is better fitted than the Beta-Binomial model for the data.

Keywords:*Child Mortality; Parameter Estimation; Kumaraswamy Distribution*.

1. Introduction

Level of child mortality of any nation is a widely accepted and sensitive sign of the social and economic progress of that nation also it also helps in evaluating the impact of various intervention programs, which have aim to improve child survival. Child mortality, the probability of death between the first and the fifth birthday has been a main concern for the Government of India over the past several decades. Many child survival programs have been commenced by Government of India. Due to these programs, child mortality has grasped remarkable improvement. However, at the current pace, is unable to attain the Millennium Development Goal (MDG) 4 -which aims to reduce Under-five Mortality (U5MR) by two thirds between 1990 and 2015. Six states, namely Kerala, Tamil Nadu, Maharashtra, Punjab, Himachal Pradesh and West Bengal are likely to achieve the goal by 2015.

Data obtained from the three rounds of National Family Health Surveys conducted in the years 1992-93, 1998-99 and 2005-06 and the Sample Registration System (1978-2010) revealed drastic turn down in seventies, stand still in nineties and then started declining again in the last decade. During 1968-70, the level of child mortality rates in India was about 190 per 1000 live births; the child mortality rate started declining in the late 1970s and until 1993 the rate of decline was substantial. The decline, however, slowed during 1993-98. The country's goal to achieve child mortality rate less than 100 per 1000 live births by the year 2000, was not achieved despite improved interventions and an increase in the overall resources. The latest mortality estimates for 2009 in India indicates that 64 per 1000 live births died before reaching the age from five years (SRS 2011).

Child mortality still remains alarmingly high in India. So it is of immense need to estimate the accurate estimate of child mortality for planning the intervention programs which have aim to improve child survival.The direct measurement of age at mortality is not reliable, because it suffers from the substantial degree of errors; such as age misreporting and digit preference. Usually, errors occur due to recall laps, which result in omission of events, misplacement of dates and distortion of reports on the duration of vital events. One of the difficulties in mortality is its non-experimental nature. Many times it is found that more than a single factor operating over a phenomenon under study. Since all the analyses in this case are based on human beings, and hence it is difficult to control a number of factor at the same time and repeats the experiment under identical conditions. These problems may be accomplished only through models, which make it possible to obtain estimates from information other than vital statistics.

Estimation of child mortality for all births has been taken place in last five years before the survey (Hill and Devid, 1989). However, the estimate obtained through this method also suffers from the problem of under reporting (Pathak et al., 1991). In these Circumstances, some of the earlier studies about child mortality by using model Chauhan (1997), Goldblatt (1989), Heligman and Pollard (1980), Krishna (1993), Ronald and Carter (1992), Keyfitz (1977) used a hyperbolic function to study the infant and child mortality. Afterward, Arnold (1993) used Pareto distribution and Krishnan (1993) applied finite range model for estimating the same. Bhuyan and Degraties (1999) suggested the use of Polya-Aeppli model to study the trend of child mortality. They used number of child deaths in the household to present the trend of child mortality.

Motivated by the fact that data related to age at child mortality may have more bias than data on a number of child loss, thus keeping in view the fact that study of the distribution of child loss is a powerful device to explain changes and variation within the population (Singh et al. 2011). Singh et al. (2011) also studied the variation of risk of child mortality. In this study, our objective is to develop some probability models as the number of child loss among females for a fixed parity. The models have been used for comparing the risk of child loss between different groups of population. Application of these models is illustrated through their application to the data from Madhya Pradesh available through the National Family Health Survey-3.

Let us assume that for i^{th} (i=1,2,….n) child born to a female, a random variable ω_i that takes value 1 if the child dies before age five and 0 otherwise. Thus $\omega_i^{'}s$ are Bernoulli variable. Now, if we consider that death of children (born to a female) below age five are independent of each other having same probability p then total number of child death X to the female is nothing but sum of independent Bernoulli variables and hence follows a binomial distribution. Hence, the distribution of X is defined as

$$P\left[X = x\right] = {}^{n}C_{x}\, p^{x}\left(1-p\right)^{n-x} \quad ; 0 \le p \le 1, ; x = 0,1,2.........n. \tag{1}$$

Where, p is the risk of child death. The probability of risk parameter (p|0 ≤ p ≤ 1) is usually assumed to be a constant from death to death. The mean and the variance of a binomially distributed random variable is then given by E(X) = np and Var(X) = np(1 − p). However, in most practical situations, it has been observed that risk does not remain as a constant from death to death but varies itself as a random variable and hence this leads to treat the risk probability as a continuous random variable P, which is bounded between 0 and 1. The resultant class of distributions is known as the class of Binomial Mixture Distributions. A Binomial mixture distribution can be symbolically denoted as Bin (n, P) ∧FP (p) where Bin (n, P) represents the Binomial distribution and FP (p) symbolizes the distribution function of the mixing distribution of the random variable P and the mixing density of P is denoted by FP (p).

One can assume one of many continuous models for the risk of child death lies within the parameter space [0, 1]. Even though the number of possible univariate continuous model defined on the standard unit interval [0, 1] are available as the mixing distribution with the success probability random variable P, the Beta distribution Beta (a, b), where a and b are the two shape parameters of the Beta distribution, is the most commonly used mixing distribution to model the random variable P due to its property of obliging wide range of shapes. Thus the Beta-Binomial (BB) distribution, represented by Bin (n, P) ∧Beta (a, b), is considered as a very versatile distribution in modeling over dispersed binomial outcome data in literature. Extensive literatures exist on the study of Beta-Binomial distribution. Theoretical properties, estimation techniques and applications of the Beta-Binomial distribution have been discussed by, for example, Skellam (1948), Chatfield and Goodhardt (1970), Griffiths (1973), Williams (1975), Haseman and Kupper (1977), Paul (1982), Tripathi, Gupta, and Gurland (1994), Gange, Munoz, Saez, and Alonso (1996), Ennis and Bi (1998), and recently Bandyopadhya, Reich, and Slate (2011), not limited but numerous.

Recently, Li, Huang, and Zhao (2011) used the Kumaraswamy double bounded (Kumaraswamy 1980), distribution as the mixing distribution of the Binomial probability of success and obtained a new Binomial Mixture distribution called the Kumaraswamy-Binomial (KB) distribution, which can be represented by Bin (n, P) ∧Kumaraswamy (ζ, θ), where ζ and θ are the two shape parameters of the Kumaraswamy distribution. The Kumaraswamy distribution which is defined on the standard unit interval [0, 1], is also known as Minimax distribution. Analogous to the Beta distribution, the Kumaraswamy distribution also has two parameters and can assume a wide variety of shapes (Jones 2009). Li et al. (2011), have used the Kumaraswamy-Binomial distribution to model binomial data and stated that both BB and KB distributions have same flexibility in modeling the over dispersed binomial data.

In view of the above, this paper's purpose to use a new probability model for the number of child loss among females for fix parity. Application of the purposed model is illustrated through their application to the data from Madhya Pradesh available from NFHS-III.

2. The model

As we discussed above, one can assume one of many continuous model for the risk of child death lies within the parameter space [0, 1], we assumed that death of children below age five years for a female having probability “p” is a random variable following Kumaraswamy distribution g(p) with probability distribution function is given by

$$g\left(p\right) = abp^{a-1}\left(1-p^{a}\right)^{b-1}; \quad a,b > 0; 0 \le p \le 1 \tag{2}$$

The conditional distribution of x given p is given by

$$g\left(x,a,b\,|\,p\right) = ab\,{}^{n}c_{x}\sum_{i=0}^{\infty}\left(-1\right)^{i}\,{}^{b-1}c_{i}\,\beta\left(a+ai+x, n-x+1\right)$$

$$; \quad a,b > 0; x = 0,1,2.........n. \tag{3}$$

The above distribution in (3) is known as Kumaraswamy-Binomial (KB) distribution where the parameter (a,b) both are shape parameter. One can get the single value of p for comparing the child death of two places; one may take a mean as an estimate of average child death of the region.

The mean and variance of the Kumaraswamy-Binomial model are.

$$E\left(x\right) = nb\beta\left(1+\frac{1}{a}, b\right)$$

$$V\left(x\right) = n^{2}b\left[\beta\left(1+\frac{2}{a}, b\right) - b\beta^{2}\left(1+\frac{1}{a}, b\right)\right] + nb\left[\beta\left(1+\frac{1}{a}, b\right) - b\beta\left(1+\frac{2}{a}, b\right)\right]$$

3. Estimation procedure

Here, we consider the maximum likelihood estimation (MLE) of the KB distribution. $X = \left(x_1, x_2, ...x_N\right)^T$ Be a random sample of size N from a KB distribution with unknown parameter vector $\Omega = \left(a,b\right)^T$. Then in view of the equation (3), the log-likelihood function for Ω can be written as follows.

$$L\left(\Omega\right) = N\,\log a + N\,\log b + \sum_{k=0}^{N}\log{}^{n}c_{x_k}\sum_{k=0}^{N}\log\left(\sum_{i=0}^{\infty}\left(-1\right)^{i}\,{}^{b-1}c_{i}\,\beta\left(a+ai+x_k, n-x_k+1\right)\right)$$

The MLE $\hat{\Omega} = \left(\hat{a}, \hat{b}\right)^T$ can be obtained either by directly maximizing the above log-likelihood function with respect to Ω or by solving the three simultaneous equations obtained by equating $\Psi\left(\Omega\right) = 0$. The set of equation obtained by $\Psi\left(\Omega\right) = 0$ can be solved for two parameters using any suitable iterative procedure like Newton Raphson’s method. Alternatively, for optimizing the log-likelihood function with respect to the involved parameters, one can use the packages like maxLik( ) of the R-software (Henningsenand Toomet 2010).

4. Results and discussion

To illustrate the application of the proposed models, data from NFHS-III, 2005-06 for Madhya Pradesh have been used. In NFHS-III, the information was collected from ever married females in the group age 15-49 years on fertility, mortality, family planning and important aspects of reproductive health, etc. Also. Information was collected on all live births to females and their survival status at time during the survey. Females having no birth during five years preceding the survey are assumed to have completed their reproductive period and only such females were considered. To avoid censored cases, we have also taken the females who are not having any birth in the last five years.

Table 1: Expected and Observed Distributions of Child Deaths to the Females in Madhya Pradesh with Parity 5

Number of dead children	Observed number of females	Expected number of females		
		Binomial	Beta-Binomial	Kumaraswamy-Binomial
0	159	153.412	160.500	160.543
1	111	116.446	105.758	105.513
2	30	35.355	36.061	36.348
3	10	5.367	7.642	7.613
4	1	0.407	0.980	0.931
5	0	0.012	0.060	0.052
Total	311	311.000	311.000	311.000
Mean	0.659	p=0.132	a=3.986	a=2,314
Variance	0.572		b=26.244	b=81.413
Chi-Square(after pooling)		5.96	1.91	2.08
p-value		0.051	0.167	0.149

Table 2: Expected and Observed Distributions of Child Deaths to the Females in Madhya Pradesh with Parity 6

Number of dead children	Observed number of females	Expected number of females		
		Binomial	Beta-Binomial	Kumaraswamy-Binomial
0	59	59.093	57.954	58.056
1	49	60.008	51.459	51.098
2	28	25.391	27.505	27.746
3	12	5.730	10.548	10.661
4	3	0.727	2.935	2.889
5	0	0.049	0.547	0.506
6	0	0.001	0.052	0.044
Total	151	151.000	151.000	151.000
Mean	1.013	p=0.169	a=2.729	a=1.907
Variance	0.842		b=13.443	b=22.933
Chi-Square(after pooing)		13.36	0.21	0.16
p-value		0.001	0.650	0.688

Table 3: Expected and Observed Distributions of Child Deaths to the Females in Madhya Pradesh with Parity 7

Number of dead children	Observed number of females	Expected number of females		
		Binomial	Beta-Binomial	Kumaraswamy-Binomial
0	23	18.058	21.3885	21.615
1	27	34.313	32.1264	31.754
2	30	27.943	24.9193	24.937
3	12	12.642	12.6973	12.869
4	3	3.432	4.5299	4.565
5	2	0.559	1.067	1.089
6	0	0.051	0.1746	0.160
7	0	0.002	0.097	0.011
Total	97	97.000	21.3885	97.000
Mean	1.495	p=0.214	a=6.145	a=3.077
Variance	1.176		b=22.635	b=81.292
Chi-Square(after pooling)		3.07	2.13	2.01
p-value		0.216	0.344	0.367

Tables 1 to 3 presents the expected frequencies along with observed frequencies for parity 5 to 7 for Madhya Pradesh females. The p-value and values of $\chi 2$ shown in the tables clearly indicate that all the three models described well the distribution of child loss to the females for fixed parity. However, the expected number of child deaths for BB and KB model is very close to the observed number of child deaths.

The advantage of the Binomial model is that the only parameters involved in the model may be obtained by simple calculation. However, the variability of the data is captured more by the BB and KB model. Table 2 and 3 clearly show that the KB model fits well as compare to BB model. So, one May also choice KB model over BB model in various real-world problems.

Acknowledgement

The authors thankfully acknowledge the criticalsuggestions and comments from the learned referee and editorial board which greatly helped us in the improvement over the paper.

References

[1] Arnold BC (1993) Pareto Distributions. Vol.5 in statistical distribution, Fairland (MD), International Co-operative Publishing house.

[2] Bandyopadhyay D, Reich BJ & Slate EH (2011), a spatial beta-binomial model for clustered count data on dental caries. *Statistical Methods in Medical Research*, 20(2), 85-102. http://dx.doi.org/10.1177/0962280210372453.

[3] Bhuyan KC &Degraties R (1999) on the probability model of child mortality pattern in North Eastern Libya. *The Turkish Journal of Population Studies*, 21, 33-37.

[4] Chatfield C &Goodhardt, GJ (1970), the beta-binomial model for consumer purchasing behaviour. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 19(3), 240-250. http://dx.doi.org/10.2307/2346328.

[5] Chauhan RK (1997), Graduation of infant Deaths by Age. *Demography India*, 26(2), 261-174.

[6] Dennis JE & Schnabel RB (1983), Numerical Methods for Unconstrained Optimization and Nonlinear Equations, *Prentice-Hall*, Englewood Cliffs, NJ, 1983.

[7] Ennis DM & Bi J (1998), the beta-binomial model: accounting for inter-trial variation in replicated difference and preference tests. *Journal of Sensory Studies*, 13(4), 389-412. http://dx.doi.org/10.1111/j.1745-459X.1998.tb00097.x.

[8] Gange SJ, Munoz A, Saez M & Alonso J (1996), Use of the beta-binomial distribution to model the effect of policy changes on appropriateness of hospital stays. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 45(3), 371-382. http://dx.doi.org/10.2307/2986094.

[9] Goldblatt PO (1989), Mortality by social class, 1971-85. *Population Trends*, No. 56, Summer.

[10] Griffiths DA (1973), Maximum likelihood estimation for the beta-binomial distribution and an application to the household distribution of the total number of cases of a disease. *Biometrics*, 29(4), 637-648. http://dx.doi.org/10.2307/2529131.

[11] Haseman JK &Kupper LL (1979), Analysis of dichotomous response data from certain toxicological experiments. *Biometrics*, 35(1), 281-293. http://dx.doi.org/10.2307/2529950.

[12] Heligman L & Pollard JH (1980) Age Pattern of Mortality, *Journal of Institute of Actuaries*, 117, 49-80. http://dx.doi.org/10.1017/S0020268100040257.

[13] Henningsena &Toomet O (2010), maxLik: a package for maximum likelihood estimation in R, *Computational Statistics*, 26 (3).

[14] Hill AG &Devid HP (1989), Measuring Child Mortality in the Third World, in N. Sources and Approaches, eds, *IUSSP Proceeding of international conference*, New Delhi, India.

[15] Jones MC (2009), Kumaraswamy's distribution: A beta-type distribution with some tractability advantages. *Statistical Methodology*, 6(1), 70-81. http://dx.doi.org/10.1016/j.stamet.2008.04.001.

[16] Keyfitz N (1977), Introduction to the Mathematics of Population with Revisions. Reading, Mssachusetts. Addition-Wesley Publishing Company.

[17] Krishnan P (1993) Mortality modeling with order Statistics, Edmonton: Population Research Laboratory, Department of Sociology, University of Alberta, Research Discussion paper No. 95.

[18] Kumaraswamy P (1980) a generalized probability density function for double-bounded random processes. *Journal of Hydrology*, 46(1), 79-88.http://dx.doi.org/10.1016/0022-1694(80)90036-0.

[19] Li XH, Huang YY, Zhao XY (2011) the Kumaraswamy Binomial Distribution. *Chinese Journal of Applied Probability and Statistics*, 27(5), 511-521.

[20] Pathak KB, Pandey A, Mishra US (1991) On Estimating current Levels of Fertility and Child Mortality from the Data on Open Birth Interval and Survival Status of the last Child, *Janasamkhya*, 9, 15-24.

[21] Paul SR (1982) Analysis of proportions of affected foetuses in teratological experiments. *Biometrics*, 38(2), 361-370.http://dx.doi.org/10.2307/2530450.

[22] Ronald D Lee, Lawrece R, Carter (1992) Modeling and Forecasting U.S. mortality. *Journal of American Statistical Association*, 87, 659-675.

[23] Singh KK, Singh Brijesh P & Singh N (2011) A Probability Model for Number of Child Death for Fixed Parity. *Demography India*, 40(2), 55-68.

[24] Singh Brijesh P, Singh N, Roy TK & Singh G (2011) on the pattern of child loss in Madhya Pradesh. India 2011, *Mortality, Health and Development (Edited by AalokRanjan)*, 75-88.

[25] Skellam JG (1948) a probability distribution derived from the binomial distribution by regarding the probability of success as variable between the sets of trials. *Journal of the Royal Statistical Society. Series B (Methodological)*, 10(2), 257-261.

[26] Tripathi RC, Gupta RC, Gurland J (1994) Estimation of parameters in the beta binomial model. *Annals of the Institute of Statistical Mathematics*, 46(2), 17-331.http://dx.doi.org/10.1007/bf01720588.

[27] Williams DA (1975) the analysis of binary responses from toxicological experiments involving reproduction and teratogenicity. *Biometrics*, 31(4), 949-952.http://dx.doi.org/10.2307/2529820.